BOOK OF ABSTRACTS

52nd Annual Meeting of the Societas Linguistica Europaea
21st – 24th August 2019
Leipzig University, Germany

Elisabeth Stark (Zurich), Sabine Stoll (Zurich), Cristina Suarez-Gomez (Balearic Islands), Inesa Šeškauskienė (Vilnius), Catherine Travis (Canberra), Beata Trawinski IDS Mannheim), Jarmila Tárnyiková (Olomouc), Johan van der Auwera (Antwerp), Guido Vanden Wyngaerd (Leuven), Arie Verhagen (Leiden), Anna Verschik (Tallinn), Jean-Christophe Verstraete (Leuven), Letizia Vezzosi (Florence), Nigel Vincent (Manchester), Jacqueline Visconti (Genoa), Søren Wichmann (MPI-EVA), Björn Wiemer (Mainz), Jacek Witkos (Poznan), Alena Witzlack-Makarevich (Kiel), Magdalena Wrembel (Poznan), Il-Il Yatziv-Malibert (INALCO), Sarka Zikanova (Prague).

SLE 2019 BOOK OF ABSTRACTS EDITOR Olga Spevak
Editor's note: Abstracts that have not been updated in due time (especially, abstracts in pdf) have not been included as well as abstracts that would damage the whole document.

TABLE OF CONTENTS

PLENARY LECTURES 19
Mira Ariel 20
Martine Robbeets 22
Stefano De Pascale 24
Karolina Grzech 26
GENERAL SESSION 28
Tanja Ackermann & Horst Simon 29
Evangelia Adamou & Yair Haendler 30
Susana Afonso & Augusto Soares da Silva 31
Wiebke Ahlers & Axel Bohmann 33
Valentina Alfarano 34
Henning Andersen 35
Lynn Anthonissen 35
Borja Ariztimuño-Lopez, Eneko Zuloaga & Dorota Krajewska 37
Sabine Arndt-Lappe & Lina Baldus 38
Sandra Auderset, Christian D. Brendel, Eric W. Campbell, Simon L. Peters, Kevin Schaefer & Albert Ventayol-Boada 39
Barend Beekhuizen 41
Kate Bellamy & Jesse Wichers Schreur 41
Thomas Belligh 43
Daria Belova & Vadim Dyachkov 45
Svetlana Berikashvili 47
Erica Biagetti & Silvia Luraghi 48
Balthasar Bickel, John Mansfield & Sabine Stoll 49
Gabriela Bîlbîie, Israel de la Fuente & Anne Abeillé 50
Barthe Bloom 52
Eva-Marie Bloom Ström 54
Charlotte Bourgoin, Gerard O'Grady, Kristin Davidse, Karen Lahousse 55
Silke Brandt, Stephanie Hargreaves & Anna Theakston 57
Jorina Brysbaert & Karen Lahousse 58
Sara Budts 60

Anouk Buyle 61
Edmond Cane 62
Giulia Cappelli, Pier Marco Bertinetto & Alessandro Lenci 64
Gerd Carling and Briana Van Epps 65
Michela Cennamo, Francesco Maria Ciconte and Luigi Andriani 65
Bożena Cetnarowska 67
Doriana Cimmino 69
Eleanor Coghill 71
Greville G. Corbett & Sebastian Fedden 71
Sonia Cristofaro 72
Carlo Dalle Ceste 72
Robert Daugs 74
Rik De Busser 75
Isabeau De Smet 76
Hendrik De Smet 78
Hilde De Vaere, Ludovic De Cuypere & Klaas Willems 79
Volodymyr Dekalo 81
Anna Maria Di Sciullo 82
Pedro Díaz Lammertyn & María Sol Sansiñena 84
Quý Ngọc Thị Đoàn 86
Laura J. Downing 87
Bridget Drinka 89
Katarzyna Dziubalska-Kołaczyk & Grzegorz Aperliński 89
Luis Eguren 90
Yoshio Endo 92
David Erschler 94
Imola-Ágnes Farkas 95
Lidia Federica Mazzitelli 97
Gertraud Fenk-Oczlon 98
Ines Fiedler & Tom Güldemann 100
Daniela Kolbe-Hanna & Natalia Filatkina 101
Markku Filppula 103
Guro Fløgstad 103
Michael Franjieh, Greville Corbett and Alexandra Grandison 103
Paola Fritz-Huechante & Antonio Machicao y Priemer 103
Livio Gaeta 105

Dorota Gaskins 106
Jerzy Gaszewski 107
David Paul Gerards 108
Alessandra Giorgi 110
Eitan Grossman, Dmitry Nikolaev & Steven Moran 111
Lutz Gunkel & Jutta M. Hartmann 113
Daniel Gutzmann & Katharina Turgay 114
Eva Hajičová, Jiří Mírovský, Kateřina Rysová and Magdaléna Rysová 116
Camiel Hamans 118
Bente Hannisdal 119
Katja Hannß 120
Maj-Britt Mosegaard Hansen 122
Paul Heggarty & Cormac Anderson 123
Melina Heinrichs, Marcel Schlechtweg & Marcel Linnenkohl 125
Lars Hellan 126
Borja Herce 128
Annette Herkenrath 129
Karina High 129
Nicole Hober 131
Hans Henrich Hock 133
Nataliia Hübler 133
Gianina Iordăchioaia 134
Britta Irslinger 136
Gerhard Jäger & Ramon Ferrer-i-Cancho 138
Katarzyna Janic & Ilja Seržant 139
Jessica Ivani 141
Hasmik Jivanyan 142
Morgane Jourdain & Karen Lahousse 142
Johannes Kabatek 144
Nataliya Kalanova 145
Anu Kalda & Mari Uusküla 146
Pavlina Kalm 148
Andra Kalnača & Ilze Lokmane 150
Egor Kashkin 152
Éva Katona 154
Nicole Katzir 155

Dan Ke ... 157
Gerrit Kentner ... 158
Peter Alexander Kerkhof, Roland Pooth, Leonid Kulikov, Laura Bruno and Jóhanna Barðdal ... 159
Václava Kettnerová & Markéta Lopatková ... 161
Liudmila Khokhlova & Boris Zakharyin ... 162
Don Killian ... 164
Anna Kisiel ... 165
Katalin É. Kiss ... 167
Linda Konnerth ... 168
Martin Konvička ... 171
Olga Krasnoukhova & Johan van der Auwera ... 171
Kateryna Krykoniuk ... 172
Nóra Kugler ... 174
Yueh Hsin Kuo ... 175
Cristina Lastres-López ... 176
Peter Lauwers and Niek Van Wettere ... 178
Sirkku Lesonen, Rasmus Steinkrauss, Minna Suni & Marjolijn Verspoor ... 179
Natalia Levshina ... 180
Barbara Lewandowska-Tomaszczyk ... 183
Silvia Luraghi ... 183
Timur Maisak ... 184
Stela Manova, Dmitri Sitchinava & Maria Shvedova ... 186
John Mansfield ... 187
Marie-Anne Markey ... 189
Antonietta Marra ... 191
Florian Matter & Rik van Gijn ... 192
Theresa Matzinger, Magdalena Schwarz & Nikolaus Ritt ... 194
Jekaterina Mažara & Sabine Stoll ... 195
Giulia Mazzola, Malte Rosemeyer & Bert Cornillie ... 197
Imke Mendoza ... 198
Thomas Menzel ... 199
Maria Messerschmidt ... 200
Susanne Maria Michaelis ... 201
Vítor Míguez ... 204
Liljana Mitkovska ... 205
Pierre-Yves Modicom ... 206

Cameron Morin 208
George Moroz 209
Céline Mounole & Denis Creissels 210
Nikita Muravyev 212
Maryam Nourzaei 213
Silva Nurmio 214
Jesús Olguín Martínez, Bernard Comrie & Eric W. Campbell 216
Nele Ots & Piia Taremaa 217
Nadine Otting 219
Manuel Padilla Cruz 221
Łukasz Pakuła & Joanna Chojnicka 223
Dafne Palú & Olaf Mikkelsen 224
Anastasia Panova 226
Ludger Paschen & Paul Heggarty 228
Doris L. Payne 230
Thomas E. Payne & Voltaire Q. Oyzon 231
Olga Pekelis 232
Yana Penkova 234
Sofía Pérez Fernández, Pedro Gras & Frank Brisard 235
Peter Petre 236
Jozefien Piersoul, Freek Van de Velde & Karlien Franco 238
Cecilia Poletto & Alessandra Giorgi 239
Marie-Caroline Pons 241
Roland A. Pooth 243
Pekka Posio & Max Wahlström 244
Monika Rind-Pawlowski 246
Maja Robbers & Harald Hammarström 247
Paula Rodríguez-Abruñeiras 249
Laura Rosseel, Jacoba Waumans & Dirk Geeraerts 251
Melanie Röthlisberger & Laura Rosseel 252
Sebastian Sauppe, Kamal K. Choudhary, Nathalie Giroud, Damián E. Blasi, Shikha Bhattamishra, Mahima Gulati, Aitor M. Egurtzegi, Ina Bornkessel-Schlesewsky, Martin Meyer & Balthasar Bickel 254
Anju Saxena, Bernard Comrie, Lars Borin & Padam Sagar 255
Marcel Schlechtweg 257

Frank Seifart, Stefan Schnell, Maria Bardají i Farré, Christian Döhler, Sonja Gipper, Katja Hannß, Katharina Haude, Birgit Hellwig, Anna Margetts, Sonja Riesberg, Harriet Sheppard, Claudia Wegener & Nikolaus P. Himmelmann ... 258
Magda Ševčíková ... 260
Christopher Shank, Anouschka Foltz & James Bragg ... 262
Marlieke Shaw ... 263
Maria Sidorova ... 265
Ivan Šimko and Teodora Vuković ... 267
Ruby Sleeman ... 267
Augusto Soares da Silva, Susana Afonso & Dafne Palú ... 269
Anna Socka & Björn Wiemer ... 271
Elena Sokur & Johanna Nichols ... 273
Michael Spagnol ... 274
William Standing ... 275
Przemysław Staniewski & Adam Gołębiowski ... 276
Thomas Stolz & Nataliya Levkovych ... 278
Adam J.R. Tallman and Sandra Auderset ... 280
Jarmila Tárnyiková ... 280
Yakov Testelets ... 282
Jacob Thaisen ... 282
Mimi Tian ... 282
Daniel Tiskin ... 284
Svetlana Toldova ... 285
Marili Tomingas ... 287
Antonios Tsiknakis ... 288
Foluke O. Unuabonah & Rotimi O. Oladipupo ... 290
An Van Linden & Lieselotte Brems ... 291
Annemarie Verkerk, Hannah Haynie, Russell Gray & Hedvig Skirgård ... 293
Åke Viberg ... 294
Rebecca Voll ... 296
Teodora Vukovic ... 298
Albert Wall ... 298
Xu Xu ... 299
Jingting Ye ... 301
Karin Zurbuchen & Mari Uusküla ... 303
POSTER PRESENTATIONS ... 305

Kirill Aksenov and Anna Klezovich 306
Heini Arjava 306
Adina Camelia Bleotu 307
Samuel Bourgeois 309
Bartosz Brzoza 310
Maria Vittoria D'Onghia 312
Vadim Dyachkov 313
Irantzu Epelde and Oroitz Jauregi 314
Riku Erkkilä & Niko Partanen 315
Laura García-Castro 316
Wojciech Guz 318
Helle Metslang & Külli Habicht 319
Charlotte Hemmings 320
Sampsa Holopainen, Veronika Milanova, Jeremy Bradley & András Czentnár 322
Victoria V. Kazakovskaya, Reili Argus & Sigal Uziel-Karl 324
Alexandra Kellner & Niko Partanen 325
Zaira Khalilova 326
Anastasia Kharlamova 328
Irina Khomchenkova 329
Ji-Eun Kim 331
Dorota Krajewska 332
Pei-Jung Kuo 333
Andra Kütt 334
Nikita Kuzin 335
Marius Zemp 337
Laura Migliori 339
Niko Partanen and Janne Saarikivi 341
Hanna Pook 341
Julia Sinitsyna 342
Anna Sjöberg & Anju Saxena 344
Natalia Stoynova 345
Yoshihisa Taguchi 346
Annika Tjuka 348
Maria D. Voeikova & Victoria Kazakovskaya 349
Judith Voß 350
Jelena Živojinović 352

WORKSHOP PRESENTATIONS 354
A comprehensive perspective on reflexive constructions 355
Alena Witzlack-Makarevich, Amos Atuhairwe and Erika Just 356
Maria Khachaturyan & Anna Volkova 357
Anna Volkova 358
Denis Creissels 359
Sebastian Dom & Leonid Kulikov 361
Lengson Ngwasi 363
Guglielmo Inglese 364
Verónica Orqueda 366
Anne C. Wolfsgruber & Iker Salaberri 367
Martin Everaert & Kārumūri V. Subbārāo 369
Zygmunt Frajzyngier 370
Bambang Kartono, Sally Wong, Eric Reuland & Martin Everaert 372
Martin Haspelmath 373
Steve Pepper 378
Pilar Valenzuela, Roberto Zariquiey and Gabriela Tello 380
John B. Lowe, Alex Francois, Martine Mazaudon & Charles Zhang 380
Thanasis Georgakopoulos & Stéphane Polis 381
Gereon A. Kaiping 382
Guillaume Segerer 383
Cara Evans, Hannah Haynie, Hedvig Skirgård and Simon Greenhill 384
Thiago Costa Chacon, Tiago Tresoldi & Johann-Mattis List 385
Abbie Hantgan-Sonko & Hiba Babiker 386
Johannes Dellert 387
Chundra Cathcart and Damián Blasi 389
Søren Wichmann 389
Tiago Tresoldi 390
Discourse particles and epistemic stance 392
Seppo Kittilä & Chingduang Yurayong 394
Svetlana Orlova, Svetlana Toldova & Maria Brykina 395
Camille Simon 397
Dorothee Beermann & Abigail Agyewaa 397
Anna Inbar & Leon Shor 398
Katsunobu Izutsu & Mitsuko Narita Izutsu 400
M.M.Jocelyne Fernandez-Vest 401

Tatiana Nikitina ... 403
Paul A. Wilson and Barbara Lewandowska-Tomaszczyk ... 404
Gerson Klumpp, Denys Teptiuk & Eda-Riin Leego ... 406
Metin Bagriacik ... 407
Vladimir Panov ... 409
Aicha Belkadi ... 410
Elevation as a deictic category. Typological and diachronic perspectives ... 412
Antoinette Schapper & Gilles Authier ... 412
Gilles Authier ... 415
Merab Chukhua and Alexander Rostovtsev-Popiel ... 416
Martha Ratliff ... 418
Yvonne Treis ... 419
Willem Adelaar ... 420
Gary Holton ... 420
Laura Arnold ... 420
Zarina Molochieva ... 422
'External' agreement with unexpected targets ... 424
Steven Kaye ... 425
Aigul Zakirova ... 427
Samira Verhees ... 429
Tania Paciaroni ... 430
Marina Chumakina and Olesya Khanina ... 430
Nina Sumbatova ... 431
Alexander Rostovtsev-Popiel ... 431
Extravagant Morphology ... 434
Luise Kempf & Stefan Hartmann ... 437
Graeme Trousdale ... 439
Anke Lensch ... 440
Christine Günther, Sven Kotowski & Ingo Plag ... 441
Leah S Bauke ... 442
Kristel Van Goethem & Muriel Norde ... 444
Ulrike Schneider ... 445
Extravagant morphology: attaching -ly to present participles ... 446
Kristin Killie ... 446
Cristiano Broccias & Enrico Torre ... 446
Maximilian Frankowsky ... 448

Barbara Schlücker ... 451
Francesca Masini & Simone Mattiola ... 452
Gestures, facial expressions and body posture in CxG accounts of language: ... 454
Core or periphery? ... 454
Alexander Bergs ... 456
Peter Uhrig ... 458
Hui-Chieh Hsu, Geert Brône & Kurt Feyaerts ... 458
Masaru Kanetani ... 459
Alan Cienki & Agni-Veatriki Stanopoulou ... 461
Elisabeth Zima ... 462
Karin Birkner & Alexandra Groß ... 463
Steven Schoonjans & Beatrix Schönherr ... 464
Ideophones and interjections ... 466
Steffen Höder ... 468
Paulette Roulon-Doko ... 469
Amy Pei-jung Lee ... 469
Martine Vanhove & Guillaume Segerer ... 470
Chiara Truppi & Patrícia Costa ... 472
Derived ideophones in Amharic and beyond ... 474
Ronny Meyer ... 474
Quint Nicolas & Noël Bernard Biagui ... 474
Françoise Rose, Mao-Line Fillon, Jennifer Krzonowski ... 475
Knowing in interaction: ... 478
Fieldwork on epistemicity and intersubjectivity ... 478
Eva Schultze-Berndt ... 480
Sonja Gipper ... 482
Erika Sandman & Seppo Kittilä ... 483
Dominique Knuchel ... 484
Matt Coler, Nicholas Q. Emlen & Edwin Banegas-Flores ... 485
Berend Hoff, Spike Gildea & Racquel-Maria Sapién ... 486
Henrik Bergqvist ... 487
References ... 487
Ditte Boeg Thomsen, Birsu Kandemirci & Silke Brandt ... 488
Marianne Mithun ... 489
Language as Network: How to conceive of a usage-based constructicon ... 491
Holger Diessel ... 492

Ewa Dąbrowska ... 493
Lars Konieczny ... 493
Maciej Borowski, Adnane Ez-zizi, Petar Milin & Dagmar Divjak ... 495
Paul Ibbotson ... 496
Anna Theakston, Heather Lemen, Kimberley Bell, Silke Brandt & Elena Lieven ... 497
Tobias Ungerer ... 498
Florent Perek and Amanda Patten ... 500
Muriel Norde & Sarah Sippach ... 500
Learners and variationist theory: New speakers and new sources of variation ... 502
Itxaso Rodríguez-Ordóñez ... 505
Jessica Kantarovich ... 507
Eider Etxebarria-Zuluaga, Itxaso Rodríguez-Ordóñez and Lorena Sainzmaza-Lecanda ... 509
Hanna Lantto ... 510
Curt Woolhiser ... 511
Noel Ó Murchadha ... 512
Managing information structure in spoken and sign languages: Formal properties and natural discourse organization ... 513
Chiara Calderone ... 516
Vadim Kimmelman ... 519
Alexandra Navarrete-González ... 520
Evgeniia Khristoforova ... 522
Shlomo Izre'el ... 523
Vydrina Alexandra ... 525
Pavel Ozerov ... 526
Nikolay Korotaev ... 528
Vera Podlesskaya ... 530
Natalia Slioussar, Natalia Tyshkevich & Sophia Popova ... 532
Massimo Moneglia & Giorgina Cantalini ... 533
Alessandro Panunzi & Valentina Saccone ... 535
Tommaso Raso, Frederico Cavalcante & Plínio Barbosa ... 536
Paradigms regained: Theoretical and empirical arguments for the reassessment of the notion of paradigm ... 540
Jenny Audring & Ray Jackendoff ... 543
Lars Heltoft ... 544
Kirsten Jeppesen Kragh & Lene Schøsler ... 546
Jaakko Leino ... 547

Peter Juul Nielsen ... 548
Dylan Glynn & Piotr Wyrośląk ... 549
Tabea Reiner ... 550
Wim Caers & Jan Nuyts ... 552
Bjarne Simmelkjær Sandgaard Hansen ... 552
Anna Kurek ... 553
Björn Wiemer, Marek Łaziński, Joanna Kwiatkowska, Ilja Egorov, Katarzyna Osior-Szot ... 554
Jakob Neels & Stefan Hartmann ... 557
Patterns of language contact within and across phylogenies ... 561
Michael Frotscher ... 563
Danny Law ... 563
Marc Tang & Michael Dunn ... 564
Christian Zimmer ... 565
Sandro Bachmann ... 567
Hermann Sonntag ... 567
Henrik Liljegren ... 568
Fedor Rozhanskiy & Elena Markus ... 570
Evgeniya Budennaya ... 571
Iker Salaberri ... 573
Roberto Zariquiey & Gabriela Tello ... 574
Phonological (in)stability and language evolution ... 575
Patrick Honeybone ... 577
Cormac Anderson ... 579
Dmitry Nikolaev ... 580
Shelece Easterday & Joan Bybee ... 581
Christian D Brendel ... 583
Leah Vandeveer ... 583
Erich R. Round ... 584
Karolina Broś ... 586
Pragmatic markers and clause peripheries ... 588
Liesbeth Degand & Ludivine Crible ... 591
Diana M. Lewis ... 592
Yinchun Bai ... 594
Nicole Nau ... 595
Vittorio Tantucci & Daniel Van Olmen ... 597
Anna Ruskan & Marta Carretero ... 598

Ton van der Wouden & Ad Foolen 600
Kaja Borthen & Elena Karagjosova 602
Ilse Depraetere & Gunther Kaltenböck 603
Heike Wiese & Oliver Bunk 604
Karin Aijmer 606
Erika Jasionytė-Mikučionienė 607
Predicative possession in a cross-linguistic perspective 609
Joanna Błaszczak 611
Jacek Witkoś 613
Piotr Cegłowski 615
Egor Tsedryk 617
Iliyana Krapova 618
Masoud Mohammadirad 620
Maria Vilkuna 622
Gréte Dalmi 624
Nikolett F. Gulyás 626
Prehistoric languages beyond the limits of comparative linguistics 628
Ljiljana Progovac & Antonio Benítez-Burraco 629
David Gil 631
Taraka Rama & Søren Wichmann 632
Michael Pleyer 634
Dan Dediu 635
Caleb Everett 637
Piotr Sobotka 638
Christine Schreyer 639
Psycholinguistic aspects of complex words 641
Pavol Štekauer, Lívia Körtvélyessy & Pavol Kačmar 645
Thomas L. Spalding & Christina L. Gagné 646
Gary Libben 647
Robert Fiorentino 648
Alexandra Bagasheva 649
Vincent Renner 651
Pierre J.L. Arnaud & Steve Pepper 652
Elizaveta Tarasova and Natalia Beliaeva 654
Simona Amenta & Marco Marelli 655
Melanie J. Bell & Martin Schäfer 657

Jesús Fernández-Domínguez 658
(Semi-)automatic retrieval of data from historical corpora: 660
Chances and challenges 660
Nikolaos Lavidas 662
Michael Percillier 664
Lara Verheyen, Sara Budts, Peter Petre and William Standing 666
Eva Zehentner, Marianne Hundt, Melanie Röthlisberger and Gerold Schneider 667
Tom S Juzek, Stefan Fisher, Pauline Krielke, Stefania Degaetano-Ortlieb & Elke Teich 668
Lauren Fonteyn 669
Sophia Voigtmann, Richard Gerbracht, Dietrich Klakow, Augustin Speyer 671
Tense, aspect and mood categories across languages 673
Evgeniya Klyagina & Peter Arkadiev 675
Åshild Næss 677
Jérémy Pasquereau & Patricia Cabredo Hofherr 678
Patrick Caudal & Eva Schultze-Berndt 680
Beatrice Nyampong, Anna Struck & Dorothee Beermann 681
Östen Dahl 683
Debra Ziegeler 684
Dmitry Gerasimov 685
Lisa Matthewson, Michael Schwan and Neda Todorović 685
Edyta Jurkiewicz-Rohrbacher 687
Miina Norvik 688
Eva van Lier, Jenny Audring & Sterre Leufkens 689
Usage-based approaches to language contacts 691
Philipp Wasserscheidt 693
Maria Frick, Dorota Gaskins, Elina Palola & Antje Endesfelder Quick 694
Geidi Kilp 695
Antje Endesfelder Quick, Stefan Hartmann Ad Backus, and Elena Lieven 696
Eeva Sippola & Hanna Lantto 697
Eline Zenner & Dorien Van De Mieroop 698
Savithry Namboodiripad 699
Nikolay Hakimov and Michael Riessler 701
V2 effects in the world's languages 702
Charlotte Galves 705
Richard Ingham 705
Pierre Larrivée 707

Anders Holmberg, Heete Sahkai & Anne Tamm 708
Aritz Irurtzun 708
Andreas Trotzke 708
Arantzazu Elordieta 710
Mélanie Jouitteau 712
Marju Kaps 713

PLENARY LECTURES

Rescuing meaning from logic

Mira Ariel
(Tel Aviv University)

Schedule: We 9.30-10.25

Key words: Gricean pragmatics; scalar implicatures; converging empirical evidence; usage-based linguistics

Grice (1989) proposed an ingenious solution to what he, and Frege before him, considered an embarrassment in natural language. Language doesn't seem to follow classical logic. For example, unlike their logical counterparts, *most* is not understood as 'more than half, and possibly all'; and *X or Y* constructions do not commit the speaker to 'possibly X, possibly Y, and possibly both X and Y'. But they should! In order to save the honor of natural language, Grice proposed that although the actual interpretations of these expressions are "nonlogical", their core semantic meanings are in fact logical. The gap between their meanings and their interpretations, as he imagined them, is bridged by (scalar) pragmatic implicatures (Horn, 1972 and onwards). Young children, presumed to be lacking full pragmatic skills, have been argued to be even more logical than adults (Noveck, 2001). In contrast, Gigerenzer (2007) argues forcefully that logic is not the yardstick by which we routinely cognize. I side with Gigerenzer, and I argue that neither language use nor natural language semantics is logic-based.

Adopting a usage-based approach to language, I present my recent corpus studies (mainly of the Santa Barbara Corpus of Spoken American English Du Bois et al., 2000-2005), as well as questionnaires and experiments with adults and with children, to argue that:

(1) Even the semantics of *most* is not logical in the classical sense; rather, *most* is semantically circumbounded (lower and upper-bounded) (Ariel, 2015);
(2) The actual use of *or* is neither 'inclusive' (logical) nor classically 'exclusive', as is commonly assumed. It has a variety of other uses, the most common one being the 'Higher-level category' use (Ariel and Mauri, 2018, Ariel et al., 2019);
(3) The core meaning of *or* is not even truth-functional – it only encodes an 'alternativity' relation between the listed options (Ariel and Mauri, in press);
(4) Children do not interpret *or* "logically", once the experimental tasks are re-designed to be more ecologically suitable, and more importantly, when the target *or* constructions match child-directed speech (Tal et al., 2019).

These findings are what we should predict based on usage-based theories of language. They all point to the conclusion that natural language semantics and pragmatics is not "spontaneously logical" (Chierchia, 2013), but proudly psycho-logical.

References

Ariel, Mira. 2015. Doubling up: Two upper bounds for scalars. *Linguistics* 53:561-610.
Ariel, Mira and Caterina Mauri. 2018. Why use *or*? *Linguistics* 56:939-994.
in press. An 'alternative' core for *or*. *Journal of pragmatics*.
Ariel, Mira, Inbal Arnon, Daniel Asherov, Alon Fishman and Nicole Katzir. 2019. Truth, *or,* interpretation. Tel Aviv University and the Hebrew University.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free Choice, and intervention*. Oxford: Oxford University Press.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson and Nii Martey. 2000-2005. Santa Barbara corpus of spoken American English, Parts 1-4: Philadelphia: Linguistic Data Consortium.

Gigerenzer, Gerd. 2007. *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.

Grice, H. Paul. 1989. *Studies in the way of words*. Cambridge, Mass.: Harvard University Press.

Horn, Laurence R. 1972. On the semantic properties of the logical operators in English. Mimeo, Indiana University Linguistics Club.

Noveck, Ira A. 2001. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition* 78:165–188.

Tal, Shira, Inbal Arnon and Mira Ariel. 2019. The acquisition of Or: a usage-based proposal. Hebrew University and Tel Aviv University.

The Transeurasian family: Wörter und Sachen und Gene

Martine Robbeets
(MPI-SHH Jena)

Schedule: Fri 10.00-10.55

The genealogical relationship of the Japonic, Koreanic, Tungusic, Mongolic and Turkic languages is among the most disputed issues of historical comparative linguistics (Poppe 1960, Miller 1971, Doerfer 1963-1975, Starostin et al. 2003, Robbeets 2005, 2015, Vovin 2005, Georg 2007, Dybo 2016). In my presentation, I will first examine whether these so-called "Transeurasian" languages are relatable within the limits of the classical historical comparative method. This will lead to a broader question about the identity and origin of the speakers: Who were the ancestral speakers of the Transeurasian languages, where did they come from and why did they move?

In order to answer these questions, I will present a bird's-eye view on the phonological, lexical, morphological and structural evidence in support of Transeurasian linguistic affinity and address some of the recent criticisms. Next, I will map the linguistic findings on inferences from genetics and archaeology, incorporating certain principles of the traditional *Wörter und Sachen* approach (Shuchardt 1912) into modern historical comparative linguistics and integrating recent advances made in the analysis of ancient DNA (Jeong, C. et al. 2019; Jeong, Ning & Wang (in press))

Whereas linguistic data underlying this presentation will be based on my previous research into Transeurasian relatedness (Robbeets 2005, 2015, 2018), the interdisciplinary data will be drawn from our current ERC-funded eurasia3angle project, in which we investigate the extent to which agriculture impacted the spread of the speakers of Transeurasian languages.

By way of result, I will suggest that it is possible to isolate a borrowing-proof core of similarities between the Transeurasian languages, for which inheritance from a common ancestor is the most parsimoneous explanation. From an interdisciplinary perspective, I will trace these common linguistic origins back to the beginnings of millet agriculture and the early days of the Amur gene pool in in North East China.

References

Dybo, A. 2016. New trends in European studies on the Altaic problem. *Journal of Language Relationship* 14(2): 71—106.

Doerfer, G. 1963–1975. *Türkische und Mongolische Elemente im Neupersischen, unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit, vols. 1-4*. Wiesbaden: Franz Steiner.

Georg, S. 2007. Review of Robbeets, Martine 2005. Is Japanese related to Korean, Tungusic, Mongolic and Turkic? *Korean Studies*, 247-278.

Jeong, C. et al. 2019. Characterizing the genetic history of admixture across inner Eurasia. *Nature Ecology and Evolution*, 1-13. doi:10.1038/s41559-019-0878-2.

Jeong, C, Ning C. and Wang, C. (in press) Transeurasian unity from a population genetic perspective. In: Robbeets M., Savelyev S. and Hübler, N. (eds.) The Oxford Guide to the Transeurasian languages. Oxford: University Press.

Miller, R. A. 1971. *Japanese and the other Altaic languages*. Chicago: The University of Chicago Press.

Robbeets, M. 2005. *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* (Turcologica 64.) Wiesbaden: Harrassowitz.

Robbeets, M. 2015. *Diachrony of verb morphology: Japanese and the Transeurasian languages*. (Trends in Linguistics Studies and Monographs 291.) Berlin: Mouton-De Gruyter.

Robbeets, M. & Bouckaert, R. (2018). Bayesian phylolinguistics reveals the internal structure of the Transeurasian family. *Journal of Linguistic Evolution* 3: 145-162.

Poppe, N. 1960. *Vergleichende Grammatik der altaischen Sprachen. Teil 1, Vergleichende Lautlehre*. (Porta Linguarum Orientalium, Neue Serie, 4.) Wiesbaden: Otto Harrassowitz.

Shuchardt, H.1912. Sachen und Wörter. *Anthropos* 7: 827-839.

Vovin, A. 2005. The end of the Altaic controversy. *Central Asiatic Journal* 49: 71-132.

Token-based vector semantics for lexical sociolectometry

Stefano De Pascale
(KU Leuven)

Schedule: Sa 9.30-9.55

Lexical sociolectometry, as introduced by Geeraerts, Grondelaers & Speelman (1999; 2003) is the aggregate-level study of lexical distances between linguistic varieties. As the distances are measured on the basis of the relative usage frequencies of synonyms, methods for semi-automated sense disambiguation of corpus materials should be helpful for scaling up the approach. This talk, then, will demonstrate the relevance of token-based vector semantics for large-scale lexical sociolectometry.

In the past decade, empirical studies of meaning have experienced an enormous leap forward thanks to advances made in computational distributional semantic analysis, i.e. a corpus-driven method for identifying meaning by encoding the textual contexts in which expressions appear as numeric vectors. Type-based vector space models have progressively been embraced in different domains of linguistic research (Dubossarsky et al., 2016; Perek, 2016; Pijpops et al., 2018), and have proven useful for the retrieval of near-synonyms in large corpora, which were then used as variants of lexical sociolinguistic variables and as input for sociolectometric calculations (Ruette et al., 2014). However, a limitation of type-based models is that all senses of a word are lumped together into one vector representation, making it harder to control for polysemy and subtle contextual distinctions.

Token-based vector space models, which have only recently received attention from the linguistic community (Heylen et al. 2015, Hilpert & Saavedra 2017) are meant to address precisely these shortcomings, by disambiguating different senses of lexical variants (Heylen et al., 2012). This technique is able to model the semantics of individual tokens of a word in a corpus and to represent them as token clouds in multidimensional vector space, with clusters of tokens revealing distinct senses of the word. By superimposing the token clouds of the lexical variants, one can distinguish which meanings are shared by near-synonyms and determine the 'semantic envelope of variation' of the lexical alternation. For instance, when comparing Netherlandic Dutch and Belgian Dutch, a type-based approach may detect the near-synonymy of *winst* and *return*, but only a token-based approach will reveal that the variant *return* is polysemous in Netherlandic Dutch between the readings 'profit' and 'return game', whereas in Belgian Dutch only the 'profit' sense is found. By isolating the cluster of tokens with the meaning 'profit' one can identify the near-synonymous tokens of the variants *winst* and *return* (see Figure). In the rest of the talk I will map out how token-based models can be put to fruitful use in an integrated sociolectometric workflow, by means of 18 concepts belonging to the lexical field of 'TRANSPORT'.

The fine-tuning of corpus-based sociolectometry illustrated here contributes to the scaling up of lexical variationist research, by providing methods for dealing with corpora whose size exceeds manual analysis. It is situated within a larger research project carried out at the Quantitative Lexicology and Variational Linguistics research unit of the University of Leuven that aims at a detailed understanding of token-based vector representations for lexical, semantic and variationist research: see https://www.arts.kuleuven.be/ling/qlvl/projects/current/nephological-semantics.

References

Dubossarsky, H., Weinshall, D., & Grossman, E. (2016). Verbs change more than nouns: A bottom-up computational approach to semantic change. In *Lingue e Linguaggio* (Vol. 15, pp. 7–27).

Geeraerts, D., Grondelaers, S., & Speelman, D. (1999). *Convergentie en divergentie in de Nederlandse woordenschat: een onderzoek naar kleding- en voetbaltermen*. Amsterdam: P.J. Meertens-

Instituut.

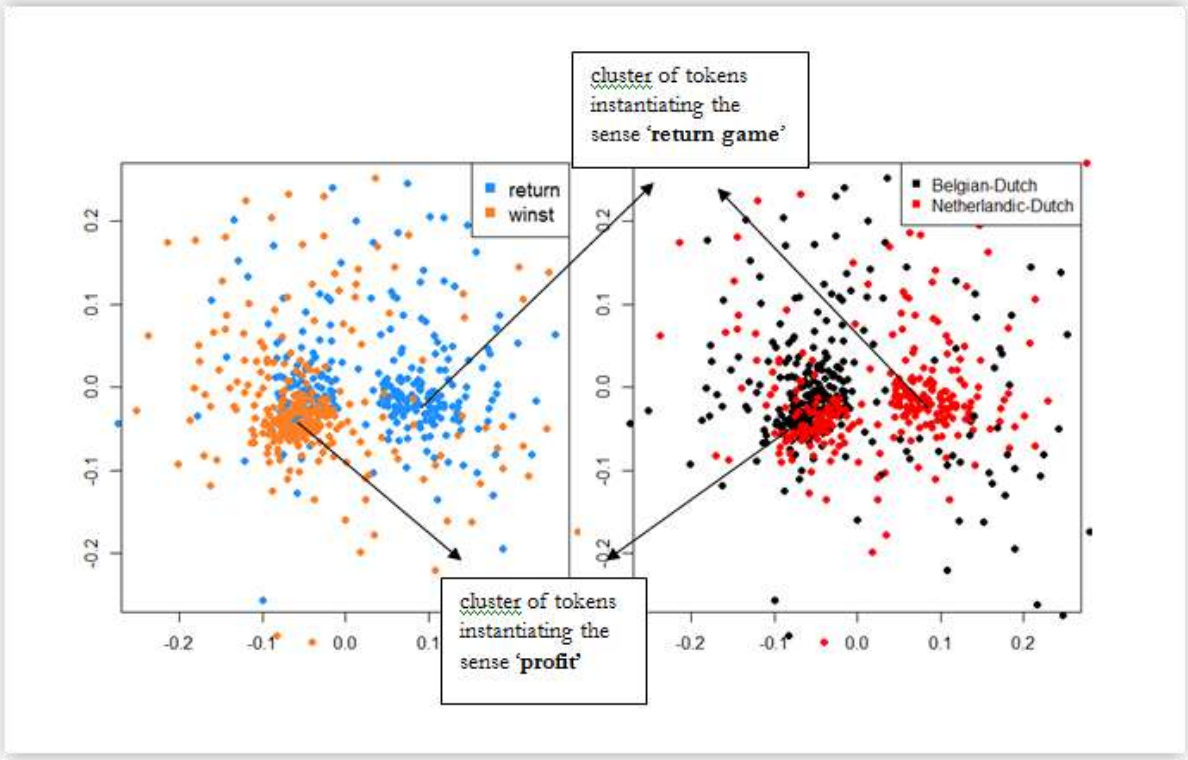
Heylen, K., Speelman, D., & Geeraerts, D. (2012). Looking at word meaning. An interactive visualization of Semantic Vector Spaces for Dutch synsets. In M. Butt, S. Carpendale, G. Penn, J. Prokic, & M. Cysouw (Eds.), *Proceedings of the EACL-2012 joint workshop of LINGVIS & UNCLH: Visualization of Language Patters and Uncovering Language History from Multilingual Resources* (pp. 16–24). Avignon, France: Association for Computational Linguistics.

Perek, F. (2016). Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics*, *54*(1), 149–188.

Pijpops, D., Speelman, D., Grondelaers, S., & Van de Velde, F. (2018). Comparing explanations for the Complexity Principle: evidence from argument realization. *Language and Cognition*, *10*(3), 514–543.

Ruette, T., Peirsman, Y., Speelman, D., & Geeraerts, D. (2014). Semantic weighting mechanisms in scalable lexical sociolectometry. In B. Szmrecsanyi & B. Wälchli (Eds.), *Aggregating Dialectology, Typology, and Register. Analysis Linguistic Variation in Text and Speech* (pp. 205–230). Berlin: de Gruyter.

Speelman, D., Grondelaers, S., & Geeraerts, D. (2003). Profile-Based Linguistic Uniformity as a Generic Method for Comparing Language Varieties. *Computers and the Humanities*, *37*(3), 317–337.



Thinking outside the evidential box: Interaction-based analysis of epistemic marking, in Quechuan and beyond

Karolina Grzech
(Stockholm University)

Schedule: Sa 10.00-10.25

Quechuan languages are widely regarded as exhibiting a grammaticalised evidential system of three 'obligatory choices of marking information source' (Aikhenvald 2004; Aikhenvald 2014; Aikhenvald 2018): direct, conjectural/ inferential and reportative. However, recent research into how Quechuan evidentials are used in discourse (e.g. Howard 2012; Grzech 2016; Hintz & Hintz 2017; Nuckolls 2018) renders a rather more complex picture. This talk contrasts the common assumptions about Quechuan evidentiality with the data from Amazonian Kichwa (Quechuan, Ecuador). I show that the 'evidential' system of the language is more accurately analysed as encoding epistemic authority (Heritage & Raymond 2005; Stivers et al. 2011) and speaker expectations. Furthermore, I show that the Amazonian Kichwa 'evidential' paradigm has its notional and functional counterparts in a number of languages across the world, and suggest how such systems are relevant to research on evidentiality and epistemicity in general.

Descriptions of Quechuan varieties show that evidential marking is not obligatory. In Huánuco (Peru), its occurs in 66% of possible contexts (Floyd 1997: 59). In South Conchuchos (Peru) - in 30% (Hintz 2007: sec. 3.5). In Amazonian Kichwa, the 'direct evidential' (*=mi*) only occurs in 6% of turns (Grzech 2016: 143). The paradigms also differ in the number of markers: certain Peruvian varieties have 5 or 6 evidentials, encoding both source of information and distinctions between individual and shared knowledge (Hintz & Hintz 2017). In Amazonian Kichwa, the 'evidentials' are altogether dissociated from the 'source of information' semantics. Consider:

(1)
uku-ma tia-k chundzulli-guna-ndi shamu-kpi, yapa**=mi** ismu-n
inside-DAT be-AG.NMLZ intestines-PL-INCL come-SWREF much**=MI** rot-3
'[if one] brings [the kill from the hunt] with the intestines, it will rot quickly'
[KICHB07AGOPEDROCHIMBO1 446]

Example (1) is consistent with its prototypical 'direct evidential'/'best possible ground' analysis of the Quechuan *=mi* (e.g. Floyd 1997; Faller 2002), since it marks information based on direct experience of the speaker. However, this analysis cannot account for (2):

(2)
chi rumi -ra payna**=mi** chura-sha, chapa-nu -shka chi -bi
D.DEM stone -ACC 3PL **=MI** put -IPFV wait -3PL-ANT D.DEM-LOC
'They have placed this stone...they've waited [having put it there]'
[el_25092014_03 048]

In this retelling of the 'Pear Story' (Chafe 1980), *=mi* occurs on a conjectural statement. The video does not show how the rock was placed on the road, and the speaker could not base the utterance on prior experience or word knowledge. Still, he used the 'direct evidential'.

In this talk, I account for the above by analysing the Amazonian Kichwa *=mi* as a marker of epistemic authority, rather than an evidential. The analysis is based on examples from a 13h corpus of

naturalistic Amazonian Kichwa discourse. I demonstrate that the Amazonian Kichwa 'evidential' system encodes meanings related to authority over knowledge, expectations of knowledge and 'Territories of Information' (Kamio 1997). Subsequently, I juxtapose the Amazonian Kichwa data with examples from other languages exhibiting epistemic markers with similar semantics, including Jaminjung/Ngaliwurru (Mirndi, Schultze-Berndt 2017) and Japanese (e.g. Hayano 2011). Finally, I suggest that the main communicative function of such systems is to facilitate 'epistemic Common Ground management' in interaction.

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: OUP

Aikhenvald, Alexandra Y. 2014. The grammar of knowledge: a cross-linguistic view of evidentials and the expression of information source. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *The Grammar of Knowledge*, 1–51. Oxford: OUP. doi:10.1093/acprof:oso/9780198701316.003.0001.

Aikhenvald, Alexandra Y. 2018. Evidentiality: The Framework. In Alexandra Y. Aikhenvald (ed.), *The Oxford Handbook of Evidentiality*, vol. 1, 1–44. Oxford: OUP. doi:10.1093/oxfordhb/9780198759515.013.1.

Chafe, Wallace L. 1980. *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, N.J: Ablex.

Faller, Martina T. 2002. *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. Stanford University. PhD Thesis.

Floyd, Rick. 1997. *La estructura categorial de los evidenciales en el quechua wanka* (Serie Lingüística Peruana). Lima: SIL International.

Grzech, Karolina. 2016. *Discourse enclitics in Tena Kichwa: A corpus-based account of information structure and epistemic meaning*. SOAS, University of London. PhD Thesis. https://eprints.soas.ac.uk/24336/.

Hayano, Kaoru. 2011. Claiming epistemic primacy: *yo*-marked assessments in Japanese. *The Morality of Knowledge in Conversation* (Studies in Interactional Sociolinguistics), 58–81. Cambridge: CUP. http://dx.doi.org/10.1017/CBO9780511921674.004.

Heritage, John & Geoffrey Raymond. 2005. The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. *Social Psychology Quarterly* 68(1). 15–38.

Hintz, Daniel J. & Diane M. Hintz. 2017. The evidential category of mutual knowledge in Quechua. *Lingua* (Essays on Evidentiality) 186–187. 88–109. doi:10.1016/j.lingua.2014.07.014.

Hintz, Diane M. 2007. *Past tense forms and their functions in South Conchucos Quechua:* Santa Barbara: University of California. https://escholarship.org/uc/item/6k19j59j (14 May, 2019).

Howard, Rosaleen. 2012. Shifting voices, shifting worlds: Evidentiality, epistemic modality and speaker perspective in Quechua oral narrative. *Pragmatics and Society* 3(2). 243–269. doi:10.1075/ps.3.2.06how.

Kamio, Akio. 1997. *Territory of Information*. Amsterdam: John Benjamins.

Nuckolls, Janis B. 2018. The Interactional and Cultural Pragmatics of Evidentiality in Pastaza Quichua. In Alexandra Y. Aikhenvald (ed.), *The Oxford Handbook of Evidentiality* (Oxford Handbooks in Linguistics), 202–221. First edition. Oxford: OUP.

Schultze-Berndt, Eva. 2017. Shared vs. Primary Epistemic Authority in Jaminjung/Ngaliwurru. *Open Linguistics* 3(1). doi:10.1515/opli-2017-0010.

Stivers, Tanya, Lorenza Mondada & Jakob Steensig. 2011. Knowledge, morality and affiliation in social interaction. In Tanya Stivers, Lorenza Mondada & Jakob Steensig (eds.), *The Morality of Knowledge in Conversation*, 3–24. Cambridge: CUP.

GENERAL SESSION

Variational pragmatics – an empirical investigation into requests in German German and Swiss German

Tanja Ackermann & Horst Simon
(Freie Universität Berlin)

Keywords: directive speech acts, pluricentricity, elicitation task, politeness

Schedule: Thu 9.30 Room 2

There is an ever-growing amount of literature on the divergent phonology, lexicon and grammar of German as spoken/written in Germany, Austria and in German-speaking Switzerland (cf., e.g., Ammon, Bickel & Lenz 2016, Dürscheid & Elspaß 2015). By contrast, the pragmatic dimension of pluricentric languages (such as German) has been studied to a much lesser extent to date (cf. Norrby & Kretzenbacher 2013: 243–244). While some studies focussing on a comparison between the German spoken in Germany and in Austria respectively do exist (cf. Muhr 2008, Warga 2008 or Norrby & Kretzenbacher 2013), contrastive pragmatic investigations that focus on the socio-cultural dimension of variation between German German and Swiss German are particularly rare (but cf. Schüpbach 2015). In the linguistic literature we find only vague hints on differences in the communicative routines. For example, Ammon, Bickel & Lenz (2016) state that some directives that are common in Germany such as *ich krieg(e) x* 'I get x' appear to be less polite to Swiss and Austrians.

In our paper we present the results from a large-scale online-survey with 585 participants, in which we investigate issues of pragmatic variation between various regions within Germany and German-speaking Switzerland. On the basis of production data as well as acceptability/adequacy judgements we will focus on differences in the expression of face-threatening speech acts. These findings will be contrasted with data compiled in five focus group interviews with Germans currently living in Zurich and Swiss participants currently living in Berlin. The interviewees' statements indicate that crucial differences regarding the perception of socio-cultural practices of language use between German German and Swiss German do exist. However, these statements are largely characterized by stereotypical prejudices. Here, the results of the online survey allow a more differentiated evaluation.

One question that was examined is whether Germans are really more direct in their verbal behaviour than Swiss as stated repeatedly in the interviews. The results of the questionnaire study give some evidence that participants from the German-speaking part of Switzerland indeed prefer more indirect strategies for requests. For instance, they dismiss explicit directives with an imperative (*Reich mir mal die Marmelade rüber!* 'Pass me the jam!') and avoid them even in communication with close friends.

Germans, by contrast, show a complex picture: while they also prefer conventionalized indirect head acts such as ability questions (*könntest du x* 'could you x') as a request strategy, they do use imperatives in directives, namely in acts with low imposition and between close interlocutors. Beside those differences regarding the distribution of request strategies on the macro level, we also find differences on the micro level (e.g. the choice of modal verbs and particles).

In sum, our results indicate that there are some important differences between Swiss and German socio-cultural practices of language use. However, these differences seem to be less fundamental than purported by frequently verbalized stereotypes.

References

Ammon, Ulrich, Hans Bickel & Alexandra N. Lenz (2016), *Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz, Deutschland, Liechtenstein, Luxemburg, Ostbelgien und Südtirol sowie Rumänien, Namibia und Mennonitensiedlungen.* 2. ed. Berlin /

Boston: de Gruyter.
Blum-Kulka, Shoshanna, Juliane House & Gabriele Kasper (1989), *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex.
Dürscheid, Christa & Stephan Elspaß (2015), Variantengrammatik des Standarddeutschen, in R. Kehrein, A. Lameli & S. Rabanus (eds), (2015), *Regionale Variation des Deutschen. Projekte und Perspektiven*, Berlin / Boston: de Gruyter, 563–584.
Muhr, Rudolf (2008), The pragmatics of a pluricentric language. A comparison between Austrian German and German German, in K. P. Schneider & A. Barron (eds), (2008), *Variational Pragmatics. A Focus on Regional Varieties in Pluricentric Languages*, Amsterdam / Philadelphia: John Benjamins, 211–244.
Norrby, Catrin & Heinz L. Kretzenbacher (2013), National variation of address in pluricentric languages: The examples of Swedish and German, in A. Soares da Silva (ed), (2013), *Pluricentricity. Language Variation and Sociocognitive Dimensions*, Berlin / Boston: Mouton de Gruyter, 243–270.
Schüpbach, Doris (2015), German or Swiss? Address and other routinised formulas in German-speaking Switzerland, in J. Hajek & Y. Slaughter (eds), (2015), *Challenging the Monolingual Mindset*, Bristol: Multilingual Matters, 63–77.
Warga, Muriel (2008), Requesting in German as a pluricentric language, in K. P. Schneider & A. Barron (eds), (2008), *Variational Pragmatics. A focus on Regional Varieties in Pluricentric Languages*, Amsterdam / Philadelphia: John Benjamins, 245–266.

When the past is near the interlocutor: Experimental evidence from Pomak

Evangelia Adamou & Yair Haendler
(French National Centre for Scientific Research (CNRS); University Paris Diderot)

Keywords: experimental; engagement; deixis; mental timeline; Pomak (Slavic)

Schedule: We 12.00 Room 15

Several experimental studies demonstrate the existence of a space-time mapping along the left-right axis following writing direction (e.g., Walker, Bergen & Núñez 2017). But what happens when the past is associated to the interlocutor's sphere in language? Evidence comes from Pomak, a Slavic dialect spoken in Greece, where the definite article for referents in the addressee's sphere (*-t-*) is also used for past reference that the speaker considers the addressee to be aware of (Adamou 2011), a value related to the notion of *engagement* (Evans, Bergqvist & San Roque 2018a,b).

Experiment 1 is a 5-point Likert-scale acceptability judgment task testing Pomak speakers' sensitivity to agreement between the temporal reference of a definite NP and the temporal reference of the predicate. Forty L1-Pomak speakers from Greece participated in this experiment (26 female; age range = 15-60, *M* = 31, *SD* = 13.49). Participants judged the acceptability of 55 auditorily-presented sentences in total. The experimental items included two conditions. In Condition 1, sentences with *t*-articles agreed with clausal tense; in Condition 2, sentences with *t*-articles combined with future clausal tense. As predicted, the sentences with *t*-articles that agreed with the past tense of the clause received significantly higher ratings than sentences with *t*-articles and future tense, confirming the importance of temporal reference.

Experiment 2 is a response time experiment adapted from Walker, Bergen & Núñez (2017). Our goal was to test whether articles in Pomak carry temporal information on their own, beyond agreement with verbal tense, and whether the linguistic association of past with the addressee's sphere through the articles affects cognitive representations. Forty participants (23 female; age range = 16-50, *M* = 30, *SD* = 8.63) listened to Pomak NPs related to the life of the experimenter (past and future). Half of the NPs had grammatical-only temporal reference (e.g., *the*$_{[t\text{-article}]}$ *friends* vs. *the*$_{[n\text{-article}]}$ *friends*), the other half had grammatical and semantic-pragmatic temporal reference. Participants responded by pressing a left key for past and a right key for future referents, or vice-versa. Each participant completed 160 experimental trials. Responses were as accurate and fast in temporally ambiguous and unambiguous items. Speed was related to the reading experiences of the participants. Additionally, the experimenter's location had a facilitation effect on the accuracy of past responses.

To explore the nature of the facilitation effect, we repeated Experiment 2 with 40 L1-Pomak participants in L2-Greek (Experiment 3). Twenty stimuli corresponded to referents in the experimenter's past (e.g., *my baby teeth*) and 20 stimuli corresponded to referents that were likely to occur in the experimenter's future life (e.g., *my wedding ring*). In this experiment, we replicated the effect found in the Pomak version for reaction times, but not the effect of the experimenter's location. This suggests that the facilitation effect is language-related and not due to general attentional processes.

Taken together these experiments reveal a cognitive representation of time that had hitherto gone unnoticed and support cognitive models that promote flexible representations of time combining cultural, linguistic, and sensorial processes (see Bylund & Athanasopoulos 2017).

Acknowledgments: This research received funding from the French National Research Agency (ANR) for the programme Empirical Foundations of Linguistics (ANR-10-LABX-0083).

References

Adamou, Evangelia (2011), Temporal uses of definite articles and demonstratives in Pomak (Slavic, Greece), *Lingua* 121(5), 871-889.

Bylund, Emanuel and Panos Athanasopoulos (2017), The Whorfian Time Warp: Representing Duration through the Language Hourglass, *Journal of Experimental Psychology: General* 146, 911-916.

Evans, Nicholas, Henrik Bergqvist, and Lila San Roque (2018a), The grammar of engagement I: Framework and initial exemplification, *Language and Cognition* 10(1), 110-140.

Evans, Nicholas R., Henrik Bergqvist, and Lila San Roque (2018b) The grammar of engagement II: Typology and diachrony, *Language and Cognition* 10(1), 141-170.

Walker, Esther, Benjamin Bergen & Rafael Núñez (2017), The spatial alignment of time: Differences in alignment of deictic and sequence time along the sagittal and lateral axes, *Acta Psychologica* 175, 13-20.

The null reflexive construction in Brazilian Portuguese

Susana Afonso & Augusto Soares da Silva
(University of Exeter & Universidade Católica Portuguesa)

Keywords: reflexive constructions, clitic deletion, constructional variation, absolute construal, Brazilian Portuguese

Schedule: Thu 11.30 Room 12

This study investigates the null reflexive construction in Brazilian Portuguese, i.e. the reflexive construction which does not exhibit a reflexive marker *se*. This research is part of a larger sociolectometrical and sociocognitive research project about *se* constructions in European Portuguese (EP) and Brazilian Portuguese (BP) aiming to measure convergence/divergence between EP and BP.

In Portuguese, the reflexive construction is a member of a network of polysemous constructions that are marked by the clitic *se* (reciprocal, middle, anticausative, passive and impersonal constructions). These constructions are posited in a transitivity continuum that range between prototypical two-participant and one-participant events. Even though reflexive constructions are closer to transitive events, co-referentiality between the participants entails a decrease in transitivity as the participants are conceptually less well distinguishable (Kemmer 1993).

In the spoken corpora of Brazilian Portuguese surveyed, the reflexive construction tends to maintain the clitic *se*, as it behaves as a real argument (Kemmer 1993), unlike other *se* constructions such as the middle and the anticausative, for example, which tend to be unmarked. Nonetheless, the reflexive constructions exhibits a counterpart without the clitic *se*. This phenomenon is usually interpreted as the result of an ongoing general tendency in BP towards the morphological loss of clitics (Galves 2001, Cyrino 2007, Carvalho 2016). We argue that the absence of the clitic offers a different conceptualisation of the reflexive event. (1) and (2) are examples of null reflexive constructions:

(1) *Largou casa, marido, que ela não Ø considera casada e veio me ver.* (BP, Fóruns)
'She left her house, husband, because she doesn't consider [herself] to be married and came to see me.'

(2) *eles Ø fechavam em casas com trinta, quarenta litros de uísque, cara, cê nũ imagina* (BP, C-Oral)
'They locked [themselves] in houses with thirty, forty litres of whisky, man, you can't imagine.'

In (1) and (2), the absence of the clitic is associated with the conceptualisation of the event with a single participant. In (1), the verb *considerar* 'to consider' is here conceptualised as a verb of state rather than a mental verb. In (2), the reflexive construction also profiles the result of the event. We hypothesise that when the reflexive event does not imply true reflexivity (i.e. events whose reflexivity are semantically unexpected as they wouldn't conventionally be performed in a reflexive fashion), the reflexive construction is more likely to be unmarked. The degree of distinguishability of the participants therefore decreases and the event is less elaborated. As a consequence, the absence of the clitic may also profile a non-energetic and *absolute* event (Langacker 1991, 2008; Maldonado 1999), the focal point being therefore the result of the event.

References

Carvalho, Janayna (2016), *A morfossintaxe do português brasileiro e sua estrutura argumental: uma investigação sobre anticausativas, médias, impessoais e a alternância agentiva*, PhD dissertation, Universidade de São Paulo.

Cyrino, Sonia (2007), Construções com SE e promoção de argumento no português brasileiro: Uma investigação diacrônica, *Revista da ABRALIN* 6(2), 85-116.

Galves, Charlotte (2001), *Ensaios sobre as gramáticas do Português*, Campinas: UNICAMP.

Kemmer, Suzanne (1993), *The middle voice*, Amsterdam: John Benjamins.

Langacker, Ronald W. (1991). *Foundations of Cognitive Grammar, Vol. 2: Descriptive Application*, Stanford: Stanford University Press.

Langacker, Ronald W. (2008), *Cognitive Grammar. A Basic Introduction*, Oxford: OUP.

Maldonado, Ricardo (1999), *A Media Voz. Problemas conceptuales del clítico SE*, México: UNAM.

"Like she says like I - like": Markers of stance in narration

Wiebke Ahlers & Axel Bohmann
(University of Osnabrück & University of Freiburg)

Keywords: stance, variation, sociolinguistics, like, pragmatics

Schedule: We 11.00 Room 2

Much of the current sociolinguistic work on stance (Holmes-Elliott/Levon 2017, Jaffe 2009, Kiesling 2009) focuses on how different linguistic forms correlate with different stances taken by interlocutors in interactions. Less attention has been paid to the question whether some linguistic innovations themselves enable new kinds of stance-taking, rather than simply being enlisted from a pool of resources.

The present study considers processes of stance-taking in a corpus of 80 participants re-narrating a children's story immediately after reading it. The speakers are residents of Austin, Texas, 18-86 years old, stratified by gender, ethnicity, and educational background. Each recording is transcribed and coded for linguistic markers of affective, epistemic and interpersonal positioning/stance (Du Bois 2007), according to a discourse-informed definition. Our study compares the deployment of *like*, both as quotative and discourse particle, with other stance-taking markers.

In a statistical analysis, we trace the development of stance-taking patterns in apparent time. We find an increase in the frequency of stance markers in the younger generation, a pattern that is driven mainly by affective and epistemic markers. Focus on individual forms shows that *like* accounts for a large proportion of this inter-generational change. Moreover, in some of the younger speakers use of *like* replaces a large number of forms attested in the older speakers. Clustering of the speaker population by stance-taking behavior reveals two groups, one of which is composed almost exclusively of young Anglo speakers, with the other being more heterogeneous. Once again, we find use of *like* to be the main distinguishing feature between these two groups.

Our analysis demonstrates how innovative linguistic features may restructure a socially meaningful speech activity over generations. We argue that *like* usage offers a more individualistic approach to storytelling, allowing speakers to shift focus from content to performance, which results in a decrease of interpersonal positioning and an increase of affective positioning over time. It is our interpretation of the data that the increasing frequency of *like* is not only implicated in, but facilitates this process.

References

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.) *Stancetaking in discourse: Subjectivity, evaluation, interaction*. John Benjamins, 139–182.

Holmes-Elliott, Sophie and Erez Levon. 2017. The substance of style: Gender, social class and interactional stance in /s/-fronting in southeast England. *Linguistics* 55(5), 1045–1072.

Jaffe, Alexandra. 2009. Introduction: The sociolinguistics of stance. in Alexandra Jaffe (ed.) *Stance: Sociolinguistic perspectives*. Oxford University Press, 3–28.

Kiesling, Scott. 2009. Style as stance: Can stance be the primary explanation for patterns of sociolinguistic variation? In Alexandra Jaffe (ed.) *Stance: sociolinguistic perspectives*. Oxford University Press, 171–194.

Indirect possessive constructions in Nalögo: type and role of possessive markers

Valentina Alfarano
(INALCO; LACITO-CNRS)

Keywords: Oceanic languages, Reefs-Santa Cruz languages, possessive constructions, possessive classifiers, typology

Schedule: Thu 16.00 Room 6

This talk aims to analyse the indirect possessive constructions in Nalögo, a virtually undescribed Oceanic language spoken on Santa Cruz Island in Solomon Islands. In Oceanic languages, possession constitutes one of the most complex domains of the grammar (Lynch *et al.* 2002: 40). In Oceanic languages, there is a basic distinction between direct and indirect possessive constructions (Lichtenberk 1983, 1985, 2009 *inter alii*). In direct constructions, the possessor is directly attached to the possessee, whereas in indirect constructions, the possessor is attached to a possessive marker. Two examples of direct and indirect possessive constructions in Nalögo are provided in (1a) and (1b), respectively.

1. a) nawë=nu
   head=1MIN.POSS
   'my head (of the possessor)'

   b) nawë ne=nu
   head GEN.POSS=1MIN.POSS
   'my head (of a fish).'

As shown in (1a), the first minimal possessive enclitic *=nu* is directly attached to the head noun, whereas in (1b), the same enclitic is attached to the general possessive marker *ne*. While the syntactic template in (1a) is typically associated to inalienable possession (*e.g.* body parts), the one in (1b) is associated to more various and fine-grained semantic relations between the two entities. The possessive markers found in indirect constructions are often defined as *possessive classifiers* in Oceanic literature in the sense that they classify the type of semantic relation between the possessor and the possessee (Lichtenberk 1983, 1985). Lichtenberk (1983) refers to this property of classifiers as *fluidity*, that is, the capability of head nouns to occur in different constructions according to the type of semantic relation encoded by the classifier. Example (1) above clearly shows this property, as the same head noun *nawë* can occur in both constructions. While in (1a), the possessee refers to the possessor's head as a body part, in (1b), it refers to the head of a fish regarded as a property of the possessor. Along with fluidity, *metaphorical extension* is another property of classifiers, *i.e.* the ability of the classifier to extend their basic semantic relation with a type of nouns to other semantically related nouns. This talk aims to analyse indirect possessive constructions in Nalögo, with a focus on the type and role of its nine possessive markers. In particular, we will see that the notion of classifier based on the properties of fluidity and metaphorical extension is not always applicable to the Nalögo possessive markers. This talk will be based on primary fieldwork data.

Abbreviations

1, 1st person; GEN, general; MIN, minimal; POSS, possessive.

References

Lichtenberk, Frantisek. 1983. Relational classifiers. *Lingua 60*: 147–176.

Lichtenberk, Frantisek. 1985. Possessive constructions in Oceanic languages and in Proto-Oceanic. In *Austronesian Linguistics at the 15th Pacific Science Congress*, ed. by Andrew Pawley and Lois Carrington, 93–140. Canberra: Pacific Linguistics.

Lichtenberk, Frantisek. 2009. Attributive possessive constructions in Oceanic. In *The expression of possession*, ed. by William B. McGregor, 249–291. Berlin and New York: Mouton de Gruyter.
Lynch, John, Malcolm Ross and Terry Crowley. 2002. *The Oceanic Languages*. Richmond, Surrey: Curzon Press.

Paradigms of paradigms

Henning Andersen
(UCLA)

The fact that paradigms of morphological exponenda are nested inside other paradigms is an old insight (though not customarily spoken of in these terms). Case paradigms are included in number paradigms: CASE] NUMBER]. Similarly, say, in the Latin verb, PERSON] NUMBER] TENSE] ASPECT] MOOD] VOICE]. Such hierarchies are corroborated in synchrony and diachrony. E.g., in Latin, deponent verbs are a morphological anomaly specified at the highest category node (active verbs with passive morphology); languages that lose CASE often retain NUMBER. Which is not to say that reanalysis cannot cut across hierarchies.

As for exponence, paradigms (selectional sets) of inflectional categories (e.g., declension, conjugation classes) are more often linear and asymmetrical with respect to one or more characteristics such as productivity or morphological type. Consider the typological gradation in the paradigm of English conjugations: (1) the productive, regular weak conjugation (*cap*, *cab*, *can*) is AGGLUTINATIVE, (2) the irregular weak conjugation (*tell*, *lose*, *build*, *teach*) is FUSIONAL, and (3) the strong conjugation (*sink*, *run*, *cling*, *drive*, *know*, *speak*) is SYMBOLIC (leaving aside the dying p.p.p. *-en* suffix). Old Church Slavonic presents a gradation across verbal categories: the aspects and tenses run the gamut from AUXILIATED (the retrospective aspect) to AGGLUTINATIVE (the imperfect, the productive aorist and present formations) to FUSIONAL (the unproductive aorist and present types) to SYMBOLIC (the suppletives).

It is no news that morphological systems are typologically diverse (Greenberg 1960). Post-Bloomfieldian descriptivists and their Word-and-Paradigm successors have promoted one-size-fits-all ’solutions’ to this ’problem’. But it appears that an approach that explicitly focuses on the typological diversity of paradigms within morphological systems holds promise of new insights, synchronic, diachronic, typological. I propose to elaborate a few examples illustrating this based on comparative Finnic and Slavic data.

References

Greenberg. 1960. A quantitative approach to the morphological typology .... *IJAL* 26.

Connecting individual and community grammars in syntactic change

Lynn Anthonissen
(University of Antwerp / LMU Munich)

Keywords: prepositional passive, Nominativus cum Infinitivo (NCI), individual variation, constructional change, Early Modern English

Schedule: Thu ~~12.00~~ 9.30 Room 14

Linguists have typically studied constructional change at the aggregate level of speech communities, yet key mechanisms of change (e.g. analogy, entrenchment, schematization) operate within the minds of individual language users; conventionalization only occurs when several such individual grammars are aligned during interaction (cf. Schmid *forthc*.). While this is often implicitly or explicitly acknowledged (e.g. Fischer 2010: 182), longitudinal studies investigating speakers' adaptiveness in ongoing change are scarce (but see, e.g., Petré and Van de Velde 2018). The present study aims to make headway in this area by zooming in on three pertinent questions: (i) Do we find evidence of constructional change across the lifespan? (ii) If so, how does change in individuals differ from intergenerational change? (iii) To what extent do related constructions co-evolve in individual usage?

In order to gain insight into the dynamics of and constraints on constructional change past adolescence, we investigate the mechanisms underlying the increasing productivity of cross-linguistically rare passives in Early Modern English (cf. Los 2009, Dreschler 2015). In particular, we focus on the prepositional passive (1), i.e. passive constructions whose subject corresponds to the logical complement of a preposition, and the NCI (Nominativus cum Infinitivo) construction (2), or passive constructions featuring a perception, cognition or utterance verb followed by a *to*-infinitival complement.

(1) [S *BE* V-*ed* (X) P] - *They were laughed at / They are cried out against*
(2) [S *BE* V-*ed* TO V] - *He is said to be a thief / He is assumed to be a thief*

Our analysis of a large set of passives (ca. 15,000 instances) drawn from the newly-established corpus of *Early Modern Multiloquent Authors* (Petré et al. 2019) provides novel insights into the three central issues defined above. As to the first issue, we show that when a construction has two different uses (e.g. the NCI's evidential vs. modalized use), their relative weight may shift significantly across the lifespan. Such trends may continue well into old age. Second, intergenerational trends are sometimes, but not always, replicated in the behaviour of individuals. Inversely, we also observe lifespan changes in some individuals that have no ramifications at the community level, e.g. idiosyncratic developments due to chunking processes. Finally, results from regression and dispersion measures indicate that the two types of passives are cognitively related, meaning changes in one construction may prime changes in the other.

References

Dreschler, Gea (2015), *Passives and the loss of verb second: A study of syntactic and information-structural factors*, Utrecht: LOT.

Fischer, Olga (2010), An analogical approach to grammaticalization, in K. Stathi, E. Gehweiler, and E. König (eds.), (2010), *Grammaticalization: Current views and issues*, Amsterdam: John Benjamins, 181–218.

Los, Bettelou (2009), The consequences of the loss of verb-second in English: Information structure and syntax in interaction, *English Language and Linguistics* 13(1), 97–125.

Petré, Peter, Lynn Anthonissen, Sara Budts, Enrique Manjavacas, Emma-Louise Silva, William Standing, and Odile A. O. Strik (2019), Early Modern Multiloquent Authors (EMMA): Designing a large-scale corpus of individuals' languages. *ICAME Journal* 43, 83–122.

Petré, Peter, and Freek Van de Velde (2018), The real-time dynamics of the individual and the community in grammaticalization. *Language* 94(4), 867–901.

Schmid, Hans-Jörg (*forthc*.), *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*, Oxford: Oxford University Press.

Against the Proto-Indo-European-Euskarian hypothesis, or why Basque continues to be a language isolate

Borja Ariztimuño-Lopez, Eneko Zuloaga & Dorota Krajewska
(University of the Basque Country, UPV/EHU)

Keywords: Basque, Proto-Indo-European, reconstruction, comparative method

Schedule: ~~Fri 17.00 Room 14~~ Thu ~~16.30~~ 10.00 Room 14

Basque linguistics has advanced significantly in the last century thanks to methods of historical linguistics, especially internal reconstruction. The reconstruction of Proto-Basque (Mitxelena 1961), or a deeper reconstruction based on root and morpheme structure (Lakarra 1995, 2016) are especially relevant. For Campbell (2011: 23), "[t]he study of the history of Basque can be an example to the study of other isolated languages".

There have been many attempts to link Basque genetically to other languages. "Classical" proposals include e.g. Caucasian, and among newer ones we find Forni (2013) linking Basque and Indo-European (cf. Gorrotxategi & Lakarra 2013 for a critique). All were rejected because of methodological and data issues (see Trask 1995, 1997, and Lakarra 2018 for summaries).

Here we analyse the latest proposal, namely the Proto-Indo-European-Euskarian hypothesis (Blevins 2018). Blevins affirms that Basque cannot descend from Proto-Indo-European, but, because of phonological similarities between Proto-Indo-European and her Proto-Basque, they can be sisters (or daughters of sister languages).

We show that Blevins's proposal cannot be accepted due to structural flaws in Basque data (we leave the Proto-Indo-European part for relevant experts).

The hypothesis suffers from the same problems which earlier attempts of linking Basque to other languages did. The following is a sample of the difficulties:

a) Loan words are mistaken for native lexicon. For example, Blevins derives *apeu*, *apego* 'decoy' from *ha-phego, but it comes from Romance (French *appeau*, Gascon *apèu*).
b) Blevins uses localized forms and relatively modern variants. For example, she states that *oixan* 'forest' < *oiso- 'ferocious, wild' + *-an* locative; *oixan* is actually a secondary variant of *oihan*.
c) Blevins employs formations transparent within Basque (often modern and documented only once) as if they were old: e.g. *hegatsu* 'winged' < *phega-so (claiming it to be related to *pegasus*). *Hegatsu* is transparent (< *hego* 'wing' and the suffix *-tsu* '-y [abundantial]'), and its first and only attestation is in Larramendi's dictionary in 1745. Moreover, the old form (found in toponyms) of the suffix *-tsu* is *-zu*.
d) Modern neologisms are used as material for reconstruction: e.g. *hirikoe* 'triglyph' < *thiri-khoi; actually from *hir(u)-ikoe ('three-line'), attested first and only in Larramendi (1745) (Urgell 2000).
e) Problems with morphology, such as arbitrary or erroneous segmentations. Blevins segments verbs such as *eduki* 'have' as *e-duki, without distinguishing the morpheme *-ki* (cf. *edeki* 'remove', *ebaki* 'cut'). Another example is *atson* 'odour, aroma' reconstructed as *ha-son, which clearly comes from *(h)ats-on 'good smell'.

In conclusion, because of these and other methodological problems the sound correspondences proposed by Blevins are often wrong, and thus cannot prove the genetic link.

References

Blevins, J. 2007, "A long lost sister of Proto-Austronesian? Proto-Ongan, mother of Jarawa and Onge of the Andaman Islands", *Oceanic Linguistics* 46:1, pp. 154-198.

Blevins, J., 2018, *Advances in Proto-Basque reconstruction with evidence for the Proto-Indo-European-Euskarian hypothesis*, New York: Routledge.

Campbell, L., 2011, "La investigación histórica de las lenguas aisladas, o ¿es raro el vasco?", in J. A. Lakarra, J. Gorrotxategi & B. Urgell (eds.), *2nd Conference of the Koldo Mitxelena Chair*, Vitoria-Gasteiz: University of the Basque Country (UPV/EHU).

Forni, G., 2013, "Evidence for Basque as an Indo-European language", *Journal of Indo-European Studies* 41:1-2, pp. 1-142.

Gorrotxategi, J. & Lakarra, J.A., 2013, "Why Basque cannot be, unfortunately, an Indo-European language", *Journal of Indo-European Studies* 41:1-2, pp. 203-237.

Lakarra, J. A., 1995, "Reconstructing the root in Pre-Proto-Basque", in J. I. Hualde, J. A. Lakarra & R. L. Trask (eds.), *Towards a History of Basque Language*, John Benjamins: Amsterdam & Philadelphia, pp. 189-206.

Lakarra, J. A., 2016, "Gramatikalizazioa, morfemen forma kanonikoak eta berreraiketa morfologikoaren bide berriak", in G. Aurrekoetxea, J. M. Makazaga & P. Salaberri (eds.), *Txipi Ormaetxea omenduz. Hire bordatxoan*, University of the Basque Country (UPV/EHU): Leioa, pp. 175-192.

Lakarra, J. A., 2018, "La prehistoria de la lengua vasca", in J. Gorrotxategi, I. Igartua & J.A. Lakarra (eds.), *Historia de la lengua vasca*, Basque Government: Vitoria-Gasteiz, pp. 23-244.

Mitxelena, K., 1961, *Fonética histórica vasca*, Gipuzkoako Foru Aldundia: Donostia/San Sebastián.

Trask, R. L., 1995, "Origins and relatives of the Basque language: Review of the evidence", in J. I. Hualde, J. A. Lakarra & R. L. Trask (eds.), *Towards a History of Basque Language*, John Benjamins: Amsterdam & Philadelphia, pp. 65-99.

Trask, R. L., 1997, *The history of Basque*, Routledge: London.

Urgell, B., 2000, *Larramendiren* Hiztegi Hirukoitz*aren osagaiez*, unpublished PhD dissertation (University of the Basque Country, UPV/EHU).

The development of subschemas in L2 Acquisition: Two case studies from German-English syntax and morphophonology

Sabine Arndt-Lappe & Lina Baldus
(Trier University)

Keywords: L2 acquisition, degrees of granularity, early generalization, subschema development

Schedule: Sat 11.00 Room 8

In usage-based approaches to language learning, it is still unclear how schema knowledge comes about. One debated issue is whether L2 learners start with item-based knowledge and gradually form more abstract schemas (a bottom-up process, as e.g. in Tomasello (2003) for L1 acquisition) or if they first form larger generalisations, refine them and then restructure the schema (a top-down process, e.g. Roehr-Brackin 2014).

To bear on this issue, we will present two empirical studies from two different areas of grammar with German L2 learners of English of different proficiency levels. The first study deals with the catenative verb construction consisting of a 'catenative verb' and a non-finite complement, which can

be a *to*-infinitival or a gerund-participial complement (e.g. *She **refused** to go, Tom **enjoys** reading books*). A sentence completion study with 1,272 German learners of English from different proficiency levels (N = 10,549) reveals large-scale overgeneralisation of the *to*-infinitival complement, but not of the *-ing* complement, especially by low-proficiency learners. In the target language, the *to*-infinitival complement schema has a higher type and token frequency than the *–ing* complement schema. In addition, relative frequency effects are confined to high-proficiency learners.

The second study looks at stress alternations in complex words ending in *-ory* (e.g. *éxploratory* vs. *explóratory* vs. *explorátory* > *explóre*). In a reading study with 30 German learners of English from different proficiency levels (N = 1,200) we find that, unlike other stress patterns, stress on the antepenultimate syllable before the suffix (e.g. *éxploratory*) is substantially overgeneralised by low-proficiency learners. In the target language, antepenultimate stress in *-ory* words is conditioned by morphological structure, as it occurs only as stress preservation from the base word (e.g. in *círculatory* > *círculate*).

Both phenomena present evidence that abstractions, in the sense of 'broader generalisations', arise early in L2 acquisition, but that relevant aspects of pertinent semantic or formal representations are less fine-grained than in the target language (cf. Ambridge et al. 2014). For instance, we find substantial evidence for overgeneralisation patterns with low-proficiency learners, which can only be explained if we assume that these learners operate on more coarse-grained schemas. Only later are schemata restructured and subschemas for the less frequent constructions are developed, taking into account more fine-grained structural properties (e.g. frequency distributions, morphological structure) that seem unavailable at lower proficiency levels.

References

Ambridge, B., J. M. Pine, C. F. Rowland, D. Freudenthal & F. Chang. 2014. Avoiding dative overgeneralisation errors: Semantics, statistics or both? *Language, Cognition and Neuroscience* 29(2). 218–243.

Roehr-Brackin, Karen. 2014. Explicit knowledge and processes from a usage-based perspective: The developmental trajectory of an instructed L2 learner. *Language Learning* 64(4). 771–808.

Tomasello, M. 2003. *Constructing a Language*: *A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Advancing the diachronic and typological study of tone with comparative lexical databases: A case study of five Mixtec varieties

Sandra Auderset, Christian D. Brendel, Eric W. Campbell, Simon L. Peters, Kevin Schaefer & Albert Ventayol-Boada
(University of California, Santa Barbara)

Keywords: database design, computational methods, diachrony, tone change, Mixtec languages

Schedule: Thu 16.30 Room 4

We present a multilingual lexical database of five fairly closely related Mixtec varieties in order to explore issues of database design focused on the representation and integration of tone. Although 60-70% of the world's languages are tonal (Yip 2002:1), tone languages are underrepresented in linguistics research in general, and especially in historical linguistics, typology, and computational approaches. As a consequence, our understanding of tone change (as opposed to segmental change)

remains poor (Campbell, in press). Computer-assisted approaches to exploring language change (e.g. List 2016) provide an opportunity to close this gap.

Mixtec languages (Otomanguean; Mexico) make extensive use of lexical and grammatical tone. Segmental sound correspondences are well established across varieties (Josserand 1983), but this is not the case for tone correspondences. Earlier work has established the latter with respect to melodies rather than individual tonemes, but often based on only a few varieties (Dürr 1987; Longacre 1957) and at a time when everything had to be done by hand. This database will allow us to explore both a toneme and a tone melody approach aided by computational methods (such as Python scripts for cognate set extraction, cf. List et. al. 2018).

Our database is designed to be both human-readable and suited to computational implementations. For every lexical entry, it contains one column each with: a) the form in practical orthography including tone diacritics, b) the form in practical orthography without tone marking, and c) the tone pattern of the form (without segmental representation) including syllable boundaries. Our workflow is set up so that columns b) and c) are automatically generated based on column a), which is entered by hand.

Translation	**SSMa**	**SSMb**	**SSMc**	**SMDa**	**SMDb**	**SMDc**
'atole'	*tutyà*	tutya	M.L	*txitxa*	txitxa	M.M
'new'	*ììn*	iin	LL	*ììn*	iin	LL
'sun'	*ndikàndiì*	ndikandii	M.L.ML	*kwàntxií*	kwantxii	L.MH

This design allows the database to serve as the basis for community- as well as research-oriented goals. Community members can use it as a dictionary based on the entries given in practical orthography with tone marks, and the column with the toneless representation is equally useful, since it allows searching for forms without having to know or type the tone diacritics. Furthermore, the tools can accommodate different orthographic conventions and even mixed conventions. The column containing just the tone patterns with the tonemes (c) provides the basis for finding tone correspondences.

The database allows for systematic comparison of segments and tone combining detailed qualitative work and computational methods to assist in establishing correspondences and identifying segmental and tonal changes. Specific goals of this project are to understand tone change in Mixtec and to learn more about Mixtec subgrouping. The database is a first step towards those goals and may serve as a model for similar lexicographic, typological, and historical linguistic projects in other language families with tone.

Acknowledgements:
We would like to thank Carmen Hernández Martínez, Gabriel Mendoza, Yésica Ramirez, Griselda Reyes Basurto, and Juvenal Solano for their collaboration and permission to use their data in this case study. This work is supported in part by NSF grant 1660355.

Abbreviations: SSM=San Sebastián del Monte Mixtec, SMD=San Martín Duraznos Mixtec; a=practical orthography with tone diacritics, b=practical orthography without tone marking, c=tone melody with syllable boundaries; H=high tone, L=low tone, M=mid tone; shaded areas entered by hand, others extracted automatically.

References

Campbell, Eric W. in press. Why is tone change still poorly understood, and how might documentation of less-studied tone languages help? In: Patience L. Epps; Na'ama Pat-El & Danny Law (eds.),

Historical Linguistics and Endangered Languages: Exploring Diversity in Language Change. New York: Routledge.

Dürr, Michael. 1987. A preliminary reconstruction of the proto-Mixtec tonal system. *Indiana*, 11, 19–61.

Josserand, Judy Kathryn. 1983. Mixtec dialect history. PhD dissertation, Tulane University.

List, Johann-Mattis. 2016. *Computer-Assisted Language Comparison: Reconciling Computational and Classical Approaches in Historical Linguistics*. Jena: Max Planck Institute for the Science of Human History.

List, Johann-Mattis; Simon Greenhill, Tiago Tresoldi & Robert Forkel. 2018. *LingPy. A Python library for historical linguistics*. Version 2.6.4. URL: http://lingpy.org. Jena: Max Planck Institute for the Science of Human History.

Longacre, Robert E. 1957. *Proto-Mixtecan*. Bloomington, IN: Indiana University Research Center in Anthropology, Folklore and Linguistics 5.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Computational tools for doing semantic typology with parallel text

Barend Beekhuizen

<pdf>

Schedule: Thu 15.00 Room 4

Code-switching in Kakheti: Gender assignment in mixed Batsbi-Georgian nominal constructions

Kate Bellamy & Jesse Wichers Schreur
(CNRS/LACITO, Leiden University & Goethe University Frankfurt)

Keywords: code-switching; grammatical gender; nominal constructions; Caucasian languages

Schedule: Thu 11.30 Room 2

In this paper we investigate how Batsbi-Georgian bilinguals assign gender to Georgian nouns appearing in otherwise Batsbi utterances. We use data from production and comprehension tasks, as well as corpus data (Wichers Schreur, 2018). Batsbi has five grammatical genders: masculine and feminine refer to humans only, whereas the three 'neuter' genders (called B, D, J after the form of their agreement targets) have a non-transparent semantic distribution (Wichers Schreur, forthc.). In contrast, Georgian has no grammatical gender. This study therefore has two main aims: (i) to extend existing research of gender systems in Caucasian languages (e.g. Corbett 1999; Chumakina & Corbett 2015; Forker 2016, 2018); and (ii) to provide new insights into the limits of code-switching in nominal constructions.

Previous codeswitching studies (e.g. Liceras et al., 2016; Parafita Couto et al., 2015; Bellamy et al., 2018), identify three gender assignment strategies in mixed nominal constructions: (i) analogy, where

the gender of the noun's translation equivalent dictates assignment; (ii) phonological cues from the ungendered language that coincide with gender assignment in the gendered language; or (iii) a default gender preference.

In the pilot director-matcher production task (adapted from Gullberg et al., 2009; N = 5 out of a total of 15), we elicited mixed nominal constructions with intended target responses of the type *baqqõ cxviri* 'big nose', containing a Georgian noun preceded by a Batsbi adjective, which is marked for gender (here *b*-). Of the 180 combinations elicited, in 163 cases (90.6%) the adjective assumes the gender of the Batsbi translation equivalent. The remaining pairs display a clear preference for D as the default gender (12/17).

These results indicate a preference for gender assignment based on analogical criteria. This is surprising given that (1) phonological cues have played a role in loanword adaptation in the history of Batsbi, and (2) the B gender is an unproductive class for loanwords (Wichers Schreur, forthc.). An ongoing forced-choice grammaticality judgement task is examining whether this pattern holds for comprehension. The results from both tasks, supplemented with corpus data, will provide novel insights for existing models of code-switching.

References

Bellamy, Kate, M. Carmen Parafita Couto and Hans Stadthagen-Gonzalez (2018), Investigating gender assignment strategies in mixed Purepecha–Spanish nominal constructions, *Languages* 3(3): 28.

Marina Chumakina and Greville G. Corbett (2015), Gender-number marking in Archi: Small is complex, in: M. Baerman, D. Brown and G.G. Corbett (eds), (2015), *Understanding and measuring morphological complexity*, Oxford: Oxford University Press.

Corbett, Greville G. (1999), Resolution rules for gender assignment in Tsakhur, in: E.V. Rakhilina and Ya. G. Testelets (eds), (1999), *Tipologija i teorija jazyka: ot opisanija k ob"jasneniju. K 60-letiju A.E. Kibrika [Typology and linguistic theory: from description to explanation. For the 60th birthday of Aleksandr Kibrik*], Moskva: Jazyki russkoj kul'tury, 224–244.

Forker, Diana (2018), Gender agreement is different, Linguistics 56(4), 865-894.

Forker, Diana (2016), Gender in Hinuq and other Nakh-Daghestanian languages, *International Journal of Language and Culture* 3(1), 90–114.

Gullberg, Marianne, Peter Indefrey and Pieter Muysken (2009), Research techniques for the study of code-switching, in B. Bullock and A. J. Toribio (eds), (2009), *The Cambridge handbook of linguistic code-switching*, Cambridge: Cambridge University Press, 21-39.

Liceras, Juana, Raquel Fernández Fuertes and Rachel Klassen (2016), Language dominance and language nativeness: The view from English-Spanish codeswitching, in R. E. Guzzardo Tamargo, C. M. Mazak and M. C. Parafita Couto (eds.), (2016), *Spanish-English Codeswitching in the Caribbean and the US*, Amsterdam: John Benjamins, 107-138.

Parafita Couto, M. Carmen, Amaia Munarriz, Irantzu Epelde, Margaret Deuchar and Beñat Oyharçabal (2015), Gender conflict resolution in Spanish–Basque mixed DPs, *Bilingualism, Language and Cognition*, 18(2), 304-323.

Wichers Schreur, Jesse (forthc.), The adaptation of borrowed nouns into the gender system of Batsbi (Nakh-Daghestanian, Georgia), in D. Forker and L. A. Grenoble (eds.), (forthc.), *Language contact in the territory of the former Soviet Union.*

Wichers Schreur, Jesse (2018*), Collection and annotation of the corpus of the Batsbi language*, paper presented at International Conference for Linguists and Caucasiologists, Tbilisi, 22-24/11/2018.

Epistemological challenges in corpus analyses of alternating constructions

Thomas Belligh
(Ghent University)

Keywords: alternating constructions, corpus linguistics, judgments on the basis of intuition, causality, cognitive subpersonal mechanisms

Schedule: We 11.30 Room 15

Over the last 20 years quantitative corpus studies have yielded many insights into the regularities involved in numerous alternations in various languages (cf. Colleman 2009, De Vaere et al. 2018, Gries 2003, Gries & Stefanowitsch 2004, Geleyn 2017, Wolk et al. 2013, among others). However, such corpus-based studies of alternating constructions face a number of epistemological challenges. This paper identifies three epistemological challenges for the corpus-based study of alternating constructions and outlines a proposal to deal with these challenges.

First, corpus studies face the difficulty to distinguish between normative language rules and regularities in language use. While both normative rules and regularities in language use can be determined from a quantitative and experimentally controlled perspective (Cowart 1997, Gibson & Fedorenko 2013, Schütze 1996, Wasow & Arnold 2005) and while normative rules emerge diachronically out of regularities in language use, rules and regularities have a different ontological, epistemic and psychological status and therefore require to be studied in their own right (Coseriu 1985, Itkonen 1978, 1997, 2006, 2016, Newmeyer 2003, Willems 2012).

Second, because correlation does not equal causation (cf. Lass 1980), caution is required in the use of causal language when describing alternations that are studied by means of a purely observational corpus approach. While correlations between constructions and parameters can be described in terms of predictions in statistical models, this does not entail that one has identified the factors that causally govern speakers' choices.

Third, linguists sometimes pay little attention to the difference between functional causes of behavior and mediating mental mechanisms (cf. De Houwer 2011, De Houwer et al. 2017, Hughes et al. 2016). The finding that there is a functional causal relation between a parameter and the choice for a construction does not directly bear on the nature of the mediating mental mechanisms responsible for this finding (cf. Bechtel 2008, Bechtel & Wright 2009). Corpus findings constitute therefore no direct evidence for postulating underlying mediating mechanisms on the subpersonal cognitive level responsible for certain outcomes (cf. Arppe et al. 2010, Sandra 1998, Stefanowitsch 2011).

These three epistemological challenges can be confronted in various ways. The paper argues that corpus analyses of alternating constructions should take into account the difference between regularities and rules and study both by means of an adequate methodology, viz. corpus research as such and a combination of corpus research and experimentally controlled introspective intuition-based judgments tasks respectively. Second, while correlational corpus findings cannot be used to make claims about causality, they can be used to generate hypotheses for subsequent behavioral experiments that can shed some light on the functional causal relationships between specific parameters and speakers' choices. Third, corpus findings can shed light on underlying mental mechanisms, if used to support or falsify independently established and explicit psycholinguistic hypotheses (cf. Gries 2003, 2017, Gilquin & Gries 2009, Stefanowitsch 2011).

References

Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., & Zeschel, A. (2010). Cognitive corpus linguistics: Five points of debate on current theory and methodology. *Corpora*, 5(1), 1-27.

Bechtel, W. (2008). *Mental mechanisms : philosophical perspectives on cognitive neuroscience*. New York (N.Y.): Erlbaum.

Bechtel, W., & Wright, C. (2009). What is psychological explanation. In P. Calvo & J. Symons (Eds.), *Routledge companion to the philosophy of psychology* (pp. 113–130). New York: Routledge.

Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences*, 31(5), 593-611.

Coseriu, E. (1985). Linguistic competence: what is it really?. *The Modern Language Review*, 80(4), xxv-xxxv.

Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: SAGE Publications

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, 6(2), 202-209.

De Houwer, J., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017). What is cognition? A functional-cognitive perspective. In S. C. Hayes and S. G. Hofmann (Eds.), *Core Processes of Cognitive Behavioral Therapies*. Oakland, CA: New Harbinger.

De Vaere, H., De Cuypere, L., & Willems, K. (2018). Alternating constructions with ditransitive "geben" in present-day German. *Corpus Linguistics and Linguistic Theory*.

Geleyn, T. (2017). Syntactic variation and diachrony. The case of the Dutch dative alternation. *Corpus Linguistics and Linguistic Theory*, 13(1), 65-96.

Gibson, E., & Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes*, 28(1-2), 88-124.

Gilquin, G., & Gries, S. T. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, 5(1), 1-26.

Gries, S. T. (2003). *Multifactorial analysis in corpus linguistics: A study of particle placement*. New York: Continuum.

Gries, S. T. (2017). *Ten lectures on quantitative approaches in cognitive linguistics : corpus-linguistic, experimental, and statistical applications*. Leiden: Brill.

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International journal of corpus linguistics*, 9(1), 97-129.

Hughes, S., De Houwer, J., & Perugini, M. (2016). The functional-cognitive framework for psychological research: Controversies and resolutions. *International Journal of Psychology*, 51, 4-14.

Itkonen, E. (1978). *Grammatical theory and metascience : a critical investigation into methodological and philosophical foundations of Autonomous linguistics*. Amsterdam: Benjamins.

Itkonen, E. (1997). The social ontology of meaning. *SKY 1997 Yearbook of the Linguistic Association of Finland*: 49–80.

Itkonen, E. (2006). Three fallacies that recur in linguistic argumentation. *Speech and Language*, 26(4): 221-225.

Itkonen, E. (2016). An assessment of (mentalist) cognitive semantics. *Public Journal of Semiotics* 7(1), 1–42.

Lass, R. (1980). *On explaining language change*. Cambridge: Cambridge university press.

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language*, 79(4), 682-707.

Sandra, D. (1998). What linguists can and can't tell you about the human mind: A reply to Croft. *Cognitive Linguistics*, 9(4), pp. 361-378.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Stefanowitsch, A. (2011). Cognitive linguistics as cognitive science. In M. Callies, W. Keller, & A. Lohofer (Eds.), *Bi-directionality in the cognitive sciences* (pp. 295–310). Amsterdam: Benjamins.

Wasow, T., & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua*, 115(11), 1481-1496.

Willems, K. (2012). Intuition, introspection and observation in linguistic inquiry. *Language sciences*, 34(6), 665-681.

Wolk, C., Bresnan, J., Rosenbach, A., & Szmrecsanyi, B. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica*, 30(3), 382-419.

Suffixes *-g-* and *-alt-* in Hill Mari: two types of middle voice

Daria Belova & Vadim Dyachkov
(Moscow State University & Institute of Linguistics (Russian Academy of Science))

Keywords: morphosyntax, middle voice, deadjectival verbs, Uralic languages, Hill Mari

Schedule: We 15.30 Room 7

The category of middle voice in Hill Mari (< Finno-Ugric) has been described in grammars and comparative works, see Galkin (1958, 1966), Salo (2006, 2015), Savatkova (2002). At the same time, the concept of middle voice has been elaborated in theoretical literature (cf. Kemmer (1993), Alexiadou (2014)). However, the middle voice and related categories are understudied in Hill Mari since their semantic properties were not explored in detail.

Our data come from fieldwork (both elicitation and corpus analysis) in the village of Kuznetsovo and some neighbouring villages (Mari El, Russia) in 2016-2018. We will argue that two productive suffixes of middle voice can be identified in Hill Mari: *-alt-* and *-g-* (along with an unproductive suffix *-n-*). It is only the *-alt-* suffix which is treated as a voice marker in traditional grammars. The *-g-* suffix is treated as a derivational affix and is not associated to the category of voice, cf. Savatkova (2002: 181, 217) stating that the *-alt-* suffix is a voice marker but the *-g-* suffix derives verbs from nouns.

The two markers differ in their semantic properties although they can be both qualified as middle voice markers. First of all, the *-g-* suffix participates in causative-inchoative alternation (in terms of Hale & Keyser (2002). Its counterpart is the transitive/causative suffix *-t-*, see (1)-(4):

(1) mən'-ən licä-em jakšar-g-en
1SG-GEN face-POSS.1SG red-MED1-PRET.3SG
'My face reddened'.

(2) mən' vas'a-m jakšar-t-em
1SG V.-ACC red-CAUS-NPST.1SG
'I disgrace Vasya (*lit*. make Vasya red)'.

(3) mašinä pədər-g-en
car break-MED1-PRET.3SG
'The car is broken'.

(4) vas'a mašinä pədər-t-en
V. car break-CAUS-PRET.3SG
'Vasya broke the car'.

Unlike the *-g-* suffix, the *-alt-* suffix mainly derives verbs from verbal stems. These verbs can have decausative (5), reflexive (6) and passive (7) meanings. The striking fact is that two middle voice suffixes are compatible within the same verb form (8):

(5) amasa (ške=ok) pač-ôlt-ən
door (self=EMPH) open-MED2-PRET.3SG
'The door opened (by itself)'.

(6) vas'a leved-ält-ən
V. cover-MED2-PRET.3SG
'Vasya covered himself'.

(7) moren kôč-alt-ôn
hare catch-MED2-PRET.3SG
'The hare was caught [by somebody]'.

(8) amasa pitər-g-ä / pitər-(g)-ält-eš
door close-MED1-NPST.3SG close-MED1-MED2-NPST.3SG
'The door was closed'.

In our talk, we will make the following claims. Using Ramchand's (2008) event structure framework, we will show that the *-g-* and *-t-* suffixes introduce the agent into the clause or co-index it to the theme. Following Levin & Rappoport Hovav's (1998) ideas, we will also give a possible semantic explanation of the fact that some stems require a middle affix specifying the cause of the state denoted by stadial level adjectives. We will support our hypothesis with typological data on deadjectival verb derivation we have collected.

On the contrary, the *-alt-* suffix operates over the event structure that already includes an agent. The suffix decreases the level of agentivity of an external verbal argument, and the latter can be removed from the syntactic structure. However, the agent is preserved in the semantic structure, which will be proved in the talk using several diagnostics (such as agent-oriented adverbs, subordinate purpose clauses and compatibility with instrument-denoting NPs). We will also consider the difference in native speakers' judgements on syntactic structures containing *alt*-verbs and will discuss the Russian influence on the properties of this marker.

The research has been supported by RBR, grant № 19-012-00627.

Abbreviations

1,3 – first, third person, ACC – accusative, GEN – genitive, EMPH – emphatic particle, CAUS – causative, MED1, MED2 – middle voice, NPST – non-past tense, PL – plural, POSS – possessive, PRET – preterite, SG – singular

References

Alexiadou, Artemis (2014), *A*ctive, middle, and passive: the morphosyntax of Voice, *Catalan Journal of Linguistics* 13, 1 - 22.

Galkin, Ivan. (1958), *Zalogi v mariyskom yazyke ["Voices in Mari"]*, Yoshkar-Ola. Galkin, Ivan (1996), *Istoricheskaya grammatika mariyskogo yazyka. Morphologiya. ["Historical grammar of Mari. Morphology"],* P.II, Yoshkar-Ola: Mari publishing house.

Hale, Ken & Keyser, Jay (2002), *Prolegomenon to a theory of argument structure*, Cambridge, MA: MIT Press.

Levin, Beth & Rappaport Hovav, Malka (1998), Building verb meanings in M. Butt and W. Geuder (eds), (1998), *The projection of arguments: lexical and compositional factors,* Stanford: CSLI Publications, 97 – 134.

Ramchand, Gillian (2008), *Verb Meaning and the Lexicon: A First Phase Syntax*, Cambridge: Cambridge University Press.

Salo, Merja (2006), The Derivational Passive and Reflexive in Mari Grammars in J. Nuorluoto (ed), (2006), *The Slavicization of the Russian North. Mechanisms and Chronology*, Helsinki, 328-340.

Salo, Merja (2015), *Passive and Reflexive Categories in Languages of the Volga Region. An Areal Typological Study*, University of Helsinki, Helsinki.

Savatkova, Anna (2002), *Gornoe narechie mariyskogo yazyka ["Hill dialect of Mari language"]*, Bibliotheca Ceremissica. T. 5. Savariae: Berzsenyi Dániel Főiskola.

Contact-induced grammatical changes in Pontic Greek spoken in Georgia

Svetlana Berikashvili
(Ilia State University of Tbilisi and Georg-August Universität of Göttingen)

Keywords: structural interference, contact-induced change, language contact, Pontic Greek, borrowing

Schedule: We 15.30 Room 8

The paper presents an empirical study on contact-induced grammatical changes in Pontic Greek (PNT), an endangered variety spoken by Pontic community of Georgia. The under-documented variety is a good candidate for a case study about language contact, because it has always been in typologically different multilingual environment. The cross-linguistic influence of contact languages, those of concatenative morphology (Turkish, Georgian) and those of non-concatenative morphology (Russian, Standard Greek) always caused the changes while transferring single words or constructions, and stimulated different processes in the language.

The discussed topics include contact-related influences of Georgian (KAT), Turkish (TUR) and Russian (RUS) as donor languages on the grammar of this genetically unrelated language variety, spoken by relatively small number of speakers. The research is based on the Pontic Corpus compiled in 2013-2016 at Bielefeld University as a result of the original fieldwork, within the framework of the project: *The impact of current transformational processes on language and ethnic identity: Urum and Pontic Greeks in Georgia*, funded by the Volkswagen Foundation and available to the academic community via the TLA archive.

The paper focuses on the synchronic outcomes of the language contact in Pontic spoken in Georgia from lexical towards structural borrowings in order to reveal the role of the source language in transference phenomena and answer the main research questions: (a) is there any sign of movement between the morphological types, i.e. acquiring of some agglutinative features by PNT; and (b) is there an advantage associated with transference between languages of the same morphological type, i.e. non-concatenative (RUS) to non-concatenative (PNT).

PNT as spoken in Georgia has different examples of structural interference: (a) shift in default gender, (b) neutralization of neuter agreement in NPs, (c) the increasing of putative cases of differential subject marking (DSM) with feminine nouns, (d) selective limitation of definiteness marking, (e) agglutination of personal suffixes in verbs, and (f) verbal agreement in number with [±animate] nouns. The talk shall be limited to the phenomena related to the developments affecting gender marking, like those that show neutralization of neuter agreement in NPs, from neuter (1) towards corresponding gender agreement (2)

(1) *to* *tranón* *odá*
DEF:N.SG.NGEN big:N.SG.NGEN room:F.SG.NGEN $_{\text{TURKISH}}$

'The big room'

(2) *atín* *tin* *káshan*
3:F.SG.ACC DEF:F.SG.ACC hot_cereal:F.SG.ACC RUSSIAN
'This cereal'

and differential subject marking which is more consistent in PNT spoken in Georgia in comparison with other Pontic varieties and is attested not only with masculines (3) but with feminines (4) as well

(3) *o* *θíon* *ípen*
DEF:M.SG.NOM uncle:M.SG.ACC say:PFV.PST:3.SG
'The uncle said'

(4) *i* *ðiaforán* *en* *polá*
DEF:F.SG.NOM difference:F.SG.ACC be:3.SG many:N.PL.NGEN
'The difference is big'

The novelty of the presentation is that it presents the original data of an understudied variety of PNT, as currently spoken by Pontic community of Georgia and it underlines structural changes that happened in Pontic after the contact with other languages in Georgia.

Comparing argument structure variation in Homeric Greek and Early Vedic: the case of *oîda* and *veda*

Erica Biagetti & Silvia Luraghi
(University of Pavia; University of Bergamo & University of Pavia)

Keywords:Argument structure constructions, accusative/genitive alternation, Homeric Greek, Early Vedic

Schedule: Fri 14.00 Room 14

In this paper, we compare the distribution of argument structure constructions of the verbs *oîda* and *veda* 'know' in H(omeric) G(reek) and E(arly) V(edic) respectively. Both verbs derive from the perfect of PIE root **wid-* 'see', and the semantic extension to cognition is explained as connected with the resultative meaning of the perfect: knowledge is regarded as a cognitive state reached through sight (Sweetser 1990: 38). *Oîda* and *veda* belong to a group of cognitive verbs that allow for construction variation in the ancient Indo-European languages and also includes verbs meaning 'hear', 'realize', 'learn'. Traditionally, the occurrence of the genitive with such verbs is viewed as connected with its partitive meaning (Delbrück 1901: 588-591), and with a low degree of transitivity (Hettrich 2006). Factors influencing the distribution of the two constructions are not clearly detectable, as they may vary depending on specific verbs (e.g. animacy plays a role only for verbs of hearing; Delbrück 1901: 588) as well as on the individual language. Focusing on *oîda* and *veda* in HG and EV reveals different patterns of distribution. While *oîda* most often takes the genitive when indicating a practical skill (*contra* Chantraine 1968-1977: 780; cf. *tóxōn* bow.GEN.PL *eὺ* well *eidṓs* know.PTCP.PF.NOM.SG 'well-skilled to use the bow' Hom. *Il.* 12.350), and the accusative when indicating intellectual knowledge, this does not appear to be the case with *veda*, which indicates practical skills infrequently, but tends to feature the accusative (cf. *pitryāṇy* paternal.ACC.PL *āyudhāni* weapon.ACC.PL *vidvān* know.PTCP.PF.NOM.SG 'skilled to use his father's weapons' *Rv* 10.8.8). Often, in the RV it is hard to attach any semantic difference to case variation (cf. *na* not *tasya* DEM.GEN *vidma* know.PF.1PL 'we

don't know this' RV 10.40.11 and *yan* DEM.ACC *na* not *vidma* know.PF.1PL 'that, we don't know' RV 5.85.8). Another difference concerns the occurrence of human participants as second arguments, which is limited to indefinite non-specific expressions in HG, while in EV it is somewhat more extended (cf. *vedā* know.PF.3SG *me* 1SG.GEN *deva(ḥ)* god.NOM 'the god knows me' RV 5.12.3). Similarities include the use of both verbs with the infinitive in the meaning 'know how to do' and the occurrence of non-indicative moods in protestations (*vidyur* know.OPT.PF.3PL *me* 1SG.DAT *asya* DEM.GEN *devā(ḥ)* god.NOM.PL 'Be the gods my witnesses of this' RV 1.23.24 and *ístō* know.IMP.PF.3PL *nûn* now *tóde* DEM.ACC *Gaîa* Earth.NOM *kaì* and *Ouranòs* Heaven.NOM 'Hereto now be Earth my witness and Heaven' Hom. *Il.* 15.36). In our paper we aim to give an in-depth account of construction variation in both languages, in order to show how corresponding patterns may point to common inheritance. Patterns that should more likely be explained through innovation, in their turn, can be shown to shed light on construction variation when set in a broader framework in the two languages, by comparing *oîda* and *veda* with other verbs that share a similar behavior.

References

Chantraine, Pierre. 1968. *Dictionnaire Étymologique de la Langue Grecque*. Paris: Klincksieck.

Delbrück, Bertold. 1901. *Vergleichende Syntax der Indogermanischen Sprachen*. Strassburg.

Hettrich, Heinrich. 2006. *Some remarks on the adverbal genitive in Rgvedic Sanskrit*. Paper read at the 13th World Sanskrit Conference, Edinburgh, July 10 – 14, 2006

Sweetser Eve. 1990. *From etymology to pragmatics: The mind-body metaphor in semantic structure and semantic change*. Cambridge: CUP.

Category clustering:
A probabilistic universal in the morphology of verbal agreement

Balthasar Bickel, John Mansfield & Sabine Stoll
(University of Zurich, University of Melbourne and University of Zurich)

Keywords: morphology, universals, typology, corpus, morpheme order

Schedule: Sat 12.30 Room 6

Several languages (e.g. Chintang, Tagalog, Murrinhpatha) challenge the expectation of strict sequential ordering in morphology. However, it is unclear whether these languages exhibit random placement of affixes or whether there are probabilistic patterns. Here we address this question for verbal agreement markers and hypothesize a probabilistic universal of *category clustering*, with two effects: (i) markers in paradigmatic opposition tend to be placed in the same morphological position (*paradigmatic alignment;* Crysman & Bonami 2016), and (ii) morphological positions tend to be categorically uniform (*featural coherence;* Stump 2001).

We first test this hypothesis in a corpus of agreement marker prefixes in Chintang, a Sino-Tibetan language where prefix placement is not constrained by categorical rules (Bickel et al. 2007). We extract 475 bigram sequences of transitive A (agent-like) markers co-occurring with a P (patient-like) marker or a negation marker from a corpus (Bickel et al. 2013). We fit a mixed-effects logistic regression model on the probability of an A marker to be placed on the left in each bigram, with fixed factors capturing (i) the difference between a co-occurring P marker vs a co-occurring negation marker (where featural coherence predicts a difference, constraining the relative placement of A markers), and (ii) the difference among A markers coding various person and number values (where paradigmatic

alignment predicts no difference). We control for persistence and priming effects (Bickel et al. 2007) through a fixed factor capturing the presence of a previous bigram with the same order, and for speaker and lexeme biases by random factors. Results show that the placement of A markers significantly depends on the category of the co-occurring prefix, with A markers appearing more likely to the left of a negation marker than to the right of a P marker, suggesting a preference for featural coherence in a P≻A≻NEG sequence. By contrast, the placement of an A marker does not significantly vary across its different person/number values, suggesting a preference for paradigmatic alignment.

We then test whether category clustering shapes the evolution of morphological structure in argument marking, predicting that although exceptions are attested, structures are much more likely to obey rather than to violate the principle. To test this, we extract information on the distribution of A and P argument markers across morphological positions from 136 languages in AUTOTYP (Bickel et al. 2017). We then compute, for each language, its paradigmatic alignment degree from the cumulative probability of alignment (e.g. 4 out of 5 A markers in the same position) given all logically possible distributions (e.g. all 5 markers in different positions). A beta regression with random intercepts for language family shows that both A and P markers evidence paradigmatic alignment degrees significantly above chance expectations. We assess feature coherence by measuring the association of A and P markers with distinct positions compared using Cramér's V. A beta regression, again with random family intercepts, shows associations to be significantly higher than chance.

Our findings suggest that category clustering may indeed be a universal principle of language, and we speculate that it facilitates language learning and processing.

References

Bickel, B, G Banjade, M Gaenszle, E Lieven, N Paudyal, IP Rai, M Rai, NK Rai & S Stoll. (2007), Free prefix ordering in Chintang. *Language* 83. 43–73.

Bickel, B, S Stoll, M Gaenszle, NK Rai, E Lieven, G Banjade, TN Bhatta, N Paudyal, J Pettigrew, IP Rai & M Rai. (2013), Audiovisual corpus of the Chintang language. DOBES Archive, http://www.mpi.nl/DOBES

Bickel, B, J Nichols, T Zakharko, A Witzlack-Makarevich, K Hildebrandt, M Rießler, L Bierkandt, F Zúñiga & JB Lowe. (2017), The AUTOTYP typological databases, version 0.1.0. GitHub, https://github.com/autotyp/autotyp-data/tree/0.1.0

Crysmann, B & O Bonami. (2016), Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52. 311–374.

Stump, G (2001), *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Syntactic and semantic constraints on embedded gapping: A crosslinguistic comparison

Gabriela Bîlbîie, Israel de la Fuente & Anne Abeillé
(University of Bucharest, Université de Lille & Université de Paris)

Keywords: gapping, embedding, factivity, acceptability judgment tasks

Schedule: We 12.30 Room 15

It is usually assumed (Hankamer 1979, Neijt 1979, Johnson 2009) that gapping differs from other kinds of ellipsis in banning embedding (1). However, Weir (2014) suggests that embedding gapping may depend on the matrix verb.

(1) *Alfonso stole the emeralds, and **I think** [**that** Mugsy the pearls]. (Hankamer 1979)

Four acceptability judgment tasks for Spanish, Romanian and English show that there is cross-linguistic variation with respect to embedded gapping, and that two constraints seem to be at work: on the semantic side, non-factive verbs embed more easily than factive ones (Kiparsky and Kiparsky 1971, Karttunen 1971), independently of ellipsis; on the syntactic side, no complementizer (with non-factive verbs) embeds more easily than a complementizer (Jaeger 2006, 2010).

To test the semantic constraint, we ran 3 experiments, using a 2x3 design (gapping, embedding-nonfactive, embedding-factive), with similar materials in each language (2). We had 56 participants for Spanish, 72 for Romanian, and 51 for English. In Spanish and Romanian, embedded gapping is as acceptable as embedded non-gapping under non-factive verbs. In English, there is an interaction between gapping and embedding (mean z-score for embedded gapping –0.8). Moreover, factivity is significant in all languages: embedded clauses under a factive verb are less acceptable than under a non-factive verb.

(2) a. [±gapping, +embed, +factive]
(S) En el bar, Pablo pidió una cerveza y **me molesta que** Juan (pidió) un whisky.
(R) La restaurant, Paul a comandat o bere şi **mă îngrijorează că** Florin (a comandat) un whisky.
(E) At the bar, Paul ordered a beer and **it bothers me that** John (ordered) a whisky.

b. [±gapping, +embed, –factive]
(S) En el bar, Pablo pidió una cerveza y **sospecho que** Juan (pidió) un whisky.
(R) La restaurant, Paul a comandat o bere şi **bănuiesc că** Florin (a comandat) un whisky.
(E) At the bar, Paul ordered a beer and **it seems that** John (ordered) a whisky.

c. [±gapping, –embed]
(S) En el bar, Pablo pidió una cerveza y Juan (pidió) un whisky.
(R) La restaurant, Paul a comandat o bere şi Florin (a comandat) un whisky.
(E) At the bar, Paul ordered a beer and John (ordered) a whisky.

To test the syntactic constraint, we ran Experiment 4 on English (49 participants), using a 2x2 design (±gapping, ±*that*) with non-factive verbs (3). It shows that the complementizer omission renders embedded gapping more acceptable.

(3) a. [±gapping, + *that*]
At the corner shop, Peter stole cigarettes and **I think that** Larry (stole) chocolates.

b. [±gapping, –*that*]
At the corner shop, Peter stole cigarettes and **I think** Larry (stole) chocolates.

We conclude that the No Embedding Constraint on gapping is affected not only by the semantic class of the embedding predicate, but also by the presence/absence of the complementizer. The difficulty of coordinating a simple clause and a complex clause may result from a more general parallelism constraint on coordination (Frazier *et al.* 2000) and the further penalty on factive verbs may come from their non-assertive nature (Hooper 1974) and/or from the QUD-incongruence (Ginzburg 2012).

References

Frazier, Lyn, Alan Munn, and Charles Clifton (2000), Processing Coordinate Structures, *Journal of Psycholinguistic Research* 29(4), 343-370.

Ginzburg, Jonathan (2012), *The Interactive Stance: meaning for conversation*, Oxford University

Press.
Hankamer, Jorge (1979), *Deletion in coordinate structures*, Garland Publishing.
Hooper, Joan, (1975), *On Assertive Predicates*, in J. P. Kimball (ed.), *Syntax and Semantics* 4, New York: Academic Press, 91-124.
Jaeger, Florian T. (2006), Redundancy and Syntactic Reduction in Spontaneous Speech, PhD thesis, Stanford University.
Jaeger, Florian T. (2010), Redundancy and reduction: Speakers manage syntactic information density, *Cognitive Psychology* 61(1), 23-62.
Johnson, Kyle (2009), Gapping is not (VP-) ellipsis, *Linguistic Inquiry* 40, 289-328.
Kartunnen, Lauri (1971), Some observations on factivity, *Papers in Linguistics* 4, 55-69.
Kiparsky, Paul, and Carol Kiparsky (1971), Fact, in D. D. Steinberg, and L. A. Jakobovits (eds.), *Semantics. An interdisciplinary reader in philosophy, linguistics and psychology*, Cambridge: Cambridge University Press, 345-369.
Neijt, Anneke (1979), *Gapping: A contribution to Sentence Grammar*, Dordrecht: Foris Publications.

ø

Multiple inheritance as the origin of the Norwegian psychologically distal demonstrative

Barthe Bloom
(Friedrich Schiller University, Jena)

Keywords: constructional change, multiple inheritance, Norwegian, preproprial article, psychological distal demonstratives

Schedule: Fri 9.00 Room 14

New demonstratives commonly arise through the reinforcement of older ones (Vindenes 2017, 145-187). Johannessen (2006), however, has discovered a demonstrative in colloquial Norwegian, whose origin cannot be explained by reinforcement: the psychologically distal demonstrative (henceforth: PDD), exemplified in (1) and (2). Following the framework of diachronic construction grammar (Noël 2007; Barðdal et al. 2015), I propose that the PDD is the result of multiple inheritance, where a construction inherits properties of multiple sources (Trousdale 2013).

(1)	*og*	***han***	***presten***		*vet*	*du*		(Bardu_ma03)
	and	3.M.SG	priest.DEF		know	you		[PDD,
	'and that priest, you know'							recognitional]
(2)	*vi*	*hadde*	*baker*	*og*	***han***	***bakeren***	*han*	*bare*
	we	had	baker	and	3.M.SG	baker.DEF	he	just
	la	*seg*	*litt*	*nedpå*				(Kalvåg_04gk)
	let	himself	bit	down				[PDD, anaphoric]
	'we had a baker and that baker he just laid down for a bit'							

The PDD has the form of the third person personal pronoun, *han* (masculine) or *hun* (feminine) and occurs adnominally with a proper noun, kinship term or common noun with human reference. It signals that the referent of the noun phrase is personally unknown to the speaker and/or addressee, and/or expresses the speaker's negative attitude towards the referent (Johannessen 2008). The adnominal position of the pronoun is known from the preproprial article as well, see (3). Despite

striking similarities, the two are distinct constructions, among other things because the PDD is stressed, while the article is not and because the preproprial article merely marks definiteness and lacks the functions of the PDD (Johannessen and Garbacz 2014).

(3) *Vi måtte hjelpe* ***han far*** *og* ***hun mor*** (Ljørdalen_ma02)
We must help M.DEF father and F.DEF mother [preproprial article]
'We must help father and mother'

This article is unlikely to be the only source for the PDD, because it would inter alia constitute a change from reduced to full form (e.g. Fischer, Norde, and Perridon 2004, 2). Instead, I argue that the PDD originates at the intersection of the preproprial article (3), the distal/neutral demonstrative (4), and the third person pronominal system (5).

(4) ***den øya***, *vet du du vokste opp midt oppå* (Tromsø_04gk)
that island.DEF know you you grew up middle on [recognitional demonstrative]
'that island, you know, where you grew up on'

(5) *Vi har jo Brunoen enda* ***han*** *er tolv år gammal* (Vos_01um)
We have yeah Bruno.DEF still he is 12 years old[anaphoric pronoun]
'We still have Bruno. He is twelve years old.'

An analysis of data from the Nordic Dialect Corpus (Johannessen et al. 2009) shows that the PDD has functions that are associated with demonstratives, e.g. recognitional reference (Diessel 1999, 105–9) in (1) and (4), and pronouns, i.e. continuing anaphoric reference in (2) and (5) (Hoop 2004, 27). In addition to the inheritance of the functional features, evidence from the form of the PDD and its distributional properties provides further support for the hypothesis. In sum, it is argued that the origin of the PDD lies in three source constructions and can best be accounted for in terms of multiple inheritance.

References

Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer, and Spike Gildea (2015), *Diachronic Construction Grammar*, Amsterdam / Philadelphia: John Benjamins.

Diessel, Holger (1999), *Demonstratives. Form, Function, and Grammaticalization*, Amsterdam / Philadelphia: John Benjamins.

Fischer, Olga, Muriel Norde, and Harry Perridon (2004), Introduction: In Search of Grammaticalization, in O. Fisher, M. Norde, and H. Perridon (eds.), (2004), *Up and down the Cline. The Nature of Grammaticalization*, Amsterdam / Philadelphia: John Benjamins, 1-16.

Hoop, Helen De (2004), On the Interpretation of Stressed Pronouns. In R. Blutner and H. Zeevat (eds.), (2004), *Optimality Theory and Pragmatics*, Hampshire / New York: Palgrave MacMillan, 25-41.

Johannessen, Janne Bondi (2006), Just Any Pronoun Anywhere? Pronouns and 'New' Demonstratives in Norwegian, in T. Solstad, A. Grønn, and D. Haug (2006), *A Festschrift for Kjell Johan Sæbø: In Partial Fulfilment of the Requirements for the Celebration of His 50th Birthday*, Oslo: University of Oslo, 91-106.

Johannessen, Janne Bondi (2008), Psykologiske Demonstrativer, in J. B. Johannessen and K. Hagen (2008), *Språk i Oslo. Ny Forskning Omkring Talespråk*, Oslo: Novus, 63-77.

Johannessen, Janne Bondi, and Piotr Garbacz, (2014), Proprial Articles. *Nordic Atlas of Language Structures (NALS) Journal* 1, 10-17.

Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor A. Åfarli, and Øystein Alexander Vangsnes (2009), The Nordic Dialect Corpus - an Advanced Research Tool, in K. Jokinen and

E. Bick (eds.), (2009), *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*. Odense: Northern European Association for Language Technology (NEALT), 73-80.

Noël, Dirk (2007), Diachronic Construction Grammar. *Functions of Language* 14 (2), 177-202.

Trousdale, Graeme (2013), Multiple Inheritance and Constructional Change, *Studies in Language* 37 (3), 491-514.

Vindenes, Urd (2017), *Complex Demonstratives and Cyclic Change in Norwegian,* University of Oslo.

The existential in Xhosa and Information Structure

Eva-Marie Bloom Ström

Schedule: Fri 9.00 Room 6

The existential in Xhosa has two forms; a 'full' form *-khona* which relates to the locative *khona* 'there' (example 0), and a form without the ending *-na* (example 0). Inflection is default with the locative noun class 17:

(ku)-khona *ne-apile,* *ku-khon(a)* *ibanana*
(SM17)-be.present CONN-5.apple SM17-be.present 9.banana
a-(ku)-kho *nto* *i-nge-kho-yo*
NEG-SM17-be.present 9.thing 9RC-NEG-be.present-REL
'There are apples, there are bananas, there isn't anything that isn't there.'

Ku-kh(o) *unyana o-wa-lusa-yo*
SM17-be.present 1a.son 1RC-SM.PST1-herd-REL
aph(e) *ku-le-k(h)aph(a)*
16DEM.PROX LOC-9DEM.PROX-5.home
'There was a shepherd in this homestead.'

The difference between the two is argued in this paper to relate to information structure. The longer form is used for contrastive focus, emphasising the presence of the logical subject, contra the assumption of the listener, or in the case of all-new information. The short form *-kho* is used otherwise. It is always used in the negative, an example of which can be seen in (0).

Previous literature on Xhosa pay little attention to existentials, and it is an under-researched field in Bantu languages in general (Bernander et al. In press). The distinction between the two forms has been noted in Xhosa and also for closely related Zulu, but is mostly claimed to be more or less due to free variation (Oosthuysen 2016) or that *-na* is a stabilizer needed for phonological reasons (Buell 2005: 95).

In line with many other languages (Milsark 1977; Salles and Matthewson 2016), the existential expresses 'there is/are' and the following nominal mostly has an indefinite, non-referential reading. The full form *-khona* appears to often be used when the referent is introduced for the first time. The following example is in the recent past perfective (0), as is the follow-up in (0), however with the short form *-kho*:

Ku-be-khon(a) *umnt(u)* *e-ndaw-eni* *ya-khe*
SM17-REC-be.present 1.person LOC-9.place-LOC 9-1POSS
'For there to be a person in his/her place.'

Ewe *ku-be-kh(o)* *umntu* *o-za-wu-thi* *ku-bonakal-e* *ukuba*
yes SM17-REC-be.present 1.person 1RC-FUT-INF-say SM17-clear-SBJV that
'Yes, for there to be a person who will make clear that/ensure that...'

A definite meaning is possible, at least in elicited examples:

Ku-kho *unyana* *wa-m* *apha.*
SM17-be.present 1a.son 1-1SG.POSS here
'There is my son here.'

This presentation analyses the differences in use between *-khona* and *-kho*, previously claimed not to have any meaning difference, in terms of information structure and speaker-hearer interaction. Its relation to definiteness is also examined, based on examples from a corpus of spoken texts as well as elicitation. Such examples include those that use agreement of the following subject, in comparison with those that use default locative agreement in the examples above.

i-khon(a) *(i)nto* *(e)-za-z(i)-push-w-a*
SM9-be.present 9.thing 10RC-SM.PST10-IPFV10-push-PASS-FV
'there was a thing that was pushed.'

References

Bernander, Rasmus, Maud Devos and Hannah Gibson. In press. Existential constructions in Bantu languages. In E.-M. Bloom Ström, H. Gibson, R. Guérois and L. Marten (eds): *Apporaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Buell, Leston. 2005. Issues in Zulu verbal morphosyntax. Doctoral disseration. Los Angeles: University of California.

Milsark, Gary. 1977. 'Toward an explanation of certain peculiarities of the existential construction of English'. Linguistic analysis 3:1-29.

Oosthuysen, Jacobus C. 2016. *The grammar of isiXhosa*. Stellenbosch: SUN MeDIA.

Salles, Raiane and Lisa Matthewson. 2016. Existentials and (in)definiteness in Pirahã. In T. Bui and R.-R. Ivan (eds): *Proceedings of the Ninth Conference on the Semantics of Under-Represented Languages in the Americas*, 127-40. Amherst, MA: GLSA.

Elucidating *it*-clefts by comparison with reduced *it*-clefts

Charlotte Bourgoin, Gerard O'Grady, Kristin Davidse, Karen Lahousse
(KU Leuven; Cardiff University; KU Leuven; & KU Leuven)

Schedule: Thu 9.00 Room 11

The traditional view of ***it*-clefts**, e.g. (1a), takes the corresponding **simple proposition**, e.g. (1b), as their closest alternate, assuming that this proposition is cleft into a biclausal structure as a way of 'grammaticalizing' (Lehman 2008) the focus-presupposition relation, which maps onto value and variable of the specification relation.

(a) a. it was "^J\/ohnny that 'stole her m\oney# (LLC) b. Johnny 'stole her m\oney

However, as pointed out by Atlas & Levinson (1981:16), *it*-cleft (1a) and simple proposition (1b) have different **presuppositional** behaviour. (1a) presupposes the open proposition 'someone stole her

money', while (1b), if we assume utterance-final information focus on *money*, either presupposes 'Johnny stole something' or has no presupposition and simply tells 'what happened' (Halliday 1967:207).

We propose that it is actually so-called "**reduced clefts**" (Declerck & Seki 1990), i.e. simple specificational clauses with subject *it*, (2a), that are pragmatically and informationally the closest variant of *it*-clefts, (2b). Both **presuppose** 'someone was banging on the ceiling' and mark *her son* with a **focus** that is informationally **contrastive** (either with –implied- lexical items or grammatical options) (Halliday 1967:207). The main difference is that clefts explicitly restate the pragmatic presupposition in their relative clause, which raises the question of the functionality of this redundancy.

(2) a. and ^how she heard re:p\eated# "^bangs on the !c\eiling# - ^thinking it was her s\on# (LLC)
b. it was her s\on who was banging the c\eiling

This paper wants to cast new light on clefts by comparative qualitative-quantative analysis with reduced clefts in a dataset of c.350 tokens extracted from the prosodically transcribed London-Lund Corpus (LLC).

Regarding the **presupposed open proposition**, we first investigate the variation in its discourse-givenness in the **preceding** text (cf. Lambrecht 2001:474) for *it*-clefts and reduced clefts in terms of Kaltenböck's (2004) distinctions evoked, inferable, and new-anchored. For the *it*-clefts, we additionally assess how closely the presupposition in the preceding text is repeated by the cleft relative clause. This allows us to draw up *prototypical profiles*, for which our starting hypothesis is as follows: reduced clefts typically have a textually evoked pragmatic presupposition, while full clefts typically have a discursively inferable presupposition restated with some variation in the cleft relative clause.

Regarding the prosodically coded **information structure**, we show that the conflation of (single) focus (marked by tonic prominence) with the value NP posited by Lambrecht (2001: 467) for *it*-clefts *(It is **champagne** that I like*) is not even the typical pattern. On the one hand, the variable typically has one or more foci, as in (1a), (3). On the other, the value NP does not always carry focus and the focus may select only a part of the value NP (Velleman *et al* 2012) as in (3), whose presupposition is 'someone's credibility is in question'. These possibilities are not available for reduced clefts.

(3) it is "^th\eir 'credi'bility that's in qu/estion (LLC)

We conclude by setting out the systems of focus-presupposition options, which are more complex for full clefts than reduced clefts, as **distinct** from the grammatical structures conveying specification.

References

Atlas, J.D., Levinson, S.C. 1981. It-clefts, informativeness, and logical form: radical pragmatics. P. Cole (Ed.), *Radical pragmatics*, 1-61. New York: Academic Press.

Declerck, R., Seki, S. 1990. Premodified reduced IT-clefts. *Lingua* 82: 15-51.

Halliday, M.A.K. 1967. Notes on transitivity and theme in English 2. *Journal of Linguistics* 3: 199-244.

Kaltenböck, G. 2004. It-*extraposition and Non-extraposition in English: a Study of Syntax in Spoken and Written Texts*. Wien: Braumüller.

Lambrecht, K. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39: 463-516.

Lehmann, C. 2008. Information structure and grammaticalization. In Elena Seoane & María José López-Couso (eds.), *Theoretical and Empirical Issues in Grammaticalization*, 207–229. Amsterdam: Benjamins.

Velleman, D. et al. 2012. It-clefts are IT (inquiry terminating) constructions. Proceedings of SALT 22: 441-460.

Putting complement clauses and false belief into context: How does understanding of beliefs influence choice of sentence structure?

Silke Brandt, Stephanie Hargreaves & Anna Theakston

(Lancaster University and ESRC International Centre for Language and Communicative Development; University of Manchester; & University of Manchester and ESRC International Centre for Language and Communicative Development)

Keywords: complement clauses, information structure, theory of mind, child language acquisition

Schedule: Thu 10.00 Room 11

When children start to produce complement clauses (e.g., *I think (that) he's nice*), they often omit the *that*-complementizer. Based on research on spontaneous speech it has been argued – for both children and adults – that these complement clauses have the same function as main clauses: they contain foregrounded information. The associated main clause (e.g., *I think*), instead, functions as an epistemic marker, which could be omitted (Diessel & Tomasello, 2001; Thompson, 2002). However, little research has investigated how complement clauses with/out *that*-complementizers are interpreted.

Another strand of research suggests that children's comprehension of complement clauses is linked to their understanding of false belief (Milligan et al., 2007). However, most of these studies have presented complement clauses in isolation and have focused on children's comprehension of the syntactic structure rather than the function. In the current study we presented complement-clause constructions with and without complementizers in a story context and tested how children's comprehension and production of these constructions were related to their general understanding of false belief and their understanding of true and false beliefs present in the stories.

We tested 25 English-speaking 4-year-olds, 25 5-year-olds and 24 adults. We used a sentence-elicitation task with a 2 (+/- *that*-complementizer) by 3 (belief: false, true, neutral) design. Participants heard a total of 18 complement-clause constructions embedded in story contexts supported by pictures: E.g. '*Dan wants to play football with Sue. But he thinks (that) she's on holiday*'. The belief was conveyed as true (*she really is on holiday*), false (*she actually isn't on holiday*), or left neutral (*there's a blue star*). The test question (*Why doesn't Dan play football with Sue?*) elicited a main-subordinate construction (*because he thinks (that) she's on holiday*) or just a subordinate clause (*because she's on holiday*). We also conducted four explicit standard false-belief tests and a short sentence-repetition test to measure children's general false-belief understanding and their verbal working memory.

Age group and belief context affected participants' responses. 4-year-olds produced the lowest percentage of main-subordinate constructions (36.83%), compared to 5-year-olds (57.46%) and adults (88.0%). However, a main effect for belief context suggests that each group used most main-subordinate constructions when the story character had a false belief, fewest when they had a true belief, with the neutral condition in-between. Whether or not the sentences in the story contained a *that*-complementizer did not influence participant's own choice of sentence structure, $\chi^2(1) = 0.013$, $p = .909$ (see Figure 1), questioning claims that omitting complementizers foregrounds subordinate propositions. Verbal working memory also affected child-participants' performance: Children with higher scores used more main-subordinate constructions. However, children's performance in false-belief tasks did not predict their choice of sentence structure.

These results suggest that, at the age of 4, children start to understand the difference between true and false beliefs when they are presented in a story context. They also start to use the appropriate

linguistic form to refer to false beliefs. However, limitations in verbal working memory still constraint their production of complement-clause constructions and their ability to linguistically encode false beliefs.

Figures

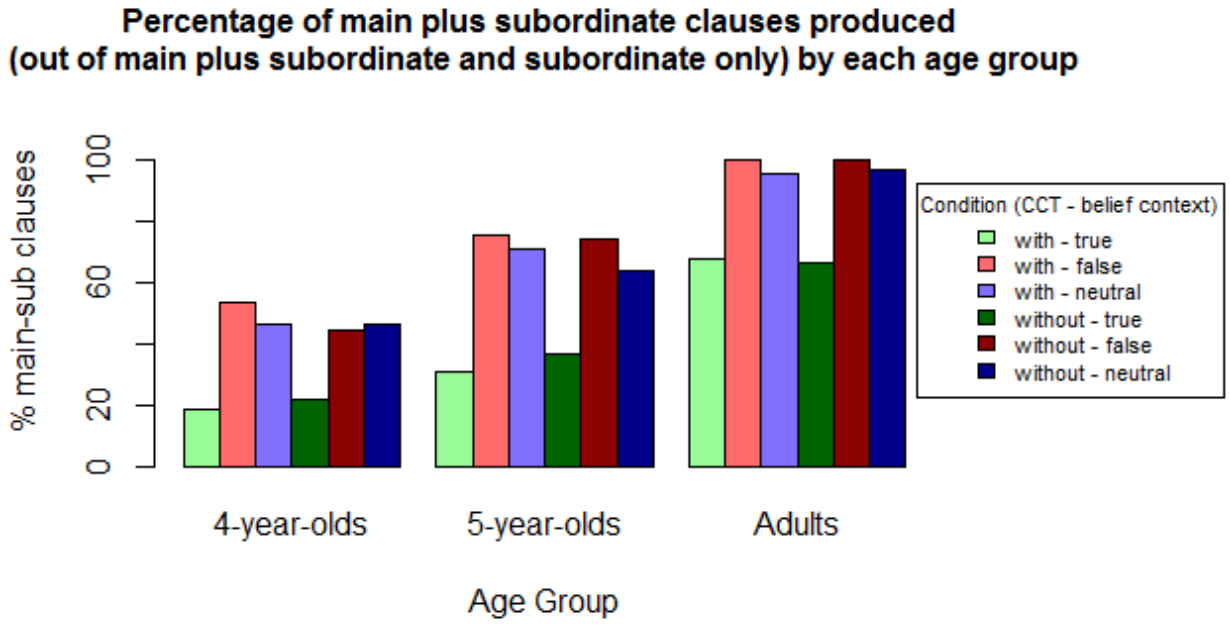


Figure 1. Percentage of main-subordinate constructions produced by each age group; grouped by condition (with/out complementizer; true, false, neutral belief).

References

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*(2), 97–141.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child development*, *78*(2), 622-646.

Thompson, S. A. (2002). "Object complements" and conversation towards a realistic account. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, *26*(1), 125-163.

Register variation in the expression of contrast in French: Frequency and use of the contrastive adverbs *au contraire*, *par contre* and *en revanche*

Jorina Brysbaert & Karen Lahousse
(KU Leuven and Research Foundation – Flanders & KU Leuven)

Keywords: contrast, adverbs, register variation, French, corpus analysis

Schedule: Fri 11.30 Room 7

French contrastive adverbs (*au contraire* 'on the contrary', *néanmoins* 'nevertheless', *en même temps* 'at the same time', etc.) have almost only been analyzed on the basis of formal (written) data. Although some linguists have pointed out that these adverbs do not all occur in the same register (see Csűry 2001, Danjou-Flaux 1980, Mellet & Ruggia 2010, etc.), a corpus analysis of the frequency and properties of a large set of contrastive adverbs in different registers is still lacking. This is a pity, because such an analysis could provide further evidence for the claim that there is a situation of diglossia in French, i.e. that spoken (informal) and written (formal) French have fundamentally different grammars (Rowlett 2013, Zribi-Hertz 2011, etc.).

We present the results of a case study on three contrastive adverbs (*au contraire* 'on the contrary', *par contre* 'on the other hand' and *en revanche* 'on the other hand') in a corpus of formal written French (*Le Monde 1998* (25.700.000 words)), informal written French (*Yahoo Contrastive Corpus of Questions and Answers*, https://fr.answers.yahoo.com/ (6.100.000 words)) and ("spontaneous") informal spoken French (http://cfpp2000.univ-paris3.fr/ (650.000 words)). We show that the register has an influence on the frequency of these three contrastive adverbs as well as on their syntactic properties.

The **FREQUENCY** of these three contrastive adverbs varies greatly depending on the register. In general, our data suggests that informal spoken French contains more contrastive adverbs than written French. In addition, *par contre* 'on the other hand' is used more often in informal French than in formal French, whereas *en revanche* 'on the other hand' is more likely to occur in formal French than in informal French. Our corpus analysis also shows that it is especially interesting to study informal written French, in order to determine the effect of the style (formal / informal) and/or the mode (written / spoken).

The **SYNTACTIC PROPERTIES** of these three contrastive adverbs vary depending on the register. In informal French, they are very often used in initial position (1), whereas in formal French, they frequently follow the conjugated verb (2) (also see Dupont 2015):

(1) *Laisse les se transformer en aide ménagère bon **par contre** il faut une limite de temps.*
'Let them turn into a housekeeper right **however** you need a time limit.' (*Yahoo Corpus*)

(2) *Nous gardons toute notre confiance à Jean Blaise pour sentir et nous faire découvrir le monde. Il faudra, **par contre**, un contrôle de gestion plus rigoureux.*
➔ 'We keep all our confidence in Jean Blaise to feel and let us discover the world. We will need, **however**, a more rigorous management control.' (*Le Monde 1998*)

In addition, we show that it is much more common to intensify the adverb *au contraire* 'on the contrary' by means of another adverb (e.g. *bien au contraire* 'quite to the contrary', *au contraire même* 'even to the contrary') in informal written French compared to informal spoken French.

References

Csűry, István (2001), *Le champ lexical de mais: Étude lexico-grammaticale des termes d'opposition du français contemporain dans un cadre textologique*, Debrecen: Kossuth Egyetemi Kiadó.

Danjou-Flaux, Nelly (1980), Au contraire, par contre, en revanche: Une évaluation de la synonymie, in *Synonymies*, Lille: Presses Universitaires de Lille, 123-148.

Dupont, Maïté (2015), Word order in English and French: The position of English and French adverbial connectors of contrast, *English Text Construction* 8(1), 88-124.

Mellet, Sylvie, and Ruggia, Simona (2010), *Quand même*, à la croisée des approches énonciatives, in M. Iliescu, H. Siller-Runggaldier, and P. Danler (eds), *XXVème Congrès international de linguistique et philologie romanes [CILPR]*, Innsbruck: Niemeyer, 201-209.

Rowlett, Paul (2013), Do French speakers really have two grammars?, *Journal of French Language Studies* 23(1), 37-57.

Zribi-Hertz, Anne (2011), Pour un modèle diglossique de description du français: Quelques implications théoriques, didactiques et méthodologiques, *Journal of French Language Studies* 21(2), 231-256.

ꕥ

Picturing prototypes: a connectionist approach to analogical attraction

Sara Budts

Schedule: Fri 11.30 Room 14

In the 1980s, cognitive science witnessed the rise of a radically usage-based framework called "connectionism", which aimed to account for the acquisition, storage and processing of all mental representations (not just language), in a neurologically plausible way. The framework has been implemented by computational models called 'neural networks' (Rumelhart & McClelland 1986). One such model, the Convolutional Neural Network (CNN), mimics the workings of the visual pathway to perform automatic image classification: it learns to map constellations of pixels to a label representing the pictured object (e.g. "car", "face").

Recently, CNNs have been applied to language too (Vanni et al. 2018), exploiting constellations of semantic and syntactic features rather than pixels to label unannotated sentences. This makes them ideal to detect prototypical usage of competing variants in linguistic alternations, especially when distinctions are subtle. If we feed a CNN with corpus sentences where the competing variants are masked, we can train it to guess which variant the sentence originally contained. By forcing it to locate the contextual features that are maximally discriminatory, the CNN not only provides pointers to all competitor's prototypical contexts, but also, crucially, to the contexts they share, which are likely to be loci of analogical attraction.

I demonstrate the potential of CNNs with an analysis of the Early Modern inclusion of DO into the paradigm of the modal auxiliaries. Because both occur in questions and negatives (1), it has been argued that the consolidation of DO-support was triggered by analogy with these auxiliaries (Warner 1993: 198).

(1) Do/will you love me? - I do/will not love you.

Still, DO-support shows various degrees of entrenchment in various contexts (Ellegård 1953, Kroch 1989). We model its consolidation as an accumulation of associations with the modals. The underlying assumption is that DO in questions is associated with different (uses of) modals than DO in negatives, etc. CNNs allow us to detect the prototypical usage of and overlap between all forms involved at various points in time, and hence chart the changing affiliations between all forms in the paradigm.
We retrieved all instances of DO and the modals from an 800 million word corpus based on EEBO-TCP and compared their prototypical and overlapping contexts in six 20-year periods between 1580 and 1700. Results indicate that late 16th-century DO prototypically occurred in affirmative sentences with a strong evidential meaning, pragmatically highly similar to factual uses of WILL. This early affirmative prototype aligns well with the earliest emphatic attestations of DO-support, as in *His*

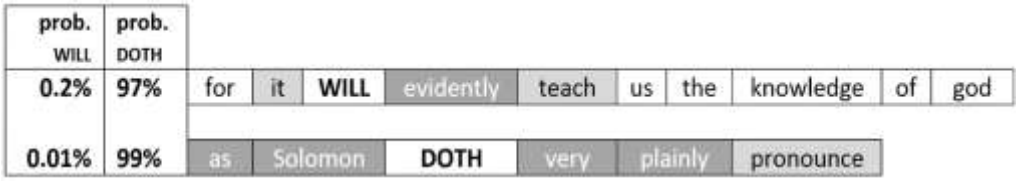


Figure 1 – The model classified both sentences as more likely to contain DOTH *than* WILL*, indicating that factual* WILL *resembles prototypical* DOTH*. The darker the background, the greater the impact of that word on the model's decision.*

sclauyn he dude dun legge (‘his cloak he did lay down’ [c.1300]).

We argue that this early similarity between evidential DO and factual WILL established an association that triggered a cascade of minor assimilations, each of which further increased overlap between DO and WILL and, by extension, the other modals, one form at a time, at the cost of the ties with lexical DO. CNNs are ideal for this purpose, as they naturally model language as a dynamic system of semi-overlapping prototypes.

References

EEBO-TCP. 2003-2019. Early English Books Online. eebo.chadwyck.com & https://www.textcreationpartnership.org/

Ellegård, A. (1953). *The auxiliary do. The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language variation and change* 1, 199-244.

Rumelhart, David & James McClelland. 1986. *Parallel Distributed Processing, Volume I: Explorations in the Microstructure of Cognition: Foundations.* Cambridge: MIT Press.

Vanni, Laurent, Mélanie Ducoffe, Damon Mayaffre, Frédéric Precioso, Dominique Longrée, Veeresh Elango, Nazly Santos, Juan Gonzalez, Luis Galdo & Carlos Aguilar. 2018. Text Deconvolution Saliency (TDS): a deep tool box for linguistic analysis. *56th Annual Meeting of the Association for Computational Linguistics*, July 2018, Melbourne.

Warner, Anthony. 1993. *English auxiliaries. Structure and history*. Cambridge: Cambridge University Press.

Don't talk like that, my dear: The functional profile of *my* in address formulae

Anouk Buyle
(University of Leuven; Research Foundation Flanders)

Keywords: address system; modulation; British English; 20th century; pragmatics

Schedule: We 11.00 Room 11

This paper investigates the use of *my* as part of address formulae. *My* has been defined as a form of orientation by reference to the speaker (Busse 2006), as part of a conventionalized formula (Nevala 2004), or as a marker of intensified intimacy and affection between two interlocutors (Nevalainen & Raumolin-Brunberg 1995). Although Busse (2006) describes the meaning of *my* in a limited number of contexts, a systematic analysis is lacking.

Our corpus consists of eight British English plays published between 1899 and 1912. For each conversational turn, address terms have been identified, as well as the relation between speaker and addressee, as in (1). The speaker-addressee relation is described in terms of the traditional power and solidarity dimensions (following Buyle & De Smet 2018), as well as an additional gender parameter. In order to account for variation between address terms in the same speaker-addressee relation (i.e. in contexts marked by the same power and solidarity values), speech acts with which the address terms occur have been identified and organized according to Leech’s classification of illocutionary functions (1983).

(1) *Yes, I tell you again,* ***my dear****, you have got yourself into a shocking mess. You've got me into a mess, and you've got yourself in a mess.* (Pinero 1899)
[speaker = Quex, addressee = Sophy, power = downward, solidarity = close, gender speaker = male, gender addressee = female, speech act = threatening/complaining, speech act type = conflictive]

The aim of this paper is to identify the functional profile of *my* by comparing the use of *my dear* and *dear*, as well as *my lord/lady* and *your lordship/ladyship*. These alternating pairs have been selected because *my* occurs primarily with these, otherwise unmodified, address terms. Results show that *my* has a complex functional profile. *My dear* is significantly more frequent than *dear* in turns marked as downward and with conflictive speech acts, while *dear* occurs more often in level interactions with collaborative and convivial speech acts. In addition, *dear* is more frequent in female-male interactions, while *my dear* is preferred in male-female interactions. Although results for the comparison between *my lord/lady* and *your lordship/ladyship* are not significant with respect to power and solidarity, *your lordship/ladyship* occurs uniquely in speaker-addressee relations marked as upward, in contrast to *my lord/lady*. Furthermore, while *my lord/lady* is most frequent with collaborative, convivial and conflictive speech acts, *your lordship/ladyship* is the preferred option with competitive speech acts (e.g. requests).

These results demonstrate that, when *my* has impact on the power dimension, the address formula with *my* construes the addressee as less powerful than the speaker. When *my* has impact on the solidarity dimension, the address formula with *my* construes the addressee as a close interlocutor. The analysis for power, solidarity and context-driven variation thus shows that the functional import of *my* varies depending on the address term it modifies, which is consistent with its function as a modulating element.

References

Busse, Beatrix (2006), *Vocative constructions in the language of Shakespeare*, Amsterdam: Benjamins.

Buyle, Anouk & Hendrik De Smet (2018), Meaning in a changing paradigm: the semantics of *you* and the pragmatics of *thou*, *Language Sciences* 68, 42-55.

Leech, Geoffrey (1983), *Principles of pragmatics*, London: Longman.

Nevala, Minna (2004), Accessing politeness axes: forms of address and terms of reference in early English correspondence, *Journal of Pragmatics* 36, 2125–2160.

Nevalainen, Terttu & Helena Raumolin-Brunberg (1995), Constraints on politeness: the pragmatics of address formulae in early English correspondence, in Andreas H. Jucker (ed.), (1995), *Historical Pragmatics. Pragmatic Developments in the History of English. Pragmatics and Beyond New Series 35*, Amsterdam/Philadelphia: Benjamins, 541-601.

Pinero, Arthur W (1899), *The gay Lord Quex: A comedy in four acts*, London: William Heinemann.

Paradigm as constructional compound

Edmond Cane
(Beijing International University, Beijing)

Keywords: paradigm; frame; construction; usage-based, schematization

Schedule: Thu 14.30 Room 5

This paper discusses the notion of the paradigm, its specific functional importance and cognitive foundation through the evaluation of one specific L2 teaching practice. The initial positing underlying this practice has been that paradigm can be viewed relying on the concepts of i) paradigms as "cohesive wholes" (Blevins 2015); ii) constructions (Goldberg, 2006), but at a compound level; iii) the concept of frame (Minsky 1974, Fillmore 1976, etc.), extended towards grammar in a particular way. The concept of frame offers a pattern for the compound level of constructions. It is also assumed that there is gradual usage-based establishment of schematic representations in line with the account and evidence provided by Abbot-Smith & Tomasello (2006), Goldberg (2006), Goldberg et al. (2004), etc.

The approach has been tried with a group of Chinese students learning Albanian as absolute beginners. The special focus has been on the noun feature of ***case***, the verb feature of ***tense*** as well as the ***numeral sequence***, included for comparative evaluation and further hypothetical propositions. The said paradigms are based on a 'pool' of more complex exemplars. Albanian case is complexly bound with number, gender and (in)definiteness. The tense too is bound with person and number.

Following the context of the language situations, the structures have been presented as case- and tense-priming paradigms, having the rest as secondary ones. The positing behind this research is that paradigms are complex networks, which, as provided for by the usage that drives the language acquisition, are complexly entrenched and wired as networks priming one feature with the others in the sub-segmentation, as well as priming another, while this is prevalent in many other respective usage contexts.

Unlike the traditional approach, the linear paradigms have been totally avoided. Instead, there have been pragmatic reductions, as relying only on the situations already entrenched by students – hence, partial paradigms. The students have been guided towards two stages of generalizations, each relying on a critical level of input as checked varyingly.

The teaching approach, the supply and the learning outcome, the students' differentiated acquisition and behavior support the account that the students' competence rests mainly on the paradigm as mental representation, a network of recurrent patterns, brought together as provided for by some piece or lining shared through the members. The oppositions are part of this frame, and emerge, as drawing from either the opposing constructions present or implied in one common situation and/or situations carrying relevant members of the paradigm, provided that such situations are significant and "heavy" enough to survive until this formation process. There is a significant parallel between the level of supply and storage, and emerging representations and behavior, each of them checked appropriately. The evidence supports the view that paradigm "packages information" and enables the mapping, thus identifying one side of the constructional pairing. The individual construction projects the other side. The paradigm also contributes to the interpretation, by bringing in the relevant members in the implied oppositional choice in the occurrence of one constriction in discourse.

References

Abbot-Smith, K. and Tomasello, M. (2006) "Exemplar-learning and schematization in a usage-based account of syntactic acquisition". In *The Linguistic Review* 23, 275-290

Blevins, James P. (2015), "Inflectional paradigms". In: Baerman, Matthew (ed.): *Oxford handbook of inflection*. Oxford: Oxford University Press: 87–111.

Goldberg, AE, Devin Casenhiser and Nitya Sethuraman. 2004 "Learning Argument Structure Generalizations." Cognitive Linguistics 14 3. 289-316.

Goldberg, Adele (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Fillmore Charles (1976). Frame semantics and the nature of language. In *Origins and Evolution of Language and Speech*, ed. by Stevan R. Harnad, Horst D. Steklis, & Jane Lancaster, 20-32. Annals of the NY Academy of Sciences, Vol. 280.

Minksy, Marvin (1975) A Framework for Representing Knowledge (html) AI memo 306, June 1974; also published in *The Psychology of Computer Vision*, P.H. Winston (ed.), McGraw-Hill, NY, 1975.

Aktionsart and verb class as predictors of argumenthood: Behavioral evidence about Italian verbs of motion

Giulia Cappelli, Pier Marco Bertinetto & Alessandro Lenci
(Scuola Normale Superiore, Scuola Normale Superiore & University of Pisa)

Keywords: argumenthood, arguments, adjuncts, goal bias, Aktionsart

Schedule: We 14.30 Room 14

The argumenthood of the locative phrases selected by motion verbs is still a widely disputed subject. By means of two behavioral experiments, we provide new evidence suggesting that motion verbs in Italian, based on their semantic profile, select for different locative PPs, whose argumenthood is non-binary.

Our first experiment adapted the two-task setting by Barbu & Toivonen (2016), requiring no meta-linguistic insight from native speakers, to tap into their intuitions about syntactic and semantic event participants; argumenthood is arguably a syntax-semantics interface phenomenon (Grimshaw 1990, Jackendoff 2002). Our stimuli consisted of 44 Italian motion verbs plus 75 non-motion distractors. We crowdsourced answers from 50 graduate linguistically-naïve native speakers, split into two groups. In the syntactic task, subjects from the first group had to complete each sentence so that it only contained information required for it to make sense (1); blank answers were allowed. In the semantic task, subjects from the second group had to provide a list of participants required for the given event to happen (2).

(1) John went _____ to France
(2) TO GO someone who goes, a destination

Each answer was then manually annotated with its thematic role, and the relative frequency of each thematic role was computed for each verb in the list.

Our second experiment was an acceptability-judgment survey on a 7-point Likert scale, using optionality as a proxy for argumenthood; despite many authors raising concerns against traditional diagnostics (Vater 1978, Tutunjian & Boland 2008, Needham & Toivonen 2011, Przepiórkowski 2017), optionality is still the most widely accepted test in the literature. We recruited 25 linguistically-naïve native speakers and elicited their judgments about the same sentences as in (1) above.
We performed an Exploratory Data Analysis via boxplots, plotting our three measured variables (semantic and syntactic thematic role frequency, optionality score) against verb Aktionsart and verb class as defined by Cennamo & Lenci (2018). Based on auxiliary selection, there exist three classes of motion verbs in Italian: Class 1 selects *be*, Class 3 *have*, Class 2 both.

Semantically, achievements have a strong preference for Goals (and non-locative fillers if belonging to Class 2), activities for Paths and non-locative fillers, accomplishments for Goals and Sources if in Class 1 and for Paths if in Class 2. Syntactically, achievements tend to occur with Goals and non-locative PPs, activities with non-locative PPs, Class 1 accomplishments with Goals and Class 2 accomplishments with any kind of PP. Interestingly, a paired-samples Wilcoxon test revealed that Goals are significantly more salient than Sources both semantically (p = 0.01) and syntactically (p =

0.00095), and Goals more than Paths in the syntactic conditions only (p = 0.02), lending support to the existence of a Goal bias (Nam 2012, Georgakopoulos et al. 2018). Optionality data reveal that achievements occur with adjunct-like PPs, activities with in-between PPs regardless of verb class, Class 1 accomplishments with adjunct-like PPs and Class 2 accomplishments with argument-like PPs. Thus, consistently with Cennamo & Lenci (2018), the argumenthood of PPs occurring with motion verbs in Italian seems to be largely predicted by the interaction between Aktionsart and verb class.

References

Barbu, R. M., & Toivonen, I. (2016). Arguments and adjuncts: at the syntax-semantics interface. *Florida Linguistics Papers 3(1)*, 13-25.

Cennamo, M., & Lenci, A. (2018). Gradience in subcategorization? Locative phrases with Italian verbs of motion. *Studia Linguistica*, 1-29.

Georgakopoulos, T., Härtl, H., & Sioupi, A. (2018). Goal realization: An empirically based comparison between English, German and Greek. *Languages in Contrast*.

Grimshaw, J. (1990). *Argument structure.* Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* New York: Oxford University Press.

Nam, S. (2012). Syntax-semantics mapping of locative arguments. *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*, 473-480.

Needham, S., & Toivonen, I. (2011). Derived arguments. *Proceedings of the LFG11 Conference*, 401-421.

Przepiórkowski, A. (2017). On the argument-adjunct distinction in the Polish "Semantic Syntax" tradition. *Cognitive Studies| Études cognitives (17)*.

Tutunjian, D., & Boland, J. (2008). Do we need a distinction between arguments and adjuncts? Evidence from psycholinguistic studies of comprehension. *Language and Linguistics Compass 2(4)*, 631-646.

Vater, H. (1978). On the possibility of distinguishing between complements and adjuncts. *Valence, semantic case and grammatical relations 1*, 21-45.

Evolutionary aspects of Indo-European gender assignment

Gerd Carling and Briana Van Epps

Schedule: We ~~11.30~~ 11.00 Room 1

<pdf>

Differential Object Marking in Old Sardinian

Michela Cennamo, Francesco Maria Ciconte and Luigi Andriani

Schedule: Thu ~~10.00~~ ~~9.30~~ 12.00 Room 14

This paper explores the differential marking of human O(bjects/DOM in old Sardinian texts from two different areas, Logudoro and Arborea, investigating its constraints, whether *semantic* (reflecting the

Individuation Hierarchy) (Silverstein 1976: 122), *syntactic* (determined by the syntactic position of the O argument in relation to the verb and in the clause) or *pragmatic* (deriving from and functioning as a marking device for topics) (Bossong 1998; Iemmolo 2010; Dalrymple & Nikolaeva 2011: 18, among others).

A preliminary study reveals a change in progress, the fixing of Definiteness as the main parameter determining the occurrence of the preposition *a* (*ad* before vowels), interacting with V1 word order and the need to disambiguate the function of the clause nuclear arguments, A and O, when highly individuated (i.e., proper/human, kinship nouns) (Lombardi 2007).

Proper names, i.e., arguments highest on the Individuation Hierarchy, are already affected by the change in the early attestations of the phenomenon, although with areal differences, reflecting different diachronic stages. In both early vernaculars human Os are marked with the preposition *a(d)* if instantiated by proper names (1a), rarely lacking it (1b). Human common (2a) and kinship (2b) nouns instead less frequently take a dedicated marker, lacking *a(d)* even when [+ DEF], [+ GIVEN] (3a)-(3b) (Putzu 2008: 417):

(1) a. *isse* ***levait*** ***a Gavini*** … (CSPS 21.5)
he took.3SG to Gavini
'He took Gavini …'

b. … *levait* *clesia* ***Arzoco*** (CSMB 100d)
took.3SG church Arzoco
'… the Church took Arzoco'

(2) a. … *vinkit* *priori Johanni* … ***assu*** ***previteru*** (CSNT 300.2)
won.3SG prior Johanni to.DOM-the priest
'The prior Johanni won … the priest'

b. *parthibi* ***a*** ***ffiios de Istefane*** [+ DEF] (CSPS 24, 1-2)
divided.1SG to.DOM sons of Istefane
'I… shared Istefane de Nussas's sons with …'

(3) a. … ***l'ockisit*** (V) ***su seruu uostru*** (O) ***su seruu meu*** (A) (CSPS 110)
him-killed.3SG the servant your the servant my
'… my servant killed your servant'

b. *Parthivi* ***fiios*** ***de Gavini*** … [+ DEF] (CSPS 296. 1-2)
shared.1SG children of Gavin …
'I shared Gavinus Troccus' children …'

In old Arborense, *a(d)* is instead well attested with human common nouns, also if [- DEF], [±REF] in later texts (4):

(4) *si alchuno homini hochirit* ***at*** ***alcuno atteru homini*** (CDLA 4.1)
if some man killed.3SG to.DOM some other man
'If a man killed another man'

Thus, old Sardinian reveals the role played by the notion of Individuation in the initial stages of the grammaticalization of the preposition *a(d)* as a DOM marker, progressively spreading from human proper names, to human, kinship and then common nouns, initially definite, subsequently indefinite. Already in 11th-13th century texts, *a(d)* does not appear to differentiate A from O arguments, but is a marker of high Individuation, its use being sensitive to the notions of definiteness and humanness.

Old Sardinian, therefore, brings interesting novel data on the rise of DOM and its possible diachronic paths, a phenomenon that will be analysed also in relation to word order changes in a V1 system (Wolfe 2015).

References

Bossong, G. 1998. Le marquage differential de l'objet dans les langues d'Europe. In: Feuillet, J. (ed.), *Actance et valence dans les langues de l'Europe*. Berlin: Mouton de Gruyter, 193-258.

Darlymple, M. & Nikolaeva, I. 2011. *Objects and information structure.* Cambridge: Cambridge University Press.

—— 2018. Grammaticalization and changes in argument linking: a case-study from old Logudorese Sardinian. In: Chini, M. & Cuzzolin, P. (eds), *Tipologia, acquisizione, grammaticalizzazione - Typology, acquisition, grammaticalization studies*. Milan: Franco Angeli, 96-117.

Iemmolo, G. 2010. Topicality and DOM: evidence from Romance and beyond. *Studies in Language* 34: 239-272.

Lombardi, A. 2007. Posizione dei clitici e ordine dei costituenti della lingua sarda medievale. In: Ledgeway, A. & Bentley, D. (eds), *Sui dialetti italoromanzi: Saggi in onore di Nigel B. Vincent*. Norfolk: Biddles, 133-148.

Meyer-Lübke W., 1902. *Zur Kenntniss der Altlogudoresischen*. Akademie der Wissenschaft, Band CXLV, Wien.

Putzu, I. E. 2008. Per uno studio dell'accusativo preposizionale in sardo antico: Emergenze dallo spoglio dei *Condaghe di San Pietro di Silki*. In: Lazzeroni, R., Banfi, E., Bernini, G., Chini, M. & Marotta, G. (eds), *Diachronica et synchronica: Studi in onore di Anna Giacalone Ramat*. Pisa: Edizione ETS, 397-428.

Silverstein, M. 1976. Hierarchy of Features and Ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*. Canberra: Australian Institute of Aboriginal Studies, 112-171.

Wolfe, S. 2015. The Old Sardinian Condaghes. A Syntactic Study. *Transactions of the Philological Society* 113.2: 177-205.

Sources

CDLA = Murgia, G. 2016. *Carta de Logu d'Arborea.* Rome: Franco Angeli.

CSMB = Virdis, P. 2003 (ed.). *Il Condaghe di Santa Maria di Bonarcado*. Nuoro: Ilisso.

CSNT = Merci, P. 1992 (ed.). *Il Condaghe di San Nicola di Trullas. Testo Logudorese inedito dei secoli XI-XIII*. Sassari: Delfino.

CSPS = Soddu, A. & Strinna, G. 2013 (eds). *Il Condaghe di San Pietro di Silki*. Nuoro: Ilisso.

Types of expressive binominal constructions: Comparing English and Polish

Bożena Cetnarowska
(University of Silesia in Katowice)

Keywords: expressive meaning, binominal constructions, juxtapositions

Schedule: We 15.00 Room 7

This talk compares types of expressive binominal NPs in English identified by Foolen (2004) with a division of expressive NN combinations in Polish proposed below.

Foolen (2004) divides expressive binominal NPs (EBNPs) in Germanic and Romance languages into two types. Type I binominals, i.e. Foolen's "impression EBNPs" as exemplified in (1), consist of

two NPs, each of them containing indefinite determiners. In the case of Type II EBNPs in (2), referred to as "attitude EBNPs", the first NP contains a definite determiner.

(1) a. *a monster of a tree* b. *an angel of a woman*
(2) a. *that moron of your brother* b. *that fool of his father*

Although Polish is an articleless language, I argue that the division shown in (1-2) is relevant to the classification of expressive NN juxtapositions in Polish (which consist of fully inflected nouns). The Polish juxtapositions in (3) call for a metaphorical interpretation, as do the English EBNPs in (1). The Polish NN combinations in (4) contain an insulting or endearing lexeme and correspond to English "attitude EBNPs" in (2).

(3) a. *drzewo potwór* (tree.NOM.SG monster.NOM.SG) 'a monster of a tree'
b. *kobieta anioł* (woman.NOM.SG angel.NOM.SG) 'an angel of a woman'
(4) a. *brat idiota* (brother.NOM.SG idiot.NOM.SG) 'an idiot of a brother'
b. *dentysta sadysta* (dentist.NOM.SG sadist.NOM.SG) 'a sadist of a dentist'

According to Foolen (2004: 93), Type I EBNPs in English occur in the predicative position (5a) while Type II EBNPs - in argument positions (5b).

(5) a. *He's a bear of a man.*
b. *That sadist of a man tortured children*.

Data from the National Corpus of Polish (NKJP) are provided to show that, similarly to English Type II EBNPs, Polish "attitude" expressive NNs occur in argument positions. They can be accompanied by the demonstrative *ten* 'this', which signals emotional distance (in 6).

(6) *Ten idiota dyrektor wyrzucił mnie z pracy!* 'That idiot of a manager fired me!'

Metaphorical (Type I) expressive NN juxtapositions in Polish occur in predicational sentences with the copula pronoun *to* 'it' (in 7a) while "attitude" (Type II) combinations sound slightly less felicitous in such sentences (see 7b).

(7) a. *Helena to kobieta anioł*. 'Helen is an angel of a woman.'
b. ?*Henryk to kierowca cham*. 'Henry is a lout of a driver.'

Polish metaphorical juxtapositions can be encountered in argument positions (in 8a). One can also come across English Type I EBNPs in such positions (contra Foolen 2004), as in (8b) (culled from COCA).

(8) a. *Spotkał kobietę anioła i się zakochał*. 'He met an angel of a woman and fell in love.'
b. *a cute innocent (…) angel of a child might see your posting*

Finally, it is argued that the application of the "IS A Condition" (Allen 1978) to Polish expressive combinations can throw some light on the controversy concerning the issue of headedness of English EBNPs (Aarts 1998, Kim and Sells 2015).

References

Aarts, Bas (1998), Binominal noun phrases in English, *Transactions of the Philological Society* 96 (1), 117–158.

Allen, Margaret R. (1978), Morphological investigations. Storrs, CT: University of Connecticut. PhD dissertation.

Foolen, Ad (2004), Expressive binominal NPs in Germanic and Romance languages, in G. Radden and K.-U. Panther (eds.), *Studies in Linguistic Motivation*, Berlin: Mouton de Gruyter, 75–100.

Kim, Jong-Bok and Peter Sells (2015), English binominal NPs: A construction-based perspective, *Journal of Linguistics* 51(1), 41–73.

COCA = Corpus of Contemporary American English. Available online at: https://corpus.byu.edu/coca
NKJP = Narodowy Korpus Języka Polskiego (National Corpus of Polish) . Available online at: http://nkjp.pl

Distributional, semantic and functional properties of adversative pragmatic markers in Italian: A corpus-based approach

Doriana Cimmino
(University of Florence)

Keywords: adversative pragmatic markers, spoken Italian, information units, textual functions, left and right peripheries

Schedule: Thu ~~15.00~~ 14.30 Room 3

Pragmatic markers can perform textual and dialogic functions: i.e. they can link portions of texts (utterances, paragraphs, etc.) as well as guide the interaction with the addressee. The complete description of their uses is still debated (see Degand et al. 2013; Fedriani/Sansò 2017) and it has been recently suggested that distributional factors play a crucial role (Beeching/Detges 2014).

This research focuses on the most frequent Italian adversative pragmatic markers (APMs) *ma* and *però*, which show numerous asymmetries (see, e.g., Bazzanella 1995, Mauri 2008). Distributionally, contrary to *ma*, *però* can occur in final, detached positions (Ferrari 2005). Semantically, they share a core opposition meaning ('but'), only *ma* can convey reformulation ('actually'), and *però* is preferred to express counter-expectation ('though'). From a functional point of view, it is maintained that only *ma* can be used as an interactive marker in initial positions (Sabatini 1997). These asymmetries still need to be checked against real data.

The study aims at providing a corpus-based account of the distributional, semantic and functional properties of APMs. For the research, 1'061 occurrences of *ma* and 525 occurrences of *però* have been extracted from DB-IPIC (Panunzi/Gregori 2012), a corpus of spontaneous speech tagged with information structure (ca. 125'000 words). The analysis has been conducted considering three parameters:

a) semantic uses, namely, opposition, reformulation, and counter-expectation (following Marconi/Bertinetto 1984);
b) distribution in prosodically identified information units (Cresti 2000, Cresti & Moneglia Forthcoming);
c) textual functions, evaluating the linking with left and right co-text in a 'discourse perspective' (Lang 2000), within and across dialogical turns.

Preliminary results emerged which are partially inconsistent with the above-mentioned literature. On the one hand, the described semantic and distributional asymmetries between *ma* and *però* are not quantitatively relevant. In fact, the specialized reformulation meaning of *ma* and counter-expectation meaning of *però* are less used and associated to dedicated positions: internal (7%) and final detached (5%), respectively. Moreover, both *ma* and *però* tend to occur in initial positions (68% and 61%, respectively). On the other hand, information and functional asymmetries have been overlooked.

When *ma* occurs in initial informationally detached positions, it performs an interactive function, while when occurring in non-detached initial positions, it can have a textual function. Differently, *però* tends to keep a textual function even in initial detached positions, and, unexpectedly, in final detached positions as well. Consider the example in (1), where *però* is used to contest an interlocutor's statement regarding the beauty of Apulia (single slash signals an information unit break, double slash signals an utterance break):

(1) eh ma /da quest'altra parte /no /*però* /eh // (DB-IPIC, ifamcv 17, 263)
'Yes but on the other side no *però* eh'

I will illustrate the above-mentioned information and functional asymmetries with audio files and F0 tracks of the utterances containing APMs in the corpus. This will allow me to discuss the correlations between information units and textual functions and to contribute to the mapping and qualitative description of APMs in left and right peripheries (Traugott 2012).

References

Bazzanella, Carla (1995), I segnali discorsivi, in L. Renzi, G. Salvi, and A. Cardinaletti (eds), (1995) *Grande grammatica italiana di consultazione*, III, Bologna: Il Mulino, 225-257.

Beeching, Kate & Detges, Ulrich (eds), (2014), *Discourse Functions at the Left and Right Periphery: Cross-linguistic Investigations of Language Use and Language Change*, Leiden: Brill.

Cresti, Emanuela (2000), *Corpus di italiano parlato*, Firenze: Accademia della Crusca.

Cresti, Emanuela & Moneglia, Massimo (Forthcoming), The Discourse Connector according to the Language into Act Theory: data from IPIC Italian, in E. Bidese, J. Casalicchio, and M. Moroni, Manuela (eds), La linguistica vista dalle Alpi. Teoria, lessicografia e multilinguismo / Linguistic views from the Alps. Language Theory, Lexicography and Multilingualism, Frankfurt am Main: Peter Lang.

Degand, Liesbeth, Cornillie Bert & Pietrandrea Paola (eds), (2013), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam: John Benjamins.

Fedriani, Chiara & Sansò, Andrea (eds), (2017), *Pragmatic Markers, Discourse Markers and Modal Particles: New perspectives*, Amsterdam: John Benjamins.

Ferrari, Angela (2005), Connettivi e struttura del testo. Oltre la semantica lessicale, in I. Korzen (ed.), (2005) *Lingua, cultura e intercultura: l'italiano e le altre lingue*, Copenaghen: Samfundslitteratur Press, 191-204.

Lang, Ewald (2000), Adversative connectors on distinct levels of discourse: a re-examination of Eve Sweeter's three-level approach, in Couper-Kuhlen E. & Kortmann B. (eds), *Cause-Condition-Concession-Contrast. Cognitive and Discourse Perspectives*, Berlin / New York: Mouton de Gruyter, 235-256.

Marconi, Diego & Bertinetto, Pier Marco (1984), Analisi di ma, *Lingua e Stile*, 19, 223-258.

Mauri, Caterina (2008), *Coordination Relations in the Languages of Europe and Beyond*, Berlin / New York: Mouton de Gruyter.

Panunzi, Alessandro & Gregori, Lorenzo (2012), DB-IPIC. An XML Database for the Representation of Information Structure in Spoken Language in H. Mello, A. Panunzi & T. Raso (eds), (2012), *Pragmatics and Prosody, Illocution, Modality, Attitude, Information Patterning and Speech Annotation*, Firenze: Firenze University Press, 133-150.

Sabatini, Francesco (1997), Pause e congiunzioni nel testo. Quel *ma* a inizio di frase... in I. Bonomi (ed.), (1997), *Norma e lingua in Italia: alcune riflessioni fra passato e presente*, Milano: Istituto Lombardo-Accademia di Scienze e Lettere, 113-146.

Traugott, Elizabeth Closs (2012), Intersubjectification and clause periphery, *English Text Construction*, 5(1), 7-28.

Grammaticalization of a past perfective marker in North-Eastern Neo-Aramaic from a verb of movement and its cross-linguistic parallels

Eleanor Coghill

Schedule: Sat 12.30 Room 1

<pdf>

Progress on Nominal classification: a Canonical Typology

Greville G. Corbett & Sebastian Fedden
(University of Surrey; University of Surrey & Université Paris 3)

Keywords: Gender, Classifiers, Canonical Typology, Features, Nominal classification

Schedule: Thu 16.30 Room 6

Fascinating new systems of nominal classification keep being found, but the tools for analysis have not kept pace. We therefore propose a typology of nominal classification, encompassing gender and the various classifiers. Earlier it made sense to oppose gender and classifiers (Dixon 1982), but the opposition cannot be maintained. Miraña has characteristics of gender *and* of classifiers (Seifart 2005); Reid's (1997) account of Ngan'gityemerri provides further evidence against a sharp divide, since classifiers can grammaticalize into gender, through intermediate types.

Relinquishing the opposition gender vs classifiers allows a clearer picture of the possibilities. We pull apart traditional gender characteristics, and traditional classifier characteristics, and see that these characteristics combine in many ways. This motivates a canonical perspective: we define the notion of canonical gender, and use this idealization as a baseline from which to calibrate the theoretical space of nominal classification. This allows us to situate the interesting combinations we find.

According to the Canonical Gender Principle "each noun has a single gender value." (Corbett & Fedden 2016: 503; cf. Dixon 1982). Under this principle there are two specific criteria:

Criterion "constant": canonically a noun takes the same gender agreement in all domains; nouns taking different agreements (hybrids) are non-canonical. German comes close to being canonical, while Russian has many hybrids.

Criterion "lexical": gender can be read unambiguously off the lexical entry. Nearest to canonical are strictly semantic systems, e.g. Bagvalal (male human / female human / other, Kibrik 2001: 64–66); then we find systems like Mawng (Singer 2016), where recategorization is readily available. Furthest from canonical are traditional numeral classifiers which can highlight different semantic aspects of referents.

Canonical agreement, on which canonical gender is based, contributes helpful criteria:

Criterion "obligatory": since canonical agreement is obligatory, this is a component of a canonical gender system, as in languages like Russian. Less canonical are those where agreement can be optional, as in Ngan'gityemerri (Reid 1997). And classifiers of various types are frequently optional.

Criterion "obligatory values": in addition to agreement being obligatory, canonical systems use the most specific feature value associated with the controller; such systems include Mian. Less canonical, since it exhibits 'superclassing', is Jingulu (Meakins & Pensalfini 2016).

Criterion "orthogonal: as with any morphosyntactic feature, gender is canonical to the extent that it is orthogonal to parts of speech. We find systems where almost all parts of speech show gender

agreement (Archi), and at the other extreme, systems where agreement is limited: in North Ambrym only the relational classifiers show relevant inflection (Franjieh 2016).

As we calibrate carefully, using these criteria, traditional gender and traditional classifier languages are close in some respects, more distant in others. This is exactly right: Russian, Archi, Mawng and Ngan'gityemerri are profoundly similar in parts of their nominal classification system, and profoundly different in others. Our typology helps reveal the great diversity of systems (greater than suspected even a few years ago), and the shared phenomena that play out as variations on similar themes.

Acknowledgment: the support of the ESRC (UK) under grant ES/R00837X/1 'Optimal categorisation: the origin and nature of gender from a psycholinguistic perspective' is gratefully acknowledged.

References

Corbett, Greville G. & Sebastian Fedden. 2016. Canonical gender. *Journal of Linguistics.* 52.495-531.

Dixon, R. M. W. 1982. Nominal classification. In: R. M. W. Dixon *Where have all the adjectives gone? and other essays in semantics and syntax*, 157–233. Berlin: De Gruyter.

Franjieh, Michael. 2016. Indirect possessive hosts in North Ambrym: Evidence for gender. *Oceanic Linguistics* 55.87-115.

Kibrik, Aleksandr E. (ed.) 2001. *Bagvalinskij jazyk: Grammatika: Teksty: Slovari.* Moscow: Nasledie. [co-editors K. I. Kazenin, E. A. Ljutikova & S. G. Tatevosov.]

Meakins, Felicity & Rob Pensalfini. 2016. Gender bender: Disagreement in Jingulu Noun Class Marking. In: Felicity Meakins & Carmel O'Shannessy (eds) *Loss and renewal: Australian languages since colonisation*, 425-450. Berlin: Mouton de Gruyter.

Reid, Nicholas. 1997. Class and classifier in Ngan'gityemerri. In: Mark Harvey & Nicholas Reid (eds) *Nominal classification in Aboriginal Australia*, 165–228. Amsterdam: John Benjamins.

Seifart, Frank. 2005. The Structure and Use of Shape-based Noun Classes in Miraña (North West Amazon). Nijmegen: Radboud University Ph.D. dissertation.

Singer, Ruth. 2016. *The Dynamics of Nominal Classification: Productive and Lexicalised Uses of Gender Agreement in Mawng* (Pacific Linguistics 642). Berlin: De Gruyter Mouton

What you see is not what you get: Implicational universals in diachronic perspective

Sonia Cristofaro

Schedule: Sat 12.00 Room 6

<pdf>

"Three [or four] is company": Number-marked pronouns in Western Oceanic languages

Carlo Dalle Ceste
(Australian National University)

Keywords: morphology, Oceanic, number, pronouns, grammaticalization

Schedule: Thu 9.00 Room 6

Several Western Oceanic languages of western Melanesia display a rich system of number distinctions in pronouns. For instance, Sursurunga has singular, dual, trial, quadral, and plural forms. Almost uniformly, trial and quadral pronouns reflect the Proto Oceanic numerals *tolu "three", and *pat[i] "four" suffixed to the plural pronouns (Lynch et al. 2002: 35).

A common feature of this system is "that the highest number below plural [...] does not denote a specific number but serves as paucal" (Ross in prep.: 37). This is reflected by Sursurunga quadral pronouns, which "refer more frequently to plural number" (Hutchisson 1986: 10), and Siar trial forms, which function as paucals (see (1)).

(1)	***Ap***	***matòl,***	***matòl***	***k-i***	***mur***	***ap***	***matò***		
	and	**1PAU.EXCL**	**1PAU.EXCL**	FOC-3SG	follow	and	**1PAU.EXCL**		
	wòt	***ó-n***	***ép***	***kuk***	***t'an***	***lakman,***	***ap***	***mèt***	***bòrbòr***
	come	OBL-POSS	ART	crab	PREP	village	and	**1PL.EXCL**	sleep

"And we, we followed and came to the village with the crabs and we (all) slept."

Frowein (2011: 143)

In (1), Siar trial forms refer to the "the narrator's own group". When the latter and another group "come together [...], the subject is represented by the plural pronoun *mèt* […], referring to all the participants" (Frowein 2011: 143).

Corbett (2000: 25–30) argues that in the languages showing both trial and quadral numbers, the first works as paucal and the second as extended paucal. However, as Palmer (2012: 444) observes, diachronic analysis suggests that the paucal meaning of trials and quadrals developed from a different configuration, reflected by Minigir, which has no quadral, and whose trial works as a "pure trial rather than a paucal" (Van Der Mark 2007: 240).

Further evidence comes from Hoava, where pronouns only distinguish between synthetic singular and plural (Davis 2003: 46). The obligatory dual is formed analytically by adding the numeral for two (*karu*) after the plural forms. This pattern is virtually unlimited, involving numerals greater than two. The resulting number-marked pronouns (greater than duals) are not obligatory, being optionally replaced by the simple plurals.

In my presentation, I will provide a typology of the different systems of number marking in pronouns, focusing on the use of these forms in context. On the one hand, some languages lacking distinct trial and quadral numbers point to the replacement of the older plural pronouns with trial and/or quadral forms, suggesting that this system of number distinctions was more common than previously acknowledged. On the other, the majority of Western Oceanic languages reflect a parent language distinguishing only two (synthetic) numbers: singular and plural. Like Hoava, speakers of the parent language could add numerals to the plural pronouns to specify the number of participants in the event/action. Possibly due to frequency, this pattern of number specification triggered the grammaticalization of the numerals between "two" and "four", resulting in additional distinctions. At some point, the semantic features underlying these innovative numbers became opaque, leading speakers to generalize the highest number below plural as paucal, which, in some cases, was further extended to general plurality.

References

Corbett, Greville (2000), *Number*, Cambridge: Cambridge University Press.

Davis, Karen (2003), *A grammar of the Hoava language, Western Solomons*, Canberra: Pacific Linguistics.

Frowein, Friedel Martin (2011), *A Grammar of Siar, an Oceanic language of New Ireland Province, Papua New Guinea*, Melbourne: La Trobe University, PhD dissertation.

Hutchisson, Don(1986), Sursurunga pronouns and the special uses of quadral number, in U. Wiesemann (ed), *Pronominal systems*, Tübingen: Gunter Narr, 1-20.

Lynch, John & Malcolm Ross & Terry Crowley (2002), *The Oceanic languages*, Richmond, Surrey: Curzon Press.

Palmer, Bill (2012), Nominal number in Meso-Melanesian, in M. Butt and T. Holloway King (eds), *Proceedings of the LFG12 conference*, Stanford: CSLI Publications, 440–460.

Ross, Malcolm (in prep), Western Oceanic revisited, Draft only.

Van Der Mark, Sheena (2007), *A Grammar of Vinitiri - an Austronesian Language of Papua New Guinea*, Melbourne: La Trobe University, PhD dissertation.

Getting its productive on: The recent life of the English *get one's* X *on* construction

Robert Daugs
(University of Kiel)

Schedule: Fri 12.30 Room 7

Constructions are learned pairings of form and meaning that show somewhat unpredictable behavior regarding either (Goldberg 1995); moreover, constructions are often productive, i.e. speakers can creatively extend their use (Hilpert 2014). The English semi-schematic [GET POSS X *on*] construction (GXO) represents such a case, see (1)–(3).

(1) But I am ready to relax by the fire and get my real simple magazine on. [Daniels and Shure, Parks and Recreation, 2009]

(2) On a side note, live-action ads seem to be the rage at the moment. Last week Need for Speed: Most Wanted and Halo 4 got their real on, and now All-Stars Battle Royale follows suit. [GloWbE, NFSUnlimited.net, 2012]

(3) One for the diary, if you feel like getting your Frenchy on, get ready for the Bastille Day French Festival. [NOW, 3AW, 2017]

The examples indicate that the GXO meets the criterion of unpredictability in that its meaning exhibits non-compositionality. The construction is also formally constrained, as it cannot convey this meaning when followed by a complement (cf. *get my mind on the game*). Based on these preliminaries, the present paper utilizes both corpus- (e.g. COCA [Davies 2008], COHA [Davies 2010], TV Corpus [Davies 2019]) well as web-based data (e.g. Google, imsdb.com, azlyrics] in order to show that the GXO has become increasingly popular over the past two decades, occurring with a wide variety of constructional slot fillers. Throughout the 20th century, the construction was almost exclusively used in its literal sense 'to put on something concrete (especially clothes)', as in get my [*shoes*/*helmet*/*makeup*]$_{N}$ on, which was then gradually extended to more figurative uses broadly denoting 'to adopt a certain mindset and act accordingly', as in get your [*groove*/*drink*/*research*]$_{N}$ on. This metaphorical extension is assumed to have facilitated the GXO's ability to coerce various lexical items (cf. Diewald 2006). Also, the construction does not seem to impose many constraints on new coinages regarding their 'sensicality' (Suttle and Goldberg 2011), as it readily overrides the meaning of nominal as well as verbal and adjectival slot fillers, changing the categories of the latter by means of conversion (cf. Booij and Audring 2018) – see e.g. *get your* [*jog*/*teach*/*eat*]$_{V}$ on and *get your* [*happy*/*interpersonal*/*fatherly*]$_{ADJ}$ on. As a result, the GXO has become rather productive. While the productivity of a construction is typically assessed in terms of its type frequency, its expansion or its

potential (Baayen 2009). The focus here, however, will not only rest on these quantitative measures, but also on how metaphorical extension interacts with category change to boost productivity and what actually induces the increase in the first place.

References

Baayen, R. Harald (2009), Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling and M. Kytö (eds), (2009), *Corpus Linguistics: An international handbook*, vol. 2, Berlin: Mouton De Gruyter, 899–919.

Booij, Geert and Jenny Audring (2018), Category change in construction morphology. In K. Van Goethem, M. Norde, E. Coussé and G. Vanderbauwhede (eds), *Category change from a constructional perspective*, Amsterdam: John Benjamins, 209–228.

Davies, Mark (2008), *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available online at https://www.english-corpora.org/coca/.

Davies, Mark (2010), *The Corpus of Historical American English (COHA): 400 million words, 1810-2009*. Available online at https://www.english-corpora.org/coha/.

Davies, Mark (2019), *The TV Corpus: 325 million words, 1950-2018*. Available online at https://www.english-corpora.org/tv/.

Diewald, Gabriele (2006), Context types in grammaticalization as constructions. Constructions, sv1. [https://journals.linguisticsociety.org/elanguage/constructions/article/view/24.html]

Goldberg, Adele (1995), C*onstructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Hilpert, Martin. (2014), *Construction grammar and its application to English*. Edinburgh: Edinburg University Press.

Suttle, Laura and Adele Goldberg (2011), The partial productivity of constructions as induction. *Linguistics* 49(6), 1237–1269.

The noun-verb distinction in Bunun

Rik De Busser
(National Chengchi University)

Keywords: Austronesian languages, Bunun, categorization, word classes, noun-verb distinction

Schedule: Thu 15.00 Room 6

This presentation discusses the nature of the noun-verb distinction in Bunun, an Austronesian language of Taiwan, where many – possibly most – lexical roots can occur in both nominal and verbal slots.

The definition of lexical categories has been a long-standing controversy in Austronesian languages (see e.g. Starosta, Pawley, and Reid 1982; Gil 1994; Kaufman 2009). To a varying degree, problems occur when one tries to apply grammatical criteria traditionally associated with word class delineation to especially Philippine-type Austronesian languages. As a consequence, it has been suggested that certain Austronesian languages have no word class distinctions (Tongan, see Broschart 1997; Riau Indonesian, see Gil 2009); that lexical roots are precategorial and word class differentiations are only realized on the grammatical level (Tagalog, see Foley 1998); or that there is a systemic mismatch between lexical and grammatical word classes (Tagalog, see Himmelmann 2008). None of the three models is easily applicable to Bunun, or any of its dialects.

Based on a corpus of narrative texts and data from directed elicitation, I will give an overview of fundamental problems associated with defining the boundaries between the categories of noun and verb in Bunun, taking into account the Austronesian context and recent developments in functional typology (e.g. Plank 2016). I will pay special attention to the influence that relatively minor morphological or syntactic discrepancies between Takbanuaz and Takivatan, two closely related dialects of Bunun, have on establishing categorial boundaries in a cross-dialectal perspective. I will suggest a solution for based on prototype clusters of salient semantic, formal and distributional criteria.

(This research has been funded by the Taiwan Ministry of Science and Technology, grant number MOST 107-2410-H-004-118-MY2)

References

Broschart, Jürgen. 1997. Why Tongan does it differently: categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1–2, 123–165.

Foley, William Auguste. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the *LFG98 Conference, Workshop on Voice and Grammatical Functions in Austronesian Languages*. Brisbane: University of Queensland, 1 July 1998.

Gil, David. 1994. The structure of Riau Indonesian. *Nordic Journal of Linguistics* 17(2), 179–200.

Gil, David. 2009. How much grammar does it take to sail a boat? In Geoffrey Sampson, D. Gil & P. Trudgill (eds.), *Language Complexity as an Evolving Variable*, 19–33. Oxford: Oxford University Press.

Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In Peter Austin & Simon Musgrave (eds.), *Voice and Grammatical Relations in Austronesian Languages*, 247–293. Stanford: CSLI Publications.

Kaufman, Daniel. 2009. Austronesian typology and the nominalist hypothesis. In K. A. Adelaar & A. K. Pawley (eds.), *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust*, 197–226. Canberra: Pacific Linguistics.

Plank, Frans. 2016. Of categories: Language-particular – comparative – universal. *Linguistic Typology* 20(2), 297–298.

Starosta, Stanley, Andrew K. Pawley & Lawrence A. Reid. 1982. The evolution of focus in Austronesian. In A. Halim, L. Carrington & S. A. Wurms (eds.), *Papers from the Third International Conference on Austronesian Linguistics, Vol. 2. Tracking the Travellers*, 145–169. Canberra: Pacific Linguistics.

Do strong and weak forms of the same verb differ in meaning?

Isabeau De Smet
(University of Leuven and FWO)

Keywords: historical linguistics, morphology, semantics, corpus linguistics, strong and weak verbs

Schedule: We 15.00 Room 2

Why some strong verbs (e.g. *zingen-zong-gezongen* ‘sing-sang-sung’) become weak (e.g. *spelen-speelde-gespeeld*, ‘play-played-played’) and others do not, is determined by (among other factors) their frequency and their phonological shape (i.a. Carroll et al. 2012; De Smet & Van de Velde forthc.). This ‘weakening’ is a gradual process: most verbs do not suddenly only show weak forms

anymore. Usually a transition period, in which the verbs shows both strong and weak forms, of at least one century takes place, before the verb is completely weak.

In our study, we zoom in on this transitional period for verbs in Dutch, investigating for different verbs which factors jointly determine the choice for either the weak or the strong form. For each different verb, the usual suspects among the variables, namely frequency and phonological shape remain more or less constant. This means that there may be room for other factors to kick in.

In the past, semantics have been largely shunned in the diachronic investigation of the determinants of strong and weak verb morphology. Still, psycholinguistic research has discovered that semantics do interact with past tense morphology: the association with the semantics of a strong verb facilitates the strong inflection for other (nonce) verbs (i.a. Ramscar 2002). Baayen & Moscoso del Prado Martin (2005) and Tabak et al. (2005) also show that the strong inflection as a whole shows a different semantic profile than the weak inflection, e.g. concerning aspect and valency. Finally, when looking at the choice between irregular and regular forms for alternating verbs like *burn* and *learn* in English, Quirk (1970) and Levin (2006) suggest that aspect plays a role with regular forms expressing more durative meaning and irregular forms more punctual meaning.

We gather weak and strong forms of Dutch verbs from a collection of historical corpora and zoom in on those verbs that show more than 25% and less than 75% weak forms. Part of these are strong or weak verbs in transition to becoming entirely weak or strong, another part is constituted by cases in which the variation is short-lived, and yet another part are cases in which the variation appears to linger on for centuries, without one inflection becoming distinctly prevalent. Especially in the latter case, the verbs can be observed to develop a meaning specialization. Examples are *plegen* (‘commit’ is weak, ‘used to’ is strong) or *stijven* (‘become stiff’ is weak, ‘make stiff’ is strong), showing a meaning distinction that was not ancestral to the period of variation.

Firstly, we will try to find out whether there are other cases of meaning specialization among these verbs. Secondly, we will investigate whether there are any general trends in those meaning specializations. In line with the previous research, we expect to find aspectual differences (more punctual meanings for strong forms and more durative meanings for weak forms) and differences in valency, but we will also look for differences in concreteness and differences between prefixed and non-prefixed verbs.

References

Baayen, Rolf Harald and Moscoso del Prado Martin, Fermin (2005), Semantic density and past-tense formation in three Germanic languages, *Language* 81, 666-698.

Carroll, Ryan and Svare, Ragnar and Salmons, Joseph (2012), Quantifying the evolutionary dynamics of German verbs, *Journal of Historical Linguistics* 2, 153-172.

De Smet, Isabeau and Freek Van de Velde (Forthcoming), Reassessing the evolution of West Germanic preterite inflection, *Diachronica*.

Levin, Magnus (2009), The formation of the preterite and the past participle, in Günter Rohdenburg and Julia Schlüter (eds) (2009), *One language, two grammars? Differences between British and American English*, Cambridge: Cambridge University Press, 60-85.

Ramscar, Michael (2002), The role of meaning in inflection: Why the past tense does not require a rule, *Cognitive Psychology* 45: 45-94.

Tabak, Wieke, Robert Schreuder and R. Harald Baayen (2005), Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch, in M. and S. Kepser (eds) (2005), *Linguistic Evidence Mouton*, 529-555.

Quirk, Randolph (1970), Aspect and variant inflexion in English verbs, *Language* 46, 300-311.

Where function meets variation

Hendrik De Smet
(KU Leuven)

Keywords: complementation; functional motivation; isomorphism; preposition; variation

Schedule: We 17.00 Room 14

Variation occurs when a language has two or more ways of achieving the same communicative goal. Traditionally, cases of variation have been approached in very different ways by two different strands of linguists. Variationists assume that variation is natural and are happy to think of it is a common phenomenon (e.g. Labov 1972; Tagliamonte 2012). Functionalists assume that variation is anomalous and expect that it is rare and – where it occurs – short-lived (Bolinger 1968; Dik 1988). The two views seem contradictory, yet both have good credentials. Variationists can argue that variation is to be expected because speech communities are never completely homogeneous. Functionalists can argue that variation should be abnormal because it makes language processing less efficient.

These conflicting views can be reconciled by considering how variation arises historically. If we think of formal categories as occupying a continuous region of functional space, variation arises when one formal category marginally extends into the functional domain of another. The resulting variation may in itself not be motivated, but its emergence (and subsequent maintenance) is typically sanctioned by analogical extensions that are. Motivation then resides in the fact that, by analogy, similar treatment is given to similar things. Because this is a form of cognitive efficiency (cf. Rosch 1978), it is fundamentally a functional principle.

This view is supported by a case study into the prepositional complements following emotion adjectives in English. For example, *upset* can combine with prepositional complements headed by *over*, *at* or *with*, as in (1a-c), as well as *by* and *about*, all potentially marking the external source of emotion.

(1) a. The whole establishment had, in fact, been *upset **over** her disappearance*. (BNC)
b. and she, Donna, was dreadfully *upset **at** the suggestion*. (BNC)
c. Paul Ritchie was *upset **with** a County Council decision to withdraw pay for trainees at the centre*. (BNC)

In a step-wise procedure, using data from the BNC, 56 frequent emotion adjectives have been identified, each taking prepositional complements headed by one or several of a list of 7 prepositions. Following extraction and analysis of all possible combinations of these adjectives and prepositions, a data set of just over 17,000 instances has been retained.

The data are consistent with the idea that variation (as Variationists maintain) is omnipresent. For example, with 7 prepositions, 21 binary alternations are possible, all of which are actually attested for at least one adjective in the data set. At the same time, the synchronic availability and historical development of specific variants is (as Functionalists maintain) also internally motivated, typically by analogical relations. There is straightforward evidence of semantic patterning in the data, in that semantically similar adjectives tend to select the same types of prepositional complements. Most strikingly, further follow-up with diachronic data from the Hansard Corpus reveals that variability is a stable property of some adjectives (even if there is turn-over in the specific complements involved). This, paradoxically, confirms that variation is not a functional anomaly but a linguistically motivated part of the system.

References

Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2: 119-127.
Dik, Simon C. 1988. Isomorfisme als functioneel verklaringsprincipe. *GLOT* 11: 87-106.
Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch and Barbara B. Lloyd (eds), *Cognition and categorization*, 27-48. Hillsdale, NJ: Lawrence Erlbaum.
Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, observation, interpretation*. Oxford: Wiley-Blackwell.

Construction or allostruction?
On the constructional status of ditransitive argument structures in German

Hilde De Vaere, Ludovic De Cuypere & Klaas Willems
(Ghent University; Ghent University and VUBrussel; & Ghent University)

Keywords: German, ditransitive argument structure, construction grammar, allostructions, corpus linguistics

Schedule: We 11.00 Room 7

Research context: There are currently several competing definitions of 'construction'. A construction is either (I) narrowly defined as a form-meaning pairing in its own right if its meaning is not compositionally derivable from the lexical items that instantiate it (Goldberg 1995); or it is (II) broadly defined as any stored or learned pattern, even if semantically predictable from its parts, as long as it occurs with sufficient frequency (Goldberg 2003, 2006). In English, both the Double Object Construction (DOC) and the Prepositional Object Construction (POC) are constructions on both accounts: both alternants have high token frequencies and are either associated with different meanings (common in Construction Grammar approaches) or with probabilistic factors that constrain the choice between DOC and POC (Bresnan et al. 2007, 2010). Typologically, English DOC has neutral alignment, whereas POC has indirective alignment (cf. Malchukov et al. 2010, Haspelmath 2013, 2015), and their canonical word orders differ. By contrast, in German both alternants – either with a Dative NP (1) or a PP (2) – have indirective alignment and both display variable word order of RECIPIENT and THEME, compare:

(1) a. *Das Hormon Insulin* *sendet* *dem Gehirn* *Sättigungssignale.*
The hormone insulin sends the:DAT brain saturation signals
'The hormone insulin sends the brain saturation signals.'

b. *Den größten Teil ihres Gehalts* *schickt* *sie* *ihren Eltern.*
The largest part of her salary sends she her parents:DAT
'Most of her salary she sends to her parents.'

(2) a. *Diese Nervenenden* *senden* *Schmerzsignale* *an unser Gehirn.*
These nerve endings send pain signals to our:ACC brain
'These nerve endings send pain signals to our brain.'

b. *Wir* *werden* *an alle Wahlberechtigten* *einen Infobrief* *schicken.*
We will to all:ACC voters an information letter send
'We will send an information letter to all voters.'

c. *Seitdem die Anlage installiert wurde,* *schickt* *sie* *fleißig Daten* *zu Andreas Wulff.*
Since the installation was installed, sends she busy data to:DAT A.Wulff
'Since the installation was installed, it has been sending a lot of data to A.Wulff.'

Research questions: This paper investigates whether the argument structures in the German ditransitive alternation are constructions in the sense of (I) and/or (II) or whether it is more appropriate to adopt a single constructional pattern with two variants. The aim is to show that combining a corpus-based analysis with an approach based on typological research of the ditransitive pattern is able to accommodate the data in a coherent way.

Methodology: We conducted a synchronic study of alternating ditransitive argument structures with the verbs *geben, schicken* and *senden* extracted from the Mannheim DeReKo corpus. A statistical analysis was performed on the basis of 1200 sentences per verb, equally divided over both alternants. The sentences were annotated for 18 morphosyntactic, semantic, and pragmatic factors, and differences in frequency of occurrence were determined for the three verbs.

Results: With regard to the narrow definition (I), we found that neither alternant is systematically associated with a particular meaning. Both alternants occur with different verb senses (concrete, abstract, propositional, cf. Bresnan et al. 2007) as well as different word orders. With regard to the broad definition (II), we concur with Stefanowitsch (2011) that frequency and entrenchment are no defining criteria for construction-hood. Preference for one alternant over the other depends on multiple factors and the individual verb (*geben* vs. *senden/schicken*). We therefore propose an account of the alternation in terms of a general three-participant AGENT-THEME-GOAL construction (cf. Bickel 2011) which contrasts with other constructions in German and which is instantiated as either of two 'allostructions' (Cappelle 2006). The GOAL argument is semantically and formally underspecified (cf. De Vaere et al. 2018). The construction's general meaning is enriched in various, quantifiable contexts of use in the sense of Levinson's pragmatic "three-layer approach" to grammar (Levinson 1995, 1997, 2000) but semantic variation correlates only to a limited extent with preferred morphosyntactic encoding.

References

Bickel, Balthasar (2011), Grammatical relations typology, in: Song, Jae Jung (ed.), *The Oxford Handbook of Linguistic Typology*, Oxford: Oxford University Press, 399-444.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen (2007), Predicting the Dative Alternation, in: Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive Foundations of Interpretation*, Chicago: The University of Chicago Press, 69–94.

Bresnan, Joan & Marilyn Ford (2010), Predicting syntax: processing dative constructions in American and Australian varieties of English, *Language* 86, 186-213.

Cappelle, Bert (2006), Particle placement and the case for "allostructions", *Constructions* SV1-7.

DeReKo, Deutsches ReferenzKorpus, Mannheim: Institut für Deutsche Sprache, http://www1.ids-mannheim.de/kl/projekte/korpora/.

De Vaere, Hilde, Ludovic De Cuypere & Klaas Willems (2018), Alternating constructions with ditransitive *geben* in present-day German, *Corpus Linguistics and Linguistic Theory*, Published Online: 2018-05-26, DOI: https://doi.org/10.1515/cllt-2017-0072.

Goldberg, Adele E. (1995), *Constructions: A Construction Grammar Approach to Argument Structure*, Chicago: The University of Chicago Press.

Goldberg, Adele E. (2003), Constructions: a new theoretical approach to language, *Trends in Cognitive Science* 7, 219-224.

Goldberg, Adele E. (2006), *Constructions at Work: The Nature of Generalization in Language*, Oxford: Oxford University Press.

Haspelmath, Martin (2013), Ditransitive Constructions: The verb 'give', in: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology, http://wals.info/chapter/105.

Haspelmath, Martin (2015), Ditransitive constructions, in: *The Annual Review of Linguistics* 1, 19-41.
Levinson, Stephen C. (1995), Three levels of meaning, in: F. Palmer (ed.) *Grammar and meaning: Essays in honour of Sir John Lyons*, Cambridge University Press, 90-115.
Levinson, Stephen C. (1997), From outer to inner space: Linguistic categories and non-linguistic thinking, in: J. Nuyts & E. Pederson (eds.), *Language and conceptualization*, Cambridge: Cambridge University Press, 13-45.
Levinson, Stephen C. (2000), *Presumptive meanings: The theory of generalized conversational implicature*, Cambridge, MA: MIT Press.
Malchukov, Andrej, Martin Haspelmath & Bernard Comrie (2010), Ditransitive constructions: A typological overview, in: Malchukov, Andrej, Martin Haspelmath & Bernard Comrie (eds.), *Studies in Ditransitive Constructions: A Comparative Handbook*, Berlin: De Gruyter Mouton, 1-64.
Stefanowitsch, Anatol (2011), Argument structure: Item-Based or Distributed? *Zeitschrift für Anglistik und Amerikanistik*, 59 (4), 369-386.

Differences in usage of German passive constructions with bekommen / erhalten / kriegen

Volodymyr Dekalo
(University of Hannover, Germany)

Keywords: passive constructions, passive auxiliary verbs, simple collexeme analysis, distinctive collexeme analysis, random forests, conditional inference trees

Schedule: Fri 15.00 Room 7

The proposed talk will discuss the usage particularities of non-canonical German passive constructions with *bekommen* / *erhalten* / *kriegen* in modern German. They represent a combination of a respective passive auxiliary verb (bekommen / erhalten / kriegen) and a schematic slot for verbal complements occurring in the past participle form (see ex. 1–3 from DWDS Core Corpus, Geyken 2007).

(1) *Sie bekam ein kleines Stück Erde zugewiesen, das sie bebauen durfte.*
(2) *Zuerst erhält die Partei mit den meisten Stimmen das erste Mandat zugesprochen.*
(3) *Er ist Arzt, das ist sein Beruf, dafür kriegt er bezahlt!*

These constructions express the passive meaning ‘to be done for smb / smth’ and can be formalized by the following general schema: [AUX.PASSIV.VERB + LEX.VERB.PP]. Accordingly, the verbs *bekommen* / *erhalten* / *kriegen* function as grammatical markers of passive meaning. Regardless of sharing similar grammatical or rather abstract content, they differ in their verbal lexeme preferences as well as in their functionality in terms of semantic and syntactic usage features.

This case study aims to ascertain the item-based, viz. distributional, differences between three passive constructions using a simple and distinctive collexeme analysis (Gries & Stefanowitsch 2004, and Stefanowitsch & Gries 2003). The simple collexeme analysis identifies the typical (or salient) verbal items occurring in the schematic LEX.VERB.PP-slot of the respective construction. Sorting top ranked, viz. highly attracted, lexical verbs into semantic classes enables to compare the semantic core potentials of passive constructions introspectively. In contrast to this, the distinctive collexeme analysis identifies the lexical verbs that exhibit a strong preference for one of three passive constructions. Correspondingly, this method allows to figure out distributional differences between the

members of a family of non-canonical German passives that shouldn't be necessarily allocated within the core potential.

In addition, this study is to determine the functional divergences between the passive constructions applying random forests & conditional inference trees (Baayen & Tagliamonte 2012, and Hothorn et al. 2006). The results of these quantitative techniques demonstrate what (bundles of) usage features (semantic class of verbal lexemes, grammatical form of passive auxiliary verb, mood, transitivity of verbal lexemes, polarity, animacy, formal realization, definiteness, grammatical number and grammatical person of subject referents) operationalized as predictors give evidence for a stronger or weaker preference of the constructions under investigation coded as response variable. The empirical basis of these surveys represent exhaustive sets of observations from the DWDS Core Corpus (Geyken 2007).

References

Baayen, Harald and Sali Tagliamonte (2012), Models, forests and trees of York English: Was/were variation as a case study for statistical practice, *Language Variation and Change* 24(2), 135-178.

Geyken, Alexander (2007), The DWDS corpus: A reference corpus for the German language of the 20th century, in C. Fellbaum (ed.), (2007), *Idioms and collocations: Corpus-based linguistic and lexicographic studies*, London / New York: Continuum, 23-41.

Gries, Stefan T. and Anatol Stefanowitsch (2004), Extending collostructional analysis: A corpus-based perspective on 'alternations', *International Journal of Corpus Linguistics* 9(1), 97-129.

Hothorn, Torsten, Kurt Hornik and Achim Zeileis (2006), Unbiased recursive partitioning: A conditional inference framework, *Journal of Computational and Graphical Statistics* 15, 651-674.

Stefanowitsch, Anatol and Stefan T. Gries (2003), Collostructions: Investigating the interaction between words and constructions, *International Journal of Corpus Linguistics* 8(2), 209-243.

Locative nouns and the pronunciation/silence of the preposition

Anna Maria Di Sciullo
(University of Quebec in Montreal and New York University)

Keywords: locative nouns, locative preposition, bare nouns, light nouns, derivations

Schedule: Thu 14.00 Room 7

We focus on the derivation of PPs with locative nouns in Italian and in Fallese, a dialect of Italian spoken in Abruzzi. We provide an analysis of the facts, independent evidence and consequences for linguistic theory.

1. In Italian and in Fallese, the locative preposition *a* must be pronounced with the locative *casa*, (1a), which, qualifies as a bare noun (Carlson 1977, Kishimoto 2000, Longobardi 2001, Collins 2007), (2). However, the determiner must be pronounced in Fallese, (1b), which indicates that a light vs. bare noun account for the optional pronunciation of the preposition with locative *home* cannot directly extend to Fallese.

(1) a. È rimasto *(a) casa. (It.)
is remained at home
"He remained at home."

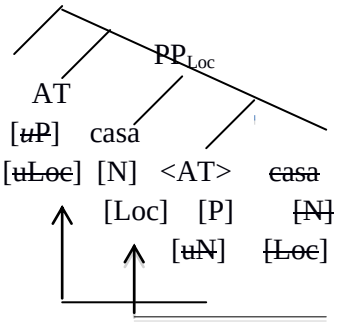
b. È rimaste *(a) la case (Fa.)
is remained at the house
"He remained home."

(2) a. È rimasto *(a) casa/*e. (It.)
is remained at home/s
"He stayed home."

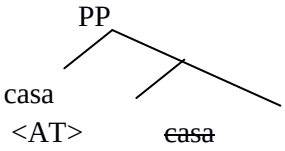
b. È rimasto a la (*bella) casa. (It.)
is remained at the nice home
"He remained at the nice home."

2. We take PPs to be phases of extended PP projections (Cinque and Rizzi 2010 and related works) and provide a feature based derivation (Chomsky 2008; Pesetsky and Torrrego 2004, Bošković 2014, Barrie and Yoo 2017) of the pronunciation vs silence (< >) of the preposition. See the partial structures in (3).

(3) a. $PP_{RelView}$



b.



We argue that locative *casa* is associated with valued nominal [N] and locative [Loc] features. The unvalued [*u*N] feature of the P is valued by the [N] feature of *casa.* The lower P Head is displaced to the higher [*u*P] Head for feature valuing, thus extending the PP*Loc* phase. The locative noun *casa* is then displaced to the Specifier position of the higher phase to value the [*u*Loc] feature of P. Given Collins' (2007) Spell-out Condition on pronunciation according to which the Specifier or the Head of a phase must be pronounced but not both, it follows that the prepositional Head is pronounced when it heads the higher RelView phase. We correctly predict that View point markers such as *proprio* (right) may modify locative *casa,* in Italian (1a) and in Fallese (1b), as well as bare noun *home,* but not light noun *home* in English, e.g. *He stayed (*right) home*.

3. Independent evidence for our analysis comes from locative pronouns *here*/*there*. In English, the preposition is silent when the pronouns moves to the Specifier position (Van Riemsdijk 1978, McCawley 1988, Kayne 2005, Collins 2007, a.o.). The preposition however can be pronounced in Fallese, (*a)ecche/(a)locche* (at here/at there) when the Specifier of the phase has no phonetic content.

4. A phase based feature driven analysis of the apparent optionality of the preposition with locative nouns and pronouns brings further support to the relation between movement and silence, as well as it supports the Strong Minimalist thesis (Chomsky 1995, 2001, 2005 *et seq.*), according to which language is the best solution to interface conditions.

Acknowledgement: This work is supported in part by funding from the FQRSC on Dynamic Interfaces, grant number 137253. www.interfaceasymmetry.uqam.ca.

References

Adger, David (2003), *Core Syntax: a Minimalist Approach*, Oxford: Oxford University Press.

Bošković, Zeljko (2014), Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis, *Linguistic Inquiry* 45: 27-89.

Barrie, Michael and Isaiah WonHo Yoo (2017), Bare nominal adjuncts, *Linguistic Inquiry* 48(4): 499-512.

Carlson, Gregory (1977), Reference to Kinds in English. Doctoral Dissertation.

Chomsky, Noam (1995), *The Minimalist Program*, Cambridge, MA: The MIT Press.
Chomsky, Noam (2001), Beyond explanatory adequacy, *MIT Occasional Papers in Linguistics 20*.
Chomsky, Noam (2005), Three factors in language design, *Linguistic Inquiry* 36: 1-22.
Chomsky, Noam (2008), On Phases. In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, eds. R. Freidin, C. P. Otero and M. L. Zubizarreta, 133–166, Cambridge, MA: MIT Press.
Cinque, Guglielmo and Luigi Rizzi, eds. (2010), *The Cartography of Syntactic Structures*, vol. 6: *Mapping Spatial PPs*, New York: Oxford University Press.
Collins, Christopher (2007), Home sweet home, In *NYU Working Papers in Linguistics* 1: 1-27.
Drummond, A., N. Hornstein, and Howard Lasnik (2010), A puzzle about P-stranding and a possible solution. Linguistic Inquiry 41: 689-692.
Kayne, Richard S. (2005), *Movement and Silence*, Oxford: OUP.
Kayne, Richard S. (2006), On parameters and on principles of pronunciation, In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, ed. Hans Broekhuis et al., 289-299, Berlin: Mouton de Gruyter.
Kishimoto, Hideki (2000), Indefinite pronouns and overt N-raising. Linguistic Inquiry 31:557-566.
Liao, Wei-wen Roger and Shi Dingxu (2013), To pronounce or not to pronounce: locating silent heads in Chinese and English. *Studies in Chinese Linguistics* 34(1):55-65.
Longobardi, Giuseppe (2001), How comparative is Semantics? A unified paramtric theory of bare nouns and proper names. Linguistic Inquiry 9:335-369.
McCawley, James D. (1988), Adverbial NPs: Bare or clad in see-through garb?, *Language* 64(3): 583-590.
Pesetsky, David and Esther Torrego (2007), The Syntax of Valuation and the Interpretability of Features, In S. Karimi, V. Samiian and W. K. Wilkins (eds.), *Phrasal and Clausal Architecture*, 262-294, Amsterdam: John Benjamins.
Riemsdijk, Henk Van (1978), *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*, Dordrecht: Foris Publications.
Uriagereka, Juan (1999), Multiple Spell-Out, in S. Epstein and N. Hornstein (eds.), (2009), *Working Minimalism*, 251-282, Cambridge, MA: MIT Press.

Estaría necesitando que alguien me haga bien: Semantic and pragmatic extensions of the conditional in *Rioplatense* Spanish

Pedro Díaz Lammertyn & María Sol Sansiñena
(University of Leuven)

Keywords: Spanish conditional, modality, evidentiality, Rioplatense Spanish, intersubjectivity

Schedule: We 15.30 Room 11

Spanish conditional forms have evolved to express a variety of modal (e.g. epistemic) and evidential (e.g. hearsay, inference) meanings, as well as attenuation and polite meanings (see Palmer 2001:105; Haverkate 2002; Azzopardi 2013; Vatrican 2013; Böhm & Hennemann 2014, among others). However, there is no comprehensive typology or semantic characterization that extends to the new uses of the Spanish conditional forms. Haverkate (2002: 25) proposes a unitary account for all the uses of the conditional arguing that they concern "assertions on counterfactual states of affairs." However,

this explanation does not account for recent uses such as (1), increasingly frequent in South American varieties, and particularly in *Rioplatense* Spanish:

(1) *Estaría necesitando que alguien me haga bien.* (Twitter)
Lit. I would be needing that someone does me good. 'I need someone that does me good'

In such undescribed uses the conditional, often in combination with progressive marking, as in (1), is used as an evidential to give rise to an ironic effect by pragmatic inferences (Fuchs 2018). The pragmatic features in (1) may be rephrased as "It seems as if I need that someone does me good," and evidence a certain disconnect between a relatively intense personal need and having to pass via an inferential path to become aware of it. Such constructions are used to express peremptory or urgent (and sometimes silly) needs, cravings or longings, negative emotions or to critical observations, often seeking or conveying alignment with the interlocutor(s). However, they can also function as dispreferred responses (Levinson 1983:333) which express subjective judgement in contexts where a previous statement or behavior is criticized as unrealistic.

The aim of this study is to explore the relationship between modal and evidential meanings of the Spanish conditional, and to inquire into their semantic and pragmatic extensions in *Rioplatense* Spanish. The study sets out to, first, describe the syntactic, semantic and pragmatic features of a set of new form-meaning pairs (i.e. constructions) such as (1), and second, to identify the mechanisms involved in the development of these constructions. The starting hypothesis is that these constructional changes follow a tendency towards increasing speech-act orientation, i.e. orientation towards the speaker, the speech situation including the interlocutor, and discourse itself (in the sense of Narrog 2012:46), and intersubjective coordination.

The data was sourced from a corpus of collected Tweets from various points in time between 2010 and 2018. The main analysis is based on Tweets from the Rioplatense variety, although data from other varieties of Spanish is included for comparison purposes. Each instance of the conditional is analyzed according to a set of relevant morphosyntactic, semantic-pragmatic and discourse-interactional criteria, which allows for the distinction of situated meanings emerging from the interaction between linguistic elements and contextual factors (Linell 2009:99).

A pilot study shows that constructions like (1) have steadily risen in frequency over the past decade, but are mostly restricted to Rioplatense Spanish and are used to express a restricted set of meanings. Moreover, these constructions show a clear preference for predicates of volition and necessity, and predicates that denote thought, argumentation or belief for the progressive slot.

References

Azzopardi, S. (2013), Valor aspecto-temporal y usos modales del condicional a la luz de la noción de gramaticalización. *Estudios de Lingüística: investigaciones, propuestas y aplicaciones*. Adrián Cabedo Nebot, Manuel Aguilar Ruiz & Elena López-Navarro Vidal, eds. Valencia: Tecnolingüística. 337–347.

Böhm, V. & Hennemann, A. (2014), The evidential use of the Spanish imperfect and the conditional in journalistic contexts. *Studia Neophilologica* 86(2), 183-200.

Haverkate, H. (2002), *The Syntax, Semantics and Pragmatics of Spanish Mood.* Amsterdam: Benjamins.

Fuchs, M. (2018), On the Conditional + Progressive marker in Spanish: meaning change in a cross-dialectal comparison. Unpublished graduate paper. Yale University.

Levinson, S. (1983), *Pragmatics*. Cambridge: Cambridge University Press.

Linell, P. 2009. Constructions in dialogue. In A. Bergs & G. Diewald (eds), *Contexts and Constructions*. [Constructional Approaches to Language 9]. Amsterdam: Benjamins, 97-110.

Narrog, H. (2012), Beyond intersubjectification: textual uses of modality and mood in subordinate clause as part of speaker-orientation. *English Text and Linguistics* 5(1), 29-52.

Palmer, F.R. (2001), *Mood and Modality*. Cambridge: Cambridge University Press.
Vatrican, A. (2013), El condicional de cortesía en español: la hipótesis como forma de atenuación. In A. Cabedo Nebot, M. J. Aguilar Ruiz, & E. López-Navarra (eds), (2013), *Estudios de lingüística: investigaciones, propuestas y aplicaciones*, Universitàt de Valencia, Valencia, 469-480.

Vietnamese Reflexivity: *minh 'body'* and *tự 'self'*- two elements and their interaction

Quý Ngọc Thị Đoàn
(Utrecht University/Hue University of Education, Vietnam)

Keywords: multiple Agree, reflexivizing operator, the anaphor *mình* 'body', the morpheme *tự* 'self', discourse antecedent

Schedule: Sat 12.00 Room 12

1. Introduction

Like many other languages, the anaphoric system of Vietnamese is not captured by the canonical binding conditions in Chomsky (1981). It has an anaphoric element *mình* 'body' that needs not be bound in the local domain, in violation of Condition A. Here I will present an analysis of *mình,* showing how it interacts with the morpheme *tự* 'self'.

2. The Puzzle

Some characteristic examples are shown in (1) and (2):

1. Họ$_i$ khen ***mình***$_{*i/s}$.
3PL praise body
'They praise me.'

2. Họ$_i$ **tự** khen ***mình***$_{i/*s}$.
3PL self praise body
'They praise themselves.'

In (1), *minh* in the object position of a simplex sentence has a discourse antecedent, the speaker (indicated with an s-subscript). In order to derive a reflexive meaning, the morpheme *tự* 'self' must be inserted preceding the verb as shown in (2). In complex sentences, in the absence of *tự*, *mình* is non-locally bound by the matrix subject and also possibly valued as the speaker as in (3). However, in the presence of *tự*, long-distance binding of *mình* is blocked, see (4).

3. John$_i$ nghĩ rằng họ$_j$ khen **mình**$_{i/*j/s}$.
John think COMP 3PL praise body
'John thought that they praised him/me'

4. John$_i$ nghĩ rằng họ$_j$ **tự** khen **mình**$_{j/*i}$.
John think COMP 3PL self praise body
'John thought that they praised themselves.'

In the literature, the binding domain of *mình* has been discussed by Fukuda (2005), Tran (2009), Ivan and Bui (2018). Yet, these authors did not give any account for the salient 1st person interpretation of *mình*.

3. The Analysis

My analysis will be modeled on Giblin (2016)'s analysis to Mandarin. I will be assuming that *mình* is unvalued for person features. The link with the antecedent (fully specified by assumption), the specifier of TP in (5), is provided by the complementizer C^0. C^0 and *mình* receive their values from $DP_{val\varphi}$ by multiple Agree (Hiraiwa 2001, 2005) as schematized in (5) and (6).

5. $[\ C^0{}_{u\varphi}\ [\ DP_{val\varphi}\ [T^0{}_{u\varphi}\ ....\ [T^0{}_{u\varphi}\ldots.\ minh_{u\varphi}\ ....\ ]]]]\ \rightarrow$
6. $[\ C^0{}_{\mathbf{val}\varphi}\ [\ DP_{\mathbf{val}\varphi}\ [T^0{}_{val\varphi}....\ [T^0{}_{val\varphi}\ldots\ minh_{\mathbf{val\varphi}}....\ ]]]]]$

The speaker interpretation of *mình* follows under the assumption that in Vietnamese the left periphery contains a silent speaker and a silent performative verb (Ross, 1970), which are visible to Agree.

That *mình* is locally bound in the presence of *tự* follows from the assumption that *tự* is a reflexivizing head with the vP as a complement. That is, *tự* is an operator reflexivizing a vP it is attached to. Formally, it serves as an operator like REFL as in (7) which applies to a 2-place predicate R (= a relation between atomic entities) and generates a 1-place predicate over sets A of atomic entities:

7. REFL :. λR. λA. ∀x∈ A [R (x, x)] (see Keenan 1988)

The anaphor *mình* cannot have a non-local interpretation in this case, hence *tự* must operate prior to whatever process yields the interpretation as a non-local subject or 1st person.

References

Chomsky, Noam (1981), *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Fukuda, Shin (2005), *Subjects and Reflexives in Vietnamese*. A Presentation at UCSD Linguistics Department Colloquium.

Giblin, Iain (2016), *Agreement Restrictions in Mandarin Long-distance Binding*. Doctoral Dissertation, MIT.

Hiraiwa, Ken (2001), On nominative-genitive conversion. *In MIT Working Papers 39: A Few From Building E-39*, eds. Ora Matushansky and Elena Guerzoni, 66-124. Cambridge, Massachusetts: MITWPL.

Hiraiwa, Ken (2005), *Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture*. Doctoral dissertation, MIT, Cambridge, Massachusetts.

Ivan, Rodica and Bui, Thuy (2018), Vietnamese Anaphora: Binding Principles and the Lack Thereof. *The 5th Meeting of the Semantics of African, Asian, and Austronesian Languages* (TripleA 5). Konstanz, Germany.

Keenan, Edward (1988), On semantics and the binding theory. *In Explaining Language Universals* J. Hawkins (ed). Blackwell. Pp. 105 – 144.

Ross, John Robert (1970), On Declarative Sentences, in:R.A. Jacobs and P.S. Rosenbaum, eds. *Readings in English Transformational Grammar*, Ginn: Waltham, Mass

Tran, Thuan (2009), *Wh-quantification in Vietnamese*. Doctoral dissertation. University of Delaware.

Non-hierarchical grammatical tone in Chichewa verb paradigms

Laura J. Downing
(University of Gothenburg)

Keywords: inflectional tone, Bantu languages, non-concatenative morphology, paradigms, dominance effects

Schedule: Thu ~~15.00~~ 14.30 Room 9

Many African tone languages have tonal morphemes and/or morphemes that introduce construction-specific grammatical tone patterns. As work like Hyman (2016), Inkelas (1998), Odden & Bickmore (2014) and Rolle (2018) shows, grammatical tone patterns have a number of properties that challenge a concatenative view of morphology. They associate a High tone or a tone pattern to a position or to positions that are not local to the sponsoring morpheme; they can also delete the tone of their morphological base; and they can introduce construction-specific tonal processes. In addition, more than one inflectional morpheme can influence the grammatical tone pattern of a word or a construction. The question then arises, as Hyman (2016) puts it, if grammatical tone patterns are in conflict, which one wins? Work like Hyman (2016), Inkelas (1998, 2018), McPherson & Heath (2016) and Rolle (2018) has taken the strong position that tonal dominance is defined by the morphosyntactic hierarchy: grammatical tones affect the base they take morphosyntactic scope over, and the hierarchically outermost dominant grammatical tonal morpheme wins.

This talk shows that grammatical tone patterns in Chichewa (Bantu N.31; Downing & Mtenje 2017; Hyman & Mtenje 1999; Kanerva 1990; Myers & Carleton 1996) are problematic for the outermost dominance principle. Indeed, often a single morpheme does not alone determine the tone pattern in a particular verb paradigm. And an innermost dominant morpheme can have an effect on the overall tone pattern, overriding the predicted effect of an outer dominant morpheme. These points can be briefly illustrated with verb paradigms related to the present habitual (PH) aspect, segmentally marked by the prefix *-ma-*. The affirmative PH suggests that the PH prefix is tonally dominant, as the contrast between High-toned roots and toneless roots is neutralized (acute accents indicate High tone): *ndí-ma-fotokóoza* 'I explain, toneless' vs. *ndí-ma-tambaláala* 'I stretch legs, High'. The root tone contrast reappears in the negative PH, however, even though the negative prefix *sí-* is tonally neutral (it simply realizes a High tone): *sí-ndí-má-fotokooza* 'I don't explain, toneless' vs. *sí-ndí-má-tambaláala* 'I don't stretch legs, High'. The intensive derivational suffix also affects the verb tone pattern. In the negative PH intensive, the tonal contrast in the root is neutralized: *sí-ndí-má-fotokoz-eéts-á* 'I don't explain a lot, toneless' vs. *sí-ndí-má-tambalal-iíts-á* 'I don't stretch legs a lot, High'. In these last two forms, we see an inside out effect of tonal dominance. The intensive conditions a dominant, neutralizing tone pattern on the stem that is different from the dominant, neutralizing tone pattern of the outer PH prefix.

In sum, Chichewa grammatical tone patterns provide a challenge to what Spencer (2004) calls a "Radical Agglutination" approach to morphology: the phonological form of inflected words is not always predicted by a one-to-one match between meaning/grammatical function, output form and hierarchical constituent structure. Instead, a combination of morphological features holistically predicts the inflectional tone pattern associated with a particular word form. I will argue that paradigm-based, realizational models like that of Spencer (2004) and Stump (2016) best account for the Chichewa data.

References

Downing, Laura J. & Al Mtenje (2017), *The Phonology of Chichewa*. Oxford University Press.

Hyman, Larry M. (2016), Morphological tonal assignments in conflict: who wins? In Enrique Palancar & Jean-Léo Leonard (eds.), *Tone and Inflection: New Facts and Perspectives*. Berlin: De Gruyter Mouton, 15-39.

Hyman, Larry M. & Al Mtenje (1999), Prosodic Morphology and tone: the case of Chichewa. In René Kager, Harry van der Hulst and Wim Zonneveld, (eds.), *The Prosody-Morphology Interface*. Cambridge University Press, 90-133.

Inkelas, Sharon (1998), The theoretical status of morphologically conditioned phonology: a case study of dominance effects. *Yearbook of Morphology 1997*, 121-155.

Inkelas, Sharon (2018), Overexponence and underexponence in morphology. In Eugene Buckley, Thera Crane & Jeff Good (eds.), *Revealing Structure: Papers in honor of Larry M. Hyman*. Stanford: CSLI Publications, 105-122.

Kanerva, Jonni (1990), *Focus and Phrasing in Chichewa Phonology*. New York: Garland Publishing.

McPherson, Laura & Jeffrey Heath (2016), Phrasal grammatical tone in the Dogon languages. *Natural Language & Linguistic Theory* 34, 593-639.

Myers, Scott & Troi Carlton (1996), Tonal transfer in Chichewa. *Phonology* 13, 39-72.

Odden, David & Lee Bickmore (2014), Melodic tone in Bantu: Overview. *Africana Linguistica* XX, 3-13.

Rolle, Nicolas (2018), Grammatical tone: typology and theory. Ph.D. dissertation, University of California, Berkeley.

Spencer, Andrew (2004), Morphology: an overview of central concepts. In Louisa Sadler & Andrew Spencer (eds.), *Projecting Morphology*. Stanford: CSLI Publications, 67-109.

Stump, Gregory (2016), *Inflectional Paradigms: Content and form at the syntax-morphology interface*. Cambridge University Press.

Updating the notion of Sprachbund: Perfects and resultatives in the Circum-Baltic "Stratified Convergence Zone"

Bridget Drinka

Schedule: We 14.30 Room 8

<pdf>

How dominant is manner of articulation in the formation of phonotactic systems of languages?

Katarzyna Dziubalska-Kołaczyk & Grzegorz Aperliński

(Adam Mickiewicz University, Poznań)

Keywords: phonotactics, MOA, Net Auditory Distance, phonotactic calculator, sonority

Schedule: Thu 10.00 Room 9

In this paper we propose a modification of the Net Auditory Distance (NAD) model of phonotactics with a view to improving its descriptive and predictive potential. In its present form, NAD relies on three parameters for the measurement of distance between segments in a cluster: manner of articulation (MOA), place of articulation (POA), and S/O (sonorant/obstruent contrast). Voicing contrast was included in the earlier version of NAD. The NAD model has so far proved to explain the phonotactics of languages better than any measure based solely on sonority. However, the predictability power of the model can be improved, especially with respect to the languages with more complex phonotactics.

The NAD model is assisted by the phonotactic calculator available online. MOA in the calculator

has five values (glide, liquid, nasal, fricative, stop), some of them subdivided language-specifically (e.g. to include affricates, or lateral and rhotic liquids). POA has the same number of values (labial, coronal, dorsal, radical, glottal) which are also subdivided language-specifically (e.g. to include labio-dental, post-alveolar or alveolar palatal place of articulation). MOAs reflect the major values of the rudimentary sonority hierarchy (Ladefoged 2006) and POAs follow Ladefoged's classification, too. S/O is a binary 0/1 contrast. We take the NAD between two segments to be the sum of distances calculated in terms of the three parameters listed above: NAD = |MOA| + |POA| + S/O. In the present version of the NAD model, MOA and POA are weighted equally. However, there is evidence that manner of articulation is be the dominant factor in the formation of phonotactic systems. According to Baumann and Wissing (2018), MOA seems to represent the strongest selection pressure, while phonation and place of articulation show less clear effects. More precisely, large differences in MOA between consonants contribute to a cluster's success in diachrony and acquisition, while large differences in POA – to the contrary. Neuro-cognitive research shows that consonants of the same MOA are processed closely together, which suggests that they are primarily categorized with reference to the manner of articulation (Mesgarani et al. 2014).

In order to embrace the above evidence we propose to introduce an extension of the MOA scale in the NAD model, to include a more comprehensive diversification of manners, based on the hierarchy of relative sonority proposed by Parker (2008: 60). Parker's hierarchy consists of 17 levels of phonetic segments of relative sonority. Extending 5 to 17 values on the MOA scale in the NAD model would automatically diminish the role of place of articulation in the formation of a cluster. The POA scale will also be modified according to the newest studies by Maddieson (p.c.) and the Voicing parameter will be re-introduced. The modified model will be tested on the phonotactic corpus of Polish and English. The results will be compared to the previously obtained cluster preferability estimates.

References

Baumann, Andreas and Daan Wissing (2018), Stabilizing determinants in the transmission of phonotactic systems: on the emergence of the Afrikaans consonant-cluster inventory, *Stellenbosch Papers in Linguistics*.

Ladefoged, Peter (2006), *A course in phonetics*, 5th edition, Boston: Heinle & Heinle.

Mesgarani, N., Cheung, C., Johnson, K., Chang, E. (2014), Phonetic feature encoding in human superior temporal gyrus, *Science* 348, 1007-1010.

Parker, Steve (2008), Sound level protrusions as physical correlates of sonority, *Journal of Phonetics* 36, 55–90.

Degree wh-questions in Spanish and the internal structure of wh-phrases

Luis Eguren
(Universidad Autónoma de Madrid)

Keywords: degree wh-forms, lexical variation, operator-variable structure, wh-phrases, Spanish

Schedule: Fri 9.00 Room 7

In this talk, I will provide a formal structural analysis of the pattern of lexical variation in Spanish adjectival degree wh-questions illustrated in (1) (on the origin, history and detailed current distribution of the wh-forms below, see Octavio de Toledo and Sánchez López 2009):

(1)	a.	*¿Cuán*	*alta*	*es*	*Ana?*	(General Spanish; register variation)
		how much	tall	is	Ann	
	b.	*¿Qué tan*	*alta*	*es*	*Ana?*	(American Spanish)
		what so	tall	is	Ann	
	c.	*¿Cómo de*	*alta*	*es*	*Ana?*	(European and American Spanish)
		how of	tall	is	Ann	
	d.	*¿Cuánto de*	*alta*	*es*	*Ana?*	(European and American Spanish)
		how much of	tall	is	Ann	

'How tall is Ann?'

I will propose that the degree wh-forms in (1) share a (basic) internal structure that syntactically maps the operator-variable semantic structure in degree wh-questions (Abrusán and Spector 2008), and thus includes (at least) two components, a wh-operator and a degree variable, heading a Quantifier Phrase (QP) and a Degree Phrase (DegP) in the extended projection of the adjective (Grimshaw 1991). As represented in (2), these two components are phonetically realized in the American Spanish wh-form *qué tan*, whereas in the rest of adjectival degree wh-expressions the degree morpheme is null:

(2) a. $[_{QP}$ qué $[_{DegP}$ tan $[_{AP}$ alta]]]
b. $[_{QP}$ cuán(to)/cómo $[_{DegP}$ Ø $[_{AP}$ (de) alta]]]]

I will also extend the analysis in (2) to the nominal wh-forms *cuántos* 'how many/much' and *qué tanto(s)* 'what so much/many' (3), and to the General Spanish manner wh-expressions *cómo* 'how' and *qué tal* 'what such' as well (4), by claiming that *qué* in *qué tantos/tal* expresses wh-quantification also in this case and *tantos*/*tal* instantiates a variable now ranging over amounts and manners, respectively, which is not pronounced in *cuántos* and *cómo*:

(3)	a.	*¿Cuántos*	*libros*	*has*	*leído?*		(General Spanish)
		how many	books	have.2SG	read		
	b.	*¿Qué tantos*	*libros*	*has*	*leído?*	(American Spanish)	
		what so many	books	have.2SG	read		

'How many books have your read?'

(4)	a.	*¿Cómo estás?*		
		how are.2SG		
	b.	¿Qué tal	estás?	
		what such	are.2SG	

'How are you?'

Building on the insight in Johnson (2009) that English question phrases like *which book* have hidden within them definite descriptions headed by a silent *the*, which has the role of a variable over individuals, I will further argue that all interrogative wh-forms in Spanish comprise a wh-quantifier and an overt or covert variable ranging over different types of entities (degrees, amounts, manners, individuals, etc.). And I will finally suggest that this may universally hold across languages, allowing us, under the copy theory of movement (Nunes 2011), and the idea in Chomsky (1993) that different parts of different copies in a wh-chain can be interpreted at LF, to capture the fact that the wh-item in a wh-chain acts semantically like a variable in its lowest position, and like the binder of that variable in its highest position, without resorting to a special semantic "trace conversion rule" (Fox 1999) turning lower copies/traces into structures that contain variables.

References

Abrusán, Marta and Benjamin Spector (2008), An interval-based semantics for degree questions: negative islands and their obviation, in N. Abner, and J. Bishop (eds), *Proceedings of the 27th*

West Coast Conference on Formal Linguistics, Somerville, Mass.: Cascadilla Proceedings Project, 17-26.
Chomsky, Noam (1993), A minimalist program for linguistic theory, in K. Hale and S. J. Keyser (eds), *The view from building 20*, Cambridge, Mass.: The MIT Press, 1-52.
Fox, Danny (1999), Reconstruction, binding theory, and the interpretation of chains, *Linguistic Inquiry* 30, 157-196.
Grimshaw, Jane (1991), Extended projection. Ms., Brandeis University.
Johnson, Kyle (2009), Why movement? Talk presented at the *Chicago Linguistics Society Meeting*.
Nunes, Jairo (2011), The Copy Theory, in C. Boeckx (ed), *The Oxford handbook of Linguistic Minimalism*, Oxford: Oxford University Press, 143-172.
Octavio de Toledo, Á. S. and C. Sánchez López (2009), Cuantificadores II. Los cuantificadores interrogativos y exclamativos, in C. Company (dir), *Sintaxis histórica de la lengua española. Segunda Parte: La frase nominal*, México: Fondo de Cultura Económica, 963-1072.

Theory of mind, non-standard questions and discourse modal particles

Yoshio Endo
(Kanda University of International University)

Keywords: cartography, expressive meaning, ToM, modality, German/Japanese

Schedule: Thu 16.30 Room 11

1. Introduction. As opposed to the view by Nuyts (2000) that modality is a non-linguistic category, Leis (2012) makes a claim that the human ability to have a Theory of Mind (ToM) is language driven, based on modal particles in German. I support the latter view by showing that particles in Japanese, which correspond to particles in German, play an active role in syntactic computations of creating non-standard questions such as "I-Can't-find-the-Value-of-x Question" (CfvQ) (Bayer and Obenauer (2011)), what Trozke (2017) calls focus of emphasis, etc. in the framework of cartography of syntactic structures (cf. Rizzi 1997).

2. **German and Japanese discourse particles**

The sentence in (1a) is an instance of CfvQ in German, where the *nur / bloß* are used in clause-medial position. In CfvQ, the speaker signals that (s)he has so far unsuccessfully tried to find an answer. This German sentence is translated into Japanese by suffixing the sentence final particle *na* to the Q-Force element *ka* and the Fin element *no* (=1b), where the expressivity can be confirmed by prefixing the expressive vocative expression *ara yada* 'Oh, no' through concord relation (cf. Endo and Haegeman 2019). This non-standard question disappears when another head element intervenes between a Force element and a particle (=1c). (The sequence *no-ka-mo-na* can only be non-expressive, as confirmed by the incompatibility with the expressive vocative *ara yada* 'oh, no' (=1d).

(1) a. Wo liegt nur / bloß meine Brille? (CfvQ)
where lies NUR/ BLOSS my glasses
'Where on earth did I put my glasses? (I have already looked everywhere)

b. Ara yada, megane-o doko-ni oita **no** **ka** **na**?
Oh, no glasses-Acc where-at put **Fin** Force Prt

c. *Ara, yada megane-o doko-ni oita **no** **ka** **mo** **na**?
Oh, no glasses-Acc where-at put **Fin** Force also Prt

d.??Ara, yada megane-o soko-ni oita **no** **ka** **mo** **na**?
Oh, no glasses-Acc there-at put **Fin** Force also Prt

3. Proposal. After introducing the basic ideas and notions used in the cartographic project, I propose that the expressive meaning of non-standard questions in Japanese we saw above can be explained by using Rizzi's (1997) idea that in the absence of an element between Force and Fin, Force and Fin are amalgamated as a single unit as shown in (2a). Because Japanese is a head-final language, the linear order of these functional elements is as shown in (2b), which is the mirror image of (2a):

(2) a. [ForceP]+[FinP]…
b. [FinP *no*]+ [ForceP *ka*]+ (Japanese)

With this in mind, I propose that there are two types of particles in Japanese, one of which is expressive and the other of which is non-expressive, where expressive particles carry the feature [+E] which locally selects the amalgamated Force-Fin pair (=3a):

(3) a.…[FinP *no*]+[ForceP *ka*] [ParticleP *na* [+E]]
←select-/
b. [FinP *no*]+[ForceP *ka*] mo (=intervener) [ParticleP *na* [+E]]

In the presence of the intervening element *mo* between the Force+Fin pair as in (3b), the particle *na* cannot locally select the Force+Fin pair, and thus the particle can only be a non-expressive type.

Acknowledgment
This research is funded by a Grant-in-Aid for Scientific Research (C) of the Japan Society for the Promotion of Science (Project 16K02639).

References

Abraham, Werner. 2012. Illocutionary force is speaker and information source concerns. In Werner Abraham and Elizabeth Leiss eds. *Modality and theory of mind elements across languages*. Berlin: Mouton de Gruyter.

Bayer, Josef and Hans-Georg Obnauer. 2011. Discourse particles, clause structure, and question types. *The Linguistic Review* 28: 449–491.

Endo, Yoshio and Liliane Haegeman. 2019. Adverbial clauses and adverbial concord. *Glossa* 4(1): 48, 1-32.

Leiss, Elizabeth. 2012. Epistemicity, evidentiality, and theory of mind. In Werner Abraham and Elizabeth Leiss eds. *Modality and theory of mind elements across languages*. Berlin: Mouton de Gruyter.

Nuyts, Jan (2000)*Epistemic Modality, Language , and conceptualization.* [Human Cognitive processing 5.] Amsterdam: John Benjamins.

Obenauer, Hans-Georg. 2004. Nonstandard wh-questions and alternative checkers in Pagotto. In Horst Lohnstein & Susanne Trissler (eds.), The Syntax and

Semantics of the Left Periphery, 343–383. Berlin: Mouton de Gruyter.Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.

Trotzeke, Andreas. 2017. *The grammar of emphasis*. Mouton de Gruyter, Berlin/New York.

An argument-adjunct asymmetry in depictive control in Ossetic

David Erschler
(Ben Gurion University of the Negev)

Keywords: Depictives; Secondary predicates; Syntax; Typology; Ossetic

Schedule: Fri 12.00 Room 6

I show that depictives in Digor and Iron Ossetic (East Iranian; the Caucasus) can be controlled by any argument, but no adjunct, of the verb, unlike in commonly studied languages, where depictives are only controlled by subjects and direct objects (1).

Introduction. Depictives are entities illustrated in (1). They are adjectives that modify noun phrases without forming a single constituent with the NP. The NP is said to control the depictive.

(1)a. *Mary*$_i$ *ate the fish*$_j$ ***raw***$_{\#i/j}$ / ***drunk***$_{i/\#j}$. b. *John*$_i$ *served Mary*$_j$ *coffee* ***drunk***$_{i/*j}$.

Semantics and morphological marking of depictives are well studied crosslinguistically, see Schultze-Berndt and Himmelmann (2004, 2005), Schroeder et al. (2008). The variation in control patterns received much less attention, although Nichols (1978) showed that in Finnish a wider variety of phrases than the subject and direct object control depictives.

Background on Ossetic. Ossetic is head final, with a moderately large case system. The case is marked on the right edge of NP. The NP is rigidly ordered and unsplittable. It lacks overt agreement (2).

(2) *asə štər worš bɐχ-t-ɐn* Iron
this big white horse-PL-DAT
'for these big white horses'

Depictives in Ossetic. Ossetic lacks resultatives (i.e. secondary predicates like *dry* in '*John wiped the table* ***dry***.'), but widely uses depictives. Depictives are obligatorily marked with the ablative, (3a), no matter what the case of the controlling NP. They don't agree with the controller in number.

(3) a. *šoslan* ***təžmɐg-*(ɐj)*** *asədi* b. **šoslan* ***bɐržond-ɐj*** *asədi* Iron
S. angry-ABL left S. tall-ABL left
'Soslan left angry.' '*Soslan left tall.'

The obligatory ablative marking excludes the possibility that what looks like a depictive and its controller are fragments of a single split NP, unlike the situation that obtains, e.g., in Slovenian, (Marušič et al. 2008). Depictives express stage level properties, see the contrast between (3a), where 'angry' is a stage level property, and (3b), with 'tall', an individual level property. Depictives are distinct from adverbs, in that only depictives carry the obligatory ablative marking. In (4), *rɐsuʁd* 'beatiful(ly)' is an adverb that modifies the writing event.

(4) *mɐdinɐ* ***rɐsuʁd(*-ɐj)*** *fins-uj* Digor
M. beautiful-ABL write-PRS.3SG
'Madina writes beautifully.'

Any verb arguments can control depictives.

(5) ablative and superessive marked controllers
soslan-ɐj$_j$ **=*mɐbɐl***$_i$ *rasug-ɐj*$_{i/j}$ *rawadɐj* Digor
S-ABL=SUP.1SG drunk-ABL befell

'I had it from Soslan (when I/Soslan was) drunk.'

Adjuncts don't control depictives; arguments win out even if the resulting reading is odd (6).

(6) *tikis*$_{i}$ ***soslan-bɐl***$_{j}$ ***<u>rasuq-ɐj</u>***$_{i/*j}$ *χussuj* Digor
cat S-SUP drunk-ABL sleeps
'The cat sleeps on Soslan (when it/*Soslan is) drunk.'

The contrast between the superessive NPs in (5), where it's an argument, and in (6), where it's an adjunct, shows that it's not the case marking, but the argument/adjunct status that matters for depictive control in Ossetic.

Conclusion. Ossetic has grammaticalized what seems a typological *rarum* --a pattern of depictive control where the difference between adjuncts and arguments plays the principal role. This reiterates the importance of these concepts in the grammar, and calls for a wider typological investigation of depictive control.

References

Himmelmann, Nikolaus P.; Schultze-Berndt, Eva. (eds.) 2005. *Secondary Predication and Adverbial Modification*. Oxford: OUP.

Marušič, Franc Tatjana Marvin, and Rok Žaucer. 2008. Depictive Secondary Predication with no PRO. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns, Roland Meyer (eds.) Formal Description of Slavic Languages. pp. 423-434. Frankfurt am Main: Peter Lang.

Nichols, Johanna. 1978. Secondary Predicates. Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society (1978), pp. 114-127

Schultze-Berndt, Eva and Himmelmann, Nikolaus P. 2004. Depictive secondary predicates in cross-linguistic perspective. *Linguistic Typology* 8. 59-131.

Schroeder, Christoph; Hentschel, Gerd & Winfried Boeder (eds.) 2008. Secondary predicates in Eastern European languages and beyond. Oldenburg: BIS-Verlag der Carl von Ossietzky Universität Oldenburg

Aspectual cognate object constructions are not light verb constructions in Hungarian

Imola-Ágnes Farkas
(Babeș-Bolyai University, Cluj-Napoca, Romania)

Keywords: pseudo-object, aspectual cognate object, light verb construction, event, result, Hungarian

Schedule: Thu 9.00 Room 7

The primary aim of the talk is to argue that the Hungarian pseudo-objects (POs) *egyet* 'one.ACC', (*egy*) *jót* '(one) good.ACC', *jókat* 'good.PL.ACC' etc. in VPs such as *sétál egyet* 'walk one.ACC//take a walk' are aspectual cognate objects (ACOs) in the sense defined by Horrocks & Stavrou (2010) and Lavidas (2013). We claim that they are light ACOs in that they have a weakened referentiality, their lexical content is bleached and they assume a more grammatical (i.e. aspectual) meaning (Simone & Masini 2014). The secondary aim is to show that these VPs cannot be considered light verb constructions (LVCs).

Hungarian POs share many of the syntactic properties of (English) ACOs: they cannot be passivized or topicalized; cannot be resumed with a (referential) pronoun; cannot be modified by a restrictive relative clause; cannot undergo *wh*-movement; and are indefinite, non-thematic and non-referential. Their aspectual contribution is also the same: they delimit the event denoted by the verb (Csirmaz 2008) despite their questionable compatibility with frame adverbials (Piñón 2001 and Farkas 2017). There are two significant differences: (i) as they do not morphologically repeat the verb and there is no tautology, modification is possible but not obligatory; and (ii) the modified PO does not have an adverbial interpretation, as we can insert a manner adverb with a meaning opposite to the one denoted by the adverbial counterpart of the adjective the PO is built on (Puigdollers Real 2008) (*Mari jót táncolt, de rosszul* 'Mary had a good dance but she did not dance well'). Furthermore, these COs have an ambiguous event-result interpretation. Their eventive behaviour is reflected in the possibility of introducing a benefactive double object (*kacsint egyet valakinek* 'give somebody a wink'), where an applicative head relates an event to an individual (Marantz 2005); and in the interpretation triggered by the modified PO primarily in denominal VPs (*fagyizni egy jót* 'have a good ice-cream eating experience'), where A modifies the event itself. The non-eventive behaviour is shown in the lack of adverbial interpretation and the possibility of pluralizing the PO (see above).

Although most of these syntactic properties are shared by LVCs (Kearns 1988, Diesing 1998, and Butt 2010), Hungarian *egyet*-VPs cannot be considered LVCs as (i) in these VPs it is not the verb but the noun that is semantically light; (ii) the interpretation of these VPs is not purely eventive but ambiguous between an event and a result reading; and (iii) *egyet*-VPs can sometimes have an LVC counterpart (*zuhanyozik egyet* vs. *zuhanyt vesz* 'take a shower'). In addition, the wide variety of LVCs (Vincze 2012) are in marked contrast to the (almost complete) absence of morphologic and semantic COs in Hungarian (**hangos kacagást kacag* 'laugh a loud laugh') and other languages (e.g. Romance, cf. Puigdollers Real 2008), hence the mechanism behind the two structures is different (Melloni & Masini 2017).

In sum, Hungarian POs are COs sharing most of the syntactic properties of ACOs. The few differences identified here stem from the fact that they are light ACOs. Moreover, these event-result constructions cannot be considered LVCs.

Acknowledgements: This research was supported by the Domus Hungarica Scientiarum et Artium research grant of the Hungarian Academy of Sciences.

References

Butt, Miriam (2010), The Light Verb Jungle: Still Hacking Away, in M. Amberber, B. Baker, and M. Harvey (eds), (2010), *Complex Predicates: Cross-linguistic Perspectives on Event Structure*, Cambridge: Cambridge University Press, 48-78.

Csirmaz, Anikó (2008), Accusative Case and Aspect, in K. É. Kiss (ed), (2008), *Event Structure and the Left Periphery. Studies on Hungarian*, Dordrecht: Springer, 159-200.

Diesing, Molly (1998), Light verbs and the syntax of aspect in Yiddish, *The Journal of Comparative Germanic Linguistics* 1(2), 119-156.

Farkas, Imola-Ágnes (2017), Miért nem tud Mari két óra alatt sétálni egyet? Az ige + *egyet* szerkezet és az *alatt* határpontos időmódosító összeférhetetlenségéről, *Nyelv- és Irodalomtudományi Közlemények* LXI(2), 119-138.

Horrocks, Geoffrey and Stavrou, Melita (2010), Morphological Aspect and the Function and Distribution of Cognate Objects Across Languages, in M. Rappaport Hovav, E. Doron, and I. Sichel (eds), (2010), *Lexical Semantics, Syntax and Event Structure*, Oxford: Oxford University Press, 284-308.

Kearns, Kate (1988), Light Verbs in English. Ms. MIT.

Lavidas, Nikolaos (2013), Null and cognate objects and changes in (in)transitivity, *Acta Linguistica Hungarica* 60(1), 69-106.

Marantz, Alec (2005), Objects out of the lexicon: objects as events. Ms. University of Viena.

Melloni, Chiara and Masini, Francesca (2017), Cognate constructions in Italian and beyond, in L. Hellan, A. Malchukov, and M. Cennamo (eds), (2017), *Contrastive Studies in Verbal Valency*, Amsterdam/ Philadelphia: John Benjamins, 219-250.

Piñón, Christopher (2001), Töprengtem egyet: azon, hogy mit jelent az *egyet*, in M. Bakró-Nagy, Z. Bánréti, and K. É. Kiss (eds), (2001), *Újabb tanulmányok a strukturális magyar nyelvtan és a nyelvtörténet köréből*, Budapest: Osiris, 182-198.

Puigdollers Real, Cristina (2008), The nature of cognate objects. A syntactic approach, in S. Blaho, C. Constantinescu, and B. Le Bruyn (eds), (2008), *Proceedings ConSOLE XVI*, 157-178.

Simone, Raffaele and Masini, Francesca (2014), On Light Nouns, in R. Simone, and F. Masini (eds), (2014), *Word Classes. Nature, typology and representations*, Amsterdam/Philadelphia: John Benjamins, 51-73.

Vincze, Veronika (2012), Light Verb Constructions in the SzegedParalellFX English-Hungarian Parallel Corpus, in N. Calzolari et al. (eds), (2012), *Proceedings of the Eighth Conference on International Language Resources and Evaluation (LREC 2012)*, Istanbul, 2381-2388.

The functions of *asang* in Lakurumau (Papua New Guinea) as an aspect and pragmatic marker

Lidia Federica Mazzitelli
(University of Cologne)

Schedule: Sat 12.00 Room 1

In my talk, I will describe the semantics of the particle *asang* in Lakurumau, a still undescribed Western Oceanic language spoken by ca. 800 people in New Ireland (Papua New Guinea). *Asang* can function as an aspect and focu marker and it has no direct cognates in the other languages closely related to Lakurumau. I base my analysis on primary fieldwork data: a corpus of Lakurumau natural speech (ca. 15 hours) and direct elicitations. *(A)sang* can occur as a particle, following a verb or a noun phrase (1, 2), or as a clitic on demonstratives and spatial deictics (3):

(1) *A* *raatai* *asang* *a* *maat*
ART man ASANG 3SG be.dead
'The man has died'

(2) *Ka* *daavui asang!*
3SG rain ASANG
'It rains! (unexpectedly)'

(3) *A vedawaai* *akamaa=sang* *di* *nga* *vedawaai pa-na* *taning*
ART language this=ASANG 1PL.INCL PROGR talk OBL-3SG today
'This (new) language in which we talk nowadays (implied: before we used another language)'

As a verbal marker, (*a)sang* expresses meanings typically related with *already*-grams in Austronesian languages (Vander Klok and Matthewson 2015), such as an earliness implication and an inchoative interpretation with stative predicates: in (1), (*a)sang* forces a dynamic interpretation of the stative *maat* 'be dead'. It can also function as a perfective marker, presenting the event as bounded, completed and anterior to subsequent events, and, when co-occurring with future grams, it acquires a prospective meaning.

As a pragmatic marker, (*a)sang* marks contrast. It can convey an 'unexpected action' or a counter-expectation (as in 2); on the noun phrase level, it signalizes that a change has occurred, affecting the *(a)sang*-marked participant. In (3), the noun phrase *a vedawaai* ***akamaa=sang*** 'this=ASANG language' implies that before another language was used, while *a vedawaai* ***akamaam*** 'this language' would imply that the language has never changed. Furthermore, (*a)sang* can convey emphasis (4) and it is used to mark elements in focus (always with the basic notion of contrast), as in (5):

(4) *Ka* *vit* *asang* *ge* *rexaas*
3SG NEG ASANG 1SG.IRR know
'I really don't know' (with implication: I was supposed to, but I do not know)

(5) *A* *laptop akamaa=sang* *ga* *daa saalim* *sam Joel*
ART laptop this=ASANG 1SG.IRR FUT sell POSS=PERS.ART J.
'It is this laptop (not any other) that I will sell to Joel'

I suggest, that the aspectual and pragmatic functions of these markers can be unified under the central semantic notion of 'change of state', which develops into an interpretation of 'new information' or 'relevant (new) information'.

References

van der Klok, Jozina and Lisa Matthewson. 2015. Distinguishing already from perfect aspect: A case study of Javanese *wis*. *Oceanic Linguistics*, 54(1):172–205

Cross-linguistic correlations between phoneme inventory size, syllable size and word length

Gertraud Fenk-Oczlon
(Alpen-Adria-Universität Klagenfurt)

Keywords: cross-linguistic correlations; phoneme inventory size; syllable size; word length;

Schedule: Sat 11.00 Room 9

Phoneme inventory size and its relationship with linguistic and extra-linguistic factors remains a hotly debated topic. Especially research reporting correlations between inventory size and extra-linguistic factors like population size (Hay & Bauer 2007) or distance of Africa (Atkinson 2011) gained a lot of interest. Relationships between inventory size and linguistic variables are reported by Nettle (1995) and Wichmann et al. (2011) who found a negative correlation between word length and inventory size; Maddieson (2006), and Fenk-Oczlon & Fenk (2008) found a positive correlation between inventory size and syllable complexity. All these studies use single words for their correlations – Nettle 50 random dictionary entries, Wichmann et al. a 40-item subset of the Swadesh list – or statistical descriptions of the permitted syllable structures in the respective languages (Maddieson, Fenk-Oczlon & Fenk). But the length of uninflected words in dictionaries or word lists, or the permitted maximum syllable complexity in individual languages do not reflect word length or syllable size in textual material.

The **objective** of the present study is therefore to examine whether a negative correlation between inventory size and word length, and a positive correlation between inventory size and syllable complexity shows also in texts.

Method: Parallel texts were obtained by an elicitation experiment: Native speakers of 56 languages from all continents (22 European, 34 Non-Indo- European) were asked to translate a matched set of 22 simple declarative sentences encoding one proposition into their mother tongue. We then calculated the average syllable complexity (in number of phonemes) and the average word length (in number of syllables and in number of phonemes) in these texts and correlated the data with the size of the language's phoneme inventories found in UPSID and/or the PHOIBLE database. For languages not represented in these databases, such as Maranungku or Mambae other sources were used.

Results:

- significant positive correlation between syllable complexity and phoneme inventory ($r = +.58$, $n = 56$, $p < .01$)
- significant negative correlation between word length in number of syllables and phoneme inventory ($r = –.38$, $n = 56$, $p < .01$)
- insignificant negative correlation between phoneme inventory size and number of phonemes per word ($r = –.14$, $n = 56$, ns**)**

Discussion: A positive significant correlation between syllable complexity and inventory size shows also when analyzing textual material. Languages with more phonemes tend to have more complex syllables. Concerning the inverse relationship between word length and phoneme inventory size in texts, our results reveal that only word length in number of syllables is significantly correlated with inventory size. Word length in number of phonemes shows only an insignificant correlation with inventory size. This rather unexpected result will be discussed referring to studies by Fenk-Oczlon & Fenk (1999) who report a negative cross-linguistic correlation between number of phonemes per syllable and number of syllables per word: The more syllables per word, the fewer phonemes per syllable. One might argue that the significant correlation between word length and inventory size found by Nettle and Wichmann et al. predominantly applies to words with roughly equal numbers of syllables but different numbers of phonemes.

References

Atkinson, Quentin D. (2011), Phonemic diversity supports a serial founder effect model of language expansion from Africa, *Science* 332, 346–349.

Fenk-Oczlon, Gertraud and August Fenk (1999.) Cognition, quantitative linguistics, and systemic typology, *Linguistic Typology* 3, 151–177.

Fenk-Oczlon, Gertraud and August Fenk (2008), Complexity trade-offs between the subsystems of language, in M. Miestamo, K. Sinnemäki and F. Karlsson (eds.), (2008), *Language complexity: Typology, contact, change*, Amsterdam / Philadelphia: John Benjamins, 43– 65.

Hay, Jennifer and Laurie Bauer (2007), Phoneme inventory size and population size, *Language* 83, 388–400.

Nettle, Daniel (1995), Segmental inventory size, word length, and communicative efficiency, *Linguistics* 33, 359–367.

Maddieson, Ian (2006), Correlating Phonological Complexity: Data and Validation, *Linguistic Typology* 10, 106–123.

Maddieson, Ian and Kristin Precoda (no date), *UCLA Phonological Segment Inventory Database*, Electronic database, University of California, Los Angeles. Accessible via interface at http://web.phonetik.uni-frankfurt.de/upsid.html

Moran, Steven, McCloy, Daniel and Wright, R. (eds.) (2014), *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at https://phoible.org

Wichmann, Søren, Rama, Taraka and Eric W. Holman (2011), Phonological diversity, word length, and population sizes across languages: The ASJP evidence, *Linguistic Typology* 15, 177–197.

ꝺ

On the diachrony of nominal classification in Guang

Ines Fiedler & Tom Güldemann
(Humboldt University Berlin; & Humboldt University Berlin and Max Planck Institute for the Science of Human History)

Keywords: Nominal classification, gender, historical reconstruction, typology, Niger-Congo

Schedule: Fri 16.30 Room 6

Guang languages are known for their elaborate nominal prefix system, but are said to show little in the way of agreement (Dakubu 1988: 82). Accordingly, the two already existing reconstructions of their nominal classification systems (Snider 1988, and Manessy 1987) relied mainly on noun form classes and included agreement forms only marginally and in a non-systematic way.

This presentation compares the synchronic gender systems of most of the 18 Guang languages in order to establish an improved reconstruction of the proto-system. This reconstruction differs methodologically from the former in two points: (i) the data base includes more languages with new data, and (ii) it is based on a more systematic cross-linguistically oriented approach to the analysis of gender systems, as proposed by Güldemann (2000) and, with a particular focus on Niger-Congo, Güldemann and Fiedler (2019). This approach consistently distinguishes the four following gender-related concepts: (a) agreement classes as the ultimate basis of the system of genders, (b) gender (classes) as reflecting the classification of nouns in the lexicon/reference domain, (c) nominal (form) classes as the partial reflex of genders hosted by the nouns themselves and (d) deriflection classes as reflecting the morphological variation of nouns triggered by number, gender etc.

The comparison reveals huge differences with respect to design and complexity of gender and declension systems in single Guang languages and across Guang. Even though a language like Foodo presents strong tendencies towards an alliterative mapping of some agreement classes with some nominal form classes (cf. example 1a where the exponents of both have the same morphological form *dV*^{Hi}- and *A*-), this is not valid for other agreement and nominal form classes. In example (1b), the nominal form class *N*- triggers the two different agreement classes *sV*^{Hi} and *bV*^{Hi}.

(1) Foodo (Benin, Mountain Oti-North)

	NF-noun-NF	NF-big-NF	AGR-IDEF	
a.	**dí**-gbá-**lì**	**dɔ̀**-ńlɛ́-**lɪ̀**	**dɔ̀**-kɔ́	‘a big market’
	á-gbá-**à**	**à**-ŋ́lɛ́-**ɛ̀**	**à**-kɔ́	‘some big markets’
b.	**ŋ**-kɔ́lˈɔ̀-**sɛ́-ɛ̀**	**ǹ**-ńlɛ́-**ɛ̀**	**sɔ̀**-kɔ́	‘some big rivers’
	ǹ-cɔ́-**ḿ**	**ǹ**-ńlɛ́-**ɛ̀**	**bɔ̀**-kɔ́	‘big water’

Any reconstruction of the nominal classification system has thus to analyze gender and deriflection systems separately, as both develop differently and independently from each other, resulting in a systemic difference in synchronic languages. Whereas the deriflection systems reflect a more conservative state of nominal classification in Guang, the gender systems of some Guang languages show a simplification towards an animate-inanimate distinction.

We argue that this new theoretical approach not only clarifies the situation in Guang, but also contributes to the reconstruction of gender systems in Niger-Congo in general.

References:

Dakubu, Mary Esther Kropp (1988), The Volta-Comoé languages, in M.E. Kropp Dakubu (ed), (1988), *The languages of Ghana*, London: Kegan Paul International (for the International African Institute), 50-90.

Güldemann, Tom and Ines Fiedler (2019), Niger-Congo "noun classes" conflate gender with deriflection, in F. Di Garbo, B. Olsson and B. Wälchli (eds), *Grammatical gender and linguistic complexity*, Berlin: Language Science Press, 97-150.

Güldemann, Tom (2000), Noun categorization systems in Non-Khoe lineages of Khoisan. *Afrikanistische Arbeitspapiere* 63:5-33.

Manessy, Gabriel (1987), La classification nominale en Proto-Guang, *Afrikanistische Arbeitspapiere* 9:5-49.

Snider, Keith (1988), The noun class system of Proto-Guang and its implications for internal classification, *Journal of African Languages and Linguistics* 10, 137-64.

The diachronic development and use of German *ich meine* and English *I mean*

Daniela Kolbe-Hanna & Natalia Filatkina
(University of Trier, Germany)

Keywords: Diachronic linguistics, German, English, comment clauses, construction grammar

Schedule: Thu 14.00 Room 3

The comment clauses *ich meine* in German and *I mean* in English can be considered similar in the main aspects of their meaning and function. Both also occur in the right periphery (RP, see (1) and (2)) and in the left periphery, (LP, see (3) and (4)), In German, *ich* and *meine* are inverted and yield *meine ich* in the RP (example (3)):

(1) ***Ich meine**, der Fehler kann nur in der mangelnden Initiative liegen.* (DWDS)
(2) ***I mean**, QP has been discussing the Accommodation Problem for the whole of my lifetime*. (BNC CC5 96)
(3) *Das sind eben die Vorboten der kommenden Ereignisse in China, **meine ich***. (DWDS)
(4) *I expect you know how to get in touch with her, in case of trouble, **I mean***. (BNC**,** FB9 2486)

Various functions of these comment clauses in both languages have been widely studied (Schiffrin 1987, Stein 1995: 239, Auer 1998, Auer and Günthner 2003, 2005, Günthner and Imo 2003, Beeching 2016 and Imo 2007), Research, however, ,mainly focuses on their use in modern times. From a construction grammar perspective, Imo (2006, 2005), assumes sentence-initial, matrix clause position as the original source construction from which the pragmatic markers have grammaticalised. Brinton (2008: 111-132), however, argues that there is no diachronic evidence for the claim that the original source of the pragmatic marker in English is a matrix clause followed by a *that*-complement clause. Instead, it is most often followed by a phrasal argument and the most frequent complement clause is infinitival. For Old German before 1050, the *daz*-complement clause is only a third frequent pattern that follows accusative objects and nominal phrases (Filatkina 2018: 219-226), Both Brinton and Imo, however, argue for grammaticalisation in sentence-initial or LP position. For English, Brinton (2007: 58-59) dates back the beginning of this process in the late Middle English, for German, Filatkina (2018: 225-226) traces it back to the earliest period of the written text tradition (750-1050),

This paper examines the use of *I mean* and *Ich meine* in the left and right periphery in corpus data from Old to Modern English and from Old to Modern German. Our goals are threefold: 1) We aim to investigate the main factors that play a role on the diachronic pathway of *I mean/Ich mein(e)* to pragmatic markers, taking into account typological differences such as evolving word-order differences between English and German, as well as other criteria. 2) With the help of the diachronic case studies and available research results from both (spoken) modern languages, we tackle the question if this path can be sufficiently described as grammaticalisation. 3) We will conclude with some new theoretical insight that appears to be relevant for the diachronic dimension of Construction Grammar.

References

Auer, Peter. (1998), Zwischen Parataxe und Hypotaxe: ‚abhängige Hauptsätze' im gesprochenen und geschriebenen Deutsch, *in Zeitschrift für Germanistische Linguistik* 26, 284-307.

Auer, Peter, and Günthner, Susanne. (2003), Die Entstehung von Diskursmarkern im Deutschen. Ein Fall von Grammatikalisierung? in *InLiSt (Interaction and linguistic structures)* 38, 1-30. https://kops.uni-konstanz.de/bitstream/handle/123456789/3832/Inlist38.pdf?sequence=1

Auer, Peter, and Günthner, Susanne. (2005), Die Entstehung von Diskursmarkern im Deutschen. Ein Fall von Grammatikalisierung? in T. Leuschner, T. Mortelmans, and S. De Groodt (eds), (2005), *Grammatikalisierung im Deutschen*. Berlin/New York: de Gruyter, 335-362.

Beeching, Kate. (2016), *Pragmatic markers in British English*. Cambridge: Cambridge University Press.

Brinton, Laurel. (2007), The development of *I mean*: Implications for the study of historical pragmatics, in S. Fitzmaurice, and I. Taavitsainen (eds), (2007), *Methods in historical pragmatics*. Berlin/New York: de Gruyter, 37-79.

Brinton, Laurel. (2008), *The comment clause in English*. Cambridge: Cambridge University Press.

Filatkina, Natilia. (2018), *Historische formelhafte Sprache. Theoretische Grundlagen und methodische Herausforderungen*. Berlin/Boston: de Gruyter.

Günthner, Susanne, and Imo, Wolfgang. (2003), Die Reanalyse von Matrixsätzen als Diskursmarker: *ich mein*-Konstruktionen im gesprochenen Deutsch, in *InLiSt (Interaction and linguistic structures)* 37, 1-31. https://kops.uni-konstanz.de/bitstream/handle/123456789/3837/Inlist37.pdf?sequence=1.

Imo Wolfgang. (2005), A Construction Grammar approach to the phrase "I mean" in spoken English, in *InLiSt* (*Interaction and Linguistic Structure*) 42. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-15607.

Imo, Wolfgang. (2006), Zur Anwendung der *Construction Grammar* auf die gesprochene Sprache: der Fall `ich mein(e)´, in V. Ágel, and M. Hennig (eds), (2006), Zugänge zur Grammatik der gesprochenen Sprache. Berlin: De Gruyter, 3–34. https://doi.org/10.1515/9783110936063.3

Imo, Wolfgang. (2007), *Construction Grammar und Gesprochene Sprache-Forschung. Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen: Niemeyer.

Schiffrin, Deborah. (1987), *Discourse markers*. Cambridge: Cambridge University Press.

Stein, Stephan. (1995), *Formelhafte Sprache. Untersuchungen zu ihren pragmatischen und kognitiven Funktionen im gegenwärtigen Deutsch*. Frankfurt a.M.: Peter Lang

Corpora

BNC = British National Corpus (http://bncweb.lancs.ac.uk/)

DWDS = DWDS (*Digitales Wörterbuch der Deutschen Sprache*) Kernkorpus 1900-1999 (https://www.dwds.de/d/k-referenz#kern)

HiFoS = Datenbank der Nachwuchsforschergruppe „Historische formelhafte Sprache und Traditionen des Formulierens (HiFoS)“ (http://www.hifos.uni-trier.de)

Dialectal evidence in establishing substratum influence: the case of ‘Celtic Englishes’

Markku Filppula

Schedule: We 15.00 Room 8

<pdf>

Applied usage-based linguistics: Variation and entrenchment in Scandoromani

Guro Fløgstad

Schedule: We 16.30 Room 8

<pdf>

Optimal categorisation: a psycholinguistic perspective on the development of gender systems

Michael Franjieh, Greville Corbett and Alexandra Grandison

Schedule: We 11.00 Room 5

<pdf>

Reflexivizing Spanish psych-verbs: Ambiguities across classes

Paola Fritz-Huechante & Antonio Machicao y Priemer
(Humboldt-Universität zu Berlin)

Keywords: Spanish, reflexive, psych-verb, inchoative-causative alternation, HPSG

Schedule: Sat 11.30 Room 12

Psych-verbs in Romance languages have been categorized in three classes (Belleti and Rizzi 1988). In one class the subject is interpreted as the experiencer (ES, class I, (1)), and in the other two classes the experiencer is the object (EO, class II, (2) & class III, (3)). Class II shows a systematic experiencer-alternation from object to subject by means of reflexivization (*molestar* > *molestar-se*), (2). The *se*-morpheme attached in the derived item in (2c) has been analyzed as: anti-causativizer (Schäfer 2008), inchoativizer (de Miguel and Fernández 2000), and recently as a left-boundary marker for the so-called inchoative states (Marín and McNally 2011).

(1) a. Luisa$_{EXP}$ *ama* a Clara$_{STM}$. 'Luisa loves Clara.'
b. Luisa$_{EXP}$ *se ama* (a sí misma). 'Luisa loves herself.'
(2) a. Clara$_{STM}$ *molesta* a Luisa$_{EXP}$. 'Clara bothers Luisa.'
b. Luisa$_{EXP}$ *se molesta* (a sí misma). 'Luisa bothers herself.'
c. Luisa$_{EXP}$ *se molesta* (por Clara$_{STM}$). 'Luisa gets bothered (by Clara).'
(3) a. A Luisa$_{EXP}$ le *gusta* Clara$_{STM}$. 'Luisa likes Clara.'
b. *Luisa$_{EXP}$ *se gusta*. Intended reading: 'Luisa likes herself.' / 'Luisa gets liked.'

We propose an analysis in HPSG in terms of reflexivization where:
• class I derives only true reflexives (e.g. co-occurring with *a sí mismo* 'himself', (1b)),
• class III does not accept the *se*-morpheme under any circumstances (Belleti and Rizzi 1988), (3b); and
• class II presents a *se*-morpheme that is ambiguous between: (a) a true reflexive (2b) and (b) an expletive (2c) (i.e. a non-thematic argument). Combining *se* with the verbal lexeme licences a syntactic argument reduction in the sense of anti-causativization (Alexiadou et al. 2015).

In addition, data shows that class II – although being able to have both readings with the *se*-morpheme ((a) & (b) above) – does not behave uniformly. Assuming that Spanish reflexive psych-verbs can be further specified into punctual psych-verbs (e.g. *asustarse* 'get frightened') and inchoative states (e.g. *molestarse* 'get bothered') (Marín and McNally 2011), the latter class but not the former shows a distinction in acceptability. For inchoative state verbs, e.g. *molestarse* 'get bothered', both interpretations are equally accepted, whereas for inchoative state verbs, e.g. *deprimirse* 'get depressed', the true reflexive interpretation is more marked. On the contrary, the punctual class remains stable with respect to acceptability judgments. Empirical evidence shows that there is a correlation between the agentivity of the stimulus and the acceptability of the true reflexive interpretation of the verb. Firstly, based on experimental data, we show how speakers' judgments vary between the two different readings of the *se*-morpheme in the psych domain. Secondly, we model the alternative readings of *se* in HPSG building on a neo-Davidsonian analysis (cf. Copestake 2006). We will show how to model the different subtypes of psych-verbs and derive the inchoativized alternants (2c) by means of a lexical rule (LR). This LR takes only the causative psych-verbs as input, not allowing the inchoativization of the subtypes such as *amar* 'love' and *gustar* 'like'.

Acknowledgments: This paper was partly funded by the German Research Association (DFG; project VE 570/1–3).

References

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer (2015), *External Arguments in Transitivity Alternations: A Layering Approach*, Oxford: Oxford University Press.

Belletti, Adriana and Luigi Rizzi (1988), Psych-verbs and θ-theory, *Natural Language and Linguistic Theory*, 6(3), 291–352.

Copestake, Ann (2006), Robust Minimal Recursion Semantics. Ms.

De Miguel, Elena and Marina Fernández (2000), El operador aspectual se, *Revista Española de Lingüística*, 30, 13–43.

Marín, Rafael and Louise McNally (2011), Inchoativity, change of state, and telicity: Evidence from Spanish reflexive psychological verbs, *Natural Language and Linguistic Theory*, 29, 467–502.

Schäfer, Florian (2008), *The syntax of (anti-)causatives: External arguments in change-of-state contexts*, Amsterdam: John Benjamins Publishing Company.

Grammaticalization and Paradigmatic Saturation: a Sisyphean Task

Livio Gaeta
(University of Turin)

Keywords: language change, grammaticalization, morphology, inflection, paradigm

Schedule: Thu 14.00 Room 5

While the role played in grammatical systems by paradigms as well as by the paradigmatic dimension in Saussurean sense can hardly be denied, its relation to grammaticalization remains disputed. Although these notions are clearly connected, the latter refers to the distributional context defining alternative choices depending on specific selective properties, while the former also refers to the particular relations which can be inferred to exist among the different forms. Accordingly, paradigms are taken to be internally structured and saturated, namely to exhaust their potential filling out all slots (cf. Gaeta & Angster 2019 for a discussion).

As for the relation to grammaticalization, two aspects need to be distinguished. First, thanks to grammaticalization new markers come into existence saturating the paradigmatic space available. This is shown by the second person suffix *-st* coming from the agglutination of the personal pronoun *d*(*u*) to the original suffix *-s* found in Old High German: *nimis du* > *nimisdu* > *nimmst* 'take.2SG', which spread as a super-stable marker (cf. Wurzel 1984) throughout all German verbs also replacing the older suffix: OHG *nām-i* > *nahm-st* 'took-2SG'. On the other hand, grammaticalization can also have a disruptive effect on paradigms insofar as it gives rise to trapped morphemes disturbing the normal inflectional behavior as shown by the adverb *dén* 'then' grammaticalized as future marker in Töitschu, a Walser German dialect of Northern Italy, bringing about introflection: *wiss* / *wiss-en* 'know.1SG / know-1PL', *wiss-dén* / *wiss-en-dén* 'know.1SG-FUT / know-1PL-FUT'.

Second, the notion of expansion context has been suggested with the aim of implementing concretely Lehmann's (1982) parameter of paradigmaticity (cf. Himmelmann 2004). In spite of its theoretical interest, few investigations have been devoted to understand precisely whether such expansion contexts follow a paradigmatic rationale in which due to saturation a given grammaticalized construction like the perfect expands to other complex constructions such as the modal construction, the passive, etc. That such a paradigmatic saturation is not a linear process is shown by the rise of mismatches due to conflicts such as for instance in the auxiliary selection of the perfect in cases like Italian *Gianni ha* / *è dovuto andare a casa* 'John had to go (lit. has / is must.PSTPTCP go.INF) home' where the conflict comes from the different selectional properties displayed by the modal taking HAVE and the main verb taking BE. In other cases, the conflict gives rise to unexpected solutions like for instance the so-called German substitutive infinitive (cf. Gaeta 2010): *Hans hat nach Hause gehen müssen* / **gemusst* 'Hans had to go home, lit. has go.INF must.INF / *must.PSTPTCP). The appearance of the infinitive results from a morpho-semantic conflict between the bounded value of the participle forced by the grammaticalization of the prefix *ge-* clashing with the modal meaning. While it provides substantive evidence for paradigmatic strength because it results from the expansion of the perfect construction saturating the paradigmatic space, this form-meaning mismatch also has a disruptive effect on paradigms as in a never-ending Sisyphean task.

References

Gaeta, Livio and Marco Angster (2019), Stripping paradigmatic relations out of the syntax, *Morphology* 29(2), 249–270.

Gaeta, Livio (2010), The Invisible Hand of Grammaticalization: West-Germanic Substitutive Infinitive and the Prefix *ge-*, In W. U. Dressler, D. Kastovsky, H. Ch. Luschützky and F. Rainer

(eds.), *Variation and Change in Morphology. Selected papers from the 13th International Morphology Meeting, Vienna February 2008*, Amsterdam / Philadelphia: John Benjamins, 89–105.

Himmelmann, Nikolaus P. (2004), Lexicalization and grammaticization: Opposite or orthogonal?, in W. Bisang, N. Himmelmann and B. Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, Berlin: Mouton de Gruyter, 21–42.

Lehmann, Christian [1982] (2015), *Thoughts on grammaticalization*, 3rd ed., Berlin: Language Science Press.

Wurzel, Wolfgang U. (1984), *Flexionsmorphologie und Natürlichkeit*, Berlin: Akademie-Verlag.

How input delivered by one parent differs from that delivered by two: implications for the bilingual acquisition of a heritage language at the preschool age

Dorota Gaskins
(University of London)

Keywords: usage-based; acquisition; heritage language; dyadic; triadic

Schedule: Sat 12.00 Room 8

Preschool children learning a heritage language are often reported to produce fewer words (Hoff et al., 2012; Thordardottir, 2014) and fewer grammatical constructions in that language (e.g. Paradis et al., 2014). This imbalance in their learning outcomes is often linked to the imbalance of the two languages in their input, especially if the heritage language is addressed to the child only by one parent at home. In fact, bilingual children are most likely to speak both of the languages presented to them if they hear them from both parents and if, at most, one parent speaks the language of the society at home (De Houwer, 2007). How different is the child's input when delivered by one speaker of a heritage language as opposed to two? Is there anything missing from the input delivered by one parent which is more clearly discernible in that delivered by two? These questions are examined in my usage-driven study of naturalistic interactions carried out between the ages of 1;5-2;5 with a child exposed to Polish and English from birth. Of these, 10 hours of data are derived from recordings made with the child's Polish-speaking mother and 10 with her mother and her maternal grandmother. The data are compared and discussed in light of the usage-based theory. Parental input is compared across the two groups in terms of frequencies, measured as total vocabulary counts recorded on video, and complexity, measured as both mean utterance length and type-token ratios.

Further examination relies on a small subset of verbs and the manner in which they are used in child-directed speech in the two types of recordings. Does the responsiveness of one caregiver to the speech of the other help to clarify the meaning of constructions? This would be expected in light of experimental research which shows that in the learning of verbs, which are a gateway to more complex constructions, children are more likely to learn a verb in an utterance '*A man is waving a balloon'* and not '*A man is waving it'* because the former provides richer semantic context for the child to rely on when deriving the meaning of the verb (Arunachalam & Waxman, 2015). Rich semantics is expected to be missing from contexts of bilingual exposure: if the phrase '*Give me the banana',* for example, was uttered by one caregiver and acted upon by the other, the meaning of the verb would be clarified in situ, a quality missing from the speech merely directed at a pre-linguistic

child in dyadic interactions. Showing links between the complexity of input heard and language acquisition could help us to understand a) the roots of child's willingness to engage with the language; b) their readiness to produce the language and c) their ability to develop some competence in it.

References

Arunachalam, Sudha & Waxman, Sandra (2015) Let's See a Boy and a Balloon: Argument Labels and Syntactic Frame in Verb Learning. *Language Acquisition* 22(2), 117-131.

De Houwer, Annick (2007) Parental language input patterns and children's bilingual use. *Applied Psycholinguistics* 29(03), 411-424.

Hoff, Erika, Cynthia Core, Sylvia Place, Rosario Rumiche, Melissa Senor, and Marisol Parra (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39: 1-27.

Paradis, Johanne, Antoine Tremblay and Martha Crago (2014) French-English bilingual children's sensitivity to child-level and language-level input factors in morphosyntactic acquisition. In T. Grüter, & J. Paradis (Eds.), *Input and Experience in Bilingual Development*. Amsterdam: John Benjamins.

Thordardottir, Elin (2014). The typical development of simultaneous bilinguals: Vocabulary, morphosyntax and language processing in two age groups of Montreal pre-schoolers. In T. Grüter, & J. Paradis (Eds.), *Input and Experience in Bilingual Development*. Amsterdam: John Benjamins.

How to get a sample of verbs for a comparative study?

Jerzy Gaszewski
(University of Wrocław)

Keywords: linguistic methodology, typology, areal linguistics, verb valency, frequency

Schedule: Fri 16.00 Room 7

Some comparative studies rely on a preselected list of items from an open class, e.g. a selected list of verbs. The choice of items for such a list is essential as all the results are contingent upon it. However, "we do not have a good way of selecting representative samples of verbs" (Haspelmath 2015: 141). Bickel et al. (2014: 488) concur: "pre-selection of predicate meanings brings with it the risk that the results are in part pre-determined". It is also hard not to agree with Hartmann et al. (2014: 467) that such a list should be "representative of the verbal lexicon". However, this is easier said than done. Typically the way to achieve this key feature of the list is by using the researchers' judgement. In this paper, I propose how to make the selection process more objective and the resultant list more representative.

One area of study that stumbles upon this problem is the research of valency and semantic roles (cf. the works quoted here) and I use specific examples from my own areal study on valency. Yet, the methodology itself may well be applied to other topics (e.g. specific subclasses of verbs, light verb constructions, etc.) investigated from areal or typological perspective.

The core idea is to combine a study-oriented criterion (e.g. two-place predicates with non-default case marking as in Say 2014) with the objective criterion of frequency of occurrence. In practice this means taking a frequency list of verbs and sieving it according to the study-oriented criterion. In this way, we obtain a list of verbs that are common in use (and thus representative of the verb class as it is used by the speakers) as well as relevant to the study.

The question that begs the answer at this point is the frequency in what language should be taken into account. The procedure can clearly be applied to a frequency list based on a reference corpus, but we rely on one language then, which is thus given unjustified preference. The solution is to re-apply the procedure to several languages and aggregate the results to produce the list for the study. These languages should be sampled from the group to be used in the study proper. For example, in my study focusing on about ten languages of Central and Eastern Europe the verb selection was based on three of them: German, Hungarian and Polish.

The paper gives a detailed account of the proposed procedure. The discussion of individual stages focuses on the array of options available at each point since different methodological decisions may be appropriate depending on the topic and scope of a particular study.

References:

Bickel, Balthasar, Taras Zakharko, Lennart Bierkandt and Alena Witzlack-Makarevich (2014), Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment, *Studies in Language* 38(3), 485 - 511.

Hartmann, Iren, Martin Haspelmath and Michael Cysouw (2014), Identifying semantic role clusters and alignment types via microrole coexpression tendencies, *Studies in Language* 38(3), 463–484.

Haspelmath, Martin (2015), Transitivity prominence, in: A. Malchukov and B. Comrie (eds), (2015) *Valency Classes in the World's Languages*, vol. 1, Berlin, Boston: de Gruyter Mouton, 131-47.

Say, Sergey (2014), Bivalent verb classes in the languages of Europe: A quantitative typological study, *Language Dynamics and Change* 4(1), 116–166.

The so-called 'partitive article' in Old Spanish and Old Portuguese

David Paul Gerards

Schedule: ~~Thu 16.00 Room 14~~ Fri 16.30 Room 14

It has repeatedly been claimed that, in some contexts, Old Spanish (OSp) *del* (1a) and its Old Portuguese (OPt) cognate *do* are so-called *partitive articles* (cf., e.g., Lapesa 1964, Mattos e Silva 2008). Partitive articles (PAs) are indefinite determiners available in some modern Gallo- (1b) and Italo-Romance varieties (1c).

(1) a. OSp. *destiénpren=la [...]* *con* ***del*** *vinagre fuerte.*
dissolve.IMP.3PL=3FSG.ACC with **PA?** vinegar strong
'Dissolve it with [Ø ?] strong vinegar.'
(*Gerardus Falconarius*, 13th century)

b. Mod.Fr. *la=dissoudre* *avec* ****(du)*** *vinaigre* *fort*
3FSG.ACC=dissolve.INF with **PA** vinegar strong

c. Mod.It. *scioglier=la* *con* ***(dell')*** *aceto* *forte*
dissolve.INF=3FSG.ACC with **PA** vinegar strong
'Dissolve it with strong vinegar.'

Typological approaches based on observations of correlations (e.g., sigmatic nominal plural marking in Spanish/Portuguese vs. absence of [unambiguous] nominal number marking in French/Italian [cf. Stark 2008, Carlier/Lamiroy 2014]; but see also Körner [1981] for a different proposal) predict that PAs are not to be expected in any variety of Spanish or Portuguese. The aim of the contribution,

therefore, is to assess whether OSp *del* and OPt *do* really do falsify established typological generalizations.

A detailed morphosyntactic and semantic analysis of 5,000 nominals (2,300 from *Corpus Diacrónico del Español*; 2,700 from 27 OSp and OPt technical treatises [13th–16th century]) will reveal that this is not the case. Instead, the respective nominals turn out to always be *bare partitives* (Hoeksema 1996), i.e., PPs embedding a superset-denoting definite. Yet, I will show that in 13th–15th-century OSp and OPt, such superset-denoting definites cannot only feature strong definite articles but also receive weakly referential, truth-conditionally indefinite *representative object interpretations* (Krifka et al. 1995; Oosterhof 2006). This is supported by the fact that representative object definites are not only attested embedded in PPs (1a), but co-vary freely with *unembedded* representative object definites (2) and bare nominals:

(2) Osp. *destiénpra=lo [...]* *con* ***el*** *binagre fuerte.*
dissolve.IMP.2SG=3NSG.ACC with **DEF.ART** vinegar strong
'Dissolve it with strong vinegar.'
(*Dancus Rex. Esc. V.II.19*, 13th century)

Typically, weakly referential definites, proposed to be contained in (1a) and (2), are associated with late stages of article grammaticalization only (cf., e.g., Mulder/Carlier 2011). The analysis proposed thus calls for further evidence. Following Epstein's (2001, 179) analysis of weakly referential uses of the Old French definite article, a plausible hypothesis is that OSp and OPt representative object definites ([1a], [2]), but not bare nominals, introduce important discourse referents. This will be investigated by measuring the length of anaphoric chains. If the hypothesis is confirmed this would also explain the data's heavy bias of weakly referential representative object definites towards substance-denoting nouns in technical treatises (> 97%).

Overall thus, the paper will not only show that OSp and OPt – contrary to some of the literature – did not feature 'partitive articles' but also formulate the question of whether the grammaticalization process of the Romance definite article should really be treated as an absolutely linear development.

References

Carlier, A., and B. Lamiroy. 2014. "The Grammaticalization of the Prepositional Partitive in Romance." In *Partitive Cases and Related Categories*, ed. by S. Luraghi, and T. Huumo, 477–520. Berlin: de Gruyter.

Epstein, R. 2001. "The Meaning of Definite Articles in Cross-Linguistic Perspective." In *Cognition in Language Use*, ed. by E. Németh, 174–189. Antwerpen: International Pragmatics Association.

Hoeksema, J. 1996. "Introduction." In *Partitives: Studies on the Syntax and Semantics of Partitive and Related Constructions*, ed. by Jacob Hoeksema, 1–24. Berlin: de Gruyter.

Körner, K.-H. 1981. "'Teilungsartikel' im Französischen und 'präpositionaler Akkusativ' im Spanischen: komplementäre Lösungen des gleichen syntaktischen Problems." In *Sprache: Formen und Strukturen*, ed. by M. Kohrt/J. Lenerz, 151–160. Tübingen: Narr.

Krifka, M., et. al. 1995. "Genericity: an Introduction." In *The Generic Book*, ed. by G. Carlson/F. J. Pelletier, 1–124. Chicago: UCP.

Lapesa, R. 1964 [2000]. "Los casos latinos: restos sintácticos y sustitutos en español." In *Estudios de morfosintaxis histórica del español*, vol. 1, ed. by R. Cano/Mª. T. Echenique, 73–122. Madrid: Gredos.

Mattos e Silva, R. V. 2008. *O português arcaico. Uma aproximação*. Lisboa: Imprensa Nacional – Casa da Moeda.

Mulder, W. de, and A. Carlier. 2011. "The Grammaticalization of Definite Articles." In *The Oxford Handbook of Grammaticalization*, ed. by B. Heine, and H. Narrog, 522–534. Oxford: Oxford

University Press.
Oosterhof, A. 2006. "Dutch Bare Plurals, Faded Partitives and Subject-Object Asymmetry." *Belgian Journal of Linguistics* 19: 59–91.
Stark, E. 2008. "The Role of the Plural System in Romance." In *The Paradox of Grammatical Change. Perspectives from Romance*, ed. by U. Detges/R. Waltereit, 57–84. Amsterdam/Philadelphia: Benjamins.

On the simple past in Northern Italian: a look from quotative parentheticals

Alessandra Giorgi
(Ca' Foscari University of Venice)

Schedule: Thu 9.30 Room 7

Problem and data: To locate an event in the past Northern Italian (NI) speakers, as opposed to Southern Italian ones, do not use the simple past – *passato remoto* – but the periphrastic form – *passato prossimo* – i.e. auxiliary have/be + past participle:

(1) Ieri Gianni **ha incontrato** Maria
Yesterday Gianni *has met* Maria
(2) *Ieri Gianni **incontrò** Maria
Yesterday Gianni *met* Maria

This property is very general, so that it might be argued that the simple past is absent from the NI tense inventory.
This cannot be the whole story, however, since there is a context where the simple past is perfectly acceptable even in NI, i.e. in quotative parentheticals:

(3) Partirò domani, **disse** Gianni
I will leave tomorrow, *said* Gianni

This kind of constructions is found in narrative texts and belongs to a special, usually written, register. However, no NI speaker has any doubt that there is a clear-cut contrast between (2), i.e. the normal usage, and (3). Moreover, in these cases the periphrastic form is even judged not felicitous.

The proposal: I will argue that (3) is possible for two different reasons, which conspire in yielding grammaticality.

The first one concerns the properties of the verbal form in itself. I will argue that the one appearing in (3) is no *past* tense at all, but a *pure perfective* inflected form. I.e., I will show that there is no *temporal* value directly expressed by means of *disse* (said) in NI. These perfective, non-temporally marked forms exist in many languages and are often identified as *aorist* forms.

The second one has to do with the syntactic structure assigned to (3) – see Collins ad Braningan (1997), Matos (2013) and Giorgi (2016). Given that, as argued above, *disse* (said) is a legitimate non-temporal verbal form in NI, the contrast between (2) and (3) must be accounted for. I will argue that this is due to the fact that these parentheticals can be syntactically defective. The relevant evidence comes from the analysis of their left periphery, together with subject inversion. Consider for instance

that high sentential adverbs, which usually appear on the left with normal, *flat*, intonation, in these cases only appear on the right:

(4) Fortunatamente è arrivato Gianni
Luckily arrived Gianni (luckily Gianni arrived)
(5) *È arrivato fortunatamente Gianni
Has arrived luckily Gianni
(6) *Partirò domani, fortunatamente **disse** Gianni
I will leave tomorrow, luckily *said* Gianni
(7) Partirò domani, **disse** fortunatamente Gianni
I will leave tomorrow, *said* luckily Gianni

The evaluative adverb *fortunatamente* (luckily) (Cinque, 1999) appears on the left of the verb, in (5), whereas on its right in (7). This kind of data shows that in these cases the left periphery is defective, or missing at all.

Conclusions: By means of this analysis this work aims at achieving two goals: one the one hand the properties of a-temporal inflected forms is further clarified, and on the other the syntax of parentheticals is better understood.

References

Cinque, Guglielmo, 1999, *Adverbs and Functional Heads*, Oxford: Oxford University Press.
Collins Chris and Paul Branigan. 1997. Quotative inversion, *Natural Language & Linguistic Theory*, Vol. 15, No. 1:1-41.
Giorgi Alessandra, 2016, Integrated parentheticals in quotations and free indirect discourse, in: A. Capone et al.(eds) *Indirect reports and pragmatics*, Springer Int. Pub.: 471-488.
Matos, Gabriela. 2013. Quotative inversion in Peninsular Portuguese and Spanish, and in English, *Catalan Journal of Linguistics*, 12: 111-130.

How different were phonological distributions? The typology of phonological segment borrowing and the Uniformitarian Assumption

Eitan Grossman, Dmitry Nikolaev & Steven Moran
(Hebrew University, Hebrew University & University of Zurich)

Keywords: typology, language contact, phonology, historical linguistics

Schedule: Sat 12.30 Room 9

Most discussions of the Uniformitarian Assumption in linguistics (Labov 1972; Lass 1997; Newmeyer 2002; Janda & Joseph 2003) stress that the Uniformitarian Assumption relates to the time-independence of the historical processes undergone by languages. However, it is also the case that linguists regularly make inferences about markedness, naturalness, or learnability on the basis of present-day distributions, particularly the cross-linguistic frequency of linguistic properties. We call this the *Implicit Uniformitarian Assumption*, according to which theories of grammar can be based on present-day distributions. This, in turn, seems to be founded on what Lass (1997: 29) has called the *Uniform Probabilities Principle*: 'The (global, crosslinguistic) likelihood of any linguistic state of affairs (structure, inventory, process, etc.) has always been roughly the same as it is now.'

In this talk, we explore the contribution of event-based factors (Bickel 2015) in shaping present-day distributions in sound systems. Our study is based on the comparison of three large-scale databases: PHOIBLE (Moran et al. 2014), for present-day distributions; BDPROTO (Marsico et al. 2018), for distributions in reconstructed proto-languages; and SEGBO (Grossman et al. 2019), a new world survey of phonological segment borrowing. At present, this survey includes over 1600 events of phonological segment borrowing in a sample of the world's languages (492 borrowing languages, 228 donor languages).

The main focus of this talk is the question: what is the contribution of phonological segment borrowing to the present-day distribution of segments in phonological inventories? The main hypothesis to be explored is that the present-day distribution of sound patterns differs from the distribution of sound patterns in the world's languages around 500–1000 years before present and that the difference can be attributed, at least in part, to event-based triggers, i.e., the contingencies of human history that brought languages into contact with each other. In other words, the hypothesis is that the Implicit Uniformitarian Assumption does not hold for sound inventories. For example, labiodental fricatives, despite their high frequency in present-day distributions, are often the result of relatively recent borrowing events, which corroborate their evolutionary lateness in phonological systems (Blasi et al. 2019). Moreover, our data points to the conclusion that pre-contact phonological inventories showed greater areal specificity, both at a macro-level (i.e., roughly continent-sized areas) and at a micro-level (i.e., of what are commonly thought of as linguistic areas, e.g., the Andes, South Asia, or Western Europe).

References

Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In B. Heine & H. Narrog (eds) *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 901–923.

Blasi, Damián, Steven Moran, Scott Moisik, Paul Widmer, Dan Dediu & Balthasar Bickel. 2019. Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363(6432), eaav3218.

Grossman, Eitan, Elad Eisen, Dmitry Nikolaev & Moran Steven. 2019. *SegBo: The world survey of phonological segment borrowing*.

Janda, Richard & Brian Joseph. 2003. On language, change, and language change – or, of history, linguistics, and historical linguistics. In B. Joseph & R. Janda (eds), *Handbook of Historical Linguistics*. Malden: Blackwell, 3–180.

Labov, William. 1972. *Sociolinguistic Patterns*. University of Pennsylvania Press.

Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.

Marsico, Egidio, Sébastien Flavier, Annemarie Verkerk & Steven Moran. 2018. BDPROTO: A database of phonological inventories from ancient and reconstructed languages. In N. Calzolari, Kh. Choukri, Ch. Cieri, Th. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, St. Piperidis & T. Tokunaga (eds), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Paris, France: European Language Resources Association (ELRA), 1654-1658.

Moran, Steven, Daniel McCloy & Richard Wright. 2014. *PHOIBLE Online*. http://phoible.org.

Newmeyer, Frederick J. 2002. Uniformitarian assumptions and language evolution research. In A. Wray (ed.), *The Transition to Language*. Oxford: Oxford University Press, 359–375.

A Comparative Study of Prepositional Object Clauses

Lutz Gunkel & Jutta M. Hartmann
(Leibniz-Institut für Deutsche Sprache, Mannheim)

Keywords: correlate, propositional proform, prepositional complement clause, syntax, subordination

Schedule: Thu 10.30 Room 7

The issue: Languages having nominal prepositional objects (POs) often allow for clausal POs too (1).

(1) a. Non dubito [**de**lla tua intelligenza]. ITA
'I don't doubt your intelligence.'
b. Non dubito [(***di**) che tu abbia capito quel libro].
'I don't doubt that you have understood that book.'
(adapted from Renzi et al. 2001:637)

Crosslinguistically there is variation with respect to the formal realization of the clausal PO. Two parameters appear to be relevant: Firstly the propensity of the preposition to combine with a clausal category, and secondly the basic VO-/OV-order. Concentrating on finite (non-interrogative) PO-clauses, we explore and analyze the main strategies as regards the formal realization of finite PO-clauses for some selected Romance (French, Italian, Spanish) and Germanic (German, Dutch, Swedish) languages.

Data: We draw on data from parallel corpora (Intercorp), reference grammars and elicited native speaker judgements. The following structural types can be observed:

S1	V+(*P)+CP	Italian	Renzi et al. (2001)	(1b)
S2	V+(P)+CP Optionality of P sometimes a matter of degree depending on verb	Spanish	Delicado Cantero (2013)	(2)
		Swedish	SAG (1999)	
S3	V+(P+DEM)+CP	French	Riegel/Pellat/Rioul (2018)	(3)
S4	*(PP)…V…CP / (PP)…V …CP Optionality of PP depending on verb	German	Breindl (1989, 2013)	(4)
		Dutch	Brockhuis/Corver(/Vos) (2015)	

When P or P+DEM occurs it must immediately precede the CP. In Dutch, the PP is tied to the clause final position, immediately preceding the verb in OV-structures. In German, in contrast, the PP is positionally more flexible.

(2) Me acuerdo [(**de**) que viniste a mi casa] SPA
'I remember that you came to my house.'
(adapted from Delicado Cantero 2013:122)
(3) Jean m'a convaincu [(**de ce**) qu'il était malade]. FRE
'Jean convinced me that he was ill.'
(adapted from Rowlett 2007:158)
(4) dass Jan sich (**darüber**) beklagt, [dass Maria ihn immer ärgert] GER
'that Jan complains about it that Mary teases him all the time'

Analysis: (S1)-(S3) The CP is embedded within a PP-shell, even when P is not overtly realized. Evidence: Clausal POs – even with covert P − and nominal (=P+NP) POs behave alike with respect to pronominaliziation (e.g. ITA *ci*, *ne* for *a*/*di*-POs, FRE *y*, *en* for *à*/*di*-POs, SPA *eso a*/*eso de* for *a*/*de*-POs, GER *da*(*r*)„P" for „P"-POs). In French, *ce* is lexically incorporated into P, so that the CP is

directly selected by P+*ce*. Evidence: P+*ce* can only be realized or omitted as a whole. (S4) The PP functions as a correlate to the CP and is present even if not overtly realized. Evidence: The (c)overt PP blocks topicalization of the CP, analogously to relative and adnominal complement clauses (cf. Müller 1995).

Prospect: We explore the impact of the VO/OV-parameter to account for the difference between the OV-languages German/Dutch and the VO-languages mentioned above. Only the OV-languages exhibit correlates, due to a conflict between two constraints: (i) arguments must be realized preverbally, but (ii) finite complement clauses are usually blocked from the prefinal position (cf. Axel 2009 for the diachrony of German DO-clauses). Evidently, no such conflict can arise in VO-languages.

References

Axel, Katrin (2009), Die Entstehung des *dass*-Satzes – ein neues Szenario, in V. Ehrich et al. (eds), (2009), *Koordination und Subordination im Deutschen*, Hamburg: Buske, 21–41.

Breindl, Eva (1989), *Präpositionalobjekte und Präpositionalobjektsätze im Deutschen*, Tübingen: Niemeyer.

Breindl, Eva (2013), Präpositionalobjektsätze, in J. Meibauer, and M. Steinbach, and H. Altmann (eds), (2013), *Satztypen im Deutschen*. Berlin: de Gruyter, 458–481.

Broekhuis, Hans/Corver, Norbert(/Vos, Riet) (2015), *Syntax of Dutch. Verbs and Verb Phrases*, vol. 1 & 2, Amsterdam: Amsterdam University Press.

Delicado Cantero, Manuel (2013), *Prepositional Clauses in Spanish. A Diachronic and Comparative Syntactic Study*, Berlin / New York: de Gruyter.

Müller, Gereon (1995), On Extraposition & Successive Cyclicity, in U. Lutz, and J. Pafel (eds), (1995), *On Extraction and Extraposition in German*, Amsterdam / Philadelphia: Benjamins, 213–243.

Renzi, Lorenzo/Salvi, Giampaolo/Cardinaletti, Anna (2001), *Grande grammatica italiana di consultazione. II. I sintagmi verbale, aggettivale, avverbale. La subordinazione*, Nuova edizione, Bologna: il Mulino.

Riegel, Martin/Pellat, Jean-Christophe/Rioul, René (2018), *Grammaire méthodique du français*, 7^{e} édition, Paris: Presses Universitaires de France.

Rowlett, Paul (2007), *The Syntax of French*. Cambridge etc: Cambridge University Press.

SAG (1999) = Teleman, Ulf/Hellberg, Staffan/Andersson, Erik (1999), *Svenska Akademiens grammatik*, vol. 4, Satser och meningar, Stockholm: Norstedt.

That moment when you realize that this construction is expressive

Daniel Gutzmann & Katharina Turgay
(University of Cologne & University of Landau)

Keywords: expressives, syntax, semantics, subordinate clause, emotions

Schedule: Thu 16.30 Room 7

Due to its technical setup and social dynamics, social media and networks are a place in which a lot of innovative and creative language can be found that caters to these specifics. In this talk, we want to present and discuss a particular construction that is used frequently in social media across a variety of

languages (we will focus on English and German here) and which we call the "expressive subordinate clause construction" (ESC, for short). These constructions typically have the form of a DP involving a noun like moment or feeling to which an attributive subordinate clause attaches. Interestingly, the DP part often is left out or abbreviated together with the subjunction.

(1) That moment when you realize it's a Monday tomorrow.
(2) Dieser Moment, wenn Dir klar wird, dass morgen Montag ist. (German)
(3) When you realize it's a Monday tomorrow.
(4) TMW you realize it's a Monday tomorrow.

Despite these constructions being syntactically ordinary DPs, they are used as standalone "utterances" (e.g. tweets or posts) and used to convey a feeling and emotion by displaying the moment/situation that triggered that feeling. Crucially, we argue that both the emotion and the situation expressed are conveyed in an expressive manner (in the sense of Kaplan 1999), thereby making the ESC akin to other DP exclamatives like The things Alex can do! Similar to these kinds of constructions, the expressive character of an ESC as in (1) gets lost if the DP is composed with other syntactic material.

(5) That moment when you realize it's a Monday tomorrow was now.
(6) The things Alex can do impress me.

While (3) just describes that there was a certain moment, (1) does not describe such a moment, it expresses that there is one and thereby expresses the emotion that is (contextually) associated with such a moment and which has to be (defeasibly) inferred by the hearer.

In our talk, we will show that the intuitive characterization of ESC as conveying (just) expressive content is supported by the observation that it exhibits all the special properties of expressive content that are assumed in the literature (scope, independence, discourse behavior etc., Gutzmann 2013). Semantically, we show that is does not refer to a particular event but instead (expressively) displays that

(7) a. there exists an event kind (as expressed by the subordinated clause),
b. and that the speaker recently experienced a contextually close event of this kind,
c. and there is emotion that (potentially contextually restricted set of) persons typically feel when experiencing an event of this kind
d. and the speaker had that emotion in that event.

Due to this double function – displaying a specific event kind as well as a highly specific emotion that is associated with that – in a rather condensed linguistic form, the ESC is perfectly tailored to the needs of communication in social media and illustrates how new forms of communication can lead to new and interesting linguistic constructions.

References

Gutzmann, Daniel (2013), Expressives and beyond. An introduction to varieties of use-conditional meaning, in D. Gutzmann, and H.-M. Gärtner (eds.) (2013): *Beyond Expressives. Explorations in Use-Conditional Meaning*, Leiden: Brill, 1–58.

Information Structure in an Annotated Parallel English-Czech Corpus: A Comparative Probe

Eva Hajičová, Jiří Mírovský, Kateřina Rysová and Magdaléna Rysová
(Charles University, Institute of Formal and Applied Linguistics
Prague)

Keywords: topic–focus articulation, information structure, word order, contrastive analysis, Prague Czech-English Dependency Treebank

Schedule: Thu 11.30 Room 11

Information structure (IS) is one of the aspects of language structure that is reflected in some way or another in most linguistic theories since the pioneering studies of Mathesius (1929) and Halliday (1967); it is more or less explicitly assumed that from the semantic and/or pragmatic points of view the function of IS is well comparable across languages (Prince 1981, Partee 1991, Steedman 2000, Krifka 2006). It would then naturally follow that in translations the IS of the source and of the target language sentences should be preserved.

Based on the Praguian theory of Topic/Focus Articulation (TFA; Sgall 1967, Sgall et al. 1973, Hajičová et al. 1998), repeatedly tested on corpus material (Mírovský et al. 2013, Rysová et al. 2015, Hajičová and Mírovský 2018), and using a parallel English(source)–Czech(target) corpus annotated both for deep syntactic structure and for TFA (PCEDT, Hajič et al. 2012), we have followed two research questions:

(i) How far does the assignment of Focus proper (= the last element of global Focus) agree in English and in Czech?

(ii) If the assignment of Focus proper differs, is the Focus-proper element in English at least a member of the (global) Focus of the Czech sentence?

Results:

The total number of automatically aligned sentences without coordination of the main predicates was 3857; there were 2514 cases (65,3%) with the same syntactic value at the last position of both source and target text and 1287 cases where there was a difference in this value.

(i) After a manual inspection of the randomly selected 120 sentences with a different syntactic label in Focus proper position, we have filtered out cases where the syntactic label differed but the target lexical item corresponded to the source one. We had then at our disposal 24 examples of an actual difference in Focus proper. Most frequently, the difference concerned the mutual position of the main predicate and its modification of time or place (in English, this modification frequently was in the Focus proper) or the position of the predicate itself (in Czech, the predicate was in the Focus proper).

(ii) As for the second question, a manual filtering of the whole set of 171 sentences ended up with a set of 30 sentences in which the element assigned Focus proper in English does not appear even in the global Focus part of the target Czech sentences. Again, most differences concerned temporal or local modifications which in the Czech sentence appeared in the Topic rather than in the global Focus. A second group of examples concerned the mutual position of the main predicate and its argument Patient: Patient was “topicalized” in Czech.

Conclusion:

Our analysis confirms that the differences in IS between the examined languages are rather rare, though certain tendencies can be observed, namely in case of the temporal and local modifications of the main predicate. We have also carefully analyzed the context in which the sentences occur and it came out that the context actually offers both interpretations of the IS structure.

Acknowledgment:
This work has been supported and using language resources and tools distributed by the LINDAT/CLARIN project of the Ministry of Education, Youth and Sports of the Czech Republic (LM2015071 and OP VVV VI CZ.02.1.01/0.0/0.0/16 013/0001781).

References

Hajič, J., Hajičová, E., Panevová, J., Sgall, P., Bojar, O., Cinková, S., Fučíková, E., Mikulová, M., Pajas, P., Popelka, J., Semecký, J., Šindlerová, J., Štěpánek, J., Toman, J., Urešová, Z., Žabokrtský, Z. (2012), Announcing Prague Czech-English Dependency Treebank 2.0. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*, European Language Resources Association, İstanbul, Turkey, ISBN 978-2-9517408-7-7, pp. 3153–3160.

Hajičová, E., Mírovský, J. (2018), Discourse Coherence Through the Lens of an Annotated Text Corpus: A Case Study. In: *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, European Language Resources Association, Paris, France, ISBN 979-10-95546-00-9, pp. 1637–1642.

Hajičová, E., Partee, B. H. and P. Sgall (1998), *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*, Dordrecht, Kluwer Academic Publishers.

Halliday, M. A. K. (1967), Notes on transitivity and theme in English. Part 2, *Journal of Linguistics 3*, 199–244.

Krifka, M. (2008), Basic Notions of Information Structure. *Acta linguistica Academica*, Vol. 55, Issue 3-4, 243–276.

Mathesius, V. (1929), Zur Satzperspektive im modernen Englisch. *Archiv für das Studium der neueren Sprachen und Literaturen*, 155(29), 202–210.

Mírovský, J., Rysová, K., Rysová, M., Hajičová, E. (2013), (Pre-)Annotation of Topic-Focus Articulation in Prague Czech-English Dependency Treebank. In: *Proceedings of the 6th International Joint Conference on Natural Language Processing, Asian Federation of Natural Language Processing*, Nagoya, Japan, ISBN 978-4-9907348-0-0, pp. 55–63.

Partee, B. H. (1991), Topic, Focus and Quantification. In: Moore, S., Wyner, A. (ed.): *Proceedings from SALT I*, Ithaka, N.Y.: Cornell University, pp. 257–280.

Prince, E. (1981), Toward a taxonomy of given/new information. In Cole, ed. *Radical Pragmatics (1981)*, 223–254.

Rysová K., Mírovský J., Hajičová E. (2015), On an apparent freedom of Czech word order. A case study. In: *14th International Workshop on Treebanks and Linguistic Theories (TLT 2015)*, Warszawa, Poland, ISBN 978-83-63159-18-4, pp. 93–105.

Sgall P. (1967), Functional Sentence Perspective in a Generative Description. In *Prague Studies in Mathematical Linguistics 2*, Prague, Academia, 203–225.

Sgall, P., Hajičová, E., Benešová, E. (1973), *Topic, Focus and Generative Semantics.* Kronberg/Taunus: Scriptor.

Steedman, M. (2000), Information structure and the syntax-phonology interface. *Linguistic Inquiry, 31*, pp. 649–689.

Another fuzzy border in morphology

Camiel Hamans
(University of Amsterdam; Adam Mickiewicz University Poznań)

Keywords: blends, morphological borders, prosodic morphology, monomorphemic word, clipped compounds

Schedule: Thu 11.30 Room 1

In his well-known Dutch introduction to linguistics *Taalwetenschap* (1959:22) E.M. Uhlenbeck claimed that the first task of morphology is to determine which words are polymorphemic and which ones are monomorphemic. Parallel to biology Uhlenbeck called monomorphemic words 'ongeleed' 'unarticulated' – without joints or segments – and polymorphemic 'geleed', 'articulated'. In this presentation it will be shown that this distinction is not absolute, as most of the borders in morphology. The examples come from Dutch, German and English.

It is known that the distinction between compounding and derivation is not as clear as the handbooks claim (cf. Bauer 2005, Trips 2009, Bauer et al. 2013 and Olsen 2014). It is also evident that words that started as derived forms may lose their internal structure in the course of the history, just as original compounds. See for instance the Dutch examples under (1) and (2)

(1)	derivation				
	lelijk	'ugly'	<	leed + suffix -lijk	'sorrow' +'ly'
	aardig	'frierndly'	<	aard+ suffix -ig 'of noble birth' + 'ly'	
(2)	composition				
	pollepel	'ladle'	<	pot + lepel	'pot' + 'spoon'
	handdoek	'towel'	<	hand + doek	'hand' + 'cloth'

For speakers of contemporary Dutch, the forms presented in (1) and (2) are unarticulated. Subsequently, it will be shown that there is a productive category of words that combine formal features of unarticulated, simplex, words and of complex words. This category consists of blends.The focus will be on blends of the type *brunch*, *motel* and *smog*.

- First, it will be shown that one must make a difference between clipped compounds and blends. Clipped compounds such as *sitcom*, combine the first parts of the two source words, whereas blends, such as *motel*, combine the left-hand part of the first source word and the right-hand part of the second source word.
- Clipped compounds appear to be compounds with a structural head, which is the right part of the clipped compound.
- Moreover, the compound stress rule applies to clipped compounds, whereas blends normally copy the tress pattern of the second source word.
- However, it appears that blends also have a structural head, which is normally the part which descends from the second source word. So, in *motel* the right-hand part *-otel* is the head. (*Hotel* is the second source word.) Headedness is a feature of complex words.
- Finally, blends copy the syllabic and prosodic structure of the second source word.

Since the second source word of many blends is a simplex word, as in *brunch*, *motel* and *smog*, the resulting blend can also be described as unarticulated. However, blend formation also resembles compounding. Consequently, blends are a borderline case between simplex, unarticulated, words and complex or articulated words.

The first task of morphology, as stipulated by Uhlenbeck, appears to be impossible. However, this is not the end of morphology, since detailed scrutiny of forms such as blends, which have been considered to be irregular and unsystematic for a long time, proves that the speakers of language systematically follow patterns.

References

Bauer, Laurie (2005), The borderline between derivation and compounding, in Wolfgang U. Dressler, Dieter Kastovsky, Otto Pfeiffer and Franz Rainer (eds.) (2005), *Morphology and its demarcations*, Amsterdam: John Benjamins, 97-108.

Bauer, Laurie, Rochelle Lieber and Ingo Plag (2013), *English Morphology*, Oxford: OUP.

Marchand, Hans (1969), *The categories and Types of Present-Day English Word-Formation*, München: Beck.

Olsen, Susan (2014), Delineating Derivation and Compounding, in Rochelle Lieber and Pavol Štekauer (eds.) (2014), *Derivational Morphology*, Oxford: OUP: 26-49.

Trips, Carola (2009), *Lexical semantics and diachronic morphology. The development of* -hood, -dom *and* -ship *in the history of English*, Tübingen: Niemeyer.

Uhlenbeck, Eugenius M. (Bob) (1959), *Taalwetenschap*, Den Haag: H.L. Smits.

Stress shift in American English

Bente Hannisdal
(University of Bergen)

Keywords: American English, stress shift, prefixes, suffixes, Rhythm Rule

Schedule: Thu 9.00 Room 9

This paper discusses stress shift in American English in two types of words: prefixed verbs and words with stress-bearing suffixes. Both types typically have stress on the last syllable, i.e. on the root and on the suffix respectively (cf. Fudge 1984). Examples are:

(1) *con'trast, dis'like, im'port, mis'lead, pro'test, re'lay, trans'mit*
(2) *ar'cade, bloc'kade, detai'nee, nomi'nee, refe'ree, pirou'ette, tat'too*

This pattern seems to be changing for many of these words, which are increasingly pronounced with stress on the first syllable. The current investigation aims to establish which lexical items are affected, to estimate the frequency of the non-traditional stress, and to identify any patterns of variation.

Previous accounts of word stress variation in English are typically based solely on dictionary analyses (e.g. Deschamps 2001, Abasq et al. 2012, Phillips 1998). This paper presents the results from a corpus-based study of spoken General American English using data from newscasts. The corpus contains approximately 4,500 tokens. The findings are correlated with factors such as speaker sex, lexical frequency, and rhythm, and compared with data from pronunciation dictionaries.

A total of 66 different verbs are found to variably or consistently occur with early stress. Many of the prefixed verbs have a corresponding noun with early stress. The majority of these are of Romance origin, and originally both the nouns and verbs had stress on the last syllable. Over time, however, the nouns underwent a shift, moving the stress to the first syllable, in line with the general Germanic stress pattern (cf. Svensson 2001), and many verbs seem to be moving in the same direction. In addition to

prefixed verbs, the analysis also includes a number of particle-verb compounds (e.g. *downplay*, *overthrow*, *undermine*, *uproot*), many of which are predominantly stressed on the first syllable.

No significant effects are observed for speaker sex or lexical frequency, but rhythm seems to play a role at intra-speaker level. Speakers who display variation in the stressing of individual verbs, tend to assign stress in accordance with the Rhythm Rule (Hayes 1984), as in:

(3) *we 'import 'oil*
(4) *the more we im'port*

Many of the stress-bearing suffixes are also of Romance origin and borrowed into English. 28 different suffixed words display variation in stress placement. For this group, two patterns emerge: the most common items (e.g. *refugee*, *nominee*, *appointee*) seem to receive early stress regardless of sentence prosody, while the less frequent words (e.g. *franchisee*, *retiree*, *tattoo*) seem to behave in correspondence with the Rhythm Rule:

(5) *He's a political ap'pointee.*
(6) *Working closely with the franchi'see. What the 'franchisee 'does.*

The stress shifts observed here suggest an ongoing change in American English. They may be seen as part of a general movement towards the Germanic preference for early stress, as well as a general adaptation to the rhythm of stress-timing, where stressed syllables occur at more or less regular intervals.

References

Abasq, V., Dabouis, Q., Descloux, E., Girard, I., Fournier, J-M., Fournier, P., Martin, M. Van Houtte, S. (2012), Multicategorical prefixed words stress behaviour: variation and frequency, PAC workshop: *The Phonology of Contemporary English: Variation and Change*, March 2012.

Deschamps, A. (2001), Stress-patterns, rules and variants: Can stress variation be accounted for? *Anglophonia* 9, 41-57.

Fudge, E. (1984), *English word-stress*, London: Allen & Unwin.

Hayes, B. (1984), The phonology of rhythm in English, *Linguistic Inquiry* 15(1), 33-74.

Phillips, B. S. (1998), Word frequency and lexical diffusion in English stress shifts, in R. M. Hogg and L. van Bergen (eds.), (2009), *Historical linguistics 1995: Volume 2: Germanic linguistics*, Amsterdam: John Benjamins, 223-232.

Svensson, A.-M. (2001), Germanic vs. French stress in disyllabic loan words in English, *Gothenburg Studies in English* 81, 123-129.

At the interface of syntax and pragmatics: Chipaya focus-marking enclitics

Katja Hannß
(University of Cologne)

Keywords: syntax; pragmatics; enclitics; Chipaya; Bolivia

Schedule: Thu 12.00 Room 11

I will discuss focus-marking enclitics of Chipaya, an endangered isolate of highland Bolivia (Adelaar 2007:19). The enclitics occur as single clitics and in doubling constructions, occupying different syntactic positions within a clause. Per clause only one enclitic can occur. I argue that the clitic

constructions and the syntactic positions together indicate the activation status of the discourse referents and relate presupposed and non-presupposed discourse referents, thus contributing to completing an assertion. The research follows assertion-centred approaches to information structure (Lambrecht 1994; Matić 2015). The data stem from the author's own fieldwork (see DobeS).

The Chipaya focus-marking enclitics are co-referential with the subject of a clause but are not obligatory. Depending on the activation status of the respective subject referent, three types of clitic constructions are attested: (1) a single enclitic without an overtly realised subject pronoun (only with active first- and second person subject referents); (2) anaphoric doubling, where an accessible first- or second person subject pronoun co-occurs with the corresponding enclitic; (3) lexical doubling, where an inactive or new lexical third person subject noun phrase and the co-referential enclitic occur together (Chafe 1976; Lambrecht 1994; for anaphoric and lexical doubling, see Belloro 2007:117). The enclitics can occupy the following syntactic positions within a clause (**X**: regular syntactic positions; (X): occasional syntactic positions).

Enclitics co-referential with	**Subject**	**Direct object**	**Indirect object**	**Modifiers**	**Negation**
1st and 2nd subject	---	**X**	(X)	(X)	**X**
3rd subject	**X**	(X)	---	(X)	**X**

Table 1: Syntactic positions of Chipaya focus-marking enclitics

The syntactic positions of the focus-marking enclitics follow from their pragmatic functions. Their main function is to establish a focus relation between the subject of the clause and the host of the enclitic (Cerrón-Palomino 2006:172) by providing new information about the subject referent and/or by establishing a new relation between the subject and the host of the enclitic. As first- and second person subject referents are necessarily presupposed, in clauses with a first- or second person subject referent the focus-marking enclitic cannot go onto the presupposed subject pronoun. Instead, it attaches to the direct object (and occasionally to indirect objects or modifiers) that provides new information about the presupposed subject referent (DuBois 1987:826). In clauses with a newly introduced or a re-introduced third person subject referent, non-presupposed subject and focus coincide and therefore newly introduced or re-introduced third person subject referents regularly receive focus-marking (Siewierska 2004:46). Only if a third person subject referent is already established in the discourse and hence presupposed can the co-referential focus-marking enclitic attach to another constituent, mostly a direct object argument. In negative clauses, however, all focus-marking enclitics regularly attach to the negation particle as this provides the greatest information update with respect to the subject referent.

Thus, the Chipaya focus-marking enclitics index the degree of activation of a subject referent and intricately interrelate presupposed and non-presupposed discourse referents, thus contributing to forming a coherent text.

Acknowledgements

Research on Chipaya is funded by the *Deutsche Forschungsgemeinschaft* (DFG; grant number: HA 6340/4-1).

References

Adelaar, Willem (2007), Threatened languages in Hispanic South America, in Matthias Brenzinger (ed.), (2007), *Language Diversity Endangered*, Berlin: Mouton de Gruyter, 9–28.

Belloro, Valeria A. (2007), *Spanish Clitic Doubling: A Study of the Syntax-Pragmatics Interface*, PhD dissertation, University of New York at Buffalo.

http://www.acsu.buffalo.edu/~rrgpage/rrg/Belloro-Spanish_Clitic_Doubling.pdf (12 May 2019).
Cerrón-Palomino, Rodolfo (2006), *El chipaya o la lengua de los hombres del agua*, Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú.
Chafe, Wallace (1976), Givenness, contrastiveness, definiteness, subjects, topics, and point of view, in Charles N. Li (ed.), (1976), *Subject and Topic*, New York/San Francisco/London: Academic Press, 27–55.
DobeS. *Dokumentation bedrohter Sprachen*. *Documentation of endangered languages*. http://dobes.mpi.nl/projects/chipaya/ (12 May 2019).
DuBois. John W (1987), The Discourse Basis of Ergativity, *Language* 63(4), 805–855.
Lambrecht, Knud (1994), *Information Structure and Sentence Form*. *Topic, Focus, and the Mental Representation of Discourse Referents*, Cambridge: Cambridge University Press.
Matić, Dejan (2015), Information structure in linguistics, in James D. Wright (ed.), (2015), *International Encyclopedia of the Social and Behavioral Sciences*, 2nd edition, Oxford: Elsevier, 95–99.
Siewierska, Anna (2004), *Person*, Cambridge: Cambridge University Press.

The role of (historical) pragmatics in the uses of response particles. The case of French

Maj-Britt Mosegaard Hansen
(The University of Manchester)

Keywords: Response particles, Historical Pragmatics, Polarity, Speech Acts, Pragmatic Orientation

Schedule: ~~Thu 16.30 Room 14~~ Fri 17.00 Room 14

Many languages use particles as minimal affirmative *vs* negative responses to a preceding utterance by a different speaker. Typologically, response particle systems come in two basic varieties, a polarity-based one and a (dis)agreement-based one (König & Siemund 2007). Currently, two main theories, one (Holmberg 2016) syntactic, and the other (Roelofsen & Farkas 2015) semantic in nature, propose to account for the distribution of response particles across languages.

The Modern French system is often thought of as polarity-based, *oui* ('yes') and *non* ('no') marking the positive *vs* negative polarity of the response. However, the language also has a second affirmative particle *si*, which marks reversal of the negative polarity of (and thus disagreement with) the utterance it responds to, cf. (1):

(1) A : Jean ne viendra pas. 'A: Jean won't come.'
B : Si(, il viendra)./Non(, il ne viendra pas). 'Yes (he will)./No (he won't).'

Actual usage, however, is more complex (Høybye 1939, Kerbrat-Orecchioni 2001, Wilmet 1976), and cannot easily be accounted for by either of the two above-mentioned theories. Thus, most saliently, speakers often prefer *oui* to either *si* or *non* when responding in agreement with syntactically negative utterances that are positively oriented at the pragmatic level, as seen in (2):

(2) A: N'êtes-vous pas la fille de X ? 'A: Aren't you X's daughter?'
B : Oui/Si. 'Yes.'

Conversely, *si* can be found in contexts that merely implicate a negative proposition. Finally, in certain context, either *non* or *si* can be used to disagree with a negative first utterance, but with different nuances of meaning.

I argue that a better understanding of the current system can be obtained by taking historical pragmatics into account. The French response particles result from lexicalization of two different constructions in Medieval French, *oui* < *oïl* < *o il* < Latin HOC ILLE (FECIT) ('this he/it (did)') vs *si* (< Latin SIC 'thus')/*non* + V. Medieval French had a second negative marker, viz. *nenni*, whose source construction *nenil* < *nen il* ('not he/it') is analogous to that of *oui*.

Based on data from Frantext (9th-21st c.), I show quantitatively that the two pairs of response markers (*oïl/nenil* vs *non/si*) originally occurred in distinct types of contexts and had different types of pragmatic import. Thus, originally, *oui/nenni* were chiefly used to respond to questions, and *si/non* to other types of speech acts. While the usage of *oui* and *non* gradually expanded to all types of contexts, *nenni* was slower to change and thus eventually ousted by *non*, whereas *si* lost its original ability to respond to positively polar utterances and became confined to its present, rather narrow, usage (Pohl 1976, Marchello-Nizia 1985). Moreover, at the outset, *oui/nenni* were preferentially used to respond to questions whose pragmatic orientation was harmonious with that of the particle, whereas *non* was typically used in pragmatically disharmonious responses. While *oui* remains a predominantly harmonious particle to this day, with *si* having become specialized for disharmonious – and often outright polemical – contexts, *non* became increasingly felicitous across all contexts.

References

Frantext electronic database: https://www.frantext.fr/

Holmberg, Anders (2016), *The Syntax of Yes and No*. Oxford: OUP

Høybye, Poul (1939), *Oui*, *si* et *non*. *Le français moderne* 7: 47-51.

Kerbrat-Orecchioni, Catherine (2001), *Oui, non, si* : un trio célèbre et méconnu. *Marges linguistiques* 2, 95-119 - http://www.marges-linguistiques.com

König, Ekkehard & Peter Siemund (2007), Speech act distinctions in grammar, in T. Shopen (ed), (2007), *Language Typology and Syntactic Description, vol. I: Clause Structure*. 2nd ed. Cambridge: CUP, 276-324.

Marchello-Nizia, Christiane (1985), *Dire le vrai. L'adverbe* si *en français médiéval*. Geneva : Droz.

Pohl, Jacques (1976), Matériaux pour l'histoire du système *oui-non-si*. *Kwartalnik neofilologiiczny* 23, 197-208.

Roelofsen, Floris & Donka Farkas (2015), Polarity particle responses as a window on the interpretation of questions and assertions. *Language* 91(2), 359-414.

Wilmet, Marc (1976), *Oui*, *si* et *non* en français moderne. *Le français moderne* 44(3), 229-251.

Bayesian phylogenies of language families: How they can go wrong — and how to get them right

Paul Heggarty & Cormac Anderson
(DLCE, Max Planck Institute for the Science of Human History, Jena)

Keywords: phylogenetics, Bayesian, language dating, methodology, Indo-European

Schedule: We ~~12.30~~ 12.00 Room 1

Sophisticated mathematical models have been devised to recover family trees of descent in the biological sciences. The same models have increasingly been applied to language data, too, on the

analogy that languages also diverge into families by descent with modification. Moreover, language divergence proceeds through *time* as a family spreads across geographical *space*. Phylogenetic analyses within a Bayesian inference framework, particularly, claim thus to be able to recover, from a family's phylogeny and appropriate calibration points, the chronology and geography of its expansion through prehistory.

These methods have faced enduring criticism and scepticism, however (e.g. Campbell 2013). Different approaches yield opposing results on the same question. Gray & Atkinson (2003) and Bouckaert et al. (2012) report a date for Proto-Indo European much older than that now found by Chang et al. (2015). At least some of these different ways of doing Bayesian phylogenetics, then, must be wrong. How are we to judge which, if any, are valid?

This talk does not focus on the Indo-European origins question itself. Rather, Indo-European serves as a cautionary tale and as an opportunity to explore how to make phylogenetic analyses more consistent, accurate and reliable. Major branches such as Romance, Slavic or Germanic represent some of the best understood cases of language divergence, over well-documented historical timeframes. These serve as benchmarks against which to assess how valid the published phylogenetic analyses really are. This assessment identifies various artefacts in phylogenetic outputs, and pinpoints the causes, whether in the data-sets or in the phylogenetic analysis models used, in order to propose solutions to both.

We first explore errors that arise from the language databases used. Most Bayesian phylogenetic analyses are based only on cognate relationships in basic lexicon (a limitation in itself). Inconsistency in handling synonyms, however, can invalidate the calculation of branch lengths and the corresponding language chronologies. The imperative of a consistent data set also entails (re)defining target meanings, and more sophisticated handling of loanword phenomena of various types.

We summarise the policies, structures and analyses needed to ensure that language databases are valid as input to phylogenetic analysis algorithms. We illustrate these as implemented by the new IE-CoR dataset, created by a consortium of over 70 language experts on 151 languages across Indo-European. This coverage provides a fuller range of cases — such as the dialect continuum of continental West Germanic, or the pervasive parallel loanwords from Persian into Iranic and Indic — that prove challenging for phylogenetic models.

Secondly, then, against this benchmark database we can also evaluate the models themselves. Bayesian phylogenetic approaches include key variables that leave scope for different results. We assess, for instance, different approaches to ancestry constraints, to rate variation from one meaning or language lineage to another, and to scoring the likelihood of parallel changes. Each assumption or parameter setting can determine differences and problems in the phylogenetic outputs. On this basis, we propose which next steps could tailor these methods more appropriately to divergence in language, to give greater confidence and consensus in this phylogenetic methodology.

References

Bouckaert, R. et al. (2012), Mapping the origins and expansion of the Indo-European language family, *Science* 337(6097), 957–60. http://doi.org/10.1126/science.1219669

Campbell, L. (2013), Historical Linguistics: An Introduction, Cambridge (Mass.): MIT Press.

Chang, W. et al. (2015), Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis, *Language* 91(1),194–244. http://doi.org/10.1353/lan.2015.0005

Gray, R.D., & Atkinson, Q.D. (2003), Language-tree divergence times support the Anatolian theory of Indo-European origin, *Nature* 426(6965), 435–439. http://doi.org/10.1038/nature02029

Phonetic variation in German supposedly identical singular-plural nouns

Melina Heinrichs, Marcel Schlechtweg & Marcel Linnenkohl
(University of Kassel)

Keywords: Number, acoustic differences, homophony, polysemy, German

Schedule: Thu 16.00 Room 9

Words such as *time* and *thyme* and suffixes such as the English plural *-s* and the genitive *-s* have normally been considered homophonous. However, in recent years, several studies have shown that these forms differ in their precise phonetic detail, e.g., in their duration (see, e.g., Gahl 2008, Plag et al. 2017). The present contribution aims at expanding research in this area by investigating whether nouns that have been regarded as identical in the singular and plural really are acoustically identical.

Eight native speakers of German (four males) participated in the production study. The material included (a) eight pairs of disyllabic, initially stressed, and monomorphemic German nouns whose singular and plural nominative forms have traditionally been considered identical (e.g., *Batzen* 'chunk(s)') and (b) eight pairs of disyllabic, initially stressed, and bimorphemic German nouns whose singular and plural genitive forms have traditionally been considered identical (*Spatzen* 'of the sparrow(s)'). The average frequency values of the four conditions did not significantly differ from each other. All frequencies were checked using the COSMAS2 (Connexor) corpus of the Institute of the German language (https://www.ids-mannheim.de/cosmas2/). The target words of the four conditions were embedded in sentences, which were read out. Each subject read a total of 32 experimental sentences, i.e., the entire test material (8 items x 2 conditions per item + 8 items x 2 conditions per item), and 64 filler sentences. The four conditions were counterbalanced using a Latin Square Design.

Repeated-measures ANOVAs by subject and by item were performed on the two dependent variables DURATION INITIAL WORD PART and DURATION FINAL WORD PART. "Initial word part" refers to the part of a word preceding the *-en*, i.e., e.g., the stem of the genitive words. "Final word part" refers to the *-en*, which was a suffix or not. The independent/fixed variables were NUMBER (singular/plural) and CASE (nominative/genitive) (both were within-subject, NUMBER was within-item and CASE was between-item). SUBJECT and ITEM were included as random variables. The analysis revealed a significant effect for DURATION INITIAL WORD PART only: Plural nouns were spoken with significantly longer duration than singular nouns ($p < .05$). The results show that plural nouns differ acoustically from their respective singular forms, although the two have been considered to be identical. That is, the "basic" singular form is uttered in a more compromised form than the plural version. Overall, the results will be interpreted against the background of models of speech production and the role of phonetic detail in the distinction between different morphosyntactic properties.

References

Gahl, Susanne (2008), *Time* and *thyme* are not homophones: The effect of lemma frequencies on word durations in spontaneous speech, *Language* 84(3), 474−496.

Plag, Ingo, Julia Homann & Gero Kunter (2017), Homophony and morphology: The acoustics of word-final S in English, *Journal of Linguistics* 53, 181−216.

A multipurpose annotation system in application to constructions in Kwa and Bantu

Lars Hellan
(Polytext AS; NTNU, Trondheim)

Keywords: Corpus annotation, grammatical parsing, serial verb constructions, extended (derived) verbs, HPSG vs. UDG

Schedule: We 14.30 Room 6

Universal Dependency Grammar (UDG, cf. https://universaldependencies.org) is a standard for grammatical annotation systems based on frameworks including parsing algorithms. Multi-verb constructions (MVCs) in Kwa and verbal extensions in Bantu challenge certain assumptions within Dependency Grammar frameworks, namely that every sentence has a unique head, and that each head category has a unique set of dependency relations defined. MVCs offer cases with more than one candidate for head status, and extended verbs offer cases of more than one set of dependency relations emanating from a given head. An annotation system suited for MVCs arguably has to provide annotation labels with scope over entire sentential constructions, and extended verbs may require annotation labels tracking correspondences and discrepancies between grammatical functions and semantic argument relations. While neither phenomenon is as yet encoded in UDG (version 2), we present a system whose grammatical formalism is Typed Feature Structures (cf. Copestake 2002) as used in HPSG, reflected in a cross-linguistic annotation set ranging over sentential construction types and valence types; the annotation set, partly described in Hellan and Dakubu 2010, Dakubu and Hellan 2017, is called 'Construction Labeling' (CL), the overall grammatical platform is called *TypeGram* (cf. Hellan 2019), and the corpus annotation platform is *TypeCraft* (cf. Beermann and Mihaylov 2014). One corpus here is https://typecraft.org/tc2/ntceditor.html#16 for the Kwa language Ga, and the example below is obtained on a search for serial verb constructions with identical subjects and aspect across the VPs, CL-encoded as sv_suAspID in the search window; a full CL specification including valence of the head verbs and properties of their arguments in the actual example is given in the 'Comment' field:

Gbekɛbii lɛ ekamɔ shi amɛwɔ
"The children have gone to sleep"

gbekɛbii			lɛ	ekamɔ		shi	amɛwɔ		
gbekɛ	bi	i	lɛ	e	ka mɔ	shi	amɛ́	wɔ	
child	offspring				lie	down		sleep	
child	*offspring*	PL	DEF	PFV	*lie*	DISTRIB	*down*	3PL.PFV	*sleep*
N			DET	V			N	V	

Comment: Sv_suAspID-v1tr-v1suAgintent_obLoc-v2intr-v2suClit-v2suTh .

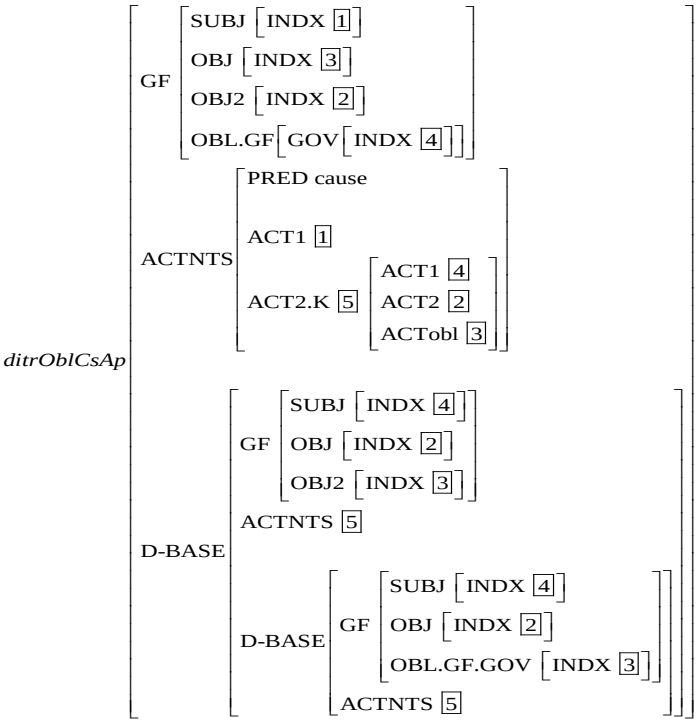
An extended verb is exemplified below, with a derivation through Applicative and subsequently Morphological Causative encoded by the CL annotation formula ditrOblCsAp (ex.Citumbuka, by Jean Chavula, pc):

Tumbikani wakamuphik**iskir**a Temwa nchunga kwa Mary

Tumbikani	wa-ka-mu-phik-**isk-ir**-a	Temwa	nchunga	kwa	Mary
Tumbikani	1SM-pst-1OM-cook-Caus-Appl-fV	Temwa	beans	'to'	Mary

'Tumbikani made Mary cook beans for Temwa'

Illustrating the connection between the CL formula and grammatical parsing for this example, the attributes ‘D-BASE’ in the parse diagram below represent the original and derived base of the resulting form (reflecting the operations encoded as resp. ‘Ap’ and ‘Cs’), the feature ‘GF’ in the uppermost part showing the grammatical relations associated with the final derived verb, while the occurrences of the feature ACTNTS display the semantic argument structure along the derivational stages, at the top level non-congruous with the GFs:



While UDG reflects ‘shallow parsing’, the present system represents so-called ‘deep parsing’; the paper assesses some advantages and disadvantages in either direction, but maintains the fruitfulness of the latter approach alongside more ‘shallow’ approaches.

References

Beermann, Dorothee and Mihaylov, Pavel 2(014), Collaborative databasing and Resource sharing for Linguists. *Languages Resources and Evaluation* 48. Dordrecht: Springer, 1-23.

Copestake, Ann (2002), *Implementing Typed Feature- Structure Grammars*. Stanford: CSLI Publications.

Dakubu, M.E. Kropp and Lars Hellan (2017), A labeling system for valency: linguistic coverage and applications. In Hellan, L., Malchukov, A., and Cennamo, M (eds) *Contrastive studies in Valency*. Amsterdam & Philadelphia: John Benjamins Publishing Co.

Hellan, Lars (2019), Construction-based Compositional Grammar. *Journal of Logic, Language and Information*. https://doi.org/10.1007/s10849-019-09284-5.

Hellan, Lars and M.E. Kropp Dakubu (2010), *Identifying verb constructions cross-linguistically*. In *Studies in the Languages of the Volta Basin* 6.3. Legon: Linguistics Department, University of Ghana.

A quantitative typological approach to the morphome

Borja Herce
(University of the Basque Country, UPV/EHU & University of Surrey, SMG)

Keywords: Morphome, database, multivariate typology, frequency

Schedule: Fri 14.00 Room 6

(Meta)morphomes (Aronoff 1994, Round 2013), are purely morphological (i.e. phonologically or semantically unjustified) formal syncretisms that are nonetheless systematic within the inflectional or paradigmatic structure of a language. The phenomenon has been investigated extensively in Romance (e.g. Maiden 2018) but hardly at all elsewhere. Typological approaches to the phenomenon are also hard to find (cf. Herce 2018). The purpose of this paper is to fill these gaps by presenting a typologization of (meta)morphomes and some generalizations based on a large cross-linguistic database of morphomes.

Concerning the first goal, in the vein of Multivariate or Distributional Typology (Bickel 2015), I will identify the variables involved in the cross-linguistic variation in the domain of morphomes. My typology will involve variables related to both form and distribution. Within the first, morphomes may differ in their i) cross-exponential recurrence (i.e. the number of different forms that instantiate the morphome) and in the ii) robustness of this phonological instantiation (e.g. average number of segments). Within the latter, morphomes may differ among others in their iii) cross-lexemic recurrence (i.e. the number/proportion of lexemes where the morphome appears overtly), iv) morphosyntactic constrainedness (see Herce forthcoming), v) morphosyntactic coherence (see Esher 2014), vi) paradigmatic recurrence (i.e. number of cells that the morphome spans) etc.

By way of illustration, a morphome in Achumawi (the so-called 'amplified stem' in De Angulo & Freeland 1930) takes the following values for the above variables: i) 11, ii) 1.55, iii) 90%, iv) 0.4, v) 33.3%, vi) 7. In addition, other possible variables will be proposed including the vii) informativity of a morphome (i.e. whether it creates additional morphosyntactic contrasts or is redundant), the viii) independence of a morphome (i.e. whether its presence is dependant on the presence of another element or not) and ix) the cross-linguistic recurrence of a morphome. Examples from natural languages will be presented to illustrate the range of variation within each of these nine variables.

Concerning the second goal (i.e. to find cross-linguistic generalizations), a sample of 100 morphomes has been compiled from the reference grammars of genetically and geographically diverse languages. Although variation is, as could be expected, rampant, some tendencies have been spotted. The more robust and revealing ones are that:

A) most morphomes involve paradigm cells which are contiguous according to their feature-values. That is, (meta)morphomes which involve completely disjunct morphosyntactic contexts (e.g. 2SG and 3PL) are comparatively rare. This might constitute a window into what counts as a possible meaning in language for formatives of any kind (Jackendoff 1985, McCreight & Chvany 1991).
B) many morphomes involve the falling-together of a frequent cell (i.e. one whose feature values have a high token frequency, e.g. 3SG.PRES, NOM.SG etc.) with a more infrequent natural class (e.g. SUJV, FUT, PL etc.). This reminds us of the role of frequency as a decisive shaping force in grammar (Haspelmath 2008).

References

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge (MA): MIT press.
Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic

diversity. *The Oxford handbook of linguistic analysis*: 901-923.
Esher, Louise. 2014. Autonomous morphology and extramorphological coherence. *Morphology* 24, 4: 325-350.
Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive linguistics* 19.1: 1-33.
Herce, Borja. 2018. Morphome interactions. Morphology.
Herce, Borja. Forthcoming. On morphemes and morphomes: exploring the distinction. Word structure.
Jackendoff, Ray. 1985. Multiple subcategorization and the θ-criterion: The case of climb. *Natural Language & Linguistic Theory* 3.3: 271-295.
Maiden, Martin. 2018. *The Romance Verb: Morphomic Structure and Diachrony*. Oxford: Oxford University Press.
McCreight, Katherine & Catherine V. Chvany. 1991. Geometric representation of paradigms in a modular theory of grammar. *Paradigms: The economy of inflection*: 91-112.
Round, Erich R. 2013. Rhizomorphomes, meromorphomes, and metamorphomes. In Matthew Baerman, Dunstan Brown & Greville Corbett (Eds.), *Understanding and Measuring Morphological Complexity*. 29-52. Oxford: Oxford University Press.

Impersonal constructions in spoken, literary, and academic Turkish

Annette Herkenrath

Schedule: Fri 16.30 Room 7

<pdf>

A Study of Pronominal Verbs in Romance: Evidence of the Antipassive in French and Italian

Karina High

Schedule: Thu ~~9.00~~ 11.30 Room 14

This study examines diachronic and synchronic evidence for the antipassive (AP) construction in French and Italian, using data from the 12th to 21st centuries and focusing on the distribution of a particular class of pronominal (PRO) verbs and their transitive (TR) counterparts. The verbs selected for this study verbs are Fr. (*s'*)*apercevoir* 'see (briefly); realize', It. *ricordar(si)* 'remind; remember', Fr. *(se) plaindre*/It. *lamentar(si)* 'pity; complain', and Fr. *(se) vanter*/It. *vantar(si)* 'praise; boast'. As shown in (1), the PRO verbs are characterized by the realization of the logical object as an oblique complement (1a), while their TR forms select a direct object (1b).

(1) a. *Dopo aver cercato dappertutto,* ***si ricordò*** ***del*** ***sogno*** *e corse in giardino, vicino al fiume,*
after have.AUX.INF search.PP everywhere REFL.3SG remember.PFV.PAST.3SG of.DEF.DET.MSG dream.SG and run.PFV.PAST.3SG in garden.SG near to.DEF.DET.MSGriver.SG

dove dormendo, l'=aveva veduta.
where sleep.GER OCL.3MSG=have.AUX.IPFV.PAST.3SG see.PP
'After having searched everywhere, he remembered the dream and ran into the garden, near the river where sleeping, he had seen it.' (Collodi 1875)

b. *Chiunque* ***ricordi*** ***<u>la</u>*** ***<u>vita</u>*** ***<u>italiana</u>***
anyone remember.SBJV.PRES.3SG DEF.DET.FSG life.SG Italian.FSG
<u>al</u> ***<u>principio</u>*** ***<u>del</u>*** ***<u>secolo</u>*** *non*
to.DEF.DET.MSG beginning.SG of.DEF.DET.MSG century.SG NEG
potrà non sottoscrivere a questo apprezzamento.
can.FUT.3SG NEG subscribe.INF to this.MSG comment.SG
'Whoever remembers the Italian life at the start of the century, cannot not subscribe to this comment.' (Salvatorelli 1943)

The above PRO verbs, which cannot be termed reflexive, reciprocal, or middle/passive, attest to the heterogeneous nature of the Romance reflexive SE. Analyses have been proposed to capture different aspects of the diversity of Romance pronominal verbs, such as Nishida (1994) for Spanish and Melis (1985; 1990a; 1990b) for French, and some have discovered elements indicative of the AP construction, such as Masullo (1992), Medová (2009). The latter do not, however, look for evidentiary support from a broader Romance nor diachronic perspective.

In terms of the overall distribution of TR and PRO verbs (n=2237) for both French and Italian, we observe a higher frequency of PRO (63.1%) than TR (36.9%). We can posit that the dominance of PRO constructions is due to greater structural flexibility, as they select finite and non-finite clausal complement more frequently, along with prepositional phrases. TR and PRO differ most in their distribution regarding nominal expressions. For TR verbs, as in (1b), the logical object is realized as the direct object (NP); for PRO verbs, as in (1a), it is realized as an oblique or, at the phrasal level, as a PP complement, headed by a preposition (*de/di*) and in which the logical object is embedded. The process involved can be described as the demotion or suppression of the logical object to an oblique, or a non-core argument, which is a defining property of the AP construction. We also observe the presence of other AP properties, such as the presence of "verbal affixation" (Polinsky 2017:7), i.e., detransitivizing affixes, such as the reflexive *se/si*. From a diachronic perspective, we find that *apercevoir* and *ricordare* start to dominate in the past two centuries, while *se plaindre* and *lamentarsi* dominate since at least the 16th century. It appears that the latter ACTIVITY verbs prefers an AP construction, while the former ACHIEVEMENT verbs prefer a TR construction. For *(se) vanter* and *vantarsi*, there is a more equal distribution since the 17th century.

This study contributes to the discussion of the Romance reflexive SE, as well as to the work on AP constructions in accusative languages by proposing AP as a Romance phenomenon with chronological depth. The comparative and diachronic perspective, while still limited in its scope, enriches current studies of AP in Romance.

References

Collodi, Carlo. 1875. *I racconti delle fate*. Florence: Paggi.

Masullo, Pascual José. 1992. Antipassive constructions in Spanish. *Romance Languages and Modern Linguistic Theory*.175–194.

Medová, Lucie. 2009. Reflexive clitics in the Slavic and Romance languages. *Doctoral diss., Princeton University*.

Melis, Ludo. 1985. Le classement des tours pronominaux dans Le Bon Usage et dans la grammaire française contemporaine in Tradition grammaticale et linguistique: Le Bon Usage de Maurice Grevisse. *Travaux de linguistique* 12.159–175.

Melis, Ludo. 1990a. *La voie pronominale. La systématique des tours pronominaux en français moderne*. Paris/Louvain-la-Neuve: Duculot.

Melis, Ludo. 1990b. Pronominal Verbs in Old and Modern French, or How prototypes can be restructured on the basis of permanent meaning effects. *Belgian Journal of Linguistics* 5.87–108.

Nishida, Chiyo. 1994. The Spanish reflexive clitic se as an aspectual class marker. *Linguistics* 32.425–458.

Polinsky, Maria. 2017. 13. Antipassive. *The Oxford Handbook of Ergativity*, ed. by Jessica Coon, Diane Massam, and Lisa deMena Travis. Oxford: Oxford University Press. http://scholar.harvard.edu/files/mpolinsky/files/antipassive_new.pdf.

Salvatorelli, Luigi. 1943. *Pensiero e azione del risorgimento*. Torino: Einaudi.

The Grammaticalization of motion verbs in the languages of the Americas and Austronesia: Beyond the encoding of motion

Nicole Hober
(University of Bremen)

Keywords: Grammaticalization, grammaticalization paths, motion verbs, Americas, Austronesia

Schedule: We 15.30 Room 6

The envisioned talk focusses on the grammaticalization of motion verbs in the languages of the Americas and Austronesia and aims to show that the genetically distinct languages from two different macro-area grammaticalize motion verbs in distinct ways. Considering theories of and research on grammaticalization (cf. Heine, Claudi, & Hünnemeyer 1991; Croft 2003; Hancil & König 2014), ten languages from each continent are investigated to demonstrate that the same source-concepts feed into different target domains, as exemplified in (1) – (4). To this end, the present contribution addresses the following research questions:

i. What grammaticalization paths do the motion verbs in the respective languages travel?
ii. What phonological and morphological developments or alternations can be observed?
iii. Are there any thus far unobserved grammatical categories that evolve from the grammaticalized motion verbs?
iv. What do the areal and genetic comparisons tell us about the universal tendencies of grammaticalization?

By resorting to various sources of data, including grammars, parallel corpora, and other descriptive resources, the research questions will be answered in-depth. The motivation for and importance of this endeavour is grounded in observations initially made by Givón (1973) who put forward the idea of the 'containment hypothesis' according to which "the development from verbs to tense/aspect/modality markers [and] the resulting grammatical meanings are largely predictable because they are already contained in the verbal meaning" (Heine et al. 1991: 109). Consequently, the potential of motion verbs for the conceptualisation of categories that go beyond motion is frequently discussed in the literature (Toedter, Zahn, & Givón 1989; Heine & Kuteva 2002; Devos & van der Wal 2014). Still, more empirical data is needed especially from areas that have been only sporadically and superficially examined. This contribution seeks to fill this gap.

To exemplify, in Yucatec Maya (Americas), the general motion verb *bíin* 'go' has gone down two grammaticalization paths, a direct one where *bíin* grammaticalized to an auxiliary verb indicating (1) future tense and an indirect one where is developed a function as a (2) quotative marker.

Yucatec Maya *bíin* 'go'

(1) GO > FUTURE

bíin *lúuk-uk-en*

FUT fall-SUBJ-B1

'I will fall.' (Lehmann 2017: 192)

(2) GO > QUOTATIVE

t-u-tàas-ah ***b'in***

COM-A3-bring-TR **QUOT**

'They say he brought it.' (Lucy [1993], 2010: 119)

In Djinang (Australonesia), on the other hand, while the general motion verb *kiri-* 'go' has also grammaticalized into a marker of future tense (3), it also functions as progressive aspect auxiliary (4).

Djinang *kiri-* 'go'

(3) GO > FUTURE

nyuni *garray-dji* ***giri-Ø***

2SG.NOM good-INSTR **go-FUT**

'You will do/live really well' (Waters 1989: 131)

(4) GO > CONTINUOUS/PROGRESSIVE

djin *ngurri-ny* ***kiri-ny*** *djadaw*

3PL.NOM sleep-RPC **PROG-**RPC [daybreak]$_{LOC}$

'They slept till daylight.' (Waters 1989: 132)

The investigation offers interesting insights to typologists, areal linguists, and researchers of grammaticalization. The overall project builds on existing work and aims towards a comprehensive understanding of the possible and predictable grammaticalization paths of motion verbs in the world's languages.

Abbreviations

A = set A pronouns (ergative function); B = set B pronouns (absolutive function); COM = completive; FUT = future; IMM = immediate (future); INCMPL = incompletive; INSTR = instrumental; LOC = locative; NOM = nominative; PL = plural; PROG = progressive; QUOT = quotative; RPC = remote past; SG = singular; SUBJ = subjunctive; TR = transitive

References

Croft, William (2003), *Typology and Universals*, 2nd edn, Cambridge: Cambridge University Press.

Devos, Maud and Jenneke van der Wal (2014), *COME and GO off the Beaten Grammaticalization Path*, Berlin: De Gruyter Mouton.

Givón, Talmy (1973), The time-axis phenomenon, *Language* 49, 890-925.

Hancil, Sylvie and Ekkehard König (2014), *Grammaticalization – Theory and Data*, Amsterdam: John Benjamins.

Heine, Bernd, Ulrike Claudi and Friedericke Hünnemeyer (1991), *Grammaticalization: A Conceptual Framework*, Chicago: University of Chicago Press.

Heine, Bernd and Tania Kuteva (2002), *World Lexicon of Grammaticalization*, Cambridge: Cambridge University Press.

Lehmann, Christian (2017), Grammaticalization of tense/aspect/mood marking in Yucatec Maya, in W. Bisang and A. Malchukov (eds), (2017), *Unity and diversity in grammaticalization scenarios*, Berlin: Language Science Press, 173–237.

Lucy, John A. ([1993], 2010), Metapragmatic presentationals: reporting speech with quotatives in Yucatec Maya, in J. A. Lucy (ed), ([1993], 2010), *Reflexive Language: Reported Speech and Metapragmatics*, Cambridge: Cambridge University Press, 91-126.

Toedter, Christa, Charlotte Zahn, and Talmy Givón (1989), *The Transformations of "go" in Pastaza Quechua: A New Channel for Voice?*, Eugene/Oregon: Typescript.

Waters, Bruce E. (1989), *Djinang and Djinba: A Grammatical and Historical Perspective* (Pacific Linguistics, Series C, 114), Canberra: Australian National University.

Convergence or not? Geography and history

Hans Henrich Hock

Schedule: We 11.00 Room 8

<pdf>

Computer-assisted approach to investigating history of the Transeurasian languages and their neighbours based on structural features

Nataliia Hübler
(Max Planck Institute for the Science of Human History, Jena)

Keywords: Transeurasian, typological feature, Bayesian tree-building, STRUCTURE, historical signal

Schedule: Thu 14.00 Room 4

The current study uses structural language data to investigate the relationships and the language history of the Transeurasian languages (Turkic, Mongolic, Tungusic, Japonic and Koreanic) and languages of the area (Uralic, Ainu, Nivkh). The workflow that switches from qualitative to quantitative analysis allows us for integrating the linguistic expertise at every stage of the study. First, data is collected from language grammars. As no tools that automatize this process have been developed yet, this step requires the understanding of the data to be collected and of the language sources, from which the data is collected. Second, an exploratory analysis of the data is performed with the tools developed for sequence analysis in genetics, such as neighbour joining. This method provides a visualization of the relationships between the languages in terms of conflicting signal along with quantification of the conflicting signal. A more sophisticated procedure that incorporates specialist expertise knowledge in the quantitative analysis is Bayesian tree-building. This method allows us to represent language history in form of a tree, additionally measuring the uncertainty around the nodes of this tree. The result of the analysis in BEAST 2 (Bouckaert et al. 2014) suggests a binary structure of the Transeurasian unity: an Altaic (Turkic, Mongolic, Tungusic) and a Japonic branch. As areal convergence cannot be fully excluded for typological data, contact events appear clearly visible in the tree: Uralic languages cluster closely together with Turkic languages (most probably due to Mari-Chuvash contacts).

Given the low distinctness of the source of the historical signal in typological data, methods directly incorporating horizontal (borrowing) transmission along with the vertical (inheritance) transmission might turn out to be more appropriate and reveal patterns in language history, not accessible for tree-building methods. By analysing our data with the software STRUCTURE (Pritchard et al. 2000), we reach five different sources of ancestry, let us call them provisorily Turkic, Tungusic, Mongolic, Japonic and Asian. These five sources can account for variation in structural features across 5 Transeurasian language families as well as Uralic languages, Ainu and Nivkh. Our analysis suggests that Asian ancestry is available in a high proportion in Ainu, Nivkh (both over 95%), Uralic (over 50%), some Japonic languages (up to 40%), Tungusic (in Manchu at almost 80%), Korean (around 30%). The Asian residue in Japonic and Tungusic languages goes in line with the assumption of the Ainu substratum in Japonic and Nivkh substratum in Tungusic languages, which corresponds to a scenario of language shift (Robbeets et al. 2017). Another striking result is the reflection of Chuvash influence on Mari (Savelyev 2018) in a shared Turkic and Asian ancestry.

In the current study, we resort to different computer-assisted methods to investigate the relationships between the languages in question, which give us insights into the sources of the historical signal in the structural features and different perspectives on the relationships between the Transeurasian languages and their neighbours. In order to look deeper into the distribution of the Asian ancestry across other languages in the region, data on other language families (Chukotko-Kamchatkan, Yeniseian, Eskimo-Aleut) needs to be collected and integrated in the analysis.

Acknowledgments

The research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 646612) granted to Martine Robbeets.

References

Bouckaert, Remco, Joseph Heled, Denise Kühnert, Tim Vaughan, Chieh-Hsi Wu, Dong Xie, Marc A. Suchard, Andrew Rambaut, and Alexei J. Drummond (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* 10, no. 4: e1003537.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Robbeets, Martine, Nataliia Neshcheret, Tao Li, Choongwon Jeong and Chuanchao Wang (2017, 11–12 July). An Asian North Pacific Coast substratum in the Transeurasian languages? Paper presented at the DLCE workshop "Language shift and substratum interference in (pre)history", Jena, Germany.

Savelyev, Alexander (2018). K utochneniju scenarija chuvashsko-marijskih kontaktov [Towards specifying the scenario of the Chuvash-Mari contacts]. To be published in the proceedings of the 11th International conference on language contact in the Volga-Ural region.

Event and argument structure in English zero-derived nominals

Gianina Iordăchioaia
(University of Stuttgart)

Keywords: zero-derived nominals, event structure, argument structure, verb classes, English

Schedule: We 14.30 Room 7

This paper is concerned with the morphosyntax of *deverbal zero-derived nominals* (ZNs, e.g., to walk > a walk-Ø), which have mostly been ignored in the literature by comparison to the suffix-based nominals (walk**ing**, examin**ation**, assign**ment**).

Following the tradition of Grimshaw (1990), syntactic models of word formation have posited a structural ambiguity between argument structure nominals as in (1a) (ASNs) and result nominals as in (1b) (RNs) in terms of inheritance of verbal event structure in ASNs as in (2a) and lack thereof in RNs as in (2b) (Marantz 1997, van Hout & Roeper 1998, Harley & Noyer 2000, Alexiadou 2001, Borer 2013 etc):

(1) a. *The instructor's* (intentional) **examination** *of the papers* took a long time. (ASN)
b. *The instructor's* (*intentional) **examination**/exam (**of the papers*) was in her bag. (RN)

(2) a. $[_{DP}$ D $[_{nP}$ n **[** ***verbal event structure*** **[**$_{vP}$ ***v*** [√ROOT]]]]] (ASN)
b. $[_{DP}$ D $[_{nP}$ n
[√ROOT]]] (RN)

In this tradition, ZNs have been considered to behave like RNs, especially since they were argued not to realize arguments (see (3)), with just a few 'exceptions'. In Borer's (2013) implementation, ZNs are derived from an uncategorized root as represented in (4) in a Distributed Morphology version (cf. (2b)).

(3) *John's* breaking/***break** *of the glass* (Marantz 1997)
(4) $[_{DP}$ D $[_{nP}$ Ø [√BREAK]]]

In this paper I argue that not all ZNs are root-derived, since we can identify coherent classes of ZNs (including Borer's 'exceptions'), which show evidence for event structure properties and argument realization as in ASNs, as supported by examples from electronic corpora (the BYU corpora):

(5) a. *their* **dispatch** *of Osama bin Laden*/[He] probably witnessed *their* **murder** *of his mother*
b. the **release** *of prisoners*/It could continue *its* **import** *of petroleum products*
c. He began *his* **climb** *of M. Everest*
d. a furious **toss** *of said wheel to the ground*/a quick **throw** *of a syringe*

(6) a. the events [...] have brought a rapid **rewrite** *of the Green story*
b. [...] useful for the ongoing **declutter** *of the legislative landscape*
c. during the **upload** *of a large file*/*Our* **bailout** *of SCF depositors* has cost us a $800 mill.

The base verbs in (5) and (6) are special in that they lexicalize a result component: see 'murder' verbs in (5a) (Levin 1993), Latinate verbs, prefix and particle verbs in (5b) and (6) (Harley 2008), and verbs that encode manner+result in (5c-d) (Beavers & Koontz-Garboden 2012). Yet, their ZNs have complex event readings, which can only come about from an event structure inherited from the verb, as also evidenced by their argument realization (see Rappaport Hovav & Levin 1998ff).

I will argue that depending on the verb class they are derived from, the possible competition with *-ing* and Latinate suffixes, and the special properties of their zero suffix (e.g., attachment to native bases, which is also found with ZNs in Romance languages like Italian and Portuguese), ZNs are in fact similar to Latinate SNs in building both ASNs and RNs as in (2a-b).

Acknowledgments: This work is supported by the DFG (German Research Foundation) research grant IO 91/1-1 to the project "Zero-derived nouns and deverbal nominalization: An empirically-oriented perspective".

References

Alexiadou, Artemis (2001), *Functional structure in nominals*, Amsterdam: John Benjamins.

Beavers, John & Andrew Koontz-Garboden (2012), Manner and result in the roots of verbal meaning, *Linguistic Inquiry* 43(3): 331-369.
Borer, Hagit (2013), *Taking form*, Oxford: Oxford University Press.
Grimshaw, Jane (1990), *Argument structure*. Cambridge, MA: MIT Press.
Harley, Heidi (2008), The bipartite structure of verbs cross-linguistically, in Silva, T. C. & H. Mello, eds., *Conferencias do V Congresso Internacional da Associacao Brasileira de Linguistica*, Belo Horizonte, Brazil: ABRALIN and FALE/UFMG, 45-84.
Harley, Harley & Rolf Noyer (2000), Formal versus encyclopedic properties of vocabulary: Evidence from nominalisations, in Peeters, B., ed., *The Lexicon-encyclopedia interface*, Amsterdam: Elsevier, 349-374.
van Hout, Angeliek & Thomas Roeper (1998), Events and aspectual structure in derivational morphology, in Harley, H., ed., *Papers from the UPenn/MIT roundtable on argument structure and aspect.* MITWPL 32, Cambridge, MA: MIT Press, 175-200.
Levin, Beth (1993), *English verb classes and alternations*, Chicago: University of Chicago Press.
Marantz, Alec (1997), No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *University of Pennsylvania working papers in linguistics* 4, 201-225.
Rappaport Hovav, Malka & Beth Levin (1998), Building verb meanings, in Butt, M. & W. Geuder, eds., *The projection of arguments*, Stanford: CSLI Publications, 97-134.

Detransitive voice in Modern Breton: Verbs with *en em*

Britta Irslinger

Schedule: We 11.30 Room 7

Modern Breton *en em* is a marker of detransitive voice, coding reflexivity (1), reciprocity (2), and various middle situation types (3).

(1) *distrujañ* 'to destroy' : *en em zistrujañ* 'to destroy oneself, to commit suicide'
(2) *karout* 'to love' : *en em garout* 'to love each other'
(3) *gwalc'hiñ* 'to wash' : *en em walchiñ* 'to wash (oneself)'
diskouez 'to show' : *en em ziskouez* 'to appear'

Originating from a verbal prefix indicating reciprocity, the marker *em*, combined with the generalized 3sg. masc. pronoun *en*, underwent a process of grammaticalisation by which its functions finally became close to those of the French reflexive marker *se*, cf. ex. (4). The latter is based on PIE **s(w)e-* like German *sich* or Russian *sebja*.

(4) [*Hogen, maouezed Kerunkun a lavare o doa gwelet Anna Tregidi o lammat er ster Avon*]
evit en em zistrujañ.
for REFL destroy.INF
'[However, the women of Kerunkun said they saw Anna Tregidi jumping into the river Avon] in order to kill herself.' (Riou 1934/1957, 80).

This development is unique both among the cognate Insular Celtic (Irslinger 2014) and the neighbouring Romance, Gemanic and Slavic languages. Haspelmath 2003 has demonstrated that in the latter languages intensifiers develop into reflexives by rather similar grammaticalization paths.

In Breton, however, a corresponding outcome was reached by a completey different process, which thus broadens the inventory of such features found in linguistic typology and areal linguistics.

Nevertheless, Breton has been neglected so far by the otherwise intensive typologic research within this field.

The present paper will undertake a first step in this direction with the help of a corpus-based study. Lacking a corpus, a I established my own one, consisting of almost half a million words and containing 133 different *en em*-verbs (i.e. 438 tokens).

text	**word count**	**types**	**tokens**	**ratio**	**per 1000 words**
Rozmor 1996	37 866	29	72	525,91	1,90
Riou 1934 (ed. 1957)	21 767	22	47	463,13	2,16
Drezen 1949	11 464	18	25	458,66	2,18
Kontadennoù c. 1870	34 189	28	121	282,55	3,54
Burel 1905	36 817	60	173	212,81	4,70
Total	**242 103**	**133**	**438**		

These verbs were analysed according to semantics, frequency and eventual French counterparts or models, and discussed against the background of typologic research.

This analysis yielded the following results:

While some examples found in the older texts dating from the end of the 19th century reflect the original use of the marker *em* as evidenced by Middle Breton, the overall tendency consists in calquing or borrowing French *se*-verbs.

However, the use of *en em* within the different texts is not consistent, both with regard to the verbs employed and their semantics and functions. Various parameters play a role, such as register, regional varieties, bilingualism of the authors or speakers as well as their choices on language use.

References

Favereau, Francis. 1997. *Grammaire du breton contemporain / Yezhadur ar brezhoneg a-vremañ.* Morlaix: Skol Vreizh.

Haspelmath, Martin, 2001. 'The European linguistic area: Standard Average European'. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.). Vol. 2, 1492–1510.

Haspelmath, Martin, 2003. 'The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison.' In Michael Tomasello (ed.), *The new psychology of language*, Vol. 2, 211–242. Mahwah, NJ: Erlbaum,

Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.) 2001. *Language typology and language universals. An international handbook*. 2 vols. Berlin: de Gruyter (HSK 20.1, 20.2).

Heine, Bernd, & Tania Kuteva. 2003. On contact-induced grammticalization. *Studies in Language* 27.3. 529–572.

Hemon, Roparz. 1954. Le verbe réfléchi en Breton. *Zeitschrift für celtische Philologie* 24. 248–263.

Hemon, Roparz. 1975. *A historical morphology and syntax of Breton.* Dublin: Dublin Institute for Advanced Studies (repr. 1984).

Irslinger, Britta. 2014. Intensifiers and reflexives in SAE, Insular Celtic and English. *Indogermanische Forschungen* 119. 159–163.

Kemmer, Suzanne, 1993. *The middle voice*. Amsterdam/Philadelphia: John Benjamins.

König, Ekkehard, 2001. 'Intensifiers and reflexive pronouns.' In Martin Haspelmath, Ekkehart König, Wulf Oesterreicher & Wolfgang Raible (eds.), Vol. 1, 747–760.

Le Roux, Pierre. 1957. *Le verbe breton (morphologie, syntaxe)*. 2nd ed. Rennes and Paris: Plihon and Champion.

Riou, Jakez. 1957. *Geotenn ar Werc'hez ha danevelloù all*. Ar Baol: Al Liamm.

Stéfanini, Jean. 1962. *La voix pronominale en ancien et en moyen français*. Gap: Ophrys.

Winford, Donald. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.

Harnessing Bayesian phylogenetics to test a Greenbergian universal

Gerhard Jäger & Ramon Ferrer-i-Cancho
(University of Tübingen & Universitat Politècnica de Catalunya)

Keywords: word order, typology, phylogenetics, dependency length minimization, adjectives

Schedule: We ~~12.00~~ 11.30 Room 1

Greenberg's (1963) Universal 17 states: *With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.*

Subsequent work has generalized this finding, claiming that Adj-N generally tends to follow VO order. Recently, Ferrer-i-Cancho (2008; 2015a) has argued that the principle of *dependency length minimization* (DLM) provides a functional motivation for Universal 17, as well as for its mirror image (Verb-final languages prefer adjective-noun order).

Dryer (1992), however, has argued against such a correlation. Using genealogically and areally stratified samples, he found no evidence for such effects.

Given that Dryer's stratified samples are modest in size, one might suspect that Dryer demonstrates absence of evidence rather than evidence of absence. Phylogenetic techniques provide us with statistical tests much more powerful than what was possible ten or twenty years ago.

We distinguish three word-order types (verb-initial, verb-medial and verb-final) since DLM predicts a bias for N-Adj in verb-initial languages and Adj-N in verb final languages with verb-medial languages as a special case.

We used a sample of 902 languages for which (a) WALS (wals.info; accessed on 1/10/19) provides the relevant information and (b) ASJP (Wichmann et al. 2018) contains sufficient information for phylogenetic inference. This sample contains languages from 181 (Glottolog-)lineages, including 76 isolates.

Using the methods described in (Jäger 2018), we extracted a lexical character matrix from ASJP word lists. For each lineage containing at least two languages, we inferred a posterior sample of phylogenies including branch lengths. The expert trees from Glottolog were used as backbone constraints.

To test for correlation between the two word-order variables, we deployed the method from (Pagel and Meade 2006). This amounts to estimating the rates of change between different word order types given our knowledge of the synchronous distribution of these variables and the phylogenetic structure of language relationships. (A similar approach is used in Dunn et al. 2011. However, unlike them we assume that the dynamics of word order evolution is invariant across lineages. Additionally, we incorporated isolated languages by assuming that their word-order characteristics are drawn from the equilibrium distribution of the estimated Markov chain.) This estimation is performed for two models:

- *uncorrelated:* The transition probabilities between two states of one variable are independent from the state of the other variable.
- *correlated:* Transition rates for one variable can depend on the state of the other variable.

Statistical model comparison revealed strong evidence in favor of the uncorrelated model (Bayes Factor > 200). We conclude that Dryer's finding holds up under modern statistical methods. Our preliminary analyses replacing adjectives by relative clauses suggest that the failure of the predictions of DLM could be due to the fact that adjectives tend to be shorter and then DLM is more likely to be beaten by other word order principles Ferrer-i-Cancho (2015b).

References

Matthew S. Dryer. The Greenbergian word order correlations. Language, 68(1):81–138, 1992.

Michael Dunn, Simon J. Greenhill, Stephen Levinson, and Russell D. Gray. Evolved structure of language shows lineage-specific trends in word-order universals. Nature, 473(7345):79–82, 2011.

Ramon Ferrer-i-Cancho. Some word order biases from limited brain resources. A mathematical approach. Advances in Complex Systems, 11(3):393–414, 2008.

Ramon Ferrer-i-Cancho. The placement of the head that minimizes online memory. A complex systems approach. Language Dynamics and Change, 5:114–137, 2015a.

Ramon Ferrer-i-Cancho. Reply to the commentary "Be careful when assuming the obvious", by P. Alday. Language Dynamics and Change, 5:147–155, 2015b.

Joseph Greenberg. Some universals of grammar with special reference to the order of meaningful elements. In Universals of Language, pages 73–113. MIT Press, Cambridge, MA, 1963.

Sebastian Höhna, Michael J. Landis, Tracy A. Heath, Bastien Boussau, Nicolas Lartillot,

Gerhard Jäger. Global-scale phylogenetic linguistic inference from lexical resources. Scientific Reports, 5, 2018. https://www.nature.com/articles/sdata2018189.

Mark Pagel and Andrew Meade. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. The American Naturalist, 167(6): 808–825, 2006.

Søren Wichmann, Eric W. Holman, and Cecil H. Brown. The ASJP database (version 18). http://asjp.clld.org/, 2018.

Towards a typology of antipassive markers in ergative languages

Katarzyna Janic & Ilja Seržant
(University of Leipzig)

Key words: antipassive, form-frequency correspondence, ergative languages, efficiency

Schedule: Fri 15.00 Room 6

Introduction: Our focus is the shape of grammatical markers used to code antipassives in ergative languages. A marker is a form that consists of phonetic segments (e.g. affix, clitic), whereas an antipassive is a construction that gives access of the A-argument to S (ergative to absolutive) and that syntactically demotes the P-argument (Polinsky 2017), without affecting the argument structure of the active verb. In our study, we exclude similar constructions such as *agent-focus constructions* (Mayan languages) or *object incorporation* (Chukotko Kamchatkan languages), which may have comparable effects to antipassives.

Database is based on a convenience sample of ca. 60 languages and covers all large linguistic areas, including Africa, Australia, Eurasia, North America and South America. We selected maximally three languages per family that are not closely related (e.g. in terms of dialects).

Methodology: We annotated antipassive constructions along the following criteria:

(i) length of the antipassive marker; length of the active marker (if there is one, else - zero);
(ii) whether or not the antipassive is more frequent than active in the respective language;
(iii) coding strategy (affix, root vowel gradation, auxiliary, etc.);
(iv) functions of antipassive (atelicization, promoting A into pivot, etc.);
(v) whether the antipassive marker is syncretic with some other functions or not;

(vi) whether the antipassive construction has selectional input restrictions;
(vii) whether the S/P agreement is retained in the antipassive (if applicable);
(viii) whether the antipassive construction may also be formed from intransitive verbs.

Theoretical background: In this study, we test the form-frequency correspondence principle (FFCP) by Haspelmath (2006, 2008a,b,c) with reference to the active-antipassive opposition, i.e. we test how corpus frequency (ii) is related to the coding strategy (iii). The mechanism underlying FFCP is based on probabilistic expectations of speakers generalized across particular contexts and utterances and stored in the memory as part of language knowledge. These expectations, in turn, provide space for communicatively more efficient coding: something that is expected may either be left uncoded or coded with less morphological material. By contrast, "(l)ess predictable information will be given more coding material" (Givón 1991: 87). Using less or no material minimizes the production efforts of speakers. Probabilistic expectations allow them managing the trade-off between the strive towards minimizing production efforts and the desire for successful communication (Kurumada & Jaeger 2015). Additionally, we take into account type frequency (vi)+(viii).

Results: Building on FFCP, we predict that the coding of the active and antipassive will reveal a coding asymmetry, i.e., if the antipassive is rarer in the language than the active, then the former will be coded by a longer marker than the respective active; and *vice versa.* Furthermore, we provide other findings emerging from the exploration of our database such as implicational correlation between the functions (iv) and the morphological type of coding of antipassives (such as ablaut, reduplication, affix, auxiliary, etc.) (iii).

References

Givón, Talmy 1991: Isomorphism in the Grammatical Code: Cognitive and Biological Considerations, *Studies in Language* 15(1), 85-114.

Haspelmath, M. 2006. Against markedness (And What to Replace It With). *Journal of Linguistics* 25(1). 25-70.

Haspelmath, Martin 2006: Against markedness (and what to replace it with), *Linguistics* 42, 25-70.

Haspelmath, Martin. 2008a: Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1), 1-33.

Haspelmath, Martin. 2008b: Creating economical patterns in language change. In: Good, Jeff (ed.), *Linguistic universals and language change.* Oxford: Oxford University Press. 185-214.

Haspelmath, Martin. 2008c: A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1), 40-63.

Kurumada, Chigusa & T. Florian Jaeger 2015: Communicative efficiency in language production: Optional case-marking in Japanese, *Journal of Memory and Language* 83, 152-178.

Polinsky, M. 2017. Antipassive. In Coon J., D. Massam & L. deMena Travis (eds.), *The Oxford Handbook of Ergativity*. Oxford: Oxford University Press. 307-331.

Demonstrative pronouns and number marking: a typological overview

Jessica Ivani
University of Kiel

Schedule: Thu 10.00 Room 6

The aim of the current study is to provide an extensive overview of the markers used to express number oppositions on demonstrative pronouns (as defined by Diessel 1999) in the languages of the world. In particular, the following research questions are addressed: a) what is the cross-linguistic distribution of number values attested (overt singular, dual, plural, others), b) what is the type of the marker used (affixes, lexical items and the like) and c) which further characteristics do these markers exhibit (e.g. cumulation with other grammatical features such as gender, and so on). The study uses data from *Tymber* (to appear), a typological database on nominal number constructions and comprises a genetically and geographically diverse sample of about 300 languages. A core focus of the study is the relation between number markers used on demonstrative pronouns and other number markers in the nominal domain (such as full nouns and independent personal pronouns). We show that languages exhibit high degree of variation, with the following high-level systems been attested:

- Demonstrative pronouns select the same markers as nouns (e.g. Cayuvava, Key 1963);
- Demonstrative pronouns use the same markers as pronouns (e.g. Edolo, Gossner 1994);
- Demonstrative pronouns show specific restricted marking or they have a richer number system than the other nominals (e.g. Lakhota, Ingham 2003).

Based on this overview, our discussion also touches upon more general questions, such as the distributional tendencies that can be observed on the broader level of the full nominal domain. From this perspective, the inherent properties of the number markers themselves (e.g. cumulation with gender, or their construction type) seem to be directly involved in the rise of the whole nominal number system.

On these grounds, we can identify three main development scenarios: a) joint development of number markers, b) spread of a generalised marker from one nominal element to the other, c) independent rise of number markers for each — or a subset of — noun type(s). A selection of case studies will be presented and discussed. These latter observations bring us a step closer towards a typology of the nominal number system diachrony, an ongoing enterprise of which the current study is part of.

References

Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Typological Studies in Language, 42. Amsterdam: John Benjamins.

Gossner, Jan. 1994. *Aspects of Edolo Grammar*. Ann Arbor: UMI. (MA thesis, University of Texas at Arlington; xiv+139pp.)

Ingham, Bruce. 2003. *Lakota*. (Languages of the World: Materials, 426.) München: Lincom. x+113pp.

Key, Harold. 1963. *Morphology of Cayuvava*. University of Texas Austin Ph.D. Dissertation. (Doctoral dissertation, Austin: University of Texas at Austin; vii+89pp.)

French causal connective "puisque"/since: information structure, position and at-issueness

Hasmik Jivanyan

Schedule: Thu 14.00 Room 11

<pdf>

Subpart inheritance links in child language: The development of *c'est* and *il y a* clefts in French

Morgane Jourdain & Karen Lahousse
(KU Leuven, University of Lille & KU Leuven)

Schedule: Thu 9.30 Room 11

Background and goal

The aim of the present study is to investigate whether *c'est* (1) and *il y a* (2) clefts in French child language inherit IS and semantic properties from their subparts, forming a network-like construct-i-con.

(1) **C'est** moi **qui** coupe en deux. (Marion, age 3, TCOF)
'It's me who cuts in half'
(2) **Il y a** tout plein de gens **qui** viennent. (Valentin, age 2, TCOF)
'There are a lot of people who are coming'

Karssenberg (2018) shows that simple *il y a* existential clauses can express the same IS articulations as *il y a* clefts, and according to Clech-Darbon, Rebuschi, and Rialland (1999), the semantic and pragmatic values of *c'est* clefts stem from the semantico-pragmatic values of their parts. In constructionist terms, there is a subpart inheritance link (Goldberg, 1995) between simple *il y a* and *c'est* clauses, and the corresponding clefts in adult language.

Methodology

We conducted an analysis of 245 clefts produced by 125 children between age 2 and 7 in the transversal corpus TCOF (ATILF, 2018), and 281 clefts by 3 children between age 1 and 3 from the longitudinal Lyon corpus (Demuth & Tremblay, 2008). We used the tests described in Karssenberg (2018) to identify clefts and the framework of Lambrecht (1994) to code IS properties.

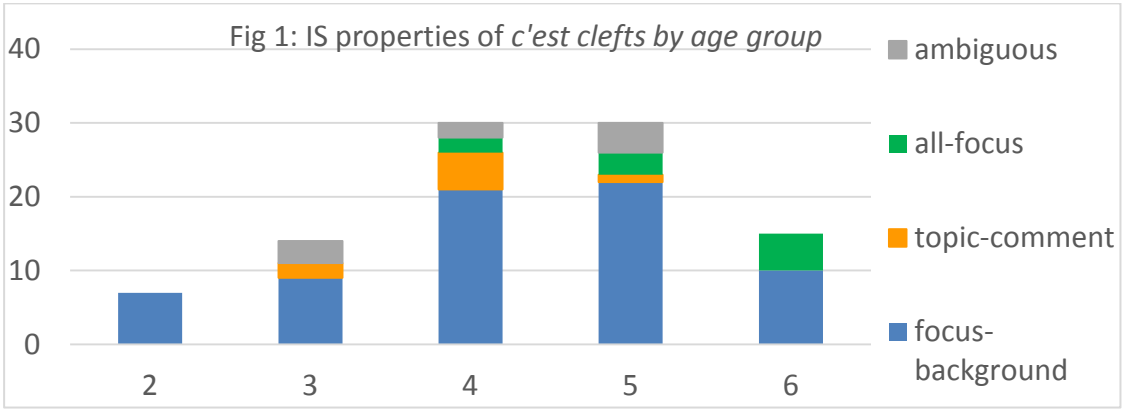
Simple c*'est* and *il y a* clauses

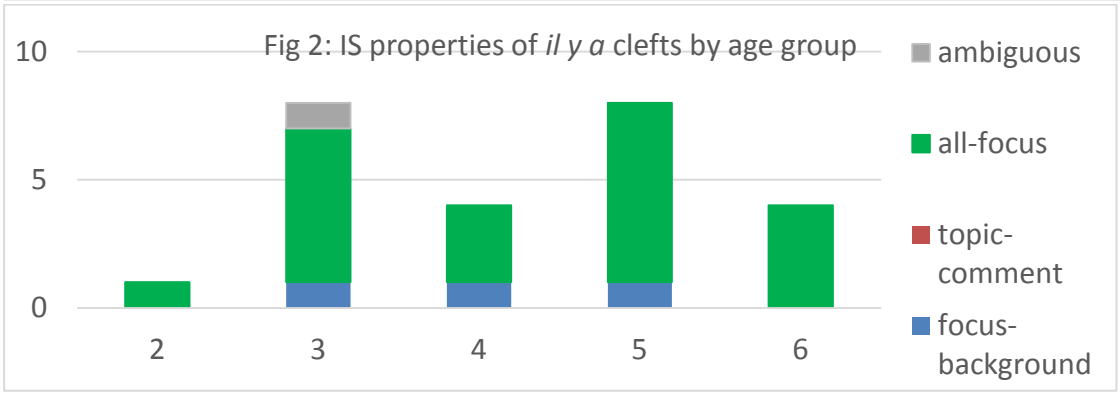
Children from the Lyon corpus start producing *c'est* clauses with identificational or predicative meanings towards age 1;10. Between age 2 and 2,5, they start producing specificational *c'est* clauses.
Children from the Lyon corpus start producing *il y a* clauses between age 2 and 2,5, which are all existential clauses.

Inheritance links and clefts

Children start producing *c'est* clefts only once they start producing specificational *c'est* clauses.
Until age 3, all *c'est* clefts are specificational and focus-background. Hence, when they start producing *c'est* clefts, they transfer the specificational properties of the simple *c'est* construction. They develop other types of articulations which do not stem from the properties of the simple *c'est* clauses really emerge from age 4 (see table 1 for the TCOF data).

Most *il y a* clefts are eventive and all-focus (see table 2 for the TCOF data), while they can have other properties in adult data (Karssenberg & Lahousse, 2018). Eventive clauses are one type of existential constructions (Leonetti, 2008), suggesting that the IS and semantic properties of *il y a* clefts are inherited from simple *il y a* clauses in child language.





Conclusion

The IS and semantic properties of French children's clefts are related to those of their simple *c'est* and *il y a* clause. This result suggests that children transfer the IS properties of *c'est* and *il y a* clauses to the corresponding clefts, supporting the theory that children's constructions form a network-like construct-i-con, and make use of subpart inheritance links with simpler constructions.

References

ATILF. (2018). TCOF : Traitement de Corpus Oraux en Français [Corpus] Retrieved 10 July 2018, from ORTOLANG (Open Resources and TOols for LANGuage)

Clech-Darbon, A., Rebuschi, G., & Rialland, A. (1999). Are there cleft sentences in French? In L. Tuller & G. Rebuschi (Eds.), *The grammar of focus* (pp. 83-118). Amsterdam: John Benjamins.

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language, 35*(1), 99-127.

Goldberg, A. (1995). *Constructions. A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Karssenberg, L. (2018). *Non-prototypical clefts in French: A corpus analysis of French il y a clefts*. Berlin: De Gruyter Mouton.

Karssenberg, L., & Lahousse, K. (2018). The information structure of French *il y a* clefts and *c'est* clefts: A corpus-based analysis. *Linguistics, 56*(3), 513-548.

Lambrecht, K. (1994). *Information structure and sentence form: topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

Leonetti, M. (2008). Definiteness effect and the role of the coda in existential constructions. In H. H. Müller & A. Klinge (Eds.), *Essays on nominal determination: from morphology to discourse management* (pp. 131-162). Amsterdam: John Benjamins Publishing Company.

New speakers and language change: the case of contemporary Galician

Johannes Kabatek

Schedule: We 16.30 Room 2

At least since the IJSL special issue on *New Speakers of Minority Languages* (IJSL 231, 2015), the notion of "new speakers" is quite common in sociolinguistics: it refers to speakers whose L1 is the language of the majority in contact situations between majority/minority languages and who consciously adopt the minority language (which is in fact a L2 for them) for everyday purposes. In some communities like the Basque country or in Wales, New Speakers play an important role in language revival. New speakers have been described from an individual perspective or as groups with particular attitudes. What is more difficult to be found are descriptions of the impact New Speakers may have on the evolution of a language.

In the case of contemporary Galician, where New Speakers participated intensively in the standardization process especially from the 1980s onwards, the phenomenon of New Speakers is not only well documented as an individual or social phenomenon, there exist also exhaustive studies on the linguistic properties of the Galcian spoken by *neofalantes*: it differs from the traditionally Castilianized urban or semi-urban varieties on all levels of linguistic structure.

This study will show how some of characteristics of *neofalantes* (reduction of the vowel system from seven to five, prosodic features, clitic positioning, changes in the verbal system, changes in the lexicon) are making their way into the rather recent Galician standard language and are now accepted and partly even adopted by speakers with traditional Galician background. Galician is regarded here as paradigmatic for the widespread regional language revival since the last decades of the 20th century, where New Speakers have played or play a role which is sometimes crucial.

The paper is based on comparison of research on Galician *neofalantes* and standard utterances from the early 1990s and on an analysis of a corpus of contemporary standard utterances (2019) from Galician radio and TV. The corpus study will be accompanied by interviews on attitudes towards the standard variety expressed by speakers with different backgrounds.

References

Bernadette O'Rourke, Joan Pujolar and Fernando Ramallo (eds): "Introduction: New Speakers of Minority Languages: The Challenging Opportunity", *International Journal of the Sociology of Language*, special issue, 1-20.

O'Rourke, Bernadette / Ramallo, Fernando (2015): "*Neofalantes* as an active minority: understanding language practices and motivations for change amongst new speakers of Galician", *IJSL* 231, 147-166.

Case marking of free-relatives based indefinites in Hill Mari

Nataliya Kalanova
(Lomonosov Moscow State University)

Keywords: indefinite pronoun, case-marking, free choice, free relative clause, grammaticalization

Schedule: We 14.30 Room 1

This study concerns two indefinite pronoun series performing the free choice function (in terms of Haspelmath 1997, Giannakidou 2001) in Hill Mari (< Finno-Ugric). The data were collected in fieldwork in the village of Kuznetsovo and its surroundings (Mari El Republic, Russia, 2018) mainly by elicitation, as well as by consulting the corpus of transcribed oral narratives (ca. 56000 tokens so far). I will focus on the pronouns with indefiniteness markers *šon* and *popaza.* Both of them have verbal origin: the source of *šon* is the verb *šoaš* 'to reach' (homonymous forms: converb and 3SG preterite), and *popaza* goes back to the verb *popazaš* 'to get' (non-past tense, 3SG). I argue that they show different strategies of morphological marking, which, as will be shown in the talk, follows from the different degree of their grammaticalization. In (Haspelmath 1997: 129–192) there is no detailed case study of a system where the inventory of free choice pronouns is formed in a similar way.

Previous descriptions of Hill Mari contain few information of these pronouns. Thus, Savatkova (2002:168 - 170) just briefly mentions the *šon*-series.

Our data show that *šon* is more grammaticalized than *popaza*. The *šon*-series allows accusative, genitive and dative markers either on the stem or on the indefiniteness marker. However, if the case affix is attached to the stem, the meaning of the sentence changes (1). The *popaza*-series permits the markers of these cases on either part without any semantic differentiation (2).

(1) a. *vas'a* *ma* *šo-n-**ə̂m*** *kačk-eš*
Vasya what reach-CVB/PRET-ACC eat-NPST.3SG
'Vasya eats just anything'

b. *vas'a* *ma-**m*** *šo-n* *kačk-eš*
Vasya what-ACC reach-CVB/PRET eat-NPST.3SG
'Vasya eats what he wants'

(2) *vas'a* *ma-**m*** *popaz-a* / *ma* *popaz-a-**m*** *näl-eš*
Vasya what-ACC get-NPST.3SG what get-NPST.3SG-ACC take-NPST.3SG
'Vasya buys just anything'.

For the *šon*-series the markers of the locative cases (illative, lative, and inessive) are not allowed on the stem. On the contrary, the marker *popaza* cannot be followed by locative case markers (3)

(3) *maša* *ma* *šo-n-**eš*** / **ma-**eš*** *šo-n* /
Masha what reach-CVB/PRET-LAT what-LAT reach-CVB/PRET
*ma-**eš*** *popaz-a*/ **ma* *popaz-a-**eš*** *risuj-a*
what-LAT get-NPST.3SG what get-NPST.3SG-LAT draw-NPST.3SG
'Masha draws just on anything (wherever, indiscriminately)'.

It is also possible to add plural or possessive markers to *šon* but not to the stem (4). At the same time, plural markers are ungrammatical on *popaza* (4), and possessive ones cannot combine with this series at all (it has some semantic reasons to be elaborated on in the talk).

(4) *vas'a* *ma* *šo-n-**vlä**-m* / **ma-**vlä**-m* *šo-n* /
Vasya what reach-CVB/PRET-PL-ACC what-PL-ACC reach-CVB/PRET
*ma-**vlä**-m* *popaz-a*/ **ma* *popaz-a-**vlä**-m* *näl-eš*

what-PL-ACC get-NPST.3SG what get-NPST.3SG-PL-ACCtake-NPST.3SG
'Vasya buys just anything'.

The research has revealed different behaviour of different cases during the grammaticalization process. Relying on the analysis of the general NP properties in Hill Mari (Pleshak 2017), I will discuss these differences in more detail. It will be also mentioned how the different degree of grammaticalization reflects in the syntactic properties of the two constructions, following the previous theoretical discussion of structurally similar constructions with free relatives (Bresnan & Grimshaw 1978, Testelets & Bylinina 2005, Assmann 2013).

The research has been supported by RFBR, grant № 19-012-00627a.

References

Assmann, Anke (2013), *A New Approach to Free Relatives*. Talk at ConSOLE XXI, Potsdam.

Bresnan, Joan W. & Grimshaw, Jane (1978), *The syntax of free relatives in English*. Linguistic Inquiry 9(3). 331–391.

Giannakidou, Anastasia (2001), *The Meaning of Free Choice*. Linguistics and Philosophy 24, 2001.

Haspelmath, Martin (1997), *Indefinite Pronouns.* Oxford: Clarendon.

Pleshak, Polina (2017), *Morfosintaksis imennoj gruppy v mokshanskom i gornomarijskom yazykah* [The morphosyntax of the NP in Moksha and Hill Mari], BA thesis, MSU, Moscow, 2017.

Savatkova, Anna (2002), *Gornoe narechie marijskogo yazyka* [Hill Mari language], Savariae.

Testelets, Yakov & Bylinina, Elizaveta. (2005), *O nekotoryh konstrukciyah so znacheniem neopredelennyh mestoimenij v russkom yazyke: amal'gamy i kvazirelyativy* [Some constructions with the semantics of indefinite pronouns in Russian: syntactic amalgams and quasi-relatives], Talk at the Seminar on Theoretical Semantics, IITP RAS.

English perception metaphors – will they get lost in translation? Examples from an English-Estonian empirical research

Anu Kalda & Mari Uusküla
(Tallinn University)

Keywords: metaphor, perception metaphor, metaphor translation, culture, eye-tracking

Schedule: We 15.00 Room 15

The topic of metaphor translation has attracted attention in linguistics, cognitive linguistics, psycholinguistics and translation studies, and it has been widely researched (Van den Broeck 1981, Newmark 1981, Glucksberg 2003, Kövecses 2003, Philip 2003, Schäffner 2004, Li 2011, Sandford 2016, etc.). However, less attention has been paid to the topic of perception metaphor translation, and therefore is an appropriate ground for further development. In translating perception metaphors from one language into another it is crucial to understand the intersection between linguistics, translation and culture, as each of these elements play an important role in any understanding of how the others behave and what their ties with each other are. On top of the cultural aspects, modern metaphor theory emphasises the usefulness of context (Wyler 1992, Olivera and Sacristán 2001, Glucksberg 2003, Philip 2003, Kövecses 2003, 2005, De Knop 2014, Kolahdouz 2015).

A cognitive empirical research with 20 participants on translating English perception metaphors was carried out. The subjects were provided with sentences that contained perception metaphors

(colour, sound, touch, taste and smell included). Data was extracted using keystroke logging software and eye-tracking technology. All participants were also asked to provide post-test feedback of the experiment. The analysis focussed both on the translation process of perception metaphors as well as on the actual translation strategies used while translating the metaphors in focus. As not all translation strategies are suitable for translating perception metaphors, the following selection of translation strategies was provided for current study:

- reproducing the perception metaphor in target text (TT) with a perception word;
- replacing the perception metaphor with a metaphor without a perception word in TT;
- converting perception metaphor to sense/paraphrase;
- deletion, if the metaphor is redundant;
- obvious mistranslation of perception metaphor (meaning gets lost in translation).

We conjectured that translators will dwell longer over the processing complexities involved in translating a metaphor than a non-metaphorical concept. It was assumed that metaphors may require increased cognitive processing in a translation situation. We also expected that universal metaphors were translated more easily than culturally specific metaphors.

The data gathered enables us to suggest that metaphors involve longer fixation time, presumably indicating a greater cognitive processing load in a translation task. Analysing translation process also indicated that digital resource consultation constitutes a considerable amount of the translation process. The study revealed that perception metaphors can become a translation problem due to the linguistic and cultural differences between languages. The analysis also showed that experience helps translators deal with metaphoric language faster. The longer the time spent in a foreign language environment, the easier metaphor comprehension seems to become. The participants also emphasised that context is important in both comprehension as well as while translating.

References

De Knop, Sabine (2014), A contrastive study of colour terms in French and German causal constructions, in L. Filipovic, and M. Pütz (eds), (2014), *Multilingual Cognition and Language Use: Processing and Typological Perspectives*, Amsterdam, Philadelphia: John Benjamins, 73–96.

Glucksberg, Sam (2003), The psycholinguistics of metaphor, *Trends in Cognitive Sciences*. 7(2), 92–96.

Kolahdouz, Mandala (2015), A Cognitive Approach to Translating Color Metaphors in English and Turkish Proverbs. Conference Paper [WWW] https://www.researchgate.net/publication/284181074 (accessed 2 April 2018)

Kövecses, Zoltán (2003), Language, figurative thought, and cross-cultural comparison, *Metaphor and Symbol* 18(4), 311–320.

Kövecses, Zoltán (2005), *Metaphor in Culture: Universality and Variation*, Cambridge: Cambridge University Press.

Li, Zhu (2011), A study on cognitive models in colour term translation between English and Chinese, *International Forum of Teaching and Studies* 7(1), 49–59.

Newmark, Peter (1981), *Approaches to Translation*. Oxford: Pergamon Press.

Olivera, Pedro A. Fuertes and Marisol Velasco Sacristán (2001), The translatability of metaphor in LSP: Application of a decision-making model, *Revista Alicantina de Estudios Ingleses* 14, 73–91.

Philip, Gillian (2003), *Collocation and Connotation: A Corpus-Based Investigation of Colour Words in English and Italian*, PhD Thesis. The University of Birmingham.

Sandford, Jodi L. (2016), Cognitive entrenchment of color categories and implicit attitudes in English, in G. Paulsen, M. Uusküla, and J. Brindle (eds), (2016), *Color Language and Color Categorization*, Newcastle upon Tyne: Cambridge Scholars Publishing, 40–61.

Schäffner, Christina (2004), Metaphor and translation: Some implications of a cognitive approach, *Journal of Pragmatics* 36(7), 1253–1269.

Van den Broeck, Raymond (1981), The limits of translatability exemplified by metaphor translation, *Poetics Today* 2(4), 73–87.

Wyler, Siegfried (1992), *Colour and Language: Colour Terms in English*, Tübingen: Gunter Narr Verlag.

Semantic extensions of the English *at*-construction

Pavlina Kalm
(University of New Mexico)

Keywords: conative, *at*-construction, semantics, semantic extensions, diachronic analysis

Schedule: We 16.30 Room 14

The semantics of the English *at*-construction (*She cut at the bread*) has received much attention in the literature. Traditionally, studies have focused on the meaning of the *at*-construction as a syntactic variant of the conative alternation with verbs of motion and contact, e.g. *cut*, *hit*, or *stab* (Dixon 1991, Dowty 1991). However, more recent accounts have attempted to explain the semantics of the *at*-construction by looking at a wider range of verbs that occur in it (Goldberg 1995, Van der Leek 1996, Broccias 2003). Goldberg (1995:63-64) argues that verbs of active perception (*look, aim*) are central to the semantics of the *at*-construction and that the semantics of the *at*-construction is 'x directs action at y'. Her constructional account is supported by Perek and Lemmens' (2010) collexeme analysis which shows that verbs with highest collostructional strengths in the *at*-construction denote events of active visual perception (*look, glance*).

This paper presents a *semantic map* (Haspelmath 2003, Croft 2003) of the *at*-construction and argues that the distribution of the construction across different verb classes in modern English (ModE) results from a network of semantic extensions that originate in *active perception verbs* (Viberg 1983). This study uses historical data from Old (OE) and Middle English (ME) to provide support for the semantic extensions that led to the current distribution of the *at*-construction in ModE.

Figure 1 shows the semantic map for the ModE *at*-construction, the OE/ME *on*-construction and OE genitive construction. This paper argues that the OE/ME *on*-construction and genitive construction were predecessors of the ModE *at*-construction. The *on*-construction was used with verbs of visual perception (e.g. *look*) in OE (Flom 1930, Lundskær-Nielsen 1993) and parallel to the modern *at*-construction, it spread to a range of verb classes, including verbs of motion and contact (Acevedo 2009, Iglesias-Rábade 2003).

Evidence from OE shows that verbs of contact (e.g. *push, pull*) and verbs of ingestion historically used a distinct genitive construction: Sbj V Obj-GEN, which alternated with the prototypical transitive construction: Sbj V Obj-ACC (Klein 1998, Díaz Vera 2001, Koike 2004). The GEN-ACC alternation was also attested with verbs of perceptual modalities that entail contact: touch and taste. I argue that verbs entailing contact in the ModE *at*-construction originated as extensions of perceptual touching and tasting. In ModE, neither *touch* nor *taste* occur in the *at*-construction. This is because during ME,

the genitive case was lost. Consequently, the GEN-ACC alternation became obsolete with many verbs such as taste and touch. In the process of the genitive case marking disappearing, some verbs of contact and ingestion replaced the genitive construction with a newly emerging *at*-construction (which succeeded the OE *on*-construction).

This study contributes to existing semantic accounts of the *at*-construction by using historical data to explain the distribution of the construction in ModE.

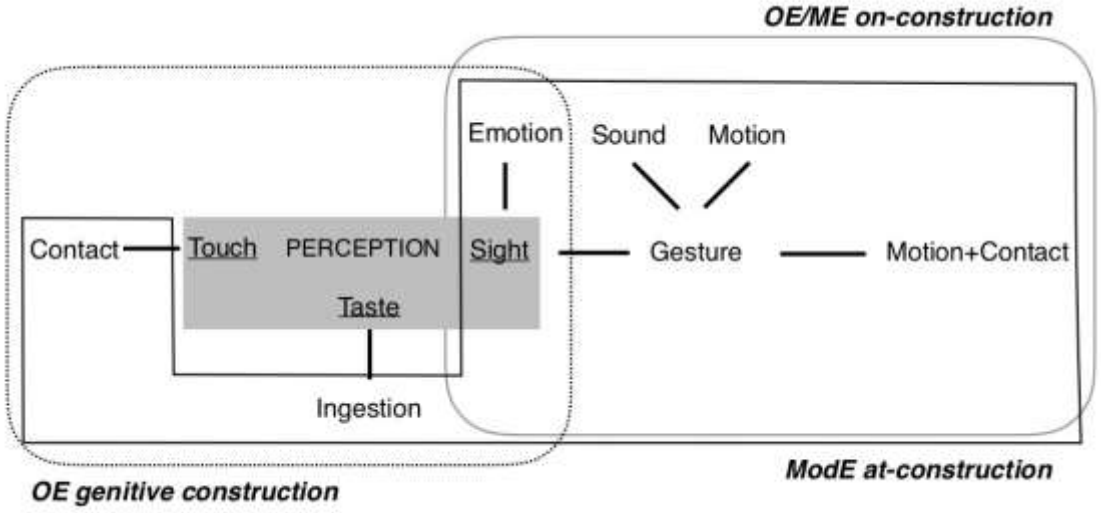


Figure 1. Semantic map

References

Acevedo, Eulalia Sosa (2009), Lexical classes and the conative construction in Old English, *Studia Anglica Posnaniensia* 45(1), 69–90.

Broccias, Cristiano (2003), *English change network: forcing changes into schemas*, Berlin: Walter de Gruyter.

Croft, William (2003), *Typology and universals*, Cambridge University Press.

Díaz Vera, Javier E (2001), Reconstructing the onomasiological structure of Old English verbs: the case of touching, tasting, and smelling, *Onomasiology Online* 2, 1–15.

Dixon, R.M.W. (1991), *A new approach to English grammar: on semantic principles*, Oxford University Press, Oxford.

Dowty, David (1991), Proto-roles and argument selection, *Language* 67(3), 547–619.

Flom, George T. (1930), *Introductory Old English grammar and reader*, Boston: D.C. Health and Co.

Goldberg, Adele E. (1995), *Constructions: A construction grammar approach to argument structure*, Chicago: University of Chicago Pres.

Haspelmath, Martin (2003), The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison, In Michael Tomasello (ed.), *The new psychology of language*, Mahwah, NJ: Erlbaum 2nd edn, 211–242.

Iglesias-Rábade, Luis (2003), A semantic study of *on*-phrases in Middle English. *Studia Neophilologica*, 75, 104-118.

Klein, Thomas Peter (1998), *The 'coarser' senses in Old English: a study of the Old English verbs of tasting, smelling, touching, and perceiving*, University of Toronto dissertation.

Koike, Takeshi (2004), *The analysis of the genitive case in Old English within a Cognitive Grammar framework*, University of Edinburgh dissertation.

Lundskær-Nielsen, T. (1993), *Prepositions in Old and Middle English*. Odense, Odense University Press.

Perek, Florent & Maarten Lemmens (2010), Getting at the meaning of the English at-construction: the case of a constructional split, http://cognitextes.revues.org/331

Van der Leek, Frederike (1996), The English conative construction: a compositional account, *Chicago Linguistic Society* 32(1), 363–378.

Viberg, Ake (1983), The verbs of perception: a typological study, *Linguistics* 21(1), 123–162.

ꕥ

The dative of attitude (*dativus ethicus*) as pragmatic marker in Latvian

Andra Kalnača & Ilze Lokmane
(University of Latvia, Riga)

Keywords: dative, personal pronoun, attitude, particle, pragmatics

Schedule: We 11.30 Room 11

The dative of attitude (term used in Haddad 2018, also *the ethical dative*, e.g. Berg-Olsen 2005) is an optional dative pronominal clitic that functions as an interpersonal pragmatic marker. It has an attitudinal and evaluative nature (Arsenijević 2013), i.e., it expresses affectedness of the speaker in a speech-act situation and reflects the speaker's opinion about entities and certain states of affairs. In terms of register, the dative of attitude normally occur in informal interaction and share similar structural, semantic and pragmatic characteristics in several languages (among others, Berg-Olsen 2005, De Knop & Mollica 2017, Haddad 2016, 2018). Such pronominal dative clitics or weak pronouns have often been likened to modal particles owing to the same syntactic, semantic and pragmatic properties. Another distinctive feature of the dative of attitude is its use in structures where the verb valency requires no dative object (De Knop & Mollica 2017).

In Latvian, the dative of attitude appears mostly in the first and second person, usually in the singular, therefore it refers either to the speaker or the hearer.

The 1st person or speaker-oriented attitude dative is mostly used with imperative prohibiting or even threatening to assert the speaker's authority over the addressee (1) but it is found also in utterances expressing solidarity or empathy towards someone (2):

(1)	*Tu*	***man***	*te*	*jokus*	*netaisi!*
	you.NOM.SG	**I.DAT**	here	fun.ACC.PL	not_make.IMP.2SG
	'Do not make fun!' (www.korpuss.lv)				

(2)	*Kur*	*tu*	***man***	*tāda*	*esi*	*gadījusies…*
	where	you.NOM.SG	**I.DAT**	such.NOM.SG	be.AUX.PRS2SG	happen.PTCP.PST.ACT
	'How come you happened here...' (www.korpuss.lv)					

The 2nd person or hearer-oriented attitude dative is used in threats (3) or to express disagreement with the opinion of the hearer (4):

(3)	*Es*	***tev***	*gan*	*gaidīšu!*
	I.NOM	**you.DAT.SG**	PRT	wait.FUT.1SG
	Kur	*tas*	*ir*	*redzēts!*
	where	it.NOM.SG	be.AUX.PRS.3	see.PTCP.PST.PASS
	'I am not going to wait! That is unheard of! (E. Vulfs)			

(4) *Kur* ***tev*** *tik* *vēlā* *rudenī*
where **you.DAT.SG** so late.LOC.SG autumn.LOC.SG
pērkons *rasies!*
thunderstorm.NOM.SG turn_on.FUT.3
'How on Earth should there be thunderstorm so late in the autumn!' (J. Janševskis)

As the different semantic roles are interconnected, the referent of the dative of attitude may be additionally construed as the possessor, the beneficiary or the experiencer. Therefore, the attitudinal and evaluative functions of the dative are often strengthened by particles (5), also (3), (6):

(5) *Gan* *tu* ***man*** *vēl* *pasvilposi!*
PTCL you.NOM.SG **I.DAT** PTCL whistle.FUT.2SG
'You will whistle for me big time!' (www.korpuss.lv)

The 3rd person attitude dative is somewhat problematic, because it is mostly used for threatening in constructions with the finite verb *rādīt* 'to show' and can therefore be interpreted as the dative of argument:

(6) *Gan* *es* ***viņam*** *rādīšu!*
PTCL I.NOM **he.DAT** show.FUT.1SG
'I will teach him a lesson alright!' (www.letonika.lv)

The presentation aims at analyzing the various types of the dative of attitude in conjunction with the syntactic constructions and the communicative types of utterances they are embedded in.

The examples are extracted from *Līdzsvarotais mūsdienu latviešu valodas tekstu korpuss (The Balanced Corpus of the Modern Latvian)* (available at http://www.korpuss.lv/id/LVK2018), fiction, mass media texts etc.

References

Arsenijević, Boban (2013), Evaluative reflexions: evaluative dative reflexive in Southeast Serbo-Croatian, in Fernández, Beatriz & Ricardo Etxepare (eds), *Variation in Datives. A Microcomparative Perspective*, Oxford: Oxford University Press, 1–21.

Berg-Olsen, Sturla (2005), *The Latvian dative and genitive: a cognitive grammar account*, Oslo: University of Oslo.

Haddad, Youssef A. (2016), Possessively Construed Attitude Dative Constructions in Lebanese Arabic, *Brill's Journal of Afroasiatic Languages and Linguistics* 8, 37–75.

Haddad, Youssef A. (2018), *The sociopragmatics of attitude datives in Levantine Arabic*, Edinburgh: Edinburgh University Press.

De Knop, Sabine & Fabio Mollica (2017), The family of German dative constructions, in Mendoza, Ibáñez de, Ruiz Francisco José et. al. (eds), *Constructing families of constructions: Analytical perspectives and theoretical challenges*, Amsterdam, Philadelphia: John Benjamins Publishing Company, 205–239.

Verbs of throwing and categorization of completive events: Evidence from Hill Mari

Egor Kashkin
(V. V. Vinogradov Russian Language Institute of RAS, Moscow)

Keywords: Uralic languages, Hill Mari, semantics, grammaticalization, verbs of throwing.

Schedule: Thu 14.30 Room 14

The research has been supported by RFBR, grant № 19-012-00627.

This paper focuses on the grammaticalization undergone by two verbs of throwing (*šuaš* and *käškäš*) in Hill Mari (< Uralic). I will discuss what semantic and collocational restrictions they have when used as completive markers and how these properties can be predicted from their basic semantics. The data were collected in fieldwork in the village of Kuznetsovo and its surroundings (Mari El Republic, Russia) by elicitation and by using the corpus of transcribed oral narratives (ca. 63.500 tokens).

The lexical typology of throwing verbs is discussed in Ivtushok (2015, 2016), but no information on their grammaticalization is provided. Maisak (2005) leaves this domain outside his research on grammaticalization patterns adopted by motion verbs. Heine and Kuteva (2004: 297–298) formulate a general pattern of their development into perfect or completive markers. Some sources contain short pieces of information on individual languages or language families, e.g. Baranova (2013: 19–21) on Kalmyk, Grashchenkov (2015: 73–74) on Turkic. Still, the patterns of such shifts remain underdescribed; neither is it studied in detail what semantic components are preserved by verbs of throwing after grammaticalization (see Sweetser (1988) and Eckardt (2006) for the theoretical discussion).

The verb *šuaš* describes one action of throwing in its primary meaning, while the verb *käškäš* asserts that several actions of throwing took place. Both predicates can be used as light verbs in complex predicates, i.e. morphosyntactically bound combinations of a lexical verb in the form of a converb and a finite light verb, see (1)–(2). Such constructions were discussed by Serebrennikov (1960: 190–199), Pengitov et al. (1961: 202–216), Bradley (2010), among others, but without much detail about semantic restrictions on their slots.

(1) *mən'* *toštə̑* *pört-əm* ***pə̑də̑rt-en*** ***šu-en-äm.***
I old house-ACC break-CVB throw-PRET-1SG
'I destroyed the old house'.

(2) *tən'* *cilä* ***port'-en*** ***käškə-š-əc!***
you all spoil-CVB throw-AOR-2SG
'You have spoiled everything!'

According to my data, *šuaš* and *käškäš* usually collocate with verbs referring to destruction or removal of the patient, e.g. *kə̑škedäš* 'to tear', *puštaš* 'to kill', *jə̑lataš* 'to burn sth. down', *jästäräš* 'to pour out, to empty', *pačkaltaš* 'to shake out', see also (1)–(3). This semantic shift follows from the categorization of throwing events as situations in which an object is removed from some location. In some cases the two verbs under consideration can collocate with predicates introducing an object less close to a prototypical patient (in terms of Dowty (1991), Malchukov (2005)), e.g. some verbs of physical impact (*mə̑škaš* 'to wash', *ut'užaš* 'to iron') or of possessor change (*və̑žalaš* 'to sell').

Having similar collocational properties, constructions with *šuaš* and *käškäš* differ in quantificational features of the object and in compatibility with uncontrollable situations. The former parameter is illustrated in (3), where *šuaš* co-occurs with either a singular or a plural object, while

käškäš requires a plural object. This can be predicted from their difference in the primary uses, namely by the fact that *käškäš* describes a pluractional event.

(3) a. *mən'* *pükšem-əm* ***šüt-en*** ***šu-š-ôm*** ***/*šüt-en*** ***käškä-š-əm.***
I knot-ACC untie-CVB throw-AOR-1SG untie-CVB throw-AOR-1SG
'I untied the knot'.

b. *mən'* *pükšem-vlä-m* ***šüt-en*** ***šu-š-ôm*** **/** ***šüt-en*** ***käškä-š-əm.***
I knot-PL-ACC untie-CVB throw-AOR-1SG untie-CVB throw-AOR-1SG
'I untied the knots'.

I will also discuss the place of throwing verbs in the system of light verbs in Hill Mari, in particular comparing their semantics and compatibility with those of the telicizers *šəndäš* 'to seat' and *šoaš* 'to reach' used in similar morphosyntactic patterns.

Abbreviations

1, 2 — 1st , 2nd person; ACC — accusative; AOR — aorist; CVB — converb; PL — plural; PRET — preterite; SG — singular.

References

Baranova, Vlada (2013), Grammaticalization and semantics of complex predicates in Kalmyk, *Journal de la Société Finno-Ougrienne* 94, 9–35.

Bradley, Jeremy (2010), *Mari converb constructions: interpretation and translation*, MA Thesis, Vienna: University of Vienna.

Dowty, David (1991), *Thematic proto-roles and argument selection*, Language 67 (3), 547–619.

Eckardt, Regine (2006), *Meaning change in grammaticalization. An enquiry into semantic analysis*, Oxford: Oxford University Press.

Grashchenkov, Pavel (2015), *T'urkskie konverby i serializacija: sintaksis, semantika, grammatikalizacija* [Turkic converbs and serialization: syntax, semantics, grammaticalization], Moscow: Jazyki slav'anskoj kul'tury.

Heine, Bernd, and Kuteva, Tania (2004), *World lexicon of grammaticalization*, Cambridge: Cambridge University Press.

Ivtushok, Elizaveta (2015), Glagoly tipa «brosat'» v leksiko-tipologicheskom osveshchenii [Verbs of throwing in a lexical-typological view], in E. Lyutikova, A. Tsimmerling, and M. Konoshenko (eds), (2015), *Typology of morphosyntactic parameters. Proceedings of the international conference «Typology of morphosyntactic parameters 2015»*, Moscow: Moscow State Pedagogical University, 114–122.

Ivtushok, Elizaveta (2016), K tipologii glagolov so znacheniem «brosat'» [On the typology of verbs of throwing], in A. Kretov (ed), (2016), *Problems of lexical-semantical typology*, vol. 3. Voronezh: Voronezh State University, 72–77.

Maisak, Timur (2005), *Tipologiya grammatikalizacii konstrukcij s glagolami dvizheniya i glagolami pozicii* [Typology of grammaticalization of the constructions with motion verbs and posture verbs], Moscow: Jazyki slav'anskih kul'tur.

Malchukov, Andrej (2005), Case pattern splits, verb types and construction competition, in M. Amberber, H. de Hoop (eds), (2005), *Competition and Variation in Natural Languages*, London, New York: Elsevier, 73–117.

Pengitov, Nikolaj, Galkin, Ivan, and Isanbaev, Nikolaj (eds), (1961), *Sovremennyj marijskij yazyk. Morfologiya* [The Modern Mari language. Morphology], Yoshkar-Ola: Marijskoje knizhnoje izdatel'stvo.

Serebrennikov, Boris (1960), *Kategorii vremeni i vida v finno-ugorskih jazykah permskoj i volzhskoj grupp* [Categories of tense and aspect in the Finno-Ugric languages of Permic and Volgaic groups], Moscow: Nauka.

Sweetser, Eve (1988), Grammaticalization and semantic bleaching, in A. Shelley, A. Jaisser, and H. Singmaster (eds), (1988), *Berkeley Linguistic Society: Proceedings of the Fourteenth annual meeting. General session and parasession on grammaticalization*, Berkeley: Berkeley Linguistics Society, 389–405.

Investigating lexical changes in the Russian language of Estonia

Éva Katona
(ELTE Eötvös Loránd Tudományegyetem, Budapest)

Keywords: Russian, Estonian, loanword, language contact, code copying

Schedule: Sat 12.30 Room 8

In a language contact situation the "most common specific type of influence is the borrowing of words" (Thomason 2001: 10). In accordance to this, in the Russian language of Estonia lexical changes are also common. In the past thirty years, after Estonia regained its independence and Estonian became the only official language of the state, the new state language became a part of the everyday lives of those with Russian as their first language as well. This everyday contact situation impacts the local variety of Russian. One of the outcomes is change in the lexicon, which is the most visible in the domains of education, government administration and culture. A good number of lexical changes have already been described, see further: Verschik (2005) and Адамсон (2009) about government administration, and Zabrodskaja (2007) for words regarding university life. Especially since the change is still happening, the list is far from complete, and there are many lexical elements awaiting documentation, and several of the documented words have not yet been tested from the point of usage or preference.

In my study I test whether the Russian Estonian speakers actually prefer those lexical elements in "fill the gap" tests that are believed to be characteristic only for the speakers in Estonia. The words in question are retrieved from my own questionnaires, interviews, some from the existing literature (Кюльмоя, 2009) and a comment section of an online magazine about this topic.

From my own corpus I chose words that are not included in Russian monolingual dictionaries, but also words people perceived as results of language contact were chosen. For my lexical test I paired up each "Estonian Russian" variant with its corresponding monolingual variant, and created sentences where the context would elicit one or the other. When both variants start with the same letter, I provided the first letter to exclude any other possibilities, otherwise the alternatives were not given. I asked my informants (N=30) to fill the gap in the sentences with the missing words. I then compared the tests with the tests of my control group who were all born and living in Russia (N=22).

Several of the word choices showed a remarkable difference between the two groups. The Estonian Russians chose hybrid words that were formed on the basis of an Estonian word, for example choosing *inventura* (12) instead of the monolingual *inventarizacija* (18) (based on the Estonian word *inventuur* for *stock-taking*), while each member of the control group wrote *inventarizacija.* There was also a remarkable presence of the Estonian borrowing *rebarbar* (4) along with *reven'* (22) and other (4) in the Estonian group, as opposed to *reven'* (13), *rebarbar* (0) and other (9) in the monolingual group (the plant is less known in Russia).

My goal is to show several of these diverging lexical elements and to argue against the terminology of "loanword" and "borrowing", in favor of the code-copying model of Johanson (2002), and to explore ways to investigate this phenomenon in a wider-scale study.

References

Johanson, Lars (2002), Contact-induced change in a code-copying framework, in Mari C. Jones & Edit Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (Contribution to the Sociology of Language 86), 285–313. Berlin: Mouton de Gruyter.

Thomason, Sarah Grey (2001), *Language contact: an introduction*. Washington, D.C: Georgetown Univ. Press.

Verschik, Anna (2005), Russian–Estonian language contacts, linguistic creativity, and convergence: New rules in the making. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication* 24(4), 413–429. doi:10.1515/mult.2005.24.4.413.

Zabrodskaja, Anastassia (2007), Code-switching and contact-incuded language change in Estonia's Russian. *Publications of the Department of the Estonian Philology of Tallinn University* 9, 13–64.

Адамсон, Инна Владимировна (2009), Динамические процессы в лексике таллиннских газет на русском языке (на примере двуязычных изданий), in Ирина Павловна Кюльмоя (ed.), *Humaniora: Lingua Russica. Активные процессы в русском языке метрополии и диаспоры. Труды по русской и славянской филологии. Лингвистика XII*, 27–42. Tartu: Tartu Ülikooli Kirjastus / Издательство Тартуского университета.

Кюльмоя, Ирина Павловна (2009), Речь русской диаспоры Эстонии: тенденции развития, in Ирина Павловна Кюльмоя (ed.), *Humaniora: Lingua Russica. Активные процессы в русском языке метрополии и диаспоры. Труды по русской и славянской филологии. Лингвистика XII*, 11–26. Tartu: Tartu Ülikooli Kirjastus / Издательство Тартуского университета.

Online sources

http://journalist.delfi.ee/news/news/obsudim-kakie-estonizmy-vas-razdrazhayut?id=36016109&com=1&reg=0&no=0&s=1

The argumentative functions of maximum expressions

Nicole Katzir
(Tel Aviv University)

Keywords: Argumentation in language; Maximum expressions; Hebrew

Schedule: We 17.00 Room 11

While minimum expressions (e.g., *at least)* have been thoroughly discussed (Kay 1992, Geurts & Nouwen 2007 inter alia), maximum expressions (e.g., *at most*) received much less attention. I focus on two Modern Hebrew expressions: *le-xol ha-yoter* 'at most', and *maksimum* 'maximally'/'at most', based on corpus data (HeTenTen). Over 100 instances of each expression have been extracted and analyzed.

The first finding is that in addition to purely quantificational constructions (1), which constitute 79.3% of all instances, maximum expressions give rise to two additional readings which parallel those

identified by Kay (1992) for English *at least*. (2) Shows an Rhetorical construction (9%) and (3) an Evaluative construction (11.1%):

1. ba-xeder ha-ze nimca'im **le-xol ha-yoter/maksimum** 100 anaʃim.
   "There are **at most** 100 people in this room."

2. ma lanu, yelidei ha-arec, ve-le-alilot dam? **le-xol ha-yoter** lamadnu alehen be-vet ha-sefer, ʃinanu et ha-pratim ve-katavnu alehen tʃuvot ba-bxinot.
   "What do we, native-born Israelis, have to do with blood libels? **At most**, we learned about them in school in history class, memorized their details and took exams about them."

(HeTenTen)

3. hiʃtaxreru me-ha-tfisa lefiha kir xayav lihiyot lavan ve-haxnisu kcat ceva la-xa'im. **maksimum**, tamid efʃar licbo'a ʃuv axar kax.
   "Let go of the thought that a wall must be white, and introduce some color into your life. Maximally (=**Worst case scenario)**, you can always paint it again later."

(HeTenTen)

I first note that the two expressions are associated with different discourse profiles, i.e., non-obligatory but frequently recurring discourse conditions that hold when some linguistic form is used for a certain discourse function (Ariel 2008). While *le-xol ha-yoter* is associated more with the Rhetorical construction, *maksimum* shows a stronger preference for the Evaluative construction.

Next, I argue for the centrality of argumentation (Anscombre & Ducrot, e.g. 1976, 1983) in the semantics of maximum expressions. I propose that such expressions are argumentative operators, where speakers construe objective states of affairs as subjectively insufficient bases for certain inferences. Specifically, my claim is that maximum expressions introduce an argument which: (i) counters the speaker's own point; (ii) is the strongest counter-argument; yet, (iii) it is insufficient to invalidate the speaker's point.

For example, in (2), the main point of the writer is that young Israelis don't remember the blood libels Jews were accused of throughout history. *le-xol ha-yoter* 'at most' introduces the argument that native-born Israelis learned about blood libels in schools. Since this argument is introduced by an Rhetorical maximum construction, it is understood as the strongest, yet an insufficient argument for the opposite conclusion than the one that the writer is advocating. Since the strongest argument against the speaker's point is insufficient, the original argument is not invalidated.

I then argue that these argumentative functions are not strictly predictable from the truth-conditional content of the expressions, which is why they constitute Constructions (Goldberg 1995).

Finally, a comparison between minimum and maximum expressions shows that despite their different truth-conditional semantics and seemingly opposite argumentative directions, both ultimately introduce a *non-optimal* argument *in favor* of the speaker's point.

References

Anscombre, Jean-Claude, and Oswald Ducrot. 1976. L'argumentation dans la langue. *Langages* 42:5-27

Anscombre, Jean-Claude, and Oswald Ducrot. 1983. *L'argumentation dans la langue*. Brussels: Mardaga.

Ariel, Mira. 2008. *Pragmatics and grammar*. Cambridge: Cambridge University Press.

Geurts, Bart, and Rick Nouwen. 2005. "At least" et al.: the semantics of scalar modifiers. *Language* 83: 533–559.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Kay, Paul. 1992. At least. In Adrienne Lehrer and Eva Feder Kittay, eds., *Frames, fields, and contrasts: new essays in semantic and lexical organization*. Hillsdale, NJ: Erlbaum.

Merging of prepositions and definite articles can be explained by frequency of use

Dan Ke
(University of Leipzig)

Keywords: frequency effect, language asymmetry, preposition, definite article, corpus linguistics

Schedule: Thu 11.30 Room 7

This corpus-based study demonstrates the frequency effect in the asymmetric distribution of the merging in prepositions and definite article pairs.

In many Indo-European languages, prepositions (P) and definite articles (D) can be merged together. But not every P-D pair can be reduced. Stolz (1990) presents the distribution in the following paradigm (The paradigm is slightly shortened from Stolz (1990: 337–338) with the Italian example added by the author):

(1) German:	an + dem -> am	an + der -> ?	an + das -> ans	
(2) French:	à + le -> au [o]	à + la -> ?	à + les -> aux [o(z)]	
(3) Spanish:	a + el -> al	a + la -> ?	a + los -> ?	a + las -> ?
(4) Portuguese:	a + o -> ao	a + a -> á	a + os -> aos	a + as -> ás
(5) Italian:	in + il -> nel	in + lo -> nello	in + l' -> nell'	in + le -> nelle

The merging phenomenon varies in languages. In German, both reduced and full forms are used, whereas in Romance languages, usually only one form is correct. Furthermore, using different forms can lead to semantic differences. Schwarz (2009) proposes that reduced articles in German are "weak articles" and that full forms are "strong articles". Weak articles indicate situational uniqueness and strong articles have a discourse anaphoric function (Schwarz 2009: 23–44).

Several reasons for the merging are discussed in the literature. Eisenberg (2006: 199) proposes that in German the phonological lightweight of the articles leads to the reduction when they are pronounced quickly. On the other hand, Cabredo Hofherr (2012) suggests that the P+D-contraction is not phonologically determined and could only be explained using language specific syntactical analyses (Cabredo Hofherr 2012: 128). Stolz (1990) and Nübling (1992) agree with the phonological requirement but note that frequency also leads to the asymmetric distribution (Stolz 1990: 337 and Nübling 1992: 161).

More generally, Bybee (2002) suggests that sequences of words that occur together very often are stored as single units, so reductive sound changes affect high frequency words first and at a faster rate. Moreover, Haspelmath (2008) proposes that all universal morphosyntactic asymmetries can be explained on the basis of frequency asymmetries, namely, more frequent patterns are coded with less material (Haspelmath 2008: 185).

This paper provides the first quantitative study on P-D merging and reduction in Indo-European languages. On the basis of corpus data, P-D pairs are examined in German, French, Spanish,

Portuguese and Italian. Preliminary results show that the factor frequency is highly significant. All reduced forms are highly frequent. In German for example, *am* (reduced from "an dem", 28199794 appearrences; data accessed on 17. June, 2017 from COSMAS I/II (*Corpus Search, Management and Analysis System*)) is the most frequent form of the preposition *an* with all definite articles. Likewise, *im*, *beim*, *vom*, and *zum* are the most frequent ones in each case. I argue that besides phonological factors, the frequency effect also serves as a general explanation for P-D merging regardless of language variability and semantic differences.

References

Bybee, Joan (2002), Phonological evidence for exemplar storage of multiword sequences. In: *Studies in Second Language Acquisition*, Vol. 24 (2), 215-221.

Cabredo Hofherr, Patricia (2012), Preposition-determiner amalgams in German and French at the syntax-morphology interface. In: P. Ackema and R. Alcorn et al. (ed.), (2012), *Comparative Germanic Syntax.* Amsterdam / Philadelphia: John Benjamins, 99-131.

Eisenberg, Peter (2006), *Der Satz. Grundris der deutschen Grammatik*, Band 2. 3. Auflage. Stuttgart / Weimar: Verlag J. B. Metzler.

Haspelmath, Martin (2008), Creating econominal morphosyntactic patterns in language change. In: J. Good (ed.), (2008), *Linguistic universals and language change*. Oxford: Oxford University Press, 185-214.

Nübling, Damaris (1992), *Klitika im Deutschen: Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Gunter Narr Verlag.

Schwarz, Florian (2009), *Two types of definites in natural language*. PhD thesis of University of Massachusetts Amherst.

Stolz, Thomas (1990), Flexion und Adpositionen, flektierte Adpositionen, adpositionelle Flexion. In: *STUF - Language Typology and Universals*, Vol. 43 (3): 334-354.

Other resource:

COSMAS I/II (*Corpus Search, Management and Analysis System*), http://www.ids-mannheim.de/cosmas2/, 1991-2016 Institut für Deutsche Sprache, Mannheim.

Repetition and reduplication in German: Balancing the thin line between rule obedience and extravagance

Gerrit Kentner
(Max Planck Institute for Empirical Aesthetics; Goethe-University Frankfurt)

Keywords: Reduplication, German, Phonology, Morphology, Prosodic Morphology

Schedule: Thu 16.30 Room 9

Repetition and reduplication are prime examples of morphological extravagance in the sense that repetitive forms display an abundance of identical phonological material. Moreover, in German, repetition and reduplication are extravagant in a different sense, namely in that they typically feature in specific registers (e.g. poetic language use, expressive speech). On the other hand, repetitive phenomena across the phonological hierarchy show an exceptionally high degree of rule obedience in terms of their (morpho)phonological form. As will be shown by way of the following examples, this

rule obedience puts strict constraints on the productivity of repetition and reduplication, thereby limiting their extravagance.
#1 Rhyme and ablaut reduplication in German (1) are productive processes of word-formation, typically used in expressive contexts, e.g. as hypocoristics (Kentner 2017). One especially extravagant feature is the variable ordering of base and reduplicant in ablaut reduplication (suffixing in e.g. *Misch$_{base}$-masch$_{red}$*, prefixing in *Quitsch$_{red}$-quatsch$_{base}$*). However, in contrast to more common types of word formation, reduplication necessitates strict adherence to a prosodic template (only monosyllables or trochees undergo reduplication) with minimal segmental non-identity (a common OCP-effect) between base and reduplicant (rhyme or ablaut).

(1) a. Hinkepinke (< hinken), Hasepase (< Hase), Popelmopel (<Popel)
hopscotch (< hobble), *sweetheart (<bunny),* *nose-picker (<bogey)*
b. Wirrwarr (<wirr), Mischmasch (<mischen), Quitschquatsch (<Quatsch)
jumble (<woozy), *hotchpotch (<mix),* *fiddlesticks (<nonsense)*

#2 Truncation+doubling in nickname formation (2a) is extravagant in that the modification of the base is produced by shortening (truncation to a light syllable thus violating morphological integrity) and subsequent doubling (producing a trochee) rather than more common affixation. As the ungrammatical examples in (2b) show, this process is strictly limited to the most natural or unmarked phonological environment and thus governed by strong (universal) constraints: Words with complex, glottal, or vocalic onsets, as well as non-cardinal vowels (e.g. front rounded vowels or diphthongs) cannot be generated in this way.

(2) a. Jojo (<Johannes), Bibi (<Birgit), Nana (<Nadine), Bobo (<Boris), Kiki (<Kirsten)
b. *Krikri (<Kristian), *Haha (<Hans), *[ʔ]i[ʔ]i (<Ina), *L[y]l[y] (<Lydia), *Maimai (<Maike)

In sum, on the basis of these examples, it is suggested that (morphological) extravagance in a given dimension (here: deviance from the normal concatenative morphology) is accompanied by stricter rule obedience in other dimensions (here: unmarked phonological structure). In spite of their marginal status, the cases presented are not deemed extra-grammatical; to the contrary, these cases demonstrate the force of the (phonological) rule system in the face of otherwise deviant constructions.

Reference

Kentner, Gerrit (2017), On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft*, 36(2), 233-277.

The Origin of Non-Canonical Case Marking of Subjects in Proto-Indo-European: Accusative, Ergative, or Semantic Alignment

Peter Alexander Kerkhof, Roland Pooth, Leonid Kulikov, Laura Bruno and Jóhanna Barðdal
Leiden University & Ghent University

Keywords: Proto-Indo-European, alignment systems, argument structure, non-nominative subjects, converse lability

Schedule: Fri 15.00 Room 14

For a long time, one of the most bewildering conundrums of Indo-European linguistics has been the issue of how to reconstruct the alignment system of this ancient language state, given the lack of distinction between S and O marking in the Proto-Indo-European neuters nouns and the problem of the Hittite ergative.

A second problem within Indo-European linguistics relates to the origin of non-canonical case marking of subjects, which has been an issue of scholarly interest since Pedersen's (1907) first attempt to derive it from ergative alignment. Barðdal & Eythórsson (2009) laid out six possible hypotheses of the emergence of non-canonically case-marked subjects in the Indo-European languages, presenting evidence against five of these hypotheses. They argue that the existence of predicates selecting for non-canonically case-marked subjects in the early Indo-European languages speaks for the assumption that alignment in Proto-Indo-European was semantically based (cf. Donohue & Wichmann 2008), corresponding to the Fluid-S type. Although such a claim is sound and in consonance with the attested case frames of the Indo-European languages, Barðdal & Eythórsson's (2009) analysis was only based on syntactic consideration and not corroborated with morphological evidence, rooted in existing knowledge of the morphology of the early Indo-European languages and their proto-stages.

We lay out a morphological analysis of the form–function correlates that have been reconstructed on the basis of the Comparative Method. We outline the morphosyntactic details of the reconstructed semantic alignment stage of Proto-Indo-European where an antipassive-like construction played a key role for the development from semantic alignment to the attested accusative system found in the Indo-European daughter languages today (Pooth 2014). This antipassive-like construction was reanalyzed as a transitive construction and the earlier agentive *-*s* marker was generalized into a subject marker, irrespective of the semantics of the subject referent, yielding an accusative system. As a part of this general process we have identified the Early PIE proto-constructions that have developed into the attested accusative and dative subject constructions, respectively. The first one involves the older *-*m* allative-marking of non-neuters which also developed into the accusative object marker. Through the construction with the *-*m* or the zero-marking of subjects of one-participant clauses with proto-middle and proto-active marking on the verb, respectively, the accusative subject construction emerged. Out of experiencer constructions involving the old locative ending *-*i*, one subconstruction of the dative subject construction arose. Therefore, we argue that non-canonical case marking of subjects is a relic of the semantically-marked experiencer and undergoer role.

Our aim with the present article is to fill two needs with one deed and to offer a unified account of these two century-long bones of contention, presenting an attempt to elucidate how non-canonical case marking of subjects are in line within the most recent discussions on PIE alignment. We show how a Fluid-S or semantic alignment model for the Proto-Indo-European ancestor language may aptly explain the presence of archaic instances of non-canonical subject marking in the ancient IE languages (cf. Barðdal et al. 2013, Danesi, Johnson & Barðdal 2017), We thus offer a hypothesis of the emergence of non-canonically case-marked subjects, grounded in early Indo-European morphology.

Acknowledgements:
This research was supported with a generous research grant to Jóhanna Barðdal (PI) from the European Research Council (EVALISA, grant nr. 313461) and to Leonid Kulikov from the National Science Centre (NCN), Poland (grant nr. UMO-2015/19/P/HS2/02028).

References:

Barðdal, Jóhanna & Thórhallur Eythórsson. 2009. The Origin of the Oblique Subject Construction: An Indo-European Comparison. *Grammatical Change in Indo-European Languages*, ed. by Vit Bubenik, John Hewson & Sarah Rose, 179–193. Amsterdam: John Benjamins.

Barðdal, Jóhanna, Valgerður Bjarnadóttir, Serena Danesi, Tonya Kim Dewey, Thórhallur Eythórsson, Chiara Fedriani & Thomas Smitherman. 2013. The Story of 'Woe'. *Journal of Indo-European Studies* 41(3–4), 321–377.

Danesi, Serena, Cynthia A. Johnson & Jóhanna Barðdal. 2017. Between the Historical Languages and the Reconstructed Language: An Alternative Approach to the Gerundive + "Dative of Agent" Construction in Indo-European. *Indogermanische Forschungen* 122, 143–188.

Donohue, Mark & Søren Wichmann (eds.). 2008. *The Typology of Semantic Alignment.* Oxford: Oxford University Press.

Pedersen, Holger. 1907. Neues und nachträgliches. *Zeitschrift für Vergleichende Sprachforschung* 40, 129–257.

Pooth, Roland A. 2014. *Die Diathesen Aktiv vs. Medium und die Verbsemantik im Vedischen der Ṛgveda-Saṃhitā.* Leiden University Ph.D. Dissertation (Leiden University Repository, cf. https://openaccess.leidenuniv.nl).

ø

The role of the reflexives in valency: Evidence from Czech

Václava Kettnerová & Markéta Lopatková
(Charles University)

Keywords: reflexive derivational morpheme, reflexive verb lemma, changes in valency

Schedule: Sat 11.00 Room 12

The reflexives in Czech – as in other Slavic languages – serve various functions (Gast & Haas 2008, Siloni 2008, and Knjazev 2007). As a result, their analysis represents a complex task, often resulting in different theories in diverse conclusions. Here we analyze functions of the reflexives with respect to valency structure of Czech verbs, based on the Functional Generative Description (Panevová 1994) and data provided by the VALLEX lexicon (Lopatková et al. 2016).

The reflexives in Czech have short (clitic) and long forms (*se/sebe*, *si/sobě, sebou*). From the perspective of valency, the crucial distinction in their functions is whether they form analogous syntactic structures as personal pronouns (i), or not (ii-iv). In case (i), the reflexives fill a valency position; depending on their position in a sentence, they have either the clitic, or the long form. These reflexives serve as a marker of reflexivity and/or reciprocity. The reflexives of type (ii-iv) have only the clitic forms; they mark (ii) the reflexive form of a verb (being a part of its inflectional paradigm, the clitic *se* serves as deagentive marker), or (iii) the reflexive lemma of a verb as a part of its derivational paradigm. Moreover, (iv) the clitic reflexives are an obligatory part of verb lemmas with reflexive tantum verbs (*bát se* 'to be afraid') and an optional part with several other verbs (*myslet si* ← *myslet* 'to think'), with no overt syntactic and/or semantic functions.

Changes in valency structure of Czech verbs associated with reflexivity and reciprocity (i) and deagentive diathesis (ii) have been extensively discussed in (Daneš et al. 1987, Panevová 1999, Panevová & Mikulová 2007). However, those changes that are brought about by the reflexive as the derivational morpheme (iii) have not been systematically described yet though they are regular to some extent; see, e.g., a conversive mapping of participants onto valency positions of the verb in (a-b), leading to changes in its surface expression.

a.	*Petr-Ø*	*nakazil-Ø*	*Jan-u*	*chřipk-ou.*
	Peter-NOM.SG.M	infected-SG.M	Jane-ACC.SG.F	flu-INS.SG.F
	'Peter infected Jane with flu.'			

b.	*Jan-a*	*se*	*nakazil-a*	*od*	*Petr-a*	*chřipk-ou*.
	Jane-NOM.SG.F	REFL	infected-SG.F	from	Peter-GEN.SG.M	flu-INS.SG.F

'Jane was infected with flu from Peter.'

In VALLEX, the following functions of the reflexive derivational morpheme are attested:

(a) decausative (*naklonit se* 'to tilt'← *naklonit* 'to give a tilt'),
(b) inherently reciprocal (*pomáhat si* 'to help each other' ← *pomáhat* 'to help'),
(c) conversive (*půjčit si* 'to borrow' ← *půjčit* 'to lend'),
(d) intransitive (*dohodnout se* 'to arrange' ← *dohodnout* 'to arrange').

We show that the reflexives (a)-(d) give rise to both a specific shift in meaning of verbs and systematic changes in their valency structure. Further, we demonstrate that these changes are regular enough to be described by formal rules associating valency structure of reflexive verbs with their irreflexive counterparts.

Acknowledgments: The research reported here is supported by the Czech Science Foundation (GAČR), grant No. 18-03984S.

References

Daneš, František, Miroslav Grepl and Zdeněk Hlavsa (eds) (1987), *Mluvnice češtiny 3, Skladba*, Praha: Academia.

Gast, Volker and Florian Haas (2008), On reciprocal and reflexive uses of anaphors in German and other European languages, in E. König, and V. Gast (eds), (2008), *Reciprocals and Reflexives: Theoretical and Typological Explorations*, Berlin / New York: Mouton de Gruyter, 307-346.

Knjazev, Jurij P. (2007), Reciprocal constructions in Russian, in V. P. Nedjalkov (ed), (2007), *Reciprocal Constructions*, Amsterdam / Philadelphia: John Benjamins, 673-707.

Lopatková, Markéta, Václava Kettnerová, Eduard Bejček, Anna Vernerová and Zdeněk Žabokrtský (2016), *Valenční slovník českých sloves VALLEX*, Praha: Karolinum.

Panevová, Jarmila (1994), Valency Frames and the Meaning of the Sentence, in P. A. Luelsdorff (ed), (1994), *The Prague School of Structural and Functional Linguistics*, Amsterdam / Philadelphia: John Benjamins, 223-243.

Panevová, Jarmila (1999), Česká recipročnı́ zájmena a slovesná valence, *Slovo a slovesnost* 60(4), 269-175.

Panevová, Jarmila Marie Mikulová (2007), On reciprocity, *The Prague Bulletin of Mathematical Linguistics* 87, 27-40.

Siloni, Tal (2008), The Syntax of Reciprocal Verbs: An Overview, in E. König, and V. Gast (eds), (2008), *Reciprocals and Reflexives: Theoretical and Typological Explorations*, Berlin / New York: Mouton de Gruyter, 451-498.

Evolution of Debitive Constructions in Western New Indo-Aryan Languages: Hindi-Urdu, Punjabi, Rajasthani and Gujarati

Liudmila Khokhlova & Boris Zakharyin
(Moscow State University)

Key words: debitive construction, gerundive, Western Indo-Aryan, case marking, syntactic evolution

Schedule: We 16.30 Room 6

This paper aimes at describing the origin and development of debitive constructions (DC) in Western New Indo-Aryan languages of India. The work is based on the functional method of linguistic analysis. The data were obtained from the narrative texts by Jain authors in Old Rajasthani and Old Gujarati starting from the 15th century and the earliest prose texts in Punjabi (Janam Sakhi) belonging to the 17th - 18th centuries. The study of Modern languages involved two methodologies: corpus analysis and elicitation. The paper will demonstrate that the increasing frequency of the Dative instead of the Instrumental Agent marking in DC was part of gradual destruction of 'passive syntax' (Comrie 1978) typical for the climactic stage of ergative development.

There were two sources of DCs in Western NIA: either Skr. gerundives with suffixes *-(ī)ya, -(i)tavya-, -anīya,* etc., (Dave 1935), or former nomina agenti in *-ana* (Bubenik 1996). The latter was most important formally, but the input of the former was also significant for DCs' syntax and semantics (Masica 1991).

Linguists were using different terms for that argument of DC which signifies the participant upon which the obligation or necessity is imposed – like, e.g., 'Experiencer' (Masica 1991), 'Goal' (Butt 2006), 'Agent' (Cardona 1965), 'Obligor Agent' (Hock 1986 and p.c.), etc. The θ-role meant is necessarily ambiguous as the referent of the corresponding NP in DCs with semantics of 'planned action' or 'moral obligation' implies more volitional properties than that one in DCs denoting 'compulsion'. Taking this ambiguity into account and understanding inaccuracy of any general term applied to θ-role in question, we arbitrarily give preference here to Hock's term 'Obligor Agent (OA)', as it simultaneously marks specificity of the role and also demonstrates historical links with gerundival constructions of OIA and MIA.

In Early Vedic OA in gerundive constructions might be marked by Dative, Genitive and Instrumental. In Classical Sanskrit OA lost its Dative marking. Genitives as OA markers were also used rarely, and only Instrumental has been preserved. The latter, thus, was dominant in any of the existing Patient-oriented constructions, including those with gerundives. Old Gujarati and Rajasthani inherited Instrumental as OA marker in their DCs, Later Instrumental OA marking was in many cases replaced by Dative.

Standard Hindi-Urdu and Rajasthani have similar OA marking (by Dative) in all constructions denoting 'obligation'. In Punjabi and sub-standard Hindi-Urdu the OA that is 'following moral rules' is combined with the OA compelled to perform the action (both are marked by Dative); they are opposed to the Ergative OA which follows its own or somebody's (including God's) plans in life. In Gujarati the OA in the constructions denoting 'moral obligation' and 'planned action' may be marked by both Instrumental-Ergative and Dative. In case of the compelled action the only possible OA marking is by Instrumental-Ergative.

References

Bubenik, Vit (1996), *The Structure and Development of Middle Indo-Aryan Dialects*, Delhi: Motilal Banarsidass.

Butt, Miriam (2006), The Dative-Ergative connection, in O. Bonami, and P. Cabredo Hofherr (eds.), (2006), *Empirical Issues in Syntax and Semantics 6*, 69–92. <http://www.cssp.cnrs.fr/eiss6 ISSN 17697158>

Cardona, George (1965), *A Gujarati Reference Grammar*, Philadelphia: The University of Pennsylvania Press.

Comrie, Bernard (1978), Ergativity, in W. P. Lehmann (ed.) (1978), *Syntactic Typology*, Austin: University of Texas.

Dave, Trimbaklal (1935), *A Study of the Gujarati Language in the 16-th Century (V.S.)*, London: The Royal Asiatic Society.

Hock, Hans H. (1986), P-Oriented Constructions in Sanskrit, in B. H. Krishnamurty (ed.), *South Asian*

Languages. Structure, Convergence and Diglossia, vol. 3, Delhi etc.: Motilal Banarsidass, 15-26.
Masica, Colin (1991), *The Indo-Aryan Languages*, Cambridge: Cambridge University Press, 351-352.

ꕥ

Towards a typology of predicative demonstratives

Don Killian
(University of Helsinki)

Keywords: demonstratives, deixis, predication, presentatives, copula

Schedule: Thu 14.00 Room 6

Despite the growing interest in the study of demonstratives, a number of demonstrative categories remain largely unstudied. This study addresses one gap, presenting a preliminary typological survey of predicative demonstratives, or non-verbal demonstrative forms employed for predicative use. Predicative demonstratives have been frequently overlooked in grammatical descriptions, and largely ignored in linguistic typology, despite their occurrence in major languages like Arabic, French, and Russian. Diessel (1997) is one of the only studies which looks at any type predicative demonstrative; his study focuses exclusively on one subtype of predicative demonstrative (identifier), however.

In grammatical descriptions, predicative demonstratives are frequently shoved into the categories of 'pronoun', 'adverb' or 'particle', often with a single example and little in the way of explanation. The Lingua Descriptive Studies Questionnaire (Comrie & Smith 1977), which forms the basis for a number of grammatical descriptions, does not mention any type of predicative demonstratives.

This study is based on data from approximately 135 genetically and geographically diverse languages which show the relevant forms, out of roughly 1100 total languages examined. The main goal is to provide a overview of the morphosyntax of predicative demonstratives and some relevant cross-linguistic tendencies, as well as the semantic subcategories that they divide into.

In terms of lexical categorization, predicative demonstratives are most commonly found in a small closed lexical class of non-verbal predicators with some verb-like properties. The prototypical construction for predicative demonstratives is a nominal or pronominal argument in combination with a distinct demonstrative form, which together make a predicate.

Predicative demonstratives are divided into four semantic-ontological classes of demonstratives, used for presentation, identification, location-existence, and to a more limited extent, general copular uses (including nominal predication). Of these categories, presentatives are by far the most common (1), followed by identifiers (2), locative-existential demonstratives (3) and (4), and the rare copular demonstrative, used for additional constructions such as nominal predication (5) and possession (6).

(1) Presentative
ínjú ***ɔ̌n-nà-Ø***
water **PROX-PRS-3SG.SBJ**
'Here's (the) water!' (Yanda Dom, Heath 2017: 113)

(2) Identifier
na-qamel *ti-qey* ***etag***
NMP-tribe POSS-3SG **IDENT.DIST**
'That's his tribal area.' (Navahaq, Dimock 2009: 183)

(3) Locative-existential

Mustafa ***ee*** *filiɓ ot-a*
Mustafa **EXIST.PROX** inside car-CL1
'Mustafa is inside the car.' (Ndut, Morgan 1996: 95)

(4) Locative-existential
Assan ***ín*** *faam*
Assan **EXIST.DIST** house
'Assan is home (and the speaker is not there).' (Ndut, Morgan 1996: 95)

(5) Copular (Nominal predication use)
yire ***wá*** *pyìyè*
they **COP.DIST** children
'They (elephants in the zoo in the far away capital city) are children.' (Supyire, Carlson 1994: 245)

(6) Copular (Possessive use)
mono ***këj*** *Tëëna yï-parai-n*
exist **ANIM:PROX:COP** Tëëna 3-knife-POSS
''Tëëna has his knife.' (Panare, Payne & Payne 2013: 311)

This study brings to light a category of demonstrative hitherto largely unstudied, encouraging future descriptive and comparative work on predicative demonstratives.

Abbreviations

3SG	3rd person singular	**NMP**	Nominal Prefix
ANIM	Animate	**POSS**	Possessive
COP	Copula	**PROX**	Proximal
DIST	Distal	**IDENT**	Identifier
PRS	Presentative	**SBJ**	Subject

References

Carlson, Robert. (1994). *A grammar of Supyire*. Berlin: Mouton de Gruyter.
Comrie, Bernard & Smith, Norval. (1977). Lingua descriptive studies: Questionnaire. *Lingua* 42, 1-72.
Diessel, Holger. (1997). Predicative demonstratives. *Berkeley Linguistics Society* 23(1), 72-82.
Dimock, Laura Gail. (2009). *A grammar of Nahavaq*. Doctoral dissertation: Victoria University of Wellington.
Heath, Jeffrey. (2017). *A grammar of Yanda Dom*. Draft Ms., Feb 2017.
Morgan, Daniel Ray. (1996). *Overview of grammatical structures of Ndut: a Cangin language of Senegal*. Master's thesis: University of Texas at Arlington, Arlington.
Payne, Thomas & Payne, Doris. (2013). *A typological grammar of Panare: a Cariban language of Venezuela*. Leiden: Brill.

Non-representational meanings of taxonomic nouns constructions in Polish

Anna Kisiel
(KU Leuven)

Keywords: taxonomic nouns, prepositional phrases, grammaticalization, discourse marker, hedging

Schedule: We 15.30 Room 14

This paper focuses on the last stage (cf. Fig.1) of the development of constructions with taxonomic nouns (henceforth TNs) in Polish. Whereas the study of this topic in Polish language has been almost non-existent (Doboszyńska-Markiewicz 2015; Kisiel 2018), there has been a flurry of studies on core Germanic (mainly English) and Romance languages (French, Italian), which have, however, not reached a consensus on central question on the number of meanings and functions to be distinguished.

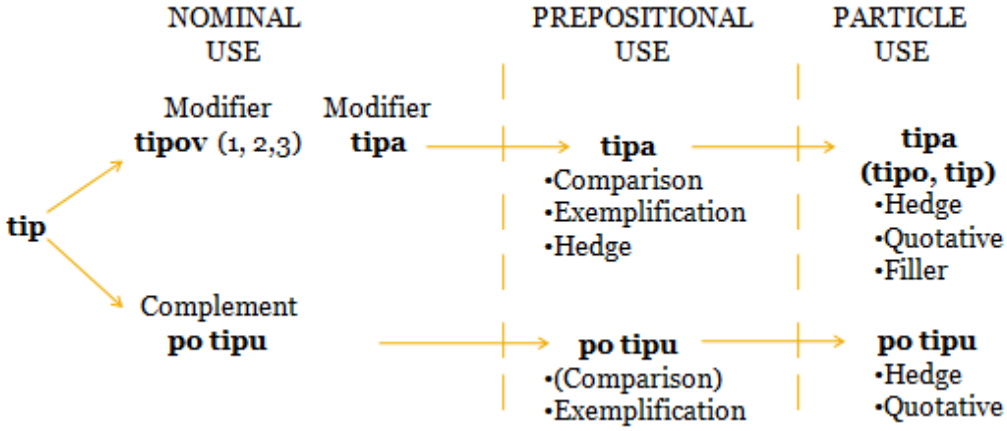


Fig.1. Hypothesized grammaticalization paths for Russian TN-constructions (Kolyaseva & Davidse 2018)

The main goal of the paper is to describe the function(s) and meaning(s) of the prepositional and case-driven linguistic units with TNs in Present-day Polish, concentrating on their particle uses. I argue that the formal (collostructional, positional, prosodic) features distinguish not only prepositional uses, *kurort*$_{\text{NOM.SG}}$ *w rodzaju bułgarskich*$_{\text{GEN.PL}}$ *Złotych*$_{\text{GEN.PL}}$ *Piasków*$_{\text{GEN.PL}}$ 'a resort like Bulgarian Golden Sands', from particle uses but also allow for delimitation of various particle functions, such as exemplification, *kurorty*$_{\text{NOM.PL}}$ ***w rodzaju*** *Złote*$_{\text{NOM.PL}}$ *Piaski*$_{\text{NOM.PL}}$ 'resorts like Golden Sands' or hedging, ***coś***$_{\text{NOM.SG}}$ ***w rodzaju*** *kłótni*$_{\text{GEN.SG}}$ 'something like a quarrel'. The particle uses differentiate Polish and Russian in a most clear way: whereas the range of particle functions of TNs constructions in Russian is very broad (see *tipa*, *po tipu* in Kolyaseva and Davidse (2018) and Kolyaseva (2018)), in Polish it is much narrower but, in return, the constructions with TNs have clear syntactic requirements in each of the functions (see above).

Hedging (Gries & David 2007; Fetzer 2010), as it can be observed also for NPs in which the TN is head (*To*$_{\text{NOM.SG.N}}$ *taki*$_{\text{NOM.SG.M}}$ *rodzaj*$_{\text{NOM.SG}}$ *zaproszenia*$_{\text{GEN.SG}}$? 'It is a kind of invitation?'), is of particular importance. I argue that in hedging function TN-based particles show the tendency to absorb other vagueness markers such as *taki* 'this' (1), *swój* 'of own kind' (*swego rodzaju* 'of a kind'), *coś* 'something' (*czy coś*$_{\text{NOM}}$ *w tym*$_{\text{LOC.SG}}$ *rodzaju* 'or something like this'). As a result of a diachronic analysis I propose that this function for TN head uses emerged only as a consequence of a rise of a discourse marker (Croft 2001).

The analysis focuses on holistic qualitative analysis, looking at formal, semantic, discursive, and, where possible, prosodic, features in individual contexts. The study is not restricted to written data (National Corpus of Polish) but also uses spoken data (Spokes) as it is now generally accepted that spoken language, particularly informal conversation, is the locus of the most radical linguistic change (Halliday 1978). Since particle uses developed in the last decades, the biggest part of the analysis can be restricted to rather modern data (19$^{\text{th}}$-20$^{\text{th}}$ century), with an exception of *rodzaj* 'kind, genre' that became a source for units with hedging function before XVIII century, cf.

(1) Te drzewo szczepiąc, urzni gałąź na łokieć grubą, jak rękojeść u siekiery, sadź gnojem świnim obmazawszy, i to w **swego rodzaju** płonkę. [B. Chmielowski, Nowe Ateny,1794] 'When you graft this tree, you have to cut a cubit thick branch, like an axe's helve, cover with pigs manure, and wrap in a kind of a net.'

References

Croft, William (2001), *Radical Construction Grammar: syntactic theory in typological perspective*, Oxford: Oxford University Press.

Doboszyńska-Markiewicz, Ktarzyna (2015), O "jakby nieostrości" i jej "swego rodzaju" operatorach (linguistic hedges) – uwagi wstępne, *Linguistica Copernicana* 12, 137-155.

Fetzer, Anita (2010), Hedges in context: Form and function of sort of and kind of, in G. Kaltenböck, W. Mihatsch, and S. Schneider (eds.), (2010), *New approaches to hedging*, Bingley: Emerald. 49-71.

Gries, Stefan Th., David, Caroline V. (2007), This is kind of/sort of interesting: Variation in hedging in English, in P. Pahta, I. Taavitsainen, T. Nevelainen, and J. Tyrkkö (eds.), (2007), *Studies in variation, contacts and change in English*, vol. 2: Towards multimedia in corpus studies, Helsinki: University of Helsinki.

Halliday, M.A.K (1978), *Language as social semiotic. The social interpretation of language and meaning*, London: Arnold.

Kisiel, Anna (2018), Polish type noun constructions: x to typ y-a, x w typie y-a, presented at the workshop *Pragmatic functions of type nouns: a crosslinguistic perspective*, Tübingen, 19/06/2018.

Kolyaseva, Alena (2018), The 'new' Russian quotative tipa: Pragmatic scope and functions, *Journal of Pragmatics* 128, 82-97.

Kolyaseva, Alena, Davidse, Kristin (2018), A typology of lexical and grammaticalized uses of Russian *tip, tipa, po tipu*, *Russian Linguistics* 42 (2), 191-220.

The spreading of two borrowed constructions across Hungarian dialects

Katalin É. Kiss
(Research Institute for Linguistics of the Hungarian Academy)

Keywords: borrowing, spreading, modals, Balkan subjunctive, Hungarian

Schedule: Sat 12.00 Room 2

Claims

Recently, the Hungarian modals *kell* (obligation), *szabad* (permission), and *lehet* (possibility/probability) have assumed two new subcategorization frames, equivalents of the Balkan subjunctive complement. They entered the language via the bilingual Hungarian population of Romania. In the case of *kell* and *szabad*, the innovations have spread across the whole language area; in the case of *lehet*, they are confined to the Hungarian dialects of Romania. The difference is derived from the assumption that the constructions have been built into the I-grammars of bilingual and monolingual Hungarian speakers in different ways.

The traditional complementation of Hungarian modals

Hungarian impersonal modals *kell* 'must/need', *szabad* 'may/is allowed', and *lehet* 'can/is possible/probable' are traditionally complemented in two ways. They can take an infinitive phrase, with the subject in the dative, and the infinitive agreeing with it. The verbal particle of the infinitive is raised before the modal:

(1)	*János-nak*	*fel*$_i$	*kell/szabad/lehet* t$_i$	*hív-ni-a*	*Mari-t.*
	John-DAT	up	must/may/can	call-INF-3SG	Mary-ACC

'John must/may/can call up Mary.'

The modals can also be complemented by a finite CP. *Kell* and *szabad* take a subjunctive clause (2a); *lehet* takes an indicative clause (2b). Particle movement cannot cross the CP (2c).

(2) a. *Nem kell,* [$_{CP}$ *hogy János fel hív-j-a Marit*]
not needs that John up call-SUBJ-3SG Mary-ACC

b. Lehet, [$_{CP}$ *hogy János fel hív-t-a Marit*]]]
possible that John up call-PAST-3SG Mary-ACC

c. **Fel*$_i$ *kell,* [$_{CP}$ *hogy János* t$_i$ *hívja Marit*]

The borrowed structures

Kell, szabad, and *lehet* have recently also assumed two versions of a new complement type: a reduced subjunctive clause. The reduced clause has no left-peripheral elements other than an optional complementizer. It has a nominative subject, and it lets the verbal particle through:

(3)a. *Fel*$_i$ *kell/szabad, hogy* t$_i$ *hív-j-a Mari-t János.*
up must/may that call-SUBJ-3SG Mary-ACC John

b. *Fel*$_i$ *kell/szabad* t$_i$ *hív-j-a Marit János.*

These structures are borrowings of the Romanian/Balkan modal construction involving a reduced subjunctive clause:

(4) *Vreau să vină Ion mâine.*
want-1SG SUBJ come Ion tomorrow
'I want Ion to come tomorrow.'

According to Dobrovie-Sorin (2001), the subjunctive particle *să* is analyzed as a complementizer by some speakers. The Hungarian (3a) and (3b) reflect the two analyses of *să.*

The differential spreading of the new constructions

The Balkan-type subjunctive complements have spread in the case of *kell* and *szabad* across the whole Hungarian language area; but they occur with *lehet* only in the Hungarian dialects of Romania. This is explained as follows. Children acquiring Hungarian in a bilingual environment asssigned to the Hungarian infinitival complements of *kell, szabad,* and *lehet* the structure of Romanian subjunctive clauses. This was made possible by the functional identity and formal similarity of the two patterns. When the borrowed structures reached monolingual Hungarian speakers, language learners parsed them as special versions of the subjunctive clause, and did no associate them with the modal *lehet,* taking an indicative complement.

Reference

Dobrovie-Sorin. C. 2001. *The Syntax of Romanian*. Mouton de Gruyter.

The inverse as a strategy to NOT index speech-act participant (SAP) objects: Evidence from Trans-Himalayan

Linda Konnerth
(Hebrew University of Jerusalem; University of Oregon)

Keywords: inverse; speech-act participants; verbal person indexation; referential hierarchy; object indexation

Schedule: We 16.30 Room 7

The 'inverse' is a marker of many different constellations of (di)transitive person scenarios where, as traditionally argued, the object outranks the A argument on a referential hierarchy. The core of the 'inverse' phenomenon has commonly been assumed to lie in the fundamental distinction between speech act participants (SAP's) and 3rd persons (e.g., DeLancey 1981; Zúñiga 2006): 'Core' inverse marking is found when 3rd person A arguments act on SAP objects. Importantly, in addition to inverse marking, these inverse verb forms generally carry person indexation indicating the SAP object. Cross-linguistically, common origins of inverse constructions are passives, cislocatives, or zero third person marking (Gildea and Zúñiga 2016). The motivation for inverse marking and the referential hierarchy (where SAP's outrank 3rd persons) has been suggested to lie in differences of topicality (Givón 1994; Thompson 1994) or agency (Silverstein 1976).

This paper argues that there exists a different type of inverse marking in scenarios involving SAP's and 3rd persons, which has implications for the nature of inverse systems. In a few Trans-Himalayan (Sino-Tibetan) languages, inverse marking crucially combines with A indexation, rather than SAP object indexation. First, in the Northwestern South-Central (SC) language Monsang, inverse marking co-occurs with A (=S) indexation (see Table 1) and can be reconstructed to antipassive origins (Konnerth under review). Second, Northeastern SC languages feature cislocative-derived inverse markers that similarly combine with A indexation (DeLancey 2001), see Sizang (1). Indexing the A argument rather than the ("higher-ranked") SAP object cannot be explained based on previous accounts that invoke referential hierarchies.

In Monsang, we find a typical (or 'canonical' (Jacques and Antonov 2014)) inverse marker occurring in 3→SAP and 2→1 (Table 1). In Sizang and other languages, we instead find inverse markers, derived from cislocatives and other sources, that occur either in first person object, general SAP (first, inclusive, second person) object, or second person object scenarios. I argue that these more and less typical inverse markers, all found distributed across quite closely related languages of the SC branch of Trans-Himalayan, are part of the same phenomenon.

I suggest that we can improve our understanding of the inverse if we examine the 'core inverse' (involving 3rd persons and SAP's) vis-à-vis the less typical inverse markers, in particular first person object marking. This has two advantages. First, there is cross-linguistic evidence for a diachronic link between first person object and 'core' inverse marking, considering data from Monsang and related languages, as well as Matses (Panoan) (Fleck 2006; cf. Konnerth under review). Second, the development of 'core' inverse marking may reflect the same functional motivation that gives rise to less typical inverse constructions: the avoidance of overt reference to SAP's, a well-documented politeness strategy (Heath 1991, 1998), also argued to motivate the development of first person object markers in Kiranti (Trans-Himalayan; Bickel and Gaenszle 2015). If sociopragmatic factors can explain the development of 'core' inverse marking (cf. DeLancey 2017), this further supports recent criticism on the explanatory force of referential hierarchies (Cristofaro 2013; Gildea and Zúñiga 2016; Cristofaro and Zúñiga 2018).

References

Bickel, Balthasar, and Martin Gaenszle. 2015. "First Person Objects, Antipassives, and the Political History of the Southern Kirant." *Journal of South Asian Languages and Linguistics* 2 (1): 63–86.

Cristofaro, Sonia. 2013. "The Referential Hierarchy: Reviewing the Evidence in Diachronic Perspective." In *Languages across Boundaries: Studies in the Memory of Anna Siewierska*, edited by Dik Bakker and Martin Haspelmath, 69–93. Berlin: de Gruyter Mouton.

Cristofaro, Sonia, and Fernando Zúñiga. 2018. "Synchronic vs. Diachronic Approaches to Typological Hierarchies." *Typological Hierarchies in Synchrony and Diachrony* 121: 3–27.

DeLancey, Scott. 1981. “An Interpretation of Split Ergativity and Related Patterns.” *Language* 57 (3): 626–57.

———. 2001. “Lectures on Functional Syntax.” presented at the LSA Summer Institute, UC Santa Barbara. http://darkwing.uoregon.edu/~delancey/sb/fs.html.

———. 2017. “On the Nature of ‘Hierarchical Alignment.’” Paper presented at the 50th Annual Meeting of the Societas Linguistica Europaea, Zürich.

Ebert, Karen H. 1991. “Inverse and Pseudo-Inverse Prefixes in Kiranti Languages: Evidence from Belhare, Athpare and Dungmali.” *Linguistics of the Tibeto-Burman Area* 14 (1): 73–92.

Fleck, David W. 2006. “Antipassive in Matses.” *Studies in Language* 30 (3): 551–73.

Gildea, Spike, and Fernando Zúñiga. 2016. “Referential Hierarchies: A New Look at Some Historical and Typological Patterns.” *Linguistics* 54 (3): 483–529.

Givón, Talmy. 1994. “The Pragmatics of De-Transitive Voice: Functional and Typological Aspects of Inversion.” In *Voice and Inversion*, edited by Talmy Givón, 3–44. Amsterdam: John Benjamins.

Konnerth, Linda. Under review. “On the nature of inverse systems: The rise of inverse marking via antipassive constructions.”

Heath, Jeffrey. 1991. “Pragmatic Disguise in Pronominal-Affix Paradigms.” In *Paradigms: The Economy of Inflection*, edited by Frans Plank, 75–89. Berlin/New York: Mouton de Gruyter.

———. 1998. “Pragmatic Skewing in 1↔ 2 Pronominal Combinations in Native American Languages.” *International Journal of American Linguistics* 64 (2): 83–104.

Jacques, Guillaume, and Anton Antonov. 2014. “Direct/Inverse Systems.” *Language and Linguistics Compass* 8 (7): 301–18.

Silverstein, Michael. 1976. “Hierarchy of Features and Ergativity.” In *Grammatical Categories in Australian Languages*, edited by Robert M. W. Dixon, 112–71. New Jersey: Humanities Press.

Stern, Theodore. 1984. “Sizang (Siyin) Chin Texts.” *Linguistics of the Tibeto-Burman Area* 8 (1): 43–58.

Thompson, Chad. 1994. “Passive and Inverse Constructions.” In *Voice and Inversion*, edited by Talmy Givón, 28:47–63. Amsterdam: John Benjamins.

Zúñiga, Fernando. 2006. *Deixis and Alignment: Inverse Systems in Indigenous Languages of the Americas*. Amsterdam/Philadelphia: John Benjamins.

Data

Table 1. Monsang postverbal negative non-future singular paradigms of *bìn* ‘beat’ (transitive) and *kàr* ‘climb’ (intransitive) (*-maʔ~-ma:* ‘negative’); inverse scenarios shaded

A O	1SG	2SG	3SG	S
1SG	----------	bìn-má:-ŋ-tʃə̀	bìn-má:-ŋ	kàr-má:-ŋ
2SG	ì-bín-má:-tʃə́	----	bìn-má:-tʃə́	kàr-má:-tʃə́
3SG	ì-bín-máʔ		bìn-máʔ	kàr-máʔ

(1) Sizang (data from Stern 1984; DeLancey 2001:132-3)

*k-**ong***	*thûk*	*kí:k*	*lâ-lê:u*	*hî:*
1-**CIS**	reply	again	once.more	FINAL

'I in turn reply **to you**.'

Obligatory grammar, optional lexicon: The limits of the concept of obligatoriness for a definition of grammatical status

Martin Konvička

Schedule: Thu 15.00 Room 14

<pdf>

On constructional diachrony: From ascription to standard negation

Olga Krasnoukhova & Johan van der Auwera
(University of Antwerp)

Key words: Negation; Ascriptive construction; Constructional change; Diachrony of negation

Schedule: Thu 12.00 Room 6

This paper reports on a diachronic development – hitherto unexplored – when a 'standard negation' construction (and its marker) has an ascriptive construction as its source. By 'standard negation' (J. Payne 1985) we refer to negation of declarative main clauses with a lexical verbal predicate. 'Ascriptive negation' (Veselinova 2013: 110) refers here to negation of clauses with a nominal or adjectival predicate. Cross-linguistically, these two environments commonly use divergent negators, for which Eriksen (2011) offers an explanation. In fact, languages that constitute counter-examples to Eriksen's (2011: 299) hypothesis are of the type we focus on here. In these languages the construction used for standard negation surfaces has an ascriptive construction. Example (1) from Matsés (Fleck 2003) illustrates this: the expression of negation 'They don't go away' is literally 'They are non-going/non-goers'. The negated lexical verb constitutes a non-finite complement taking an affirmative copula (see also Miestamo 2005: 410).

(1) Matsés (Panoan; Fleck 2003: 1063)
nid-en-quio *ic-e-c*
go-NEG-AUG be-NPST-INDIC
'They [the bats] don't go away.' (Lit. 'They are non-going')

Thus it is the (negative) property of the referent and not the (negative) property of the state of affairs that is encoded in the construction. While the latter is known from the 'Negative Existential Cycle' development (Croft 1991), the former phenomenon has not received attention in diachronic studies to date.

Data from the Tukanoan languages, taken as a case-study, give us a telling distribution: a construction similar to (1) is found in this family either solely for ascriptive negation ('*John is the one who is non-big*') (in Máihĩkì, S. Farmer, p.c., and Barasano, Jones & Jones 1991: 127-130), or for both ascriptive and standard negation with habitual semantics (Ecuadorian Secoya, A. Schwarz, p.c.), or only for standard negation with habitual or abilitative semantics (Ecuadorian Siona, Bruil 2014: 214,193). That a standard negation construction can have another construction as its historical source is explicitly noted by Fleck (2003: 1064) for the Panoan language Matsés. In Matsés (1), predicate adjective copular constructions and standard negation still resemble each other syntactically and semantically, although the main elements of each construction have already developed divergent characteristics.

Summarizing, the paper contributes by showing that this type of standard negation construction observed in some South American languages and beyond (see Miestamo 2005:410) can be explained by its origin in a different source construction. This diachronic construction grammar approach has

proved relevant for various linguistic phenomena (Van de Velde 2013 et al, Barðdal & Gildea 2015, *inter alia*). We propose its relevance for explaining a negation pattern that hitherto remained underexplored.

Abbreviations:

AUG = 'augmentative'; INDIC = 'indicative'; NEG = 'negative'; NPST = 'non-past'.

References

Barðdal, Johanna & Spike. Gildea. 2015. Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds). *Diachronic Construction Grammar*, 1-50. Amsterdam: John Benjamins.

Bruil, Martine. 2014. Clause-typing and evidentiality in Ecuadorian Siona. Doctoral dissertation, Rijksuniversiteit Leiden.

Croft, William. 1991. The evolution of negation. *Journal of Linguistics* 27: 1-27.

Eriksen, Pål Kristian. 2011. 'To not be' or not 'to not be': The typology of negation of non-verbal predicates. *Studies in Language* 35(2): 275-310.

Fleck, David William. 2003. A grammar of Matses. Doctoral dissertation, Rice University.

Jones, Wendell & Paula Jones. 1991. *Barasano Syntax.* (Studies in the Languages of Colombia, 2.) Arlington, TX: SIL International and University of Texas at Arlington.

Miestamo, Matti. 2005. *Standard negation. The negation of declarative verbal main clauses in a typological perspective.* Berlin: Mouton de Gruyter.

Payne, John R. 1985. Negation. In Timothy Shopen (ed.) *Language typology and syntactic Description*, Vol 1. 'Clause structure', 197-242. Cambridge: Cambridge University Press.

Van de Velde, Freek, Hendrik de Smet & Lobke Ghesquìre. 2013. On multiple source constructions in language change. *Studies in Language* 37(3): 473-489.

Veselinova, Ljuba. 2013. Negative existentials: A cross-linguistic study. *Rivista di Linguistica* 25(1): 107-145.

The role of valency, type and token frequency in word formation grammar: Evidence from Persian

Kateryna Krykoniuk
(Cardiff University)

Keywords: word formation, valency, type and frequency, formal morphological analysis, the Persian language

Schedule: Fri 14.30 Room 6

Token and type frequencies are important characteristics of a language's grammatical system. Token frequency considers the number of times a unit appears in running text (Bybee 2007: 9), whereas type frequency refers to the number of distinct items that can occur in the open slot of a construction (Bybee 2007: 14). For emergent probabilistic grammars, token and type frequencies offer insight into the leading grammatical forces of language changes. Furthermore, both frequencies are the determining factor for the structure of grammatical knowledge in the mind (Croft 2007: 499).

More specifically, morphological token and type frequencies reveal interesting aspects of the relations between lexicon and grammar. Some work has been done on inflectional representations and

their frequency (e.g. Bybee 1985, Bybee et al. 1994), considering how various morphological phenomena trigger each other. On the other hand, a significant contribution to uderstanding the role of type and token frequency in derivational morphology has been been made by Laudanna & Burani (1995), Baayen et al. (1997), Hay (2001), Hay & Baayen (2002) and Järvikivi et al. (2006).

Furthermore, whereas determining token frequency is a fairly straightforward procedure due to the fact that it concerns actual instances of units in a text/corpus, computing type frequency is a more challenging task. The first difficulty lies in finding the optimal way to identify a construction/pattern. Other problems involve setting a frequency range so that a given type can be assigned to a high- or low-frequency group. One possible solution to these problems comes from the formal morphological analysis developed by Russian linguists (e.g. Bratchikov et al. 1958, Tyschenko 2003). In this approach, every multimorphemic lexeme is re-written using a metalanguage, specifically designed for this purpose. Particularly, all lexical morphemes are encoded by the morphological class to which they belong (e.g. *N* – noun, *Aj* – adjective, *V* – verb present tense root), and affixes are denoted unchanged by means of the Latin alphabet. This is illustrated in Persian, for example, where *khoshāli* ('gladness') is coded as *N"=Aj'+i={Aj+N}+i*, *kooshā* ('diligent') as *Aj'=V+ā*, etc. Consistent implementation leads to the formation of a morphological metacorpus, the further analysis of which reveals patterns and regularities in the morphemes' arrangement. This morphological metacorpus can be interpreted in the light of a usage-based approach, where a word formation pattern can be seen as equivalent to a construction, i.e. a type.

In this paper, formal morphological analysis is used to identify word formation patterns in Persian and calculate their type frequencies in two levels of generalization. Using a morphological metacorpus of 3273 multimorphemic nouns in Persian (as developed in Krykoniuk 2014), a frequency dictionary of Persian (Miller & Aghajanian-Stewart 2017) and the Hamshahri Corpus (AleAhmad et al. 2009), I will explore the relationship between type and token frequency, as well as between type frequency and valency. In a broader sense, I will discuss what these variables reveal about the word formation structure of present-day Persian. In addition, I will show how a further generalization of word formation patterns may contribute to a theory of grammatical typology.

References

AleAhmad, Abolfazl, Amiri, Hadi, Rahgozar, Masoud, and Oroumchian, Farhad (2009) 'Hamshahri: A standard Persian text collection', *Journal of Knowledge-Based Systems, 22 (5)*, 382–387.

Baayen, Harold & Dijkstra, Ton & Schreuder, Robert (1997), Singulars and plurals in Dutch: Evidence for a parallel dual route model, *Journal of Memory and Language*, 36, 94-117.

Bybee, Joan & Perkins, Revere & Pagliuca, William (1994), The evolution of grammar. Tense, aspect and modality in the languages of the world. Chicago: University of Chicago Press.

Bybee, Joan (1985), Morphology: A study of the relation between meaning and form. Amsterdam: Benjamins.

Bybee, Joan (2007), Frequency of use and the organization of language. New York: Oxford University Press.

Bybee, Joan (2013), Usage-based theory and exemplar representations of constructions in Thomas Hoffman and Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, Oxford/ Oxford University Press.

Croft, William (2007), Construction grammar in Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, Oxford: Oxford University Press, 463-508.

Hay, Jannifer & Bayyan, Harold (2002), Parsing and productivity in Geert Booij and Jaap van Marle (eds.), *Yearbook of morphology 2001*, Dordrecht: Klewer Academic, 203-235.

Hay, Jannifer (2001), Lexical frequency in morphology: Is everything relative? *Linguistics 39 (6)*, 1041-1070.

Järvikivi, Juhani & Bertram, Raymond & Niemi, Jussi (2006), Affixal salience and the processing of derivational morphology: The role of suffix allomorphy, *Language and Cognitive Processes, 21 (4)*, 394-431.

Laudanna, Alessandro & Burani, Cristina (1995), Distributional properties of derivational affixes: Implications for processing, in L. B. Feldman (ed.), Morphological aspects of lamguage processing, Hillsdale, NJ: Lawrence Erlbaum Associates, 345-364.

Miller, Corey, & Karineh Aghajanian-Stewart. *A Frequency Dictionary of Persian : Core Vocabulary for Learners*. Routledge Frequency Dictionaries. Milton Park, Abingdon, Oxon: Routledge, 2017.

Tyschenko, Kostynatyn (2003). La morfologia formale del verbo italiano in *Lingua e letteratura italiana dentroe fuori la penisola*, Atti del III Congresso degli Italianisti Europei, Cracovia, 521-527.

Братчиков И. Л. О структуре словаря и кодировке информации для машинного перевода / И.Л. Братчиков, С.Я. Фитиалов, Г.С. Цейтин. − Л.: Изд-во ЛГУ, 1958, С. 61-87. – (Материалы по машинному переводу).

(Bratchikov, I., Fitialov, S., Tseitin, G. (1958). O structure slovarya i kodirovke informatsii dlya mashynnoho perevoda. Leningrad: Izd-vo LGU, pp. 61-87).

Криконюк К.М. Формальна морфологія сучасного перського словотвору: Катерина Миколаївна Криконюк. – Київ, 2014.

(Krykoniuk, K. (2014). Formal morphology of word formation in present-day Persian (thesis). Kyiv, Kyiv National Taras Shevchenko University).

Subjectivized or objectivized construal – intersubjective sharing of experience

Nóra Kugler
(Eötvös Loránd University, Budapest)

Keywords: experimental study, everyday epistemology, epistemic-inferential modality, evidentiality, subjectification, metapragmatic awareness

Schedule: We 15.00 Room 11

The paper discusses the results of two empirical studies on inference making from visual stimuli. The conceptual tools of the analysis come from Cognitive Grammar (cf. Langacker 2006, 2017).

The first experiment (Design 1) had informants watch a 30-second video clip showing a single person in an everyday situation with ambiguous interpretation. Informants were asked to narrate what had happened in the video clip, and to answer questions about the inferences and beliefs they had just shared by reflecting on their own inferential processes. The two discourse types showed up characteristic differences in the frequency of subjectivized and objectivized construal.

The main methodological implication of this study is that in Hungarian (a language lacking systemic paradigmatic oppositions) it seems to be impossible to separate expressions of evidentiality (observations, 'shared knowledge', etc.) and expressions of assumptions and beliefs lying with the conceptualizer (epistemic-inferential modality). Metadiscourses about inferential processes further suggest that the basis of inference may be accessible and explicable to varying degrees along a continuum. One problem with the experimental design was that the informants had reason to assume that the investigator knew as much or even more about the events in the video clip than they did.

Therefore, they may have perceived the interview as a task situation rather than a neutral setting for sharing their experiences.

The questionnaire study of Kugler and Pomázi (2019) (Design 2) asked informants to answer questions about visible elements of a photograph, and about further pieces of information that could only be reached by inferences supported by the picture. The results suggest that the only reliable method for learning about the choices made by informants is to elicit metadiscursive reflections from them.

Acknowledgements

This paper was supported by the Hungarian National Research, Development and Innovation Office, under project number 129040, "Verbal constructions in Hungarian. Research in usage-based construction grammar", led by Gábor Tolcsvai Nagy.

References

Kugler, Nóra – Pomázi, Bence (2019), A látással összefüggő, kiértékelést kifejező episztemikus predikáció alternatív megkonstruálási módjai [Alternative construals of seeing related epistemic grounding predicates], in Laczkó, Krisztina and Tátrai, Szilárd (eds), *Kontextualizáció és metapragmatikai tudatosság* [Contextualization and metapragmatic awareness], Budapest: Eötvös Collegium.

Langacker, Ronald W. (2006), Subjectification, grammaticization, and conceptual archetypes, in A. Athanasiadou, C. Canakis, and B. Cornillie (eds), *Subjectification: Various paths to subjectivity*, Berlin / New York: Mouton de Gruyter, 17–40.

Langacker, Ronald W. (2017), Evidentiality in Cognitive Grammar, in J. I. Marín Arrese, G. Haßler, and M. Carretero (eds), *Evidentiality revisited: Cognitive grammar, functional and discourse-pragmatic perspectives*. Amsterdam / Philadelphia: John Benjamins, 13–55.

From Deontic Modality to Conditionality via Performativity: A Diachronic Investigation into Bi in Chinese

Yueh Hsin Kuo
(University of Edinburgh)

Keywords: conditional, modal, subjectification, deontic, advice

Schedule: Thu 16.00 Room 11

This paper explores how a deontic modal may develop into a conditional connective, using 必*bi* in Chinese as an example and the Invited Inferencing Theory of Semantic Change (Traugott & Dasher 2002) as the framework. It will be argued that this change results from an interaction of *teleological*, *conditional*, and *performative* meanings.

The dictionary *Hanyu Da Zidian* notes that *bi*, commonly classified as an epistemic or deontic modal, is also a conditional connective:

王必無人，臣願奉璧往使

wong	**bi**	wu	ren,	chen	yuan	feng	bi
lord	if	have.not	person,	servant	willing	take	jade.disc
wong.shi							
go.on.a.diplomatic.mission							

'If my lord doesn't have anyone (to send on this diplomatic mission), your humble servant is willing to take the jade disc and serve as your ambassador.' 史记*Shiji*（1st c. BC）

So far, no research has looked into exactly how *bi* as a modal developed into a connective. Traugott (1985) offers a starting point: she observes that modals, especially epistemic and optative ones, commonly give rise to conditional connectives, which is motivated by their similarities in meaning. van der Auwera & Plungian (1998), building on Traugott (1985), claim that conditionality has *epistemic possibility* and *necessity* as its sources. However, since then, Traugott's brief remark has remained largely unsubstantiated by actual data and in-depth analysis: under exactly what kind of condition does such a development take place?

This paper draws on corpus data from the Center of Chinese Linguistics Corpus to show that when *deontic bi* is used as a *teleological* modal and to perform an indirect speech act of *advice*, it is possible for inferences of conditionality to arise, if there is a following clause that is interpretable as the goal teleological *bi* is with respect to. This is because both teleological modals and conditionals can perform *advice* indirectly, and by Horn (1984)'s R-heuristic for inferencing ('more is intended than said'), in using teleological *bi* as an indirect act of *advice* the speaker may be taken to utter a piece of conditional *advice* coded by *bi*. That is, teleological meaning can turn into conditional meaning via performative meaning.

This research thus not only fleshes out Traugott's observation with more detail by accounting for how conditional *bi* arose from teleological *bi*, but also modifies van der Auwera & Plungian's claim by showing *deonticity* is also a possible source of conditionality.

References

Horn, Laurance. R. (1984), Toward a new taxonomy for pragmatic inference: Q-based and R based implicature. In Deborah (ed.), (1984), *Meaning, form and use in context: linguistic applications,* Washington DC: Georgetown University Press, 11-42.

Traugott, Elizabeth C. & Richard Dasher (2002), *Regularity in Semantic Change*. Cambridge: CUP.

Traugott, Elizabeth C. 1985. Conditional markers. In Haiman (ed.), (1985), *Iconitiy in Syntax,* Amsterdam: Benjamins, 289-307.

van der Auwera, Johan. and Vladimir. A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2(1), 79-124.

Corpus and dictionary

汉语大字典*Hanyu Da Cidian*.

Available online at http://words.sinica.edu.tw/sou/sou.html.

Center for Chinese Linguistics Corpus.

Available online at http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp.

Exploring the functions of conditional constructions at the right and left periphery: Evidence from English and French

Cristina Lastres-López
(University of Santiago de Compostela)

Keywords: conditionals, right periphery, left periphery, English, French

Schedule: Thu 14.30 Room 11

Conditionals are usually defined as constructions which exhibit cause-consequence patterns. Prior studies examining the discourse-pragmatic properties of conditionals have shown, however, that these constructions may express a wider range of functions in discourse (Ford and Thompson 1986, Ford 1997, Declerck and Reed 2001, Dancygier and Sweetser 2005, Warchal 2010, Lastres-López 2019, among others); but very few studies have adopted a corpus-based methodology to explore the pragmatic richness of conditional clauses and their functions at the right and left periphery of the sentence.

This presentation intends to contribute to fill this gap by analysing conditional constructions introduced by *if* and *si* in spoken English and French. Two different registers will be analysed – parliamentary discourse and face to face-conversations – which stand at the two extremes of the formal-informal continuum. My aim is to explore (i) the discourse-pragmatic functions of conditionals in both languages, and the extent to which such functions show cross-linguistic similarities and/or divergences; and (ii) the impact of the position of the *if*-/*si*-clause at the right or left periphery of the sentence on its discourse-pragmatic function.

Taking prototypical conditionals as a point of departure, a path of pragmaticalization will be proposed, along the lines of the three metafunctions distinguished by Halliday and Matthiessen (2014), namely: ideational, interpersonal and textual. In this suggested cline, ideational conditionals, as in (1) and (2), move beyond their prototypical conditional meaning to acquire more pragmaticalized functions, as illustrated in (3) and (4).

(1) If it's a really nice day we could walk <ICE-GB: S1A-006#301:1:B>

(2) Si tu manges trop de pizzas tu t'en lasses <C-ORAL-ROM ffamcv01>

If you eat too many pizzas, you get tired of them

(3) Oh those are the uhm those are Alastair Black's if you're interested <ICE-GB: S1A-070#127:1:A>

(4) Mais je suis consciente que j'ai fait des efforts aussi si tu veux comme tu disais tout à l'heure ils étaient peut-être inconscients <C-ORAL-ROM ffamcv01>

But I am aware that I have made efforts as well, if you want, as you were saying a bit earlier, they were maybe unaware

Data are extracted from the British component of the International Corpus of English (ICE-GB) (Nelson *et al*. 2002) and the French component of the Integrated Reference Corpora for Spoken Romance Languages (C-ORAL-ROM) (Cresti and Moneglia 2005), for conversations; and from the Hansard Corpus, including both the British Parliament Hansard Corpus and the French component from Hansard Corpus of the Canadian Parliament, for parliamentary discourse. Corpus findings confirm the diversity of pragmatic functions of conditional clauses in spoken discourse in English and French, and show correlations between certain functions and the position of the *if*-/*si*-clause at the right or left periphery of the sentence, especially with certain types of interpersonal conditionals.

References

Cresti, Emanuela and Moneglia, Massimo (2005) (Eds), *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam / Philadelphia: John Benjamins.

Dancygier, Barbara and Sweetser, Eve (2005), *Mental spaces in grammar: Conditional constructions*. Cambridge: Cambridge University Press.

Declerck, Renaat and Reed, Susan (2001), *Conditionals: A comprehensive empirical analysis*. Berlin: Mouton de Gruyter.

Ford, Cecilia (1997), Speaking conditionally: Some contexts for *if*-clauses in conversation, in A. Athanasiadou, and R. Dirven (eds), *On conditionals again*. Amsterdam / Philadelphia: John Benjamins, 387-413.

Ford, Cecilia and Thompson, Sandra (1986), Conditions in discourse: A text-based study from English, in E. C. Traugott, A. Meulen, J. S. Reilly, and C. A. Ferguson (eds), *On conditionals.* Cambridge: Cambridge University Press, 353-372.

Halliday, Michael and Matthiessen, Christian (2014), *Halliday's Introduction to Functional Grammar*. London: Routledge.

Lastres-López, Cristina (2019), Conditionals in spoken courtroom and parliamentary discourse in English, French and Spanish: A contrastive analysis, in T. Fanego, and P. Rodríguez-Puente (eds), (2019), *Corpus-based research on variation in English legal discourse*, Amsterdam / Philadelphia: John Benjamins, 51-78.

Nelson, Gerald, Wallis, Sean, and Aarts, Bas (2002), *Exploring Natural Language: Working with the British component of the International Corpus of English*. Amsterdam / Philadelphia: John Benjamins.

Warchal, Krystyna (2010), Moulding interpersonal relations through conditional clauses: Consensus-building strategies in written academic discourse, *Journal of English for Academic Purposes*, 9, 140-150.

From fictive motion to fictive change: interactions, parallels and asymmetries

Peter Lauwers and Niek Van Wettere

Schedule: We 12.00 Room 14

This talk deals with the expression of change in constructions built around change of state copulas such as French *devenir* and *se faire* 'become'. In principle, these constructions denote a process during which a subject referent undergoes some kind of change :

Edmond Edmont has become famous.

However, since the work on fictive or virtual change (a.o. Fauconnier 1985, Matsumoto 1996, Talmy 2000 (a.o.), Brandt 2009), we know that, alongside virtual motion (e.g. *This fence goes from the plateau to the valley* ; Talmy 2000: 99), change of state predicates may express no change at all (French Ten Ten Corpus):

(1) Une route bitumée sur les trente premiers kilomètres et qui **devient** très vite une piste très périlleuse.
'A paved road for the first thirty kilometres and which very quickly becomes a very dangerous track'.

(2) Mais plus l'objet est lourd, plus il a du mal […]. Une chaise, une table passe encore, mais une armoire **devient** un peu trop compliqué.
'But the heavier the object, the more difficulties he has: […] A chair, a table, that is one thing, but a wardrobe becomes a little too complicated'.

(3) Le fondement des règles morales doit en fait s'appuyer sur un regard positif et juste du corps humain […]. Les règles morales **deviennent** alors comme un itinéraire […].
'The foundation of moral rules must in fact be based on a positive and correct view of the human body […]. Moral rules then become like a journey […]'.

(4) Alors que son liquide écarlate se repandait par terre il se cacha derrière un grand rocher […]. Puis tout **devint** flou ... Sa vu baissa, ses paupières lourde se refermèrent [sic]

'As his scarlet liquid spread on the ground he hid behind a large rock [...]. Then everything became blurry... His sight went down, his heavy eyelids closed up'.

(5) Mes étudiants **deviennent** de plus en plus jeunes.
'My students are getting younger and younger'.

Based on this negative working definition and by means of random samples of 400 corpus instances of each verb (French Ten Ten Corpus; Frantext), we aim to recognize different forms of virtual change (in a broad sense), providing an answer to the following questions:

(i) to what extent does fictive change involve motion, like in (1) ?
(ii) what are the other mechanisms leading to virtual change effects? Fictive change may result from a comparison with an absolute (cf. in Japanese, Matsumoto 1996) or contextual (2) reference point, or from any kind of temporality inherent to discourse (3). Change can also be highly restricted to one particular experiencer (4), as very often with *se faire*.
(i) what is the impact of fictive change on non-predicational uses of the copula ? For this, we will also examine Dutch *worden*, which is more prone to specificational usage :

De volgende halte wordt Brussel-Zuid.
'The next stop??becomes (-> is) Bruxelles-Midi'

Particular attention will be paid to interactions, parallels and asymmetries between fictive motion and change (cf. Matsumoto 1996).

References

Brandt, L. 2009. "Subjectivity in the act of representing: the case for subjective motion and change". *Phenomenology and the Cognitive Sciences* (online publication: www.springerlink.com/index/f7h1l25u75537024.pdf).

Fauconnier, G. 1985. *Mental spaces: Aspects of meaning construction in natural language*. New York, NY, US: Cambridge University Press.

Kilgarriff, A., Baisa, V., Bu˘sta, J., Jakub´ı˘cek, M.,Michelfeit, J.,Rychl´y, P. and Suchomel, V. 2014. "The Sketch Engine: ten years on". *Lexicography*, 1. 7–36.

Matsumoto, Y. 1996. Subjective-change expressions in Japanese and their cognitive and linguistic bases. In: Fauconnier, G., Sweetser, E. (Eds.), *Spaces, Worlds, and Grammar*. University of Chicago Press, Chicago, pp. 124–156.

Talmy, L. 2000. *Toward A Cognitive Semantics*, vol. 1. MIT Press, Cambridge, MA.

Formulaic versus productive constructions in L2 Finnish learners

Sirkku Lesonen, Rasmus Steinkrauss, Minna Suni & Marjolijn Verspoor
(University of Groningen and University of Jyväskylä; University of Groningen; University of Jyväskylä; & University of Groningen and University of Pannonia)

Keywords: L2 development, usage-based language learning, constructions, L2 Finnish, item-based language learning

Schedule: Sat 11.30 Room 8

This paper traces the development of two different constructions in four Finnish L2 learners. Based on studies on L1 development (Dąbrowska and Lieven 2005.), it has been assumed that L2 learners

develop their constructions from lexically specific, formulaic expressions into a more productive, abstract schema (see Ellis 2002). Indeed, the role of item-based expressions in L2 learning has been demonstrated in several studies (e.g. Mellow 2006, Eskildsen 2012). However, schematic constructions also have appeared right from the beginning of L2 learning (Roehr-Brackin 2014).

Given this discussion, the aim of this paper is to investigate the role of both formulaic, and more productive expressions in L2 learning. The following research question was formulated: Do the *haluta* 'want' and *tykätä* 'like' constructions of four Finnish L2 learners develop from lexically specific to more productive constructions over time? In order to answer to this question, we collected free response data weekly from four learners over the course of 9 months. In total, 272 utterances were included in the analysis. Because our main interest was to find out to what extent our participants' constructions are productive at the different phases in development, we calculated the number of different forms of the main verb of the construction, i.e. *haluta* or *tykätä*, and the number of their complements. If the learner uses only one or very few forms of the verb initially, and the complements do not show a great deal of variation, the construction is considered to be learned in an item-based fashion and the learner is not assumed to have (yet) developed an abstract and productive schema.

Our results show that, contrary to the default assumption, both formulaic and schematic patterns play a role in L2 use both at early and later stages of development, and that learners exhibit individual learning paths for the same constructions. Some learners start with lexically-specific expressions rooted in a specific communicative function, while the other learners' constructions exhibit a greater number of variable instantiations initially. We also show that for some learners, chunk-like expressions do not only characterize the early stages of L2 development but are also used later on.

Besides this inter-individual variability, our results point to intra-individual variability in development too: a learner may develop two superficially similar constructions in fairly different ways probably because of the actual usage-events she/he has experienced. In this paper, the role of initial conditions and L2 instruction in this speaker-dependent development will be discussed.

References

Dąbrowska, Ewa & Elena Lieven (2005), Towards a lexically specific grammar of children's question constructions, *Cognitive Linguistics* 16(3), 437–474.

Ellis, Nick C. (2002), Frequency effects in language processing, *Studies in Second Language Acquisition* 24(2), 143–188.

Mellow, J. Dean (2006.), The emergence of second language syntax: A case study of the acquisition of relative clauses, *Applied Linguistics* 27(4), 645–70.

Roehr-Brackin, Karen (2014), Explicit knowledge and processes from a usage-based perspective: The developmental trajectory of an instructed L2 learner, *Language Learning* 64(4), 771–808.

Word order variability: A quantitative usage-based approach

Natalia Levshina
(Leipzig University)

Keywords: word order, Universal Dependencies, entropy, mixed-effects regression, corpora

Schedule: Sat 11.00 Room 6

This study examines word order variability in the world's languages. In comparison with word order correlations (e.g. Greenberg 1963; Vennemann 1974; Lehmann 1978; Dryer 1992; Dunn et al. 2011), word order variability has received relatively little attention in typology, with a few notable exceptions (e.g. Givón 1984: Ch. 6; Payne 1992; Bakker 1998; Siewierska 1998). Unlike in the previous studies,

which are usually based on reference grammars and individual examples, I take a radical corpus-driven perspective, measuring variability as entropy (Shannon 1948) of the head and dependents (or two co-dependents) in syntactic dependencies found in corpora. I also argue on the basis of statistical evidence that the variability of different word order patterns is explained by a multitude of usage-based and processing factors, rather than by a single parameter, such as branching direction (Dryer 1992) or configurationality (Hale 1983). This is a first attempt to systematize these factors.

The case studies are based on the Universal Dependencies (Nivre et al. 2017) and Leipzig Corpora Collection (Goldhahn et al. 2012), processed with the help of the UDPipe software (Straka & Straková 2017). In the first part, I present a classification of word order variability in more than 60 languages available in the Universal Dependencies corpora. It reveals the continuum of average head-dependent entropy shown in (1):

(1) Synthetic > Analytic/Isolating > strictly OV languages

This cline is explained in terms of processing constraints (Hawkins 2014) and the distinction between 'tight fit' and 'loose fit' languages (Hawkins 1986; Müller-Gotama 1994). Also, a trade-off between morphological and word order information plays a role (e.g. Siewierska 1998).

Next, I will explain why some patterns exhibit higher variability than others, and present the results of several mixed-effects regression models, which demonstrate the importance of the following factors:

- *Disambiguation needs:* There is a negative, albeit non-linear, correlation between the proportion of confusable objects and subjects in a language and word order entropy of subjects and objects (cf. Sapir 1921: 66; Blake 2001: 15).
- *Entrenchment:* According to the usage-based perspective (cf. Bybee 2002), word combinations that occur very frequently will also be more entrenched and have more rigid word order. This explains low syntagmatic variability of highly frequent grammaticalized elements (Lehmann 2015: 167). I also show that lexically specific adverbials and obliques that are more strongly associated with their verbal heads exhibit on average less positional variability than less strongly associated ones.
- *Optimization of processing:* In particular, the position of clauses in VO languages is less flexible than the position of other elements. This reflects Hawkins' (2014) principles "Minimize Domains" and "Maximize Online Processing" at the level of language-internal variation.
- *Multiple functions of dependent elements:* For example, adverbial modifiers and clauses exhibit high positional variability because they perform a variety of functions (causal, temporal, epistemic, etc.), which are associated with different positions due to iconicity or discourse-processing constraints (e.g. Diessel 2001).

References

Bakker, Dik (1998). Flexibility and consistency in word order patterns in the languages of Europe. In: Anna Siewierska (ed.), *Constituent Order in the Languages of Europe*, 383–419. Berlin: Mouton de Gruyter.

Blake, Barry J. (2001). *Case*. Cambridge: Cambridge University Press.

Bybee, Joan (2002). Sequentiality as the basis of constituent structure. In: Talmy Givón & Bertram F. Malle (eds.), *The Evolution of Language out of Pre-Language*, 109–32. Amsterdam/Philadelphia: John Benjamins.

Diessel, Holger (2001). The ordering distribution of the main and subordinate clauses: A typological study. *Language* 77(3): 433–455.

Dryer, Matthew S. (1992). The Greenbergian word order correlations. *Language* 68: 81–138.

Dunn, Michael, Simon Greenhill, Stephen C. Levinson & Russel Gray (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473: 79–82.

Givón, Talmy (1984). *Syntax: A Functional-Typological Introduction. Vol. I*. Amsterdam/ Philadelphia: John Benjamins.

Goldhahn, Dirk, Thomas Eckart & Uwe Quasthoff (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. In: *Proceedings of the 8th International Language Ressources and Evaluation (LREC'12), 2012.*

Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg, Joseph H. (ed.), *Universals of Human Language*, 73–113. Cambridge, Mass: MIT Press.

Hale, Ken (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1: 5–47.

Hawkins, John A. (1986). *A Comparative Typology of English and German: Unifying the Contrasts*. Austin: University of Austin Press.

Hawkins, John A. (1994). *A Performance Theory of Order and Constituency*. (Cambridge Studies in Linguistics, 73.) Cambridge: Cambridge University Press.

Hawkins, John A. (2014). *Cross-Linguistic Variation and Efficiency*. Oxford: Oxford University Press.

Lehmann, Winfred P. (1978). The great underlying ground-plans. In: Winfred P Lehmann (ed.), *Syntactic Typology: Studies in the Phenomenology of Language*. Austin: University of Texas Press.

Lehmann, Christian (2015). *Thoughts on Grammaticalization*. 3rd edn. Berlin: Language Science Press.

Müller-Gotama, Franz (1994). *Grammatical Relations: A Cross-Linguistic Perspective on their Syntax and Semantics*. Berlin: Mouton de Gruyter.

Nivre, Joakim, Željko Agić, Lars Ahrenberg et al. (2017). Universal dependencies 2.0 – CoNLL 2017 shared task development and test data. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics, Charles University. http://hdl.handle.net/11234/1-2184. See also http://universaldependencies.org/ (last access 14.12.2017).

Payne, Doris L. (ed.) (1992). *Pragmatics of Word Order Flexibility*. Amsterdam: John Benjamins.

Sapir, Edward (1921). *Language, an introduction to the study of speech*. New York: Harcourt, Brace and Co.

Shannon, Claude E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27 (3): 379–423.

Siewierska, Anna (1998). Variation in major constituent order: a global and a European perspective. In: Anna Siewierska (ed.), *Constituent Order in the Languages of Europe*, 475–551. Berlin: Mouton de Gruyter.

Straka, Milan, & Jana Straková (2017). *Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe*. In: *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, Vancouver, Canada, August 2017.

Vennemann, Theo (1974). Theoretical word order studies: results and problems. *Papiere zur Linguistik* 7: 5–25.

Apprehensive and avertive meanings and their functions of English and Polish

Barbara Lewandowska-Tomaszczyk

Schedule: We 12.30 Room 14

<pdf>

Perception verbs in Ancient Greek

Silvia Luraghi
(University of Pavia)

Schedule: Fri 14.30 Room 14

The construction of the verb *akoúō* 'hear' consitutes a puzzling issue for Ancient Greek descriptive grammars. This verb may take either an accusative or a genitive second argument, but why exactly one of the two cases is chosen is not always clear: while animate NPs consistently occur in the genitive, inanimates can occur in either case. Reference to the partitive or to the ablative meanings of the genitive (Chantraine 1953; Schwyzer/Debrunner 1950) does not seem to account for the data, and describing the alternation in terms of high (accusative) vs. low (genitive) affectedness or content (accusative) vs. source (genitive) of the information is not compelling in light of occurrences such as (1), in which an accusative and a genitive NP are coordinated.

(1) *mukēthmoû* *t'* *ēkousa* *boôn* ... *oiôn* *te*
lowing.GEN PTC hear.AOR.1SG cow.GEN.PL sheep.GEN.PL PTC
blēkhēn
bleating.ACC
'I heard the lowing of the cattle ... and the bleating of the sheep.' (Hom. *Odyssey* 12.265-266)

Crucially, the issue is usually approached in the framework of case variation, and *akoúō* is compared with other verbs that display the genitive/accusative alternation, rather than with other perception verbs. I argue that better insights can be gained by approaching the issue from the perspective of typological work on perception verbs (Viberg 1984, Evans/Wilkins 2000). Once one considers the groups of 'experiencer based' perception verbs (with a subject experiencer and a stimulus second argument), some common patterns emerge that single out different perception modalities. Verbs of smelling, touching and tasting (*osmáō*, *osphraínomai* 'smell', *háptomai, psaúō* 'touch, feel by touching', *geúomai* 'taste', Humbert 1986:273; Viti 2017) consistently take genitive stimuli, while verbs that indicate sight (*dérkomai*, *horáō*, *eîdon*, *blépō*) take accusative stimuli. Verbs that admit variation besides *akoúō* are other verbs of hearing, e.g. *klúō*. This distribution singles out verbs of sight (never taking the genitive), and is consistent with cross-linguistic data underlying Viberg's (1984:36) Modality hierarchy in (2).

(2) The modality hierarchy: sight > hearing > touch > ┌ smell
└ taste

The distribution of cases with hearing verbs reflects the in-between position of this perception modality and points to a fluctuation in its construal. Further comparison with other types of

experiential predicates also reveals a split. Genitive stimuli occur with verbs of bodily sensation (*peináō* 'be hungry', *dipsáō* 'be thirsty') and satiation (dative stimuli with this last group are beyond the scope of this paper), while accusative stimuli occur with verbs of mental activity, e.g. *noéō*, *oíomai*, *phronéō* 'think'. Of special interest is the verb *oîda* 'know', etymologically connected with *eîdon* 'see', which increasingly features accusative stimuli, with genitive stimuli virtually limited to occurrences where the verb indicates a skill rather than knowledge in Homeric Greek. Hence different cases featured by verbs surveyed here point toward a distinction between cognition and bodily sensations that underlies the hierarchy of perceptual modalities, and connects sight with cognition, while hearing has connections with both other experiential fields. These findings are in accordance with findings from genetically unrelated languages that have been described in typological research on perception verbs.

References

Chantraine, Pierre. 1953. *Grammaire Homérique* II: *Syntaxe*. Paris: Klincksieck.

Evans, Nicholas and David Wilkins. 2000. In the mind's ear: the semantic extension of perception verbs in Australian languages. *Language* 76/3: 546-592.

Humbert, Jean. 1986. *Syntaxe greque.* Paris: Klinksiek.

Schwyzer, Eduard and Albert Debrunner. 1950. *Griechische Grammatik* II, München: C.H. Beck.

Viberg, Åke. 1984. The verbs of perception: a typological study. *Linguistics* 21: 123-162.

Viti, Carlotta. 2017. Semantic and cognitive factors of argument marking in ancient Indo-European languages. *Diachronica* 34/3: 368–419.

Numeral classifiers in Udi: A unique contact-induced development among Nakh-Daghestanian?

Timur Maisak
(Institute of Linguistics & HSE University, Russia)

Keywords: numeral classifiers, language contact, Nakh-Daghestanian, Iranian, Turkic

Schedule: Sat 11.30 Room 2

Nakh-Daghestanian (a.k.a. East Caucasian) languages are head-final, with rich, mostly dependent-marking, morphology. Nouns inflect for number (singular and plural) and case. In most languages, there is a category of gender (noun class) which displays itself in agreement. Noun phrase dependents typically include adjectives, demonstratives, genitives, numerals and participial clauses. With numerals, nouns tend to occur in the singular. One thing which seems to be definitely lacking in the languages of the family are classifiers, including numeral classifiers. To my knowledge, the existence of such grammatical devices has never been described for any Nakh-Daghestanian language (neither in individual grammar sketches, nor in general cross-linguistic studies like Aikhenvald 2000 or Gil 2013). However, a small-inventory numeral classifier system does exist in at least one of the languages of the family, namely Udi, a language of the Lezgic branch. The paper will give an overview of the use and origin of the Udi classifiers, mostly based on the available texts in the Nizh dialect.

In Udi numeral phrases, the function word *dänä* can be employed with both human and non-human nouns, in the postposition to the numeral (1, 2). There is also a dedicated classifier for humans *tan*

'person', which is only used with numerals as well, but only as the noun phrase head (3), not in the "numeral – classifier – noun" construction.

(1) ***χib*** ***dänä*** *χüjär,* ***qo*** ***dänä*** *ʁar*
[three CLF daughter] [five CLSF son]
{My mother had eight children:} 'three daughters and five sons'.

(2) *göj-n-aχun* ***χib*** ***dänä*** *eʿšI□e* *bist:a*
sky-OBL-ABL [three CLSF apple]=3SG fall+PRS
'Three apples fall from the sky...'

(3) *beš* *ajl-in* *boš* ***uʿq*** ***tan****□jan*
we:GEN family-GEN inside [six PERSON]=1PL
'In our family, there are six people (lit. we are six)...'

Udi, which is historically spoken in northern Azerbaijan, has been heavily affected by language contact with genealogically unrelated languages, especially Turkic (Azerbaijani) and Iranian. Certain morphosyntactic properties of Udi can be definitely explained by contact influence, and the existence of numeral classifiers is no exception here. The small-inventory classifier system of Udi is exactly like those described by Stilo (2018) as typical of the Araxes-Iran area, covering the South Caucasus (Georgia, Armenia, Azerbaijan), Northern Iran, Northern Iraq, and Eastern Turkey. Both *dänä* (< 'seed, grain') and *tan* are Iranian loans, the first of which also functions as a neutral numeral classifier in Azerbaijani and Turkish. Although Stilo (2018) does not mention Udi or any other Nakh-Daghestanian languages, he includes Udi in the languages of the Araxes-Iran zone (see Stilo 2015).

In the paper, I will present the results of a corpus study of the Udi classifiers according to a set of parameters: in particular, the degree of obligatoriness (presence vs. absence in numeral phrases) and the frequency of occurrence with different numerals (e.g. 'one' vs. numerals higher than 'one') and with different nouns (e.g. animate vs. inanimate). Given that geographically Udi belongs to the periphery of the Araxes-Iran zone, the study will show whether it does also display 'fade-out' effects (Stilo's term) in the behaviour of classifiers.

References

Aikhenvald, Alexandra Y. (2000), *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Gil, David (2013), Numeral Classifiers, in Matthew S. Dryer, and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/55, Accessed on 2019-01-14.)

Stilo, Donald L. (2015), An introduction to the Atlas of the Araxes-Iran Linguistic Area, in Bernard Comrie, and Lucía Golluscio (eds.), (2015), *Language contact and documentation. Contacto lingüístico y documentación*. Berlin etc.: Walter de Gruyter, 343–355.

Stilo, Donald L. (2018), Numeral classifier systems in the Araxes-Iran linguistic area, in William B. McGregor, and Søren Wichmann (eds.), (2018), *The diachrony of classification systems*. Amsterdam: John Benjamins, 135–164.

The concept of young animal (Proto-Slavic **nt*-stems) and diminutivization in modern Slavic languages

Stela Manova, Dmitri Sitchinava & Maria Shvedova
(University of Vienna; Russian Language Institute & Higher School of Economics, Moscow; & Kyiv National Linguistic University)

Keywords: typology, diminutivization, derivation, inflection, Slavic

Schedule: We 15.00 Room 1

The literature on diminutivization in the world's languages reports a diachronic link between diminutive suffixes and linguistic items expressing the relation between 'the adult and the young' (Grandi 2011). Likewise, Ivanova-Mircheva & Xaralampiev (1999) claimed that in colloquial Old Bulgarian (Old Church Slavic in some sources) all diminutives belonged to the so-called *nt*-stems, an inflection class that comprised the young of animals. A few inflectional forms of a *nt*-noun follow:

(1) *tel-ę* 'calf-nom.sg'
tel-ęte 'calf-gen/loc.sg'

Ivanova-Mircheva & Xaralampiev's claim explains the existence of inflectional diminutives (2) in modern Bulgarian (Bg.):

(2) Bg. *meč-ka* 'bear' → *meč-e* 'teddy-bear & bear-dim' (*-e* is inflection)

Compare with Bg. *meč-ence*, either 'bear-dim' or 'bear-dim-dim', the latter meaning if *meč-ence* is seen as derived from *meč-e*. In other words, the theme *-ęt-* (1) developed into the very productive diminutive (i.e. derivational) suffix *-enc(e)* in modern Bulgarian, cf. *meč-enc-ence*.

Based on the Proto-Slavic **nt*-stems, Slavic nouns for young animals and their diminutive forms provide an excellent testing ground for investigation of the relation between the concept of young animal and diminutivization. That is, the goal of our research is to contribute to the better understanding of diminutivization as a morphological process.

We have collected and analyzed large sets of nouns for animals, diminutives of animals, nouns for young animals and diminutives of young animals (sample in (3)) in a number of Slavic languages. We will demonstrate that those forms, if seen as paradigmatically organized, allow for a number of interesting observations, both language-specific and typological. As shown in (3), Bulgarian has collapsed the difference between young animals and diminutives, but languages such as Russian (R.), Czech (Cz.) and Slovak (Sl.) keep the two groups of nouns strictly separate:

(3)	'stork'	DIM	young X	DIM of young X
	Bg. *štărkel*	*štărkel-če*	*štărkel-če*	*štărkel-č-ence*
	R. *aist*	*aist-ik*	*aist-enok*	*aist-enoč-ek*
	Cz. *čáp*	*čáp-ek*	*čáp-ě*	*čáp-átko*
	Sl. *bocian*	*bocian-ik*	*bocian-ča*	*bocian-č-atko*

An interesting position in the overall picture has Ukrainian. In this language, there are two different paradigms for young animals and diminutives, i.e. Ukrainian resembles Russian, Czech and Slovak (3) in this respect, but it is like Bulgarian (4) when it comes to the use of the suffixes for young animals as diminutivizers:

(4) a. Ukrainian
štan-en'-ata 'trousers-dim-pl'
(*-en'(a)* derives young animals, as in *koš-en'a* 'kitten')
b. Bulgarian
pantalon-če 'trousers-dim', plural *pantalon-č-eta*
(recall *štărkel-če* 'stork-dim & stork-young' in (3))

The type of diminutivization illustrated in (4) seems to be typical of plural nouns in Ukrainian, is not really developed in Russian and is impossible in Czech and Slovak. For a better understanding of the relation between young animals and diminutives, we also ran psycholinguistic experiments with native speakers.

Based on a contrastive paradigmatic analysis of the data, electronic corpora and the psycholinguistic experiments made, we conclude that the nature of the relation 'young animal-diminutive' is more complex than reported in the literature so far and, importantly, the direction of that relation is different in the different languages (contra Grandi 2011 and other scholars).

References

Grandi, Nicola (2011), Renewal and innovation in the emergence of Indo-European evaluative morphology, in L. Körtvélyessy, and P. Štekauer (eds.), (2011), *Diminutives and Augmentatives in the Languages of the World, Lexis: e-J. Engl. Lexicology* 6, 5–25.

Ivanova-Mircheva, Dora & Ivan Xaralampiev (1999), *Istorija na bălgarskija ezik / A History of the Bulgarian Language*, Veliko Tărnovo: Faber.

One stem or two? A corpus study of bipartite verb stems in Murrinhpatha

John Mansfield
(University of Melbourne)

Keywords: bipartite stems, complex predicates, grammaticalisation, predictability

Schedule: Fri 12.30 Room 6

Many languages around the world exhibit BIPARTITE VERB STEMS of various types (Bickel & Nichols, 2007, p. 187ff.; DeLancey, 2009). A recurrent problem for such verbs is deciding whether they should be treated as having a two lexical stems, one stem plus a particle, or a simplex (though discontinuous) stem (1–3). In this study I use corpus data to explore the problem in an Australian language, Murrinhpatha. Using combinatoric predictability measurements, I show that Murrinhpatha bipartite stems are distributed along a smooth cline from two-stem compounds to stem-particle constructions. As a whole, the Murrinhpatha verb system therefore does not belong to either type, but neither do we find natural cut-offs for dividing the lexicon into distinct types, at least based on distributional criteria. I propose that this corpus method should be extended to other languages with bipartite stems, to find out whether this is a general property of such systems.

(1) *Two-stem compound* (Klamath: DeLancey, 2009, p. 6)
ye=qew'i-wapk
with.foot=break.in.two-FUT
'will stomp and break in two'

(2) *Stem and particle* (Nimboran: Inkelas, 1993, p. 574)

patiái-rár-bá-k-u
hold.pig.feast.PL-PRT-LOC2-PST-1
'We held a pig feast above.'

(3) *Simplex stem* (Udi: Harris, 2002, p. 6)
bu-za-q'-sa
$love_1$-1SG-$love_2$-PRS
'I love her/him.'

Like several other languages of north-western Australia, Murrinhpatha has bipartite verb stems in which a small set of 'classifiers' combines with a large set of 'coverbs' (4) (Mansfield, 2019; McGregor, 2002; Nordlinger, 2015, 2017). Previous work on the bipartite stem nexus draws on semantic and syntactic evidence, such as compositionality and boundedness (e.g. Bril, 2007; Schultze-Berndt, 2000). But in this study I set these parameters aside, and focus on purely distributional criteria: Which elements occur in many combinations, and which are specific to a certain construction?

(4) *Classifier–Coverb* (Murrinhpatha)
pa-nhi-ngkapurl-nu
$slash_{CLS}$-2SG.O-sun.$rise_{COV}$-FUT
'The sun will come up on you.'

The distributional characteristics of the three types (1–3) can be schematised as in Figure 1. I investigate the prevalence of these types in a 100,000-word corpus of naturalistic Murrinhpatha discourse. However, rather than using a one/many distinction as in Figure 1, I use a gradient measure of combinatoric predictability, following established corpus methods for measuring co-occurrence patterns (e.g. Biber, Conrad, & Reppen, 1998). By this measure, Murrinhpatha bipartite verb stems are arrayed along a smooth cline from two-stem compounds to stem-particle constructions, with a few discontinuous simplex lexemes. Figure 2 illustrates the distribution of combinatoric predictability in 2,965 verb tokens extracted from the corpus. Starting either from coverbs or classifiers, we see that predictability of the second element has a normal distribution, with tokens clustered around the mean. I argue that this distribution presents a severe challenge for characterising the system according to recognised linguistic categories.

Figure 1. Schematic distributional types of bipartite stems.

(a) COV→CLS predictability (b) CLS→COV predictability

(c) COV→CLS x CLS→COV predictability

Figure 2. Distributions of Murrinhpatha bipartite stem predictability, measured by conditional entropy.

References

Biber, D., Conrad, S., & Reppen, R. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Bickel, B., & Nichols, J. 2007. Inflectional morphology. In T. Shopen (Ed.), *Language typology and syntactic description* (Second edition, Vol. 3, pp. 169–240). Cambridge: Cambridge University Press.

Bril, I. 2007. Nexus and juncture types of complex predicates in Oceanic languages: Functions and semantics. *Language and Linguistics*, *8*(1), 267–310.

DeLancey, S. 2009. Bipartite verbs in languages of western North America. In A. Filchenko & O. Potatina (Eds.), *Time and Space in Languages of Various Typology*. Tomsk: Tomsk State Pedagogical University.

Harris, A. C. 2002. *Endoclitics and the origins of Udi morphosyntax*. Oxford: Oxford University Press.

Inkelas, S. 1993. Nimboran Position Class Morphology. *Natural Language and Linguistic Theory*, *11*, 559–624.

Mansfield, J. B. 2019. *Murrinpatha morphology and phonology*. De Gruyter Mouton.

McGregor, W. 2002. *Verb classification in Australian languages*. Berlin: Mouton de Gruyter.

Nordlinger, R. 2015. Inflection in Murrinh-Patha. In M. Baerman (Ed.), *The Oxford handbook of inflection* (pp. 491–519). Oxford: Oxford University Press.

Nordlinger, R. 2017. The languages of the Daly River region (Northern Australia). In M. Fortescue, M. Mithun, & N. Evans (Eds.), *Oxford handbook of polysynthesis*. Oxford: Oxford University Press.

Schultze-Berndt, E. 2000. *Simple and complex verbs in Jaminjung* (PhD thesis). Katholieke Universiteit Nijmegen, Nijmegen.

The speed of change: Is grammaticalization faster when it has an analogical model?

Marie-Anne Markey
(KU Leuven)

Keywords: grammaticalization, analogy, language change, English, prepositions

Schedule: Fri 12.00 Room 14

While analogy recently gained popularity as one of the motivations and/or mechanisms for language change, its explanatory potential is still controversial (De Smet, 2016; Fischer, 2013). In this paper, we address the question whether the speed of language change can be adduced as evidence of analogy.

It has been proposed that a construction will change faster if it can model itself after a construction that has already taken the same path of change. Aaron (2016) provides evidence for such an effect from changes in Spanish modifiers. She finds a clear path of change, with a recognizable bridging context, for the first modifier to emerge but not for the ones that develop later. Instead, the latter seem to "borrow the path" of the earlier modifier (Aaron, 2016, p. 57). A similar mechanism has been suggested by Hoffmann (2004) for the grammaticalization of low-frequency complex prepositions. At the same time, the idea has been called into question. Brems (2007) argues against Hoffmann's version of analogy-driven grammaticalization, while Plank (2015) proposes that similar changes will proceed at a similar pace.

In order to test whether analogy can speed up language change, we carry out a corpus-study of three adnumeral markers in English: *about*, *above* and *over*. These prepositions also appear in decidedly unprepositional positions, when they modify numerical expressions, occurring for instance inside subjects, as in (1), or following prepositions, as in (2).

(1) ***About 151 causes*** *were standing for trial* (1851, Hansard Corpus)

(2) *The mail consisted* ***of over 3,400 bags*** (1911, Hansard Corpus)

About and *above* developed adnumeral uses in Middle English. *Over* became an adnumeral only in the 19th century, when an extensive adnumeral paradigm was already well-established. It is therefore predicted that *over* could develop much faster than *about* and *above*.

For *over*, data have been analysed from the *Hansard Corpus* for the British Parliament (n=1295). The results are consistent with the hypothesis, in that the development of *over* is by all appearances very fast indeed. First, there is no evidence for bridging contexts since the first attestations of adnumeral *over* already show a nearly complete range of usage contexts. Second, the change seems to have happened quickly: *over* enters the full syntactic range of adnumeral behaviour in under three decades and, in fact, before the major frequency gains in its use as adnumeral take place.

For *about* and *above*, preliminary results based on the *Penn-Parsed Corpus of Middle English (2nd ed.)* show a long stage of sporadic adnumeral use during (especially) Late Middle English, with use in some contexts (notably inside subjects and following other prepositions) remaining for a long time very rare. Further data analysis is needed, however, to confirm these results, systematically analysing the use of adnumeral *about* and *above* also in the *Innsbruck Middle English Prose Corpus*, to get better coverage of Late Middle English, and *Early English Books Online*, for the Early Modern period.

References

Aaron, J. E. (2016). The road already traveled: Constructional analogy in lexico-syntactic change. *Studies in Language*, *40*(1), 26–62. https://doi.org/info:doi/10.1075/sl.40.1.02aar

Alexander, Marc and Mark Davies. (2015-) *Hansard Corpus 1803-2005*. Available online at http://www.hansard-corpus.org.

Brems, L. (2007). The Grammaticalization of Small Size Nouns: Reconsidering Frequency and Analogy. *Journal of English Linguistics*, *35*(4), 293–324. https://doi.org/10.1177/0075424207307597

Davies, Mark. (2017) *Early English Books Online. Part of the SAMUELS project*. Available online at https://www.english-corpora.org/eebo/.

De Smet, H. (2016). How gradual change progresses: The interaction between convention and innovation. *Language Variation and Change*, *28*(1), 83–102. https://doi.org/10.1017/S0954394515000186

Fischer, O. (2013). An inquiry into unidirectionality as a foundational element of grammaticalization: On the role played by analogy and the synchronic grammar system in processes of language change. *Studies in Language*, *37*(3), 515–533. https://doi.org/info:doi/10.1075/sl.37.3.03fis

Hoffmann, S. (2004). Are low-frequency complex prepositions grammaticalized? In H. Lindquist & C. Mair (Eds.), *Corpus Approaches to Grammaticalization in English* (pp. 171–210). Amsterdam: John Benjamins Publishing.

Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. (2004) *The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)*. Department of Linguistics, University of Pennsylvania. CD-ROM, first edition, release 3 (http://www.ling.upenn.edu/ppche-release-2016/PPCEME-RELEASE-3).

Kroch, Anthony, and Ann Taylor. (2000) *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2).* Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, release 4 (http://www.ling.upenn.edu/ppche-release-2016/PPCME2-RELEASE-4).

Plank, F. (2015). Time for change. In C. Viti (Ed.), *Syntactic change and syntactic reconstruction* (pp. 61–92). Amsterdam: Benjamins.

Future tense reference and language contact. The reorganized verbal system of the Slavic of Molise

Antonietta Marra
(University of Cagliari)

Keywords: Language contact, future tense, modality, Slavic, grammaticalization

Schedule: Sat 12.30 Room 2

It is well known that future tense (FT) forms appear particularly unstable in verbal systems of many Indo-European languages (Banfi 1985). Also in Balkan region there is a strong variability (described synchronically and diachronically by, among others, Banfi 1985, Dahl, 2000, Tomić 2003, Vaillant 1966), which highlights the dynamic evolution of future forms.

In my presentation, I will show the evolution of FT in Slavic of Molise (SML), a Croatian dialect settled in a Romance environment in the South of Italy since the 16th century (Rešetar 1908). In SML, language contact played a further role in the restructuring process of FT forms, determining an enriched system with two periphrasis (while in other Balkan languages just one of them has been chosen), in addition to other forms borrowed from Romance languages of the area.
Nowadays in SML FT can be expressed principally with three forms:

a) *tit* 'want'+infinitive: *ču pisat* (lit. 'I want to write');
b) *imat* 'have/have to'+infinitive: *mam pisat* (lit. 'I have/have to write');
c) present tense: *sutr pisam list* (lit. 'tomorrow I write a letter').

We can find similar structures in language varieties of the Balkan area:

- (a) is very close to FT in Standard Croatian (with differences in full or clitic forms of the auxiliary); moreover, periphrastic future forms with 'want' auxiliaries can be found also in other Slavic languages of the area and in Greek;
- (b) is limited, nowadays, to Northern Albania (Banfi 1985, Dahl 2000), since typical Balkan form of future is a periphrastic structure with 'want' auxiliary (Vaillant 1966: 109), with different models of periphrasis (Tomić 2003).

Both periphrastic forms preserve modal meaning in addition to temporal one.

Even though (b) is present in some varieties of Balkan area, we have to explain this form in SML as the output of a separate evolution. Štokavian dialects (the group SML belongs to) presented *tit* auxiliary structure since the 13th century, that is much earlier than the arrival on the Italian territory.

Imati periphrastic future in SML, then, seems to be influenced by Romance languages contact, as in the same area of Southern Italy we can find outputs of a Latin sequence *habeo ad*+infinitive ('have to+infinitive') recorded and described by Rohlfs (1968 [1949]: 335) as a sort of FT, where a little obligation modality is still expressed.

My data confirm this hypothesis and offer evidences of the change in progress.

Distribution of the two periphrastic futures produced by habitual speakers show that the more conservative speakers use more often 'want' periphrasis with a solid temporal meaning, limiting 'have/have to' constructions to contexts with stronger modal meaning. On the contrary, less fluent, younger or innovative speakers tend to increase the use of 'have/have to' periphrasis, also in contexts where there is no obligation or necessity meaning.

The distribution of the two periphrastic futures (and also of the present tense form (c)) in different age groups of habitual speakers show the ongoing grammaticalization process of 'have/have to' periphrasis and present tense form, which are expanding their role in FT reference.

References

Banfi, Emanuele (1985), Isogrammatismo e calco linguistico: il cambio del futuro nelle lingue balcaniche, in L. Agostiniani/P. Bellucci Maffei/M. Paoli (eds), *Linguistica storica e cambiamento linguistico, SLI, Atti del XVI Congresso Internazionale di Studi, Firenze 7-9 Maggio 1982*, Roma: Bulzoni, 147-159.

Breu, Walter (1992), Das italokroatische Verbsystem zwischen slavischem Erbe und kontaktbedingter Entwicklung, in T. Reuther (ed.) *Slavistische Beiträge Band 292*, Slavistische Linguistik 1991: Referate des XV Konstanzer Slavistischen Arbeitstreffens, München: Sagner, 93-122.

Dahl, Östen (2000), The grammar of future time reference in European languages, in Ö. Dahl (ed), *Tense and Aspect in the Languages of Europe. Empirical Approaches to Language Typology*, EUROTYP 20–6, Berlin / New York: Mouton de Gruyter, 309-328.

Rešetar, Milan (1908), Die serbokroatischen Kolonien Süditaliens, *Anzeiger* 2 (also in: *Vorläufige Berichte der Balkankommission* 12).

Rohlfs, Gerhard (1968 [1949]), Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia, Torino: Einaudi.

Tomić, Olga Mišeska (2004), The syntax of the Balkan Slavic future tenses, *Lingua* 114(4), 517-542.

Vaillant, André (1966), *Grammaire comparée des Langue Slaves. Vol.3: Le Verbe*, Paris: Klincksiek.

Vaillant, André (1977), *Grammaire comparée des Langue Slaves. Vol. 5: La Syntaxe*, Paris: Klincksiek.

"Linking" morphology in Tupían, Cariban, and Macro-Jê languages

Florian Matter & Rik van Gijn
(University of Berne & University of Zurich)

Keywords: Morphology; South American languages; Je; Cariban; Tupian

Schedule: We 12.30 Room 8

Tupían, Cariban, and Macro-Jê are three of South America's larger language families. The former two are firmly established, but Macro-Jê with Jê proper as core has only gained traction more recently, with some of its suggested members still putative. Some languages of these families share an interesting morphosyntactic pattern: "linking" prefixes, which connect verbs, possessed nouns, and postpositions with immediately preceding arguments:

(1) Tapirapé

a. t–ãʔir–a
3–son.of.man–REF
'his son' (Neiva Praça 2007:57)

b. korinãkaʔi wãriniãiʔi r–ãʔir–a
K. W. LK–son–REF
'Korinãka'i is the son of Wãriniãy'i' (Neiva Praça 2007:33)

(2) Panare

a. n–ama–jah kən
3–knock.down–REC ANIM.DIST
'He/she knocked him/her down.' (Payne and Payne 2013:200)

b. toman j–ama–jah kən
T. LK–knock.down–REC ANIM.DIST
'He/she knocked Tom down.' (Payne and Payne 2013:202)

(3) Apinajé

a. əbri pa prɛ kəm arẽ
then 1.NOM PST 3.DAT tell
'And I said to her, […]' (de Oliveira 2005:321)

b. iŋ–mə mɛbɔj j–arẽ pa ku–ba
1–DAT something LK–tell 1 3–hear
'Say something for me to listen.' (de Oliveira 2005:164)

In Tapirapé (Tupían), third person possessors are expressed by *t-* on the possessum (1a), but when expressed by a preceding NP, with *r-* (1b). This is called the **linker**, connecting dependent and head. In Panare (Cariban), the same pattern can be found for P arguments of transitive verbs, with *n-* appearing with no overt NP (2a) and *j-* otherwise (2b). Similarly, Apinajé (Jê) shows alternation in transitive verbs between Ø and *j-* (3).

These patterns are at the core of the TuCaJê hypothesis, suggesting a distant relationship between the three families (Rodrigues 2000, 2009). Not much other evidence exists besides some purported lexical and person marking correspondences (Rodrigues 1985). What is more, while linking mor- phology is a stable feature of Tupí-Guaraní, it is attested in only a few Tupían languages outside the Tupí-Guaraní branch, and is less robustly present in Cariban and Jê proper (even less so in peripheral Macro-Jê languages). This raises the question whether the similarities between the linking systems in the three families are strong enough to warrant postulating a deep-time TuCaJê connection.

In order to answer this question, we tested a representative sample of languages from all three families for the presence, form, morphosyntactic distribution, and diachronic origin of such markers, as well as the languages' geographic distribution. While function and distribution, and partly also form, are similar or identical across the three language families, the historical origins look rather dif- ferent. Only few Cariban languages have linkers, but vowel alternations in extant languages allow reconstruction of a linker **j-*, possibly stemming from an older third person marker (Meira, Gildea, et al. 2010). For Jê, the source of the alternation likely lies in root-initial consonant mutation(s), supporting Salanova (2009). Moreover, while some peripheral Macro-Jê languages with linkers show promising but sparse sound correspondence (Ribeiro 2012), others have functionally and formally different patterns (Ribeiro & Voort 2010). For Tupían, a case can be made for reconstructing relational marking to Proto-Tupí (Rodrigues & Cabral 2012), although this is contested (Meira & Drude 2013).

These diachronic scenarios speak against relational markers as an inherited feature and are there- fore not indicative of a genealogical relationship. However, the similarities in function and distribu- tion might be indicative of contact between speakers of proto-languages at an earlier stage, possibly accompanied by similar syntactic structure supporting the development of linking morphology.

References

de Oliveira, Christiane Cunha (2005). "The language of the Apinajé people of Central Brazil". Ph.D. dissertation. Oregon: University of Oregon.

Meira, Sérgio & Sebastian Drude (2013). "Sobre a origem histórica dos "prefixos relacionais" das línguas Tupí-Guaraní". In: *Cadernos de Etnolinguística* 5.1.

Meira, Sérgio, Spike Gildea & B. J. Hoff (2010). "On the Origin of Ablaut in the Cariban Family".

In: *International Journal of American Linguistics* 76.4: 477–515.
Neiva Praça, Wilkíria (2007). "Morfossintaxe da Língua Tapirapé (Família Tupí-Guaraní)". Ph.D. dissertation. Universidade de Brasília.
Payne, Thomas E. & Doris L. Payne (2013). *A Typological Grammar of Panare: A Cariban Language of Venezuela*. Brill's Studies in the Indigenous Languages of the Americas. Leiden: Brill.
Ribeiro, Eduardo Rivail (2012). "A Grammar of Karajá". Ph.D. dissertation. University of Chicago.
Ribeiro, Eduardo Rivail & Hein van der Voort (2010). "Nimuendajú was right: The inclusion of the Jabutí language family in the Macro-Jê stock". In: *International Journal of American Linguistics* 76.4: 517–570.
Rodrigues, Aryon D. (1985). "Evidence for Tupi-Carib relationships". In: *South American Indian languages: Retrospect and Prospect*. Harriet E. Manelis Klein & Louisa R. Stark (eds.). University of Texas Press: 371–404.
Rodrigues, Aryon D. (2000). "'Ge-Pano-Carib' x 'Jê-Tupí-Karib': sobre relaciones lingüísticas prehistóricas en Sudamérica". In: *Actas del I congreso de lenguas indígenas de Sudamérica*. Vol. 1: 95–104.
— (2009). "A case of affinity among Tupí, Karíb, and Macro-Jê". In: *Revista Brasileira de Linguística Antropológica* 1.1: 137–162.
Rodrigues, Aryon D. & Ana Suelly Arruda Câmara Cabral (2012). "Tupían". In: *The Indigenous Languages of South America: A Comprehensive Guide*. Lyle Campbell & Verónica Grondona (eds.). Vol. 2. The World of Linguistics. Berlin/New York: Mouton de Gruyter: 495–574.
Salanova, Andrés Pablo (2009). "Não existem prefixos relacionais nas línguas Jê". In: *Línguas e Culturas Macro-Jê*. Silvia Lucia B. Braggio & Sinval M. de Sousa Filho (eds.). Vol. 1. Goiânia: Gráfica e Editora Vieira: 259–271.

Middle English Open Syllable Lengthening produced canonical word form shapes

Theresa Matzinger, Magdalena Schwarz & Nikolaus Ritt
(University of Vienna)

Keywords: open syllable lengthening, schwa loss, Middle English, sound change, word frequencies

Schedule: Thu 9.30 Room 9

Our paper asks a new question about a sound change known as Middle English Open Syllable Lengthening (OSL; Ritt 1994, Myungsook 1993, Lahiri & Dresher 1999). This durational change affected non-high short vowels in open disyllables such as *name* /namə/ > /na:m(ə)/, *hope* /hopə/ > /hɔ:p(ə)/, or *beaver* /bevər/ > /bɛ:v(ə)r/. Crucially, it is consistently reflected only in words like *name* or *hope*, whose second syllables ended in schwa and were lost. However, words like *beaver*, which have retained their final syllable, reflect OSL only sporadically, and those that do are of a specific subtype: their second syllable begins with an obstruent, ends in a sonorant, and admits schwa syncope (as in ModE [bi:vr̩] for *beaver*). This has given rise to the theory (Minkova 1982, Bermudez-Otero 1998) that OSL compensated for weight loss in post-tonic syllables.

While the compensatory account of OSL is descriptively adequate, it still raises the question what the motivation of the compensation was. Our paper addresses that question and tests the hypothesis that the lengthenings in /CVCə/ items that lost their schwas made them conform, in terms of weight, to

the majority of monosyllabic word forms that existed at the time of the change (Mailhammer, Kruger & Makiyama 2015). Extending a previous study based on the PPCME, we report a quantitative analysis of words forms attested in the LAEME corpus, which covers the period in which schwa loss and OSL unfolded. We extracted major class word forms that were not inputs or outputs of OSL, and determined the frequency and the morphological structure of types such as CVC, CVCC, CVVC, CVCV, CVVCV, CVCCV, etc. Specifically, we looked (a) at the relative (type and token) frequencies of morphologically simple CVC and CVVC monosyllables, and (b) at the relative frequencies of morphologically simple CVCV(C) and CVVCV(C) vs. morphologically complex CVC+V(C) and CVVC+V(C) disyllables. We show that at the time when schwa loss and OSL began to spread, CVVC and CVCC forms were indeed significantly more frequent than CVC types among monosyllabic major class words, but only if the vowels were mid or low. Among words with high vowels, CVC items were not less frequent than CVVC items. For disyllables, preliminary results suggest that the majority of items with long vowels in their first syllables were morphologically complex, i.e. CVVC+V(C).

Thus, the way in which OSL was implemented had two effects. In monosyllables resulting from schwa loss, the consistent lengthening of non-high vowels adapted their word shapes to conform to the majority pattern. In stable disyllables, on the other hand, lengthening would have produced word forms shapes that were typical of complex rather than simple word forms, and may have failed to affect them for that reason.

Apart from potentially deepening our understanding of OSL, our presentation proposes that the implementation of specific sound changes may be affected by preferences for word form shapes to conform to canonical patterns, and to assume patterns that signal their morphological structure.

References

Bermúdez-Otero, Ricardo (1998), Prosodic optimization: the Middle English length adjustment, *English Language and Linguistics* 2, 169-197.

Kim, Myungsook (1993), On lengthening in the open syllables of Middle English, *Lingua* 91, 261-277.

Lahiri, Aditi &, B. Elan Dresher (1999), Open syllable lengthening in West Germanic, *Language* 75, 678-719.

Mailhammer, Robert, Kruger, William, & Alexander Makiyama (2015), Type frequency influences phonological generalizations: eliminating stressed open syllables with short vowels in West Germanic, *Journal of Germanic Linguistics* 27, 205-237.

Minkova, Donka (1982), The environment for open syllable lengthening in Middle English, *Folia Linguistica Historica* 3, 29-58.

Ritt, Nikolaus (1994), *Quantity adjustment. Vowel lengthening and shortening in Early Middle English*, Cambridge: Cambridge University Press.

Development of verb morphology in Russian: From item-specificity to proficient use

Jekaterina Mažara & Sabine Stoll
(University of Zurich)

Keywords: language acquisition, longitudinal study, aspect, Russian

Schedule: Thu 10.00 Room 1

Usage-based accounts of language acquisition propose an item-specific stage during which grammatical forms and syntactic constructions appear with a limited number of lexical items (e.g. Lieven et al. 1997; Tomasello 2003). The Aspect Hypothesis (see Shirai & Andersen 1995) posits that early aspect use is characterized by a strong correlation between aspect and tense. It was shown for several languages that children preferably use *imperfectives+present* and *perfectives+past tense*; these correlations are also present in adults, but to a lesser degree (for an overview: Li & Shirai 2000).

Here, we analyze the development of the verb-system in Russian preschoolers. In Russian, aspect is a central feature; each verb is obligatorily either imperfective or perfective. Perfective verbs are the marked members of the binary opposition. The goal of this study is to (i) establish phases in the development of verb use and (ii) compare the lexical and grammatical flexibility of verb-forms in the two aspects during these phases. Our **data** stems from a longitudinal corpus of weekly recordings (1h) of four Russian children (recorded for 24-36 months, aged 1;3-1;11 at the start of recordings). The corpus contains 2 million morphologically annotated words including target children and surrounding speakers.

In **analysis 1**, we establish phases in acquisition by conducting a segmented regression on the growth curves of the verb-form increase in each child's inventory. The analysis of the form-use in the individual phases reveals that during phase 1, children use forms item-specifically in correspondence with predictions made by the Aspect Hypothesis. Perfective verbs tend to occur in the past and as imperatives, while imperfectives occur mostly with non-past and infinitive morphology.

In **analysis 2**, we focus on the development during phase 2 by computing the entropy of lexical items and grammatical markers used with verbs of both aspects over time. This allows us to evaluate how quickly children approach the flexibility of use displayed by surrounding adults. For verb-forms of both aspects we see a quick increase in entropy towards the level of surrounding adults and a rapid development of form-use across the entire paradigm, implying that the children in our sample start using verbs of both aspects with a variety of forms early during phase 2; they rapidly abandon the item-specific tendencies of their early production.

Our results confirm earlier findings of an association between aspect and tense in the earliest phase of language acquisition (Stoll & Gries 2009). However, we show that Russian children display a high degree of flexibility of combinations of stem+morphemes across the entire paradigm early on and are not restricted to the use of certain forms with a specific aspect. The Aspect Hypothesis holds only for the earliest form use and might not reflect the greater saliency of the combination of *perfective+past* or *imperfective+present*, but instead merely mirror adult input distributions. Children quickly converge towards the distributions found in adult input. We discuss whether this can be analysed as an effect of semantic saliency or merely of distributional cues in the input.

References

Li, Ping, and Yasuhiro Shirai. 2000. *The Acquisition of Lexical and Grammatical Aspect*. Walter de Gruyter.

Lieven, Elena V. M., Julian M. Pine, and Gillian Baldwin. 1997. Lexically-Based Learning and Early Grammatical Development. *Journal of Child Language* 24.1: 187-219.

Shirai, Yasuhiro. and Andersen, Roger. W. 1995. The acquisition of tense-aspect morphology: a prototype account. *Language*, 71:743–762.

Stoll, Sabine, and Stefan Th. Gries. 2009. How to measure development in corpora? An association-strength approach to characterizing development in corpora. *Journal of Child Language*, 36:1075–1090.

Tomasello, Michael. 2003. Constructing a Language: A Usage-Based Theory of Language Acquisition. *Harvard University Press.*

The role of co-referentiality in Spanish *que*-deletion

Giulia Mazzola, Malte Rosemeyer & Bert Cornillie

Schedule: Fri 16.00 Room 14

In classical Spanish, *que* 'that'-deletion, e.g. (1), was quite frequent with a wide range of predicates, whereas in contemporary Spanish, *que* can only be omitted in subjunctive complement clauses mainly following volitional verbs.

(1) Por ende vos rogamos [Ø] le dedes entera fee y creencia. (anon. 1497)
'Thus we beg you [that] you give him full trust and belief'

Pountain (2015) suggests that *que*-deletion is a Latin calque typical of formal genres in the 16th century, which slowly disappears in the 17th century, and that the type of matrix verb governing *que* has an important influence on the alternation. Blas Arroyo & Miralles (2015, 2016) indicate that omission is especially common in lower prestige authors.

This paper proposes a different parameter to be the main predictor of the alternation, namely the use of the subject in the subordinate clause (Spanish is a pro-drop language) and whether or not the subject is co-referential with the indirect object of the main clause (e.g. co-referentiality of *vos* 'you' and the subject of *dar* in (1)). In line with Givón's (2001) complementation scale, we claim that co-referentiality between subject and indirect object leads to increased syntactic integration, which then allows for the deletion of the complementizer.

The parameter of co-referentiality is related to covert expression of the subordinate subject, as with ditransitive *rogar* 'request' in (1). The *Esbozo* grammar (RAE 1973) mentions that deletion is possible with ditransitive verbs exclusively when the subordinate clause has no explicit preverbal subject. Blas Arroyo and Miralles (2015) confirm this tendency for the non-ditransitive verb *creer* ('to believe'), emphasizing the adjacency of the main and the subordinate predicates as a basis for omitting the complementizer. However, it has been shown that with many other verbs deletion was favored by an overt subject, which is then claimed to mark the boundaries between the clauses, making the presence of the conjunction redundant (see also Pountain 2015:79). Hence, both absence and presence of an overt subject seem to play a role in *que*-deletion depending on whether it concerns a ditransitive construction or not.

We conduct a logistic regression analysis of some 3000 tokens from the historical CODEA+-corpus (GITHE, 2015) in order to predict the relationship between the dependent binary variable (*que* vs ø) and multiple independent variables (presence and position of overt subjects, presence and type of objects of the main verb --clitic pronoun, lexical expression, stressed pronoun, doubling-- that co-refer with the subject of the subordinate verb), so as to account for the origins, spread and decline of *que*-deletion.

Our results indicate a clear distribution of *que*-deletion along a co-referentiality cline, from marking the indirect object (with different types: clitic/lexical/doubling) without overtly marking the subordinate subject to absence of an indirect object with presence of an overt subject. Thus, the paper highlights that an explanatory analysis of *que*-deletion needs to go beyond the parameter of the lexical semantics of the matrix verb (volition, fear), analyzing also the prominence of its objects.

References

Blas Arroyo, J. L., & Porcar Miralles. (2015). Tiene tanto temor a la mar que creo no lo hará: variación en la sintaxis de las completivas en los Siglos de Oro. In *Actas del X Congreso Internacional de Historia de la lengua española.* Zaragoza.

Blas Arroyo, J. L., & Porcar Miralles, M. (2016). Un marcador sociolingüístico en la sintaxis del Siglo de Oro: patrones de variación y cambio lingüístico en completivas dependientes de predicados doxásticos. *Revista internacional de lingüística iberoamericana*, (28), 157–185.

GITHE. (2015). Codea+ 2015. Retrieved 10 January 2019, from http://www.corpuscodea.es/

Givón, T. (2001). Verbal complements and clause union. In *Syntax: an introduction* (Rev. ed., Vol. 2). Amsterdam: Benjamins.

Pountain, C. (2015). Que-deletion: the rise and fall of a syntactic fashion. In *Estudios dedicados a David Mackenzie* (p. 17). Universidade de Santiago de Compostela.

Real Academia Española. Comisión de Gramática. (1973). *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.

Integrating "Exotic Objects" in Baltic and Slavic

Imke Mendoza
(University of Salzburg)

Keywords: Historical Linguistics, Languagec Conctact, Differential Object Marking, Slavic, Baltic, Old Russian

Schedule: Sat 11.00 Room 2

One of the features of the Eastern Circum-Baltic Area is the nominative object construction (NOC). In certain syntactic environments such as impersonal constructions, non-finite predications or imperative constructions, the object of a NOC shows the nominative case instead of the accusative (or genitive). Examples are given in (1)-(3):

(1) Old Russian, birch bark, Staraja Russa No. 40 (www.gramoty.ru)
davati mi doci
give.INF 1SG.DAT daughter.NOM
'I have to get my daughter married'

(2) East High Lith. dial. (Ambrazas 2001, 391)
Reikia šienas grėbti
need.3PRS hay.NOM rake.INF
'It is necessary to rake the hay'

(3) Finnish (Timberlake 1974, 158)
Minun täytyy kirjoittaa kirje.
1SG.GEN necessary write.INF1 letter.NOM
'It is necessary for me to write the letter'

NOCs are attested in West Finnic, East Baltic and in some historical and dialectal varieties of Russian. Baltic and Slavic are genetically closely related, West Finnic is completely unrelated to either one. In West Finnic, NOCs occur on a regular basis, whereas they are rather infrequent in Baltic. Standard Russian has lost this construction altogether, but it can still be found in some northwestern dialects. From an Indo-European perspective, NOCs are quite exotic; there are no attestations in other Indo-European languages outside of Baltic and Russian. Their origin in the latter languages is without much doubt the result of early contact of the Baltic and Slavic populations with speakers of Finnic languages.

Constructions with a nominative object share many common features in said languages. However, there are also striking differences. This is particularly true of the so-called "animacy constraint" (Timberlake 1974), according to which animate objects are less prone to take the nominative case than non-animate objects. In West Finnic and Baltic this restriction concerns only personal pronouns. In Russian, however, it is far more prominent and also applies to nouns that belong to the animate (sub)gender. This is illustrated by (4), which constitutes a syntactic environment very similar to (1), but, unlike the latter, has its direct object marked by the accusative, since *voevoda* 'voivode' is animate:

(4) Old Russian (AAĖ, 29)

i	*mně*	*poslatь*	*svoego*	*voevodu*	*sъ*	*tvoimъ* *voevodoju*
and	1SG.DAT	send.INF	REFL.ACC	voivode.ACC	with	your.INS voivode.INS

'and I have to send my voivode together with your voivode'

In my presentation, I will discuss the similarities and differences of NOCs between Russian, East Baltic and West Finnic from the perspective of the overarching questions of the workshop. I will show that the pivotal factor in the language specific development of NOCs is not the genetic affiliation of the languages in question, but rather their idiosyncratic grammatical make-up, such as the existence of a grammatical category of animacy.

References

AAĖ = Akty arxeografičeskoj ėkspedicii (1836), in: *Akty, sobrannye v bibliotekaxъ i archivaxъ Rossijskoj Imperii arxoegrafičeskoju ėkspedicieju Imperatorskoj Akademii Nauk* 1,. Sankt-Peterburg, 1294-1598.

Ambrazas, Vytautas (2001), On the Development of the Nominative Object in East Baltic, in Ö. Dahl and M. Koptjevskaja-Tamm (eds.), *Circum-Baltic Languages. Vol. 2: Grammar and Typology*, Amsterdam / Philadelphia: John Benjamins, 391-412.

Timberlake, Alan (1974), *The Nominative Object in Baltic, Slavic and West Finnic*, München: Sagner.

Inflectional complexity of Slavic "small languages"

Thomas Menzel
(Sorbisches Institut Bautzen)

Keywords: Slavic languages, Upper and Lower Sorbian, Inflectional complexity, Dual number, Small languages

Schedule: We 17.00 Room 1

The label "small language" is employed here in accordance with sociolinguistic features, which are described by Trudgill (2011): an isolated, low contact language with tense social networks, which is actually spoken by a small number of speakers with highly convergent informational background and shows only an insignificant number of L2-learners. It is frequently suggested that languages, which pertain to these attributes, reveal high grammatical complexity in their inflectional structures. Nevertheless, the correlation of "small languages" and "grammatical complexity" is helpful only in the case, that we dispose of reliable knowledge about the sociolinguistic situation of the given linguistic community in the past. Deductions as "there is a language with complex structure, so it must have been an isolated small language somewhere in the past" lead directly into a vicious circle.

My concern is on West Slavic languages – Upper and Lower Sorbian in East Germany as compared with Polish. Their contemporary inflectional systems show significant differences with respect to inflectional complexity. There are three items to be dealt with in this paper: 1) Both Sorbian languages have preserved a dual number, but not Polish. However, dual inflections seem to be more conservative in Lower Sorbian than in Upper Sorbian, the latter being less complex. 2) A virile : non-virile subgender opposition is realized in the nominative plural of nouns by inflections in Polish and Upper Sorbian, but not in the literary variety of Lower Sorbian. Lower Sorbian seems to be less complex at this point. 3) Paradigm structures contrast virile and non-virile subgender in the accusative plural (resp. dual) in Polish and Upper Sorbian, but masculine-animate and non-masculine-animate subgender in Lower Sorbian. Taking into account, that subgender is realized only in certain syntagmatic contexts for plural, but generally for dual, the inflectional structures of Lower Sorbian seem to be the most complex once more.

The diachronic development of these inflectional structures may be traced back to the 14th/15th (Polish) and 17th centuries (Sorbian). But what do we know about the social situations of the speakers of Polish, Upper and Lower Sorbian in the past? The two Sorbian languages are affected by intense language contact with German, but they did never appear as subjects of L2-learning on a large scale. In this sense, they mate with the sociolinguistic features of the "small language"-label better than Polish. However, the Sorbian languages differ significantly in inflectional complexity. If we want to test the correlation between sociolinguistic "small language"-features and grammatical complexity, we need to discover social differences in the history of the three linguistic communities which could have given rise to distinct grammatical developments. Therefore, the two main issues in this paper will be 1) to show that differences in inflectional complexity of the three languages are not random, and 2) that there are indeed sociolinguistic differences in the histories of Sorbian and Polish, but also between Upper and Lower Sorbian respectively, which can serve as reference points for the evaluation of inflectional complexity.

References

Trudgill, Peter (2011), *Sociolinguistic typology. Social determinants of linguistic complexity*, Oxford: University press.

Increasing the valency of motion verbs

Maria Messerschmidt
(University of Copenhagen)

Keywords: voice, valency-change, motion verbs, typology, applicative

Schedule: We 15.00 Room 6

Typologies of voice or valency-change usually consider two main types of valency-increasing operations: causatives and applicatives (e.g. Haspelmath & Müller-Bardey 2004, Kulikov 2011). But a number of constructions in various languages seem to fall between these two categories, e.g. the portative construction, exemplified in (1b), which typically takes intransitive motion verbs as a base. Conceptually the portative construction contains an element of (co-motional) causation, making it similar to a causative construction. But unlike the causative construction, exemplified in 1c, the it is not the *causer* argument that is introduced to the argument structure, but rather the *causee*, since the

prototypically inanimate *causee* could not be the agentive subject of the base intransitive construction containing a motion verb. Nor is the portative an example of a prototypical applicative construction.

(1) Caddo (Melnar 1998: 170)
a. *ci-ʔa=d(ih)-ʔaʔ*
1A-go-FUT
'I will go.'
b. *ci-ni-ʔa=d(ih)-ʔaʔ*
1A-PORT-go-FUT
'I will take it.'
Purepecha (Capistrán-Garza 2015: 152)
c. *ni-tára-a-ka=ni* *tumpí-ni*
go-CAUS-FUT-1=1SG.S boy-O
'I will make the boy go.'

In this talk, I will present the results of a cross-linguistic study of 50 genealogically diverse languages which looks at the various ways the valency of motion verbs can be increased. Valency-increasing operations are almost always lexically restricted to certain verbs or verb classes and interact with these in different ways. The talk therefore focuses on a single class of verbs which can provide good examples of a greater variety of operation types than we are perhaps used to considering. The valency-increasing operations in the sample languages have been analysed and classified in order to discover the range of operation types that occur with motion verbs and determine whether the common categories of causative and applicative are sufficient to describe this range operation types.

Based on the results of the study, I will argue that we should expand our standard inventory of valency-increasing operation types beyond causatives and applicatives and consider interactions with verb classes, differences in argument structure and complex event structure as well as conceptual differences between the operation types in out classifications.

References

Capistrán-Garza, A. (2015) *Multiple object constructions in P'orhépecha: argument realization and valence-affecting morphology*. Leiden; Boston: Brill.

Haspelmath, M. & Müller-Bardey, T. (2004) Valency change. In G. E. Booij, C. Lehmann, & J. Mugdan (eds.) *Morphologie / Morphology. Ein Internationales Handbuch Zur Flexion und Wortbildung* (An International Handbook on Inflection and WordFormation). Berlin/Boston: De Gruyter, pp. 1130–1145.

Kulikov, L. (2011) Voice Typology. In J. J. Song (ed.) *The Oxford Handbook of Linguistic Typology*. Oxford University Press.

Melnar L.R. (1998) *Caddo verb morphology*. Ph.D. dissertation. Chicago: University of Chicago.

Differential place marking in creole languages

Susanne Maria Michaelis
(Leipzig University & MPI-SHH, Jena)

Schedule: We 12.30 Room 6

It seems to be a robust empirical observation that in locational constructions, different kinds of landmarks receive differently elaborated coding (e.g. Comrie 1986, Aristar 1997, Haspelmath 2019).

In the examples (1)-(3) from French, the place name (*Place Vendôme*), the (typical landmark) inanimate common noun ('market'), and the human noun ('grandmother') are coded differently in that in (1) there is no allative marker, in (2) there is a shorter marker, and in (3) there is a longer marker:

Place name
(1) *Je vais Ø Place Vendôme.*
'I'm going to Place Vendôme.'

Inanimate common noun
(2) *Je vais **au** marché.*
'I'm going to the market.'

Human noun
(3) *Je vais **chez** ma grand-mère.*
'I'm going to my grandmother.'

Such coding patterns are not random, but can be subsumed under the following scale in that in a given language, human nouns tend to be coded with more (or at least the same amount of) linguistic material than inanimate common nouns, and place names tend to be coded with the least linguistic material (often zero):

Scale of differential place marking
(4) human nouns > inanimate common nouns > place names (Haspelmath 2019, Stolz et al. 2014)

That locational markers with human landmarks tend to be longest compared to inanimate nouns and place names can be seen as a functional response to the need to highlight rarer, less predictable constructions, because human nouns serve as landmarks least often, while place names are very often landmarks (with inanimate nouns being in between).

In this talk, I claim that data from high-contact languages, such as pidgins and creoles, support this universal generalization (see Michaelis & APiCS Consortium 2013), as illustrated by Seychelles Creole:

Place name
(5) *Apre ou 'n al **Ø** Sent Ann.*
then 2SG PRF go Saint Anne
'Then you went to Sainte Anne.' (Michaelis & Rosalie 2013)

Inanimate common noun
(6) *Mon al **dan** bwa.*
1SG go in forest
'I go into the forest.' (ibid.)

Human noun
(7) (...) *mon al **kot** sa zonm la.*
(...) 1SG go at DEM man there
'(...) I went to this man.' (Bollée & Rosalie 1994: 152)

Other creole languages have similar patterns, always in line with the differential place marking scale. The following table shows the allative marking in some languages (even Tok Pisin, where all landmarks take the same marker *long*, is covered by the scale).

	Place name	Inanimate common noun	Human noun
Guadeloupean Creole	Ø	*anba*	*aka*
Hawai'i Creole	Ø	*tu*	*by*
Juba Arabic	Ø/*fi*	Ø/*fi*/*le*	*le*
Kriol	Ø	*la*/*langa*	*langa*
Tok Pisin	*long*	*long*	*long*

Interestingly, creoles tend to not continue the markers of their lexifiers, but instead they show freshly grammaticalized markers from lexical material of the lexifiers (Guad. *anba* < French *en bas* 'underneath', *aka*(*z*) < *à* (*la*) *case* 'at (home)'; Kriol *langa*, Tok Pisin *long* < English *along*). Despite of the considerable amount of restructuring and refunctionalization processes due to heavy language contact in the course of creolization, these relatively young languages support the universal trend of differential place marking.

References

Aristar, Anthony Rodrigues. 1997. Marking and hierarchy types and the grammaticalization of case-markers. *Studies in Language* 21(2). 313–368. doi:10.1075/sl.21.2.04ari.

Bollée, Annegret & Rosalie, Marcel. 1994. *Parol ek memwar*. *Récits de vie des Seychelles*. Hamburg: Buske.

Comrie, Bernard. 1986. Markedness, grammar, people, and the world. In Fred R. Eckman, Edith A. Moravcsik & Jessica R. Wirth (eds.), *Markedness*, 85–106. New York: Plenum.

Haspelmath, Martin. 2019. Differential place marking and differential object marking. STUF–Language Typology and Universals.

Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog & Elif Bamyacı. 2014. Coding causal-noncausal verb alternations: A form-frequency correspondence explanation. *Journal of Linguistics* 50(3). 587–625.

Michaelis, Susanne Maria & Marcel Rosalie. 2013. Seychelles Creole structure dataset. In Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Michaelis, Susanne Maria & the APiCS Consortium. 2013. Going to named places. In Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). 2013. *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

Stolz, Thomas, Sander Lestrade & Christel Stolz. 2014. *The crosslinguistics of zero-marking of spatial relations.* Berlin: De Gruyter Mouton.

Epistemic modality and grammatical mood: An investigation of Galician adverbs

Vítor Míguez
(Instituto da Lingua Galega, University of Santiago de Compostela)

Keywords: adverbs, epistemic modality, Galician, indicative, subjunctive

Schedule: We 14.30 Room 11

Epistemic modality is a functional domain concerned with the degree of likelihood of a state of affairs. Particularly, it involves an estimation of probability made by the speaker (Nuyts 2001). Grammatical mood is a formal device of the verb in many languages, associated to semantic as well as syntactic distinctions (see e.g. Malchukov & Xrakovskij 2016). Although Romance languages are experiencing a loss of subjunctive temporal forms (Harris, 1974), the use of the subjunctive with epistemic adverbs seems to be increasing (Cornillie, 2016; Houle & Martínez Gómez 2011). In Romance languages the subjunctive is usually restricted to subordination contexts, but some exceptions exist. When an epistemic adverb precedes a finite verb, grammatical mood may alternate between indicative and subjunctive, as in series (1) from Galician.

(1) a. *Probablemente* esta **será** a última vez que visite esta cidade.
'This will probably be (IND) the last time I will visit this city.' (CORGA, PROBNARR0585)
(2) b. Mañá *probablemente* **sexan** só tres os que estean traballando tralo mostrador.
'Tomorrow there will probably be (SUBJ) only three people working behind the counter.' (CORGA, PROBNARR0054)

In both (1a) and (1b) the verb *ser* 'be' appears after the adverb *probablemente* 'probably' and refers to a situation in the future. However, in (1a) the future indicative is used whereas in (1b) the present subjunctive appears instead. Interestingly, both forms could be swapped (adjusting number agreement) and the result would be perfectly grammatical. To account for this apparently free alternation, several syntagmatic factors have been invoked, and it has been found that negation, past temporal reference, and adjacency between adverb and verb favor the use of subjunctive in Spanish (Finanger 2011 Yelin & Czerwionka 2017). Some authors (e.g. Givón 2001: 314–315) argue that the use of the subjunctive entails a lower degree of certainty as compared to the use of the indicative. Nevertheless, in Yelin and Czerwionka's (2017) experiment no significant difference was found between certain and not certain contexts regarding mood selection.

This paper examines the behavior of four Galician adverbs with different epistemic values (*posiblemente*, *probablemente*, *seguramente*, and *quizais*) in relation to mood selection. Samples were extracted from CORGA, the biggest corpus for present-day Galician, and split into two written genres. Each sample contains 100 occurrences of an adverb in a particular genre. Analytical categories include polarity, verb number and person, temporal reference and mood of the verb form, position of the adverb within the clause, control and dynamicity.

The results show that time reference, epistemic strength of the adverb, and dynamicity are useful predictors of mood selection: the subjunctive is favored by non-past time reference, weak epistemic adverbs and non-dynamic states of affairs. A functional explanation is sought in terms of (non-)expectedness, according to which the use of the subjunctive cancels the expectation of actualization of the state of affairs. This proposal captures common intuitions about the mitigating effects of the subjunctive and explains why non-past time reference and weaker epistemic adverbs favor its use in the context of epistemic adverbs.

References

CORGA = Centro Ramón Piñeiro para a Investigación en Humanidades (2019), *Corpus de Referencia do Galego Actual* [3.0], http://corpus.cirp.gal/corga/

Cornillie, Bert (2016), Acerca de la locución epistémica *tal vez* en el Siglo de las Luces: innovación y especialización, in M. Guzmán & D. Sáez (eds.), *Márgenes y centros en el español del siglo XVIII*, Valencia: Tirant Humanidades, 183–199.

Finanger, Elizabeth (2011), Time Reference and Lexical Effects in Mood Choice Following Spanish Epistemic Adverb *quizá(s)*: A Dialectal Comparison, in J. Michnowicz & R. Dodsworth (eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, Somerville, MA: Cascadilla Proceedings Project, 90–102, http://www.lingref.com/ cpp/wss/5/paper2509.pdf

Givón, Talmy (2001), *Syntax: An Introduction*, 2nd ed., Amsterdam: John Benjamins.

Harris, Martin (1974), The Subjunctive mood as a changing category in Romance, in J. M. Anderson & C. Jones (eds.), *Historical linguistics: Syntax, morphology, internal and comparative reconstruction*, vol. 2, Amsterdam: North-Holland, 169–188.

Houle, Leah, & Martínez Gómez, Rebeca (2011), A Closer Look at *quizá(s)*: Grammaticalization and an Epistemic Adverb, in Ortiz-López, Luis A, (Ed.), *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, Somerville, MA: Cascadilla Proceedings Project, 296–304, http://www.lingref.com/cpp/hls/13/paper2497.pdf

Malchukov, Andrej, & Xrakovskij, Viktor S. (2016), The linguistic interaction of mood with modality and other categories, in J, Nuyts & J, van der Auwera (eds.), *The Oxford Handbook of Modality and Mood*, Oxford: Oxford University Press, 196–220.

Nuyts, Jan (2001), *Epistemic Modality, Language, and Conceptualization: A Cognitive-Pragmatic Perspective*, Amsterdam: John Benjamins.

Yelin, Boris, & Czerwionka, Lori (2017), *Capaz* and *quizás* in Argentine Spanish: Epistemic Adverbs with Mood Variability, *Hispania* 100 (1), 30–46, doi:10.1353/hpn.2017.0003

The role of the dative in the Macedonian dative-predicative structures

Liljana Mitkovska
(University of Ss Cyril and Methodius, Skopje; Silesian University, Katowice)

Keywords: dative, copula, predicative, experiencer, affectedness

Schedule: Fri 14.00 Room 7

This paper deals with the Macedonian sentences containing a copula, non-verbal subject complement and a nominal constituent in dative, e.g. *Na Ana i e strav.* 'Ana is afraid.' *Ineresno mi e tuka.* 'I find it interesting here./I'm having fun here.' These constructions, called dative-predicative structures (DPS), exist in all Slavic languages with varied scope of use and may express various types of situations, ranging from physical to psychological and cognitive states. They have been studied extensively in Russian (c.f. Say 2013 and Zimmerling 2018), but there are few comprehensive studies for the South Slavic languages (for a recent contribution to the topic see the papers in *Русский язык за рубежом* № 5, 2018). This paper examines the type of words that can function as subject complements in DPS in Macedonian, comparing it with its neighboring languages Bulgarian and Serbian. It also discusses the types of situations in which DPS are used and the role of the dative in creation of the meaning of the construction,

This research is based on examples collected from literature and journalistic texts, as well as internet sites and interactive forums. In total 263 subject complement types were attested, represented with different frequency. They mostly comprise complements in *–o* that have a corresponding adjective (190 entities), but there are also adverbs and adverbial phrases (*Dosta mi e.* 'I've had enough.', *Bez vrska mi e ovde.* 'I find it odd here'), as well as some nouns (*žal mi e* 'I'm sorry'). The two main goals of the research are: (1) to determine what type of situations are expressed in Macedonian with a DPS, and (2) what syntactic patterns are possible in each of them. Such research should indicate the status of the Macedonian DPS construction and its development tendencies in relation to the neighbouring languages. It has been noted that in South Slavic languages DPS is more productive for expressing opinion (*Apsurdno mi e da se raspravam.* 'I find it absurd to argue.'), rather than physical and psychological states. However, it seems there is a different tendency in Bulgarian, as Gradinarova 2010 has shown on examples from internet forums. Examples like *Gordo mi e* 'I feel proud', *Gnevno mi e* 'I am furious' seem to be commonly used, though not conforming to the Bulgarian standard norm. We have not noticed such trend in Macedonian.

To fully understand the meaning and functioning of the DPS in the language they are compared to the corresponding clauses without a dative constituent (*Interesno e tuka*. 'It is interesting here.'). The comparison indicates that the dative in DPS has a similar role as in the other dative constructions of the so called 'free dative', sharing with them the feature of coding "a passive and unintentional experience" (Wierzbicka 1986: 408).

References

Gradinarova, Alla (2010), Bezličnie konstrukcii s datelnim subekta i predikativom na -o v ruskom i bolgarskom jazikah [Impersonal constructions with a dative subject and a predicative in –o in Russiona nad Bulgarian], *Bolgarskaja rusistika* 3–4, 34–55.

Say, Sergey (2013), On the nature of dative arguments in Russian constructions with «predicatives», in K. I. Chahine (ed.), (2013), *Current Studies in Slavic Linguistics*, Amsterdam: John Benjamins, 225-245.

Wierzbicka, Anna (1986), The meaning of a case: A study of the Polish dative,in R.D. Brecht and J.S. Levine (eds), *Case in Slavic*, Columbus/OH: Slavica, 386–426.

Zimmerling, Anton V. (2018), Predikativi i predikati sostojanija v russkom jazike [Predicatives and predicates expressing states in the Russian language], *Slavistična revija*, 66(1), 45–64.

German modal particles in exclamative utterances

Pierre-Yves Modicom
(U. Bordeaux Montaigne)

Keywords: particles, sentence-type, illocution, exclamation, German

Schedule: We 16.30 Room 11

Modal particles (MPs) of German are short, semi-grammaticalized (Abraham 1991) items of adverbial or adjectival origin marking intersubjective modality (mostly epistemic : Doherty 1985, Abraham 2012). They mark the speaker's anticipation of the hearer's ratification of the content of the utterance (Waltereit 2006). MPs are widely seen as interacting strongly with sentence mood, usually in the sense of illocutionary specification (Jacobs 1991, Waltereit 2006, Müller 2014, Gutzmann 2015, Alm et al. 2018). Although they have been extensively studied in the past 50 years, their use in exclamative

sentences has received less attention, especially when it comes to the link between MPs and sentence-mood (the most important exception being Kwon 2005). This talk provides a corpus-based systematic panorama of MP usage in German exclamatives, thus filling a gap in the literature. Data are drawn either from my personal research corpus composed of 20 hours of transcribved radio interviews, or from the corpus-based dictionary of particles by Métrich et al. (2002).

The first part is devoted to a discussion of the frequent view that exclamative sentences are either unaddressed or at least demand no ratification by the hearer. Based on proposals by Morel (1995), I defend the view that exclamative sentences do include an intersubjective feature paving the way for expectations about hearer ratification (disc. Zaefferer 2001, Larrory 2017, Pfeiffer 2017, d'Avis 2017). The second part deals with the exclamative usage of canonical MPs *doch*, *ja* and *schon*. I show that doch and ja involve both anticipations of ratification and the speaker substituting herself to the hearer in this ratifying role. *Schon* appears to behave slightly differently and to occur mainly in structures derived from rhetorical questions. The third and last part is concerned with *vielleicht* and *aber*, MP-similar elements that do not exhibit MP features outside of exclamatives. I show that their contribution is directly interpretable on the basis of their argumentative function in assertions. In the case of *vielleicht*, this involves restrictions on predicate types and contrastive focus marked by specific prosodic means. For *aber*, the exclamative use is in line with the original meaning of addition or surplus leading to a qualitative change. This leads me to discuss whether those two items should really be regarded as MPs.

References

Abraham, Werner (1991b), The grammaticization of the German modal particles. In Traugott, Elizabeth Closs & Heine, Bernd (eds.), *Approaches to Grammaticalization. Volume II. Types of grammatical markers* (Typological Studies in Language 19:2). Amsterdam: John Benjamins. 331-379.

Abraham, Werner (2012), Illocutive force is speaker and information source concern. What type of syntax does the representation of speaker deixis require? Templates vs. Derivational structure? Abraham, Werner & Elisabeth Leiss (eds.), *Modality and Theory of Mind elements across langages*. Berlin: De Gruyter. 67-108.

Alm, Maria, Janina Behr & Kerstin Fischer (2018), Modal particles and sentence type restrictions: A construction grammar perspective. *Glossa: A Journal of General Linguistics*, 3(1). 133. DOI: http://doi.org/10.5334/gjgl.702

d'Avis, Franz (2017), Exclamatives, Normalcy Conditions and Common Ground. *Revue de Sémantique et de Pragmatique* 40. 17-34.

Doherty, Monika (1985), Epistemische Bedeutung. Berlin : Akademieverlag. (=Studia Grammatica, 23).

Gutzmann, Daniel (2015), *Use-conditional meaning: Studies in multidimensional semantics* (Oxford Studies in Semantics and Pragmatics 6). Oxford: Oxford University Press.

Jacobs, Joachim (1991), On the semantics of discourse particles. Abraham, Werner (ed.), Discourse particles : Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German.Amsterdam : John Benjamins. 141-162.

Kwon, Min-Jae. 2005. *Modalpartikeln und Satzmodus: Untersuchngen zur Syntax, Semantik und Pragmatik der deutschen Modalpartikeln.* Munich: PhD dissertation Luwig-Maximilians-Universität München.

Larrory-Wunder, Anne (2017), Exclamation et intersubjectivité en allemand. *Revue de Sémantique et de Pragmatique* 40. 79-98.

Métrich, René, Faucher, Eugène & Courdier, Gilbert (1995-2002), *Les Invariables Difficiles. Dictionnaire allemand-français des particules, connecteurs, interjections et autres mots de la communication.* Nancy: Nouveaux Cahiers d'allemand. (4 vol.)

Morel, Mary-Annick. 1995. L'intonation exclamative dans l'oral spontané. *Faits de Langue* 6. 63-70.

Müller, Sonja. 2014. *Modalpartikeln* (=Kurze Einführungen in die germanistische Linguistik 17). Heidelberg: Winter.

Pfeiffer, Martin. 2017. The deictic dimension of exclamations: On the use of wh-exclamatives in German face-to-face interaction. *Revue de Sémantique et de Pragmatique* 40. 35-57.

Waltereit, Richard. 2006. *Abtönung. Zur Pragmatik und historischen Semantik von Modalpartikeln und ihren funktionalen Äquivalenten in romanischen Sprachen* (Beihefte zur Zeitschrift für romanische Philologie 338). Tübingen: Niemeyer.

Zaefferer, Dietmar. 2001. Deconstructing a Classical Classification : A Typological Look at Searle's Concept of Illocution Type. *Revue Internationale de Philosophie* 2. 209-225.

Towards a unified theory of multiple modals in English

Cameron Morin
(Ecole Normale Supérieure Paris-Saclay; University of Paris-Diderot)

Keywords: English, syntax, dialects, modality, research proposal

Schedule: We 15.30 Room 2

Multiple modals in English (MM) (*will can, might could, shouldn't ought* to) are a surprising series of non-standard dialectal structures found at least in certain varieties of the Southern United States (Di Paolo 1986; Nagle 1994; Butters 1996) and the British Isles (among others Brown 1991). Although they are a relatively understudied topic in research on modality and syntactic variation, MMs pose a number of theoretical issues, the most salient being the precise nature of their structure, e.g. is there a "true" modal, and if so, where is it, and what is the other element? This specific problem has been exclusively tackled and barely solved in dialects of the American South (Battistella 1995; Nagle 1997; Elsman & Dubinsky 2009; Hasty 2012). Moreover, new research on the topic with respect to Britain was only initiated a few years ago in Scotland (Bour 2014), and it has complexified the issue of building an integrated theory of these dialectal features.

This paper argues that we should overcome the present state of things and look forward to a unified theory of multiple modals in English. It aims to give a well-deserved overview of what MMs are, what has been found on them, and most importantly what remains to be done. It includes the results of a field survey I conducted for my MSc in the Scottish Borders (Hawick, 2018), which elicited judgment data from 60 respondents to investigate their claimed usage of double modals (DMs) and the ways they could manipulate them to create negatives, interrogatives and question-tags. The results have led me to formulate a provisional Speaker's Choice hypothesis, according to which DMs result from the free selection of a true modal and the recategorisation of the remaining constituent as an idiosyncratic adverb. As part of my current PhD research proposal, I am interested in developing a constructional approach to this hypothesis, to be tested again in the future for a comprehensive synchronic account of multiple modals cross-dialectally.

References

Battistella, Edwin L. (1995), The syntax of double modal constructions, *Linguistica Atlantica* 17, 19–44.

Bour, Anthony (2014), *Description of multiple modality in contemporary Scotland: Double and triple modals in the Scottish Borders*, Freiburg, Germany: University of Freiburg dissertation.

Brown, Keith (1991), Double modals in Hawick Scots, in Peter Trudgill & John K. Chambers (eds.), *Dialects of English*, London and New York: Longman, 74–103.

Di Paolo, Marianna (1986), *A study of double modals in Texas English*, Austin, TX: University of Texas dissertation.Elsman, Minta & Stanley Dubinsky (2009), Double modal syntactic patterns as single modal interactions, *University of Pennsylvania Working Papers in Linguistics* 15(1), 74–83.

Hasty, J. Daniel (2012), We might should oughta take a second look at this: A syntactic re-analysis of double modals in southern United States English, *Lingua* 122, 1716–1738.

Nagle, Stephen J. (1994), The English Double Modal Conspiracy, *Diachronica* 11(2), 199–212.

Nagle, Stephen J. (1997), What is double about double modals? in Raymond Hickey & Stanislav Puppel (eds.), *Language history and linguistic modelling*, Berlin: Mouton de Gruyter, 513–526.

Analysing phonological systems: On Bayesian typological research

George Moroz
(Moscow, Linguistic Convergence Laboratory NRU HSE)

Keywords: typology, sampling, phonology, Bayesian statistics

Schedule: Sat 12.00 Room 9

Most typical typological research consists of the following steps: formulating a problem, defining the values of the variable of interest, sampling languages, followed by an investigation of how the variable of interest's values are distributed among selected languages. However, there are a lot of different biases that influence language sampling (see e. g. (Widmann 2001, Baker 2010)). The assumption that languages are independent entities in some statistical population is more than doubtful. In my talk, I propose to avoid sampling and analyse all obtained data using Bayesian philosophy.

Are some phonological segments typical or rare from a typological perspective? To answer this question, I will use the PHOIBLE database Moran et al. (2014) as a source and some segments of interest (vowels i, a, u) as an example. The typical typological research will take into account only languages from the sample and treat them as independent observations. In my analysis, I take into consideration all available data from the database, grouped by macro language family. For each macro family I counted the ratio of languages with the segment of interest. Using empirical Bayesian shrinkage towards a Beta prior Robbins (1985), it is possible to use the mean ratio as a prior and then estimate, how far the state of each macro family is removed from the value obtained from the whole dataset. The results are presented in Figure 1. Each data point is a macro family or an isolate language (capitalised in the Figure 1), and all points are coloured by the number of languages in the group. The dashed line is just a reference line $y = x$. If a macro family or isolate language is located above the line, than the original ratio is lower then our typological expectation, and vice versa.

It is not surprising that most languages are below the line, so they have a greater ratio than we expected. It is also not surprising to see a grey cluster with zero original ratio on the left — these are isolates or language groups with only a few languages, so the probability that some of them lack

vowels of interest is rather high. The most interesting patterns are shown on the bottom part of the graph above the line: they unexpectedly don't have the vowels of interest, which appears to be a feature of the language family.

So I briefly showed a Bayesian analysis of the phonological systems based on language families from the PHOIBLE database. This analysis shows that, rather than creating a sample of languages, it can be fruitful to try to use as much data as we have. Empirical Bayesian analysis of such data can show what is typical and what is not, and which language families are a-typical. In Frequentism, sampling should be used once, so if we want to replicate the result of previous research we need to create a new sample. Using the Bayesian approach, it is possible to update our beliefs on the subject of interest. So adding and editing existing data make it possible to use previous results on the subject of interest.

This work was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

References

Baker, D. (2010). Language sampling. In J. J. Song (Ed.), *The Oxford Handbook of Linguistic Typology*. Oxford University Press.

Moran, Steven, Daniel McCloy,, Richard Wright (Eds.) (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Robbins, Herbert (1985). An empirical bayes approach to statistics. In *Herbert Robbins Selected Papers*, pp. 41–47. Springer.

Widmann, Thomas Martin (2001). *Language sampling for typological studies.* Master's thesis, University of Aarhus.

Figure 1: Bayes shrinkage for three vowels, based on the PHOIBLE database

The psychological verbs of Basque in typological and diachronic perspective

Céline Mounole & Denis Creissels
(Université de Pau et des Pays de l'Adour and Iker UMR 5478; Université Lumière Lyon 2 and DDL UMR 5596)

Keywords: psychological verbs, ambitransitivity, Basque, typology, diachrony

Schedule: Thu ~~11.30~~ 9.00 Room 14

Four main types of coding frames are available for the psychological verbs of Basque:

– absolutive experiencer, oblique stimulus (1)
– ergative experiencer, absolutive stimulus (2)
– absolutive stimulus, dative experiencer (3)
– ergative stimulus, absolutive experiencer (4)

(1) Jon ehizaz aspertu da.
Jon:ABS hunting:SG:INSTR get.bored:CPL be:PRES:A3SG

‘Jon got bored of hunting.’

(2) Jonek zure portaera gaitzesten du.
Jon:ERG 2SG:GEN behavior:ABS disapprove:ICPL have:PRES:A3SG:E3SG
‘Jon disapproves your behavior.’

(3) Joni txokolatea gustatzen zaio.
Jon:DAT chocolate:SG:ABS please:ICPL be:PRES:A3SG:D3SG
‘Jon likes chocolate.’

(4) Zure portaerak Jon asaldatzen du.
2SG:GEN behavior:ERG Jon:ABS perturb:ICPL have:PRES:A3SG:E3SG
‘Your behavior perturbs Jon.’

Several types of alternations are attested for the psychological verbs of Basque. None of them involves a change in the verb stem (changes in the auxiliary are an automatic consequence of differences in the transitivity of the construction). The most widespread alternation is that between the ExpAbs-StimObl and StimErg-ExpAbs coding frames, cf. (5) to be compared with (1).

(5) Ehizak Jon aspertu du.
hunting:SG:ERG Jon:ABS bore:CPL have:PRES:A3SG:E3SG

The first part of the presentation discusses the valency properties of the psychological verbs of Basque in typological perspective. Basque is remarkable not only for the systematic use of the ambitransitivity strategy (as opposed to the causativization and decausativization strategies) in noncausal-causal pairs such as ‘break (intr.) / break (tr.) (cf. **Ispilua *puskatu* da** ‘The mirror broke’ vs. **Jonek ispilua *puskatu* du** ‘Jon broke the mirror’), but also for the extension of the ambitransitivity strategy to the expression of the perspectivization of psychological events. In particular, with the psychological verbs of Basque, causative derivation is not used to express a mere change in perspectivization, but to introduce an indirect causer, as in (6).

(6) Mikelek Jon asaldarazi du.
Mikel:ERG Jon perturb:CAUS:CPL have:PRES:A3SG:E3SG
‘Mikel acted in such a way that Jon was perturbed.’

The second part of the presentation discusses the evolution of the valency properties of psychological verbs in the history of Basque. In the general context of a tendency to extend the use of the ergative case (cf. Mounole 2011, Creissels & Mounole 2017), the coding frames ExpAbs- StimObl and StimAbs-ExpDat, which were predominent in the 16th and 17th centuries have declined since then, whereas the coding frames ExpErg-Abs and StimErg-ExpAbs have progressed.

Abbreviations

A absolutive index, ABS absolutive, CAUS causative, CPL completive, D dative index, DAT dative, E ergative index, ERG ergative, GEN genitive, ICPL incompletive, INSTR instrumental, SG singular, CPL completive, PRES present

References

Creissels, Denis & Mounole Céline. 2017. Non canonical valency patterns in Basque, variation and evolution. In Lars Hellan, Andrej Malchukov & Michela Cennamo (eds), *Contrastive Studies in Verbal Valency*. John Benjamins. 151-174

Etxepare, Ricardo. 2003. Valency and argument structure in the Basque verb. In José Ignacio Hualde & Jon Urtiz de Urbina (eds.), *A grammar of Basque*. Mouton de Gruyter.

Fernández Fernández, Beatriz & Jon Ortiz de Urbina. 2010. *Datiboa hiztegian* [The dative in the lexicon]. Euskal Herriko Unibertsitateko Argitalpen Zerbitzua.

Mounole, Céline. 2011. *Le verbe basque ancien: étude philologique et diachronique*. Ph Dissertation. Universidad del País Vasco & Université Bordeaux Montaigne.

When ‘after’ means ‘when’: the puzzle of the *-məkə* converb in Hill Mari

Nikita Muravyev
(National Research University Higher School of Economics, Moscow)

Keywords: converbs, temporal clauses, Anteriority, Simultaneity, Hill Mari

Schedule: Fri 12.00 Room 7

This research was conducted in terms of the project supported by Russian Science Foundation, grant No. 18-78-10128. The data were collected during fieldwork in Kuznetsovo village (2017-2019). Most of the examples are elicited, some examples are taken from the corpus of transcribed oral narratives (ca. 56.000 tokens).

Typological works on converb forms distinguish between **specialized**, i.e. expressing one meaning, and **contextual** converb type, i.e. expressing one of several meanings depending on the context, cf. (Nedjalkov 1995, Nedjalkov 1998). Anterior converb in *-məkə* in Hill Mari (< Mari < Finno-Ugric < Uralic) appears to be a more complex case. As shown in the examples below, this converb relates to the **Mixed converbs of contextual Anteriority** class (Nedjalkov 1998: 437-439), expressing the meanings of Anteriority (1) and Condition (2).

(1)	*[el’ektr’ičəstvə-m*	***pər-t-əmə̑kə̑]***	*əl-əmäš*	*kuštə̑lg-en*	
	electricity-ACC	enter-CAUS-CVB.ANT	live-NMLZ	become.easier-PRET	
	‘After the electricity was invented life became easier.’				
(2)	***[tol-mə̑kə̑-štə̑]***	*kogo-n*	*susu*	*li-äm*	*ə̑l-ne-žə*
	come-CVB.ANT-POSS.3PL	big-ADV	happy	become-NPST.1SG	be-DES-3SG
	‘If they come, I will be very happy.’				

However, in a particular context, where the two actions are performed by different subjects and the main event is durative, the default Anteriority (‘after’) meaning yields the Simultaneity (‘when’) interpretation which is not expected for the converbs of this class. Interestingly enough, the temporal framing subordinate verb in this case has to describe someone’s entering but not leaving the stage of the main event (3), cf. ungrammaticality of (4).

(3)	***[tol-mə̑k-em]***	*maša*	*lem-əm*	*kačk-eš*	*ə̑l’ə̑*
	come-CVB.ANT-1SG	Mary	soup-ACC	eat-NPST.3SG	RETR.IPFV
	‘When I came Mary was eating soup.’				
(4)	****[ke-mək-em]***	*maša*	*lem-əm*	*kačk-eš*	*ə̑l’ə̑*
	leave-CVB.ANT-1SG	Mary	soup-ACC	eat-NPST.3SG	RETR.IPFV
	Exp.: ‘When/After I left Mary was eating soup.’				

This contrast is strikingly consistent throughout the elicited data including a number of setups with different lexical choices of the main and the dependent predicate and different arrangements of the participants on the scene. Supplied with consultants’ comments it suggests the analysis of ‘when’ contexts as containing an implicit event of the main event discovery (cf. *when Pete came Mary was sitting ~ after Pete came* ***he saw that*** *Mary was sitting*). Using the framework of Mental Spaces Theory (Fauconnier 1997, Fauconnier, Sweetser 1996, Dancygier, Sweetser 2005) this idea can be represented through setting up a perception space which is temporally distanced from the dependent

event of coming. Thus, despite the actual coincidence of the two events on the time axis, what is asserted is not the 'when' relation but the 'after' relation between the dependent event and the main event discovery. This solution allows to maintain the default Anteriority meaning for all temporal uses of *-məkə* without having to posit an unexplainable polysemy relation. In the meantime, it raises the question concerning the role of perceptual accessibility of the events in temporal adverbial constructions in general which is yet to be clarified.

References

Dancygier, Barbara, and Sweetser, Eve (2005), *Mental spaces in grammar: Conditional constructions*. Cambridge University Press.

Fauconnier, Gilles (1997), *Cognitive mappings for language and thought*. Cambridge: Cambridge University Press.

Fauconnier, Gilles, and Sweetser, Eve. (eds.), (1996), *Spaces, worlds, and grammar*. Chicago & London: University of Chicago Press.

Nedjalkov, Igor V. (1998), Converbs in the languages of Europe, in J. van der Auwera, Dónall P. Ó Baoill (eds.). *Adverbial Constructions in the Languages of Europe. Empirical approaches to language typology* (20-3). Berlin: Mouton de Gruyter, 421–456.

Nedjalkov, Vladimir P. (1995), Some typological parameters of converbs, in M. Haspelmath, E. König (eds). *Converbs in Cross-Linguistic Perspective*, Berlin: Mouton de Gruyter, 97–136.

Abbreviations:

ACC - accusative, ADV - adverbializer, ANT - anteriority, CAUS - causative, CVB - converb, DES - desiderative, IPFV - imperfective, NMLZ - nominalizer, NPST - non-past, PL - plural, POSS - possessive, PRET - preterite, RETR - retrospective shift, SG - singular

The emergence of definiteness in New Western Iranian languages: Extending the typology of definiteness

Maryam Nourzaei
(Uppsala University, Bamberg University)

Keywords: New Western Iranian languages, definiteness, typology, grammaticalization, and diminutive

Schedule: Thu 14.30 Room 6

Generally, definiteness is unmarked in Iranian languages. The aim of the present paper is to present the development of grammatical markers for definiteness in New Western Iranian languages of the Zagros Mountains, which are unique in this regard. The paper concentrates on a sample of six languages spoken in Iran: Balochi, Koroshi, Kurdish, colloquial Persian, Sistani, and Shirazi. The data was extracted from a quantitative analysis of digital corpora of connected spontaneous speech (Nourzaei et al. 2015, Nourzaei 2017, Öpengin 2016) and new recordings of a questionnaire based on six mini narratives, which targeted specific functional domains of definiteness modeled after Becker (2017). The suffixes *-ok, -ak(a), (e)k, ak,* and *–ū (k)* have been attested as definiteness markers in these languages. The degree of grammaticalization differs across the languages; for instance, the suffix *-ū(k)* in Shirazi still preserves some traces of its original diminutive sense alongside its new definiteness meaning. I will demonstrate this point with more examples during my presentation.

While the grammaticalization of definite markers has been a central issue in grammaticalization theory, the regularly cited cases (e.g., Romance languages) involve the development of erstwhile demonstratives, or linking particles such as relatives, into articles (Lyons 1999, Himmelmann 2001). In the Iranian languages under consideration, there is no obvious candidate for these markers among known Iranian demonstratives or linking particles. More interestingly, we witness a distinct development whereby an original derivational suffix with a diminutive sense changes into a suffixal marker of discourse identifiability.

Given the possibility that these markers do not originate from demonstratives or linking particles in these languages, but from a diminutive, the question becomes: What does the functional distribution of these suffixes in these languages reveal about historical processes in the grammaticalization of definiteness?

The paper concludes that there is no obvious candidate for these markers among the known Iranian demonstratives or linking particles. Instead, the most plausible candidate appears to be an older diminutive suffix. This appears to be the first solidly attested case of development. In addition, the various markers (*-ok, -ak(a), (e)k, ak, and –ū (k)*) are all historically related, which allows us to reconstruct the following historical development of New Iranian definiteness: diminutive > endearment > proximity > identifiability > definiteness.

References

Becker, Laura. 2017. Definiteness in languages with and without articles: A parallel corpus study. Association of Linguistic Typology Conference, Canberra, December 2017.

Himmelmann, Nikolaus. 2001."Articles." In: M. Haspelmath, E. König, W. Oesterreicher and W. Raible (eds.). Language Typology and Language Universals. An International Handbook, Vol. 1. Berlin: de Gruyter, 831-841.

Lyons, Christopher. 1999. Definiteness. Cambridge University Press

Nourzaei, Maryam, Carina Jahani, Erik Anonby & Abbas Ali Ahangar 2015. Koroshi. *A Corpus-based Grammatical Description*. Uppsala: Acta Universitatis Upsaliensis.

Nourzaei, Maryam. 2017. *Participant Reference in Three Balochi Dialects: Male and Female Narrations of Folktales and Biographical Tales*. Uppsala: Acta Universitatis Upsaliensis.

Öpengin, Ergin. 2016. *The Mukri variety of Central Kurdish. Grammar, texts and lexicon*. Wiesbaden: Harrassowitz.

The typology of singulatives: Connecting number inflection, nominalizing derivation and evaluative semantics

Silva Nurmio
(University of Helsinki)

Keywords: singulatives, mass nouns, diminutives, grammatical number, inflection/derivation

Schedule: Thu 9.30 Room 6

The singulative is a category that denotes a (single) unit and is formed by adding a morphological marker to a non-unit-denoting base, e.g. a plural, mass noun, collective, or a form not marked for number at all (general number). In some languages, singulatives form part of the inflectional number system, e.g. Welsh (Indo-European) *moch* 'pigs', singulative *moch-yn* [pigs-SING] 'a pig'; these two

forms are in a singular/plural opposition, with the singular overtly marked (Nurmio 2017). But in many instances the formation of singulatives is closer to derivation, e.g. Russian (Indo-European) *gorox* 'pea(s)' (mass noun), *goroš-ina* 'a pea' (singulative).

In addition to forming unit nouns from mass nouns, the singulative can have other functions beyond inflectional number. As already noted by Jurafsky (1996), the same markers can have both diminutive and singulative functions in several languages. For example, the Welsh suffix *-yn* above forms diminutives (and, by extension, hypocoristic terms) when added to singular count bases, e.g. *bachgen* 'boy', *bachgenn-yn* 'a little/dear boy'. Singulative markers can also turn non-nominal bases into singular count nouns, e.g. Welsh *unigol* 'individual' (adj.), *unigol-yn* [individual-SING] 'an individual (person)'. As the Welsh examples show, derivational functions can occur alongside the inflectional number-marking function in a single language.

This category has never previously been subject to a full typological study, despite its occurrence in a wide variety of language families and its essential contribution to typology, and especially the typology of unitization. As part of my project I am building an as-comprehensive-as possible database of singulatives. Due to the lack of a definition informed by a broad typological picture, singulatives have often gone unrecorded in language descriptions. The first aim of this paper is to provide an operational definition, which clarifies the status of the singulative as a morphological category. The literature so far (e.g. Grimm 2018, Haspelmath & Karjus 2017, Acquaviva 2016) has mostly focused on inflectional singulative markers, which are rare in comparison to inflectional plural markers cross-linguistically. Including derivational markers, productive and unproductive, in the definition uncovers more examples of singulatives and brings to light the nature of this category as one that spans inflection and derivation.

Secondly, I look at the kinds of number systems that singulatives occur in: languages with inflectional singular/plural, those with general number, etc. Singulatives based on general number forms are found in many Afro-Asiatic languages, e.g. Bayso *lúban* 'lion(s)' (general), *lúban-titi* 'a lion' (singulative) (Corbett 2000: 11). I also argue, following Grimm (2012, 2018), that inverse number marking should be included as a type of singulative. In such languages, a single marker can mark singular or plural, depending on the number value of the base.

This typological work brings to light new connections between derivation and inflection and between number morphology and evaluative semantics. It is also central to the study of how languages denote units or packaged meanings ('a coffee', etc.) from mass nouns, and to discussions of how evaluative meanings relate to other morphological phenomena.

References

Acquaviva, Paolo (2016). Singulatives, in Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *HSK Word-Formation. An International Handbook of the Languages of Europe*, Berlin: Mouton de Gruyter, 1171–1183.

Corbett, Greville G. (2000). *Number*. Cambridge: CUP.

Grimm, Scott (2012). Individuation and inverse number marking in Dagaare. In Diane Massam (ed.), *Count and Mass Across Languages*, Oxford: OUP, 75–98.

Grimm, Scott (2018). Grammatical number and the scale of individuation. *Language* 94(3), 527–574.

Haspelmath, Martin & Andres Karjus (2017). Explaining asymmetries in number marking: singulatives, pluratives, and usage frequency, *Linguistics* 55(6), 1213–1235.

Jurafsky, Daniel (1996). Universal tendencies in the semantics of the diminutive. *Language* 72(3), 533–578.

Nurmio, Silva (2017). Collective nouns in Welsh: a noun category or a plural allomorph? *Transactions of the Philological Society* 115(1), 58–78.

‘As soon as’ clauses: A typological study of immediate temporally subsequent events

Jesús Olguín Martínez, Bernard Comrie & Eric W. Campbell
(University of California, Santa Barbara)

Keywords: Adverbial clause, complex sentence, typology, subordination, temporality

Schedule: Thu 11.30 Room 6

Temporally subsequent events have gained increasing interest recently. While some studies have concentrated on the deranking and balancing status of these complex sentence constructions (e.g. Cristofaro 2003), others have addressed their degree of semantic explicitness (e.g. Martowicz 2011; Olguín-Martínez, Campbell & Comrie 2018).

Temporally subsequent events may express different “degrees of subsequence” (extent of time lapse). However, all the studies mentioned above have concentrated primarily on constructions that can be used irrespective of extent of time lapse. The present work makes inroads into the study of the expression of temporal subsequence relations of specific time lapse ranges through an investigation of immediate temporally subsequent events, a.k.a. ‘as soon as’ clauses. In doing so, we use a genetically balanced sample of 300 languages based on *WALS* (Haspelmath et al. 2005).

The aims of this presentation are two-fold. First, we explore the range of clause-linking devices used to convey immediate temporal subsequence. In this regard, we show that immediate temporal subsequence is mainly encoded by: (i) adverbials, as in (1), where the adverbial suffix *-wa* ‘just’ is required to encode the sense of immediate subsequence; (ii) general coordinating devices, as in (2) with the coordinating conjunction *bɔ̀* ‘and’, where the sense of immediate subsequence is conveyed by means of verb copying; and (iii) dedicated subordinating devices (e.g. subordinating conjunctions, specialized converbs), as in (3), with the free subordinating conjunction *henc’or* ‘as soon as’. We also introduce other less common devices attested in the sample, such as relative clauses and quantifiers meaning ‘all’.

Awtuw (Sepik/Rama: Papua New Guinea; Feldman 1986: 77)

(1) nom æye rokra-tay-ka-**wa**-re, rey ra-tæw-e.
1PL.SBJ food cook-finish-PERF-just-OBJ 3SG.SBJ.M eat-begin-PST
‘As soon as we had finished cooking the food.’

Fongbe (Niger-Congo/Kwa: Benin; Lefebvre & Brousseau 2002: 172; cf. Fiedler 2014)

(2) **wá** Kɔ́kú **wá** **bɔ̀** Àsíbá **yì.**
arrive Koku arrive and Asiba leave
‘As soon as Koku arrived, Asiba left.’

Western Armenian (Indo-European/Armenian: Armenian; Dum-Tragut 2009: 434)

(3) **henc’or** hasn-em tun-ĕ, k’ez k-zangahar-em.
as_soon_as arrive-PRS.1SG house-the you.DAT FUT-phone-PRS.1SG
‘As soon as Koku arrived, Asiba left.’

Second, we explore the cross-linguistic distribution of the clause-linking devices shown above. The present study reports on 53 languages that overtly express immediate temporally subsequent events. Although explicit encoding of immediate temporal subsequence is found in languages of all macro-areas, our study suggests a geographical skewing between the areas of Australia, with scarce occurrences, and the areas of Eurasia, Africa, and North America, which host the majority of cases in our sample. While subordinating conjunctions occur throughout the sample, they are most frequent in languages of Eurasia and Africa. Specialized converbs are most common in Eurasia. Finally, adverbs

meaning 'immediately' as the basic encoding of immediate subsequence are relatively frequent in South America.

This work is one of the first of many steps towards a finer-grained typology of the strategies used for the expression of specific time lapses in temporal subsequence, and the distribution of those strategies in the world's languages.

References

Cristofaro, Sonia. 2003. *Subordination*. Oxford/New York: Oxford University Press.

Dum-Tragut, Jasmine. 2009. *Armenian*. Amsterdam/Philadelphia: John Benjamins.

Feldman, Harry. 1986. *A Grammar of Awtuw*. (Pacific Linguistics, Series B, 94.) Canberra: Australian National University.

Fiedler, Ines. 2014. Why Are 'As Soon As' Clauses Marked for Predicate-Centered Focus. Handout, Humboldt-Universität zu Berlin.

Haspelmath, Martin, Dryer, Matthew S., Gil, David & Comrie, Bernard (eds.). 2005. *World Atlas of Language Structures*. Oxford: Oxford University Press.

Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A Grammar of Fongbe*. Berlin: Mouton de Gruyter.

Martowicz, Anna. 2011. The Origin and Functioning of Circumstantial Clause Linkers: A Cross-Linguistic Study. Ph. D. Dissertation. University of Edinburgh.

Olguín-Martínez, Jesús, Campbell, Eric W. & Comrie, Bernard. 2018. Temporally Subsequent Events in the World's Languages. Handout of paper presented at the Encuentro Internacional de Lingüística en el Noroeste. University of Sonora, Hermosillo. November 14–16.

Perceptual study of language chunks in spontaneous Estonian

Nele Ots & Piia Taremaa
(Goethe University of Frankfurt & University of Tartu)

Keywords: speech perception, language chunk, prosody, clause structure, experimental approach, Estonian

Schedule: Thu 12.00 Room 9

The study investigates the perception of *language chunks* in spontaneously spoken Estonian. Chunking is a universal cognitive mechanism to expand the limitations of working memory and to efficiently process different kind of information, including continuous speech flow (Baddeley and Hitch, 1974; Carpenter and Just, 1989; Cowan et al., 2005; Gathercole et al., 1993; Miller, 1956). The study assumes that the language chunks can be defined by their boundaries.

Earlier research has demonstrated that the comprehension of continuous speech flow is influenced by the abstract syntactic knowledge (e.g., Green, 2017; Caplan and Waters, 1999). The syntactic relations, in turn, are to a large degree recovered from the prosody (e.g., Watson and Gibson, 2005). A number of studies investigating expert as well as non-expert perception of prosodic boundaries have found interacting influence from the syntactic information (the part-of-speech of a word, presence of the clause boundary and the clause type) and from the prosodic cues (pauses, intonational boundary tones and word-final lengthening; e.g., Christodoulides et al., 2018; Cole et al, 2010; Lelandais et al., 2018, Roux et al., 2016). Thus, syntactic information and the prosody are expected to be tightly interrelated in the language chunking.

We will present results from a speech perception experiment with native speakers of Estonian. The subjects (N=51) listened to the spontaneously spoken excerpts (N=396) and were instructed to mark between the words where they heard some sort of juncture. They were provided no additional explanation what the juncture or the boundary would be. A number of acoustic (duration, pitch, intensity) and linguistic factors (presence of clause boundary, syntactic phrase) were defined as the predictors of the boundary perception.

The univariate general linear mixed effects analyses of the boundary marks showed a significant influence of the categorical linguistic predictors, such as clause boundaries and type of the syntactic phrase (subject, object etc.), but also strong effect of continuous predictors, such as duration of the word-final unstressed syllable rhyme and pitch difference across the word boundaries. The multivariate analyses based on the conditional random forests and inference trees demonstrated that the most important predictors of the boundary perception are the presence of a clause boundary, final lengthening and an acoustic pause.

The study indicates that both prosodic (pause, word-final lengthening, pitch difference across the word boundaries) and linguistic (clause boundary) cues are important for determining the boundaries of speech chunks in native non-expert listening. Our results corroborate earlier findings showing syntactic and phonetic influences on language chunking but more importantly, they demonstrate an intricate relationship between the linguistic knowledge and speech acoustics in the language chunking.

References

Baddeley, A. D., and G. J. Hitch, (1974), Working Memory, in Bower, G. H. (ed.), *The Psychology of Learning and Motivation*, 8, New York: Academic Press, 47–89.

Carpenter, P. A., and M. A. Just, (1989), The Role of Working Memory in Language Comprehension, in Klahr, D. and Kotovsky, K. (eds.), *Complex Information Processing: The Impact of Herbert A. Simon*, New York: Psychology Press, 47–89.

Cole, J., Y. Mo, and S. Baek, (2010), The Role of Syntactic Structure in Guiding Prosody Perception with Ordinary Listeners and Everyday Speech, *Language and Cognitive Processes* 25 (7–9): 1141–1177.

Cowan, N., E. M. Elliott, J. S. Saults, C. C. Morey, S. Mattox, A. Hismjatullina, and A. R. Conway, (2005), On the Capacity of Attention: Its Estimation and Its Role in Working Memory and Cognitive Aptitudes, *Cognitive Psychology* 51 (1): 42–100.

Christodoulides, G., A. C. Simon and I. Didirková, (2018), Perception of Prosodic Boundaries by Naïve and Expert Listeners in French. Modelling and Automatic Annotation, in *Proceedings of the 9th Speech Prosody Conference*, Poznań, Poland, 13-16 June, 2018.

Gathercole, S. E., and A. D. Baddeley, (1993), *Working Memory and Language*, Hove/East Sussex/New York: Psychology Press.

Green, C., (2017), Usage-based linguistics and the magic number four, *Cognitive Linguistics* 28 (2): 209–237.

Caplan, D. and G. S. Waters, (1999), Verbal working memory and sentence comprehension, *Behavioural and Brain Sciences* 22, 77–12.

Lelandais, M. and G. Ferré, (2018), Perception of prosodic boundaries by naïve listeners in three different types of subordinate syntactic constructions, in *Proceedings of the 9th International Conference on Speech Prosody 2018*, 13-16 June 2018, Poznan, Poland.

Miller, G. A., 1956, The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychological Review* 63, 81–9710.1037/h0043158.

Roux, G., R. Bertrand, A. Ghio, C. Astésano, 2016, Naïve listeners' perception of prominence and boundary in French spontaneous speech, in *Speech Prosody*, May 2016, Boston, United States. Speech Prosody, 2016, <hal-01462259>.

Watson, D. and E. Gibson, (2005), Intonational phrasing and constituency in language comprehension, *Studia Linguistica* 59(2/3), 279–300.

ꕥ

AABB reduplicated plurals in Mandarin Chinese

Nadine Otting
(The University of Hong Kong PhD graduate)

Keywords: Chinese, reduplication, plural, morphosyntax, phonology

Schedule: Sat 11.30 Room 9

Among AABB reduplications in Mandarin Chinese, there is a group of words that indicate plurality. These AABB plurals include not only nouns and verbs (Zhang 2015), but also adjectives. Examples are given in (1).

(1) a. nánnanǘnǔ
red.man-woman
'men and women'

b. láilaiqùqù
red.come-go
'come and go repeatedly'

c. gāogaodīdī
red.tall-short
'tall and short'

These words all indicate a vague cardinality of a large number of entities or events and cannot be quantized exactly. This study investigates the derivation of such AABB plurals. Working from the theoretical assumption that a reduplicative affix (see, e.g., Marantz 1982, Broselow & McCarthy 1983, McCarthy & Prince 1995) is responsible for the pairing between AABB surface form and plural meaning, the internal syntactic representation and the phonological realization of these words are examined.

The syntactic representation of AABB plurals combines the reduplicative morpheme with the base for reduplication. The base for reduplication consists of coordinate monosyllabic constituents. This coordinate relation is motivated by the semantic parallelism between the A and B constituents and the possibility of some AABB plurals to occur in reversed BBAA order. The constituents are argued to be roots in the sense of Distributed Morphology (Halle & Marantz 1993, and subsequent work) due to the categorial variability and flexibility of AABB plurals. This cross-categorical nature also supports a root status of the reduplicative plural affix. Moreover, a root analysis of the reduplicative affix captures its semantic denotation and derivational nature.

The theoretical implication of this syntactic analysis of AABB plurals is that plural marking is not restricted to the Number projection (Ritter 1991, 1992), nor is it exhaustively captured by a syntactic typology of heads and adjuncts (Wiltschko, 2008). Instead, AABB plurals support the proposal that some affixes belong in the root domain (Creemers, Don, & Fenger 2017, De Belder 2011, Lowenstamm 2010, 2015).

Phonologically, AABB plurals resemble AABB reduplicated adjectives that express a degree meaning in Mandarin Chinese. Following Lee-Kim (2016), we assume that the coordinate base for reduplication compels double exponence of the reduplicative affix. Total reduplication then

successfully accounts for the exact match in segmental material of the bases and reduplicants. However, the AABB words exhibit tonal peculiarities, namely tone neutralization of the second syllable and the obligatory presence of tone on the BB syllables. We implement the idea that the metrical structure of these words has an impact on tone realization (Lee-Kim 2016, Sui 2013, 2016) and develop a comprehensive Optimality Theoretic (McCarthy & Prince 1993) analysis of these AABB words.

This phonological analysis of AABB words in Mandarin Chinese provides cross-linguistic support for the interaction of tone and metrical prominence (De Lacy 2002).

References

Broselow, Ellen, & McCarthy, John J. (1983), A theory of internal reduplication, *The Linguistic Review*, *3*, 25–88.

Creemers, Ava, Don, Jan, & Fenger, Paula (2017), Some affixes are roots, others are heads, *Natural Language and Linguistic Theory*, 1–40. (Advance online publication)

De Belder, Marijke (2011), *Roots and affixes: Eliminating lexical categories from syntax*, Doctoral dissertation, Universiteit Utrecht.

De Lacy, Paul (2002), The interaction of tone and stress in Optimality Theory, *Phonology*, 19, 1–32.

Halle, Morris, & Marantz, Alec (1993), Distributed Morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvian Bromberger*, Cambridge: MIT Press, 111–175.

Lee-Kim, Sang-Im (2016), Syntax-based phonological asymmetries: The case of adjectival reduplication in Mandarin Chinese, *Lingua*, 179, 1–23.

Lowenstamm, Jean (2010), *Derivational affixes as roots: Phasal spellout meets English stress shift*. (Ms. Université Paris-Diderot & CNRS)

Lowenstamm, Jean (2015), Derivational affixes as roots: Phasal spell-out meets English stress shift. In Artemis Alexiadou, Hagit Borer, & Florian Schäfer (Eds.), *The syntax of roots and the roots of syntax*, Oxford, UK: Oxford University Press, 230–258.

Marantz, Alec (1982), Re reduplication. *Linguist Inquiry*, 13 (3), 435–482.

McCarthy, John J., & Prince, Alan (1993), Generalized alignment. In Geert Booij & Jaap Van Marle (Eds.), *Yearbook of morphology 1993*, Dordrecht: Springer, 79–153.

McCarthy, John J., & Prince, Alan (1995), Faithfulness and reduplicative identity. In Jill Beckmann, Suzanne Urbanczyk, & Laura Walsh Dickey (Eds.*), University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*, Amherst: Graduate Linguistic Student Association. (Available on Rutgers Optimality Archive, 249–384. (http://roa.rutgers.edu, ROA: 60)

Ritter, Elizabeth (1991), Two functional categories in noun phrases: Evidence from Hebrew. In Susan Rothstein (Ed.), *Syntax and semantics 25: Perspectives on phrase structure*, New York: Academic Press, 37–62.

Ritter, Elizabeth (1992), Cross-linguistic evidence for Number Phrase, *Canadian Journal of Linguistics*, 37, 197–218.

Sui, Yanyan (2013), *Phonological and phonetic evidence for trochaic metrical structure in Mandarin Chinese*, Doctoral dissertation, University of Pennsylvania.

Sui, Yanyan (2016), The interaction of metrical structure and tone in Standard Chinese. In Jeffrey Heinz, Rob Goedemans, & Harry van der Hulst (Eds.), *Dimensions of phonological stress*, Cambridge: Cambridge University Press, 101–122.

Wiltschko, Martina (2008), The syntax of non-inflectional plural marking, *Natural Language and Linguistic Theory*, 26, 639–694.

Zhang, Niina Ning (2015), The morphological expression of plurality and pluractionality in Mandarin, *Lingua*, 165, 1–27.

Towards a relevance-theoretic procedural account of pejorative morphology

Manuel Padilla Cruz
(University of Seville)

Keywords: relevance theory, morphology, procedural meaning, ad hoc concepts, affective attitude representation

Schedule: Thu 12.00 Room 1

Relevance theorists argue that the decoded *logical form* of an utterance is pragmatically enriched in a number of ways (Sperber and Wilson 1995; Carston 2002). In addition to disambiguation of constituents and word senses, and recovery of unarticulated constituents, personal pronouns, deictics and demonstratives (Blakemore 1992; Wilson and Sperber 1993; Scott 2013, 2016; Wilson 2016), and inflections for tense and aspect (Amenós Pons 2011; Escandell Vidal and Leonetti 2011; de Saussure 2012; Moeschler 2016) trigger inferences contributing to the delimitation of the proposition expressed. In turn, the denotation of the concepts encoded by content words is narrowed, broadened or both, and *ad hoc concepts* are constructed (Carston 2002, 2013; Wilson and Carston 2007). Enactment of those inferences is possible thanks to various *procedures* –i.e. processing instructions– and yields an enriched propositional form amounting to the *lower-level explicature* of the utterance (Sperber and Wilson 1995).

Lower-level explicatures may be further embedded into conceptual schemas capturing the action that speakers perform by speaking, their psychological states while speaking or about what they say, the sort and amount of evidence for it, or their degree of commitment to it. Construction of such schemas, which are labeled *higher-level explicatures*, is also assisted by other procedures encoded by mood indicators, word order, evidential adverbs –'obviously', 'evidently', etc.– *hearsay* adverbs –'allegedly', 'reportedly', etc.– *hearsay* particles –'*-tte*' in Japanese or 'ré' in Sissala– parenthetical phrases –'I/they say/claim', etc.– interjections, intonation or gestures (Blass 1989, 1990; Ifantidou 1993a, 1993b, 2001; Wilson and Sperber 1993; Itani 1998; Wharton 2003, 2009; Wilson and Wharton 2006; Wilson 2012, 2016). However, conceptual schemas capturing speakers' psychological states can also be constructed and take under their scope propositional constituents as a result of the procedures encoded by elements like expletives or insults (Padilla Cruz 2018, in press).

In Romance languages and other languages like Greek or Jordanian Arabic, diminutive and augmentative morphemes can be argued to also encode procedures yielding, on the one hand, ad hoc concepts and, on the other hand, schemas capturing speakers' attitudes towards the referents of the nouns, adjectives and adverbs to which they are attached. This presentation will extend this procedural analysis to another type of evaluative morphology: pejorative morphemes. It will argue that their semantics is, like that of diminutive and augmentative morphemes, procedural. However, as opposed to those morphemes, pejorative ones solely enact procedures whose output is the representation of a range of negative or disparaging attitudes towards the referents of the concepts encoded by the lexical items to which they are attached. Accordingly, in contrast to diminutive and augmentative morphemes, which may be claimed to encode a *meta-procedure* activating procedures for ad hoc-concept

construction or construction of affective attitude schemas, this presentation will portray pejorative morphemes as monosemic procedural elements (Carston 2016), which simply encode one procedure: construction of an attitudinal description confined to a particular propositional constituent. This analysis enables to place pejorative morphemes within the heterogeneous group of the so-called *expressives* (Blakemore 2015).

References

Amenós Pons, José (2011), Cross-linguistic variation in procedural expressions: Semantics and pragmatics, in M. V. Escandell Vidal, M. Leonetti, and A. Ahern (eds.), (2011), *Procedural Meaning: Problems and Perspectives*, Bingley: Emerald Group Publishing, 235-266.

Blakemore, Diane (1992), *Understanding Utterances*. *An Introduction to Pragmatics*, Oxford: Blackwell.

Blakemore, Diane (2015), Slurs and expletives: A case against a general account of expressive meaning, *Language Sciences* 52, 22-35.

Blass, Regina (1989), Pragmatic effects of co-ordination: The case of 'and' in Sissala, *UCL Working Papers in Linguistics* 1, 32-51.

Blass, Regina (1990), *Relevance Relations in Discourse: A Study with Special Reference to Sissala*, Cambridge: Cambridge University Press.

Carston, Robyn (2002), *Thoughts and Utterances. The Pragmatics of Explicit Communication*, Oxford: Blackwell.

Carston, Robyn (2013), Word meaning, what is said and explicatures, in C. Penco, and F. Domaneschi (eds.), (2013), *What Is Said and What Is Not*, Stanford: CSLI, 235-266.

Carston, Robyn (2016), The heterogeneity of procedural meaning, *Lingua* 175-176, 154-166.

de Saussure, Louis (2012), Temporal reference in discourse, in K. Allan, and K. Jaszczolt (eds.), (2012), *The Cambridge Handbook of Pragmatics*, Cambridge: Cambridge University Press, 423-446.

Escandell Vidal, María V., and Manuel Leonetti (2011), On the rigidity of procedural meaning, in M. V. Escandell Vidal, M. Leonetti, and A. Ahern (eds.), (2011), *Procedural Meaning: Problems and Perspectives*. Bingley: Emerald Group Publishing, 81-102.

Ifantidou, Elly (1993a), Sentential adverbs and relevance, *Lingua* 90, 69-89.

Ifantidou, Elly (1993b). Parentheticals and relevance, *UCL Working Papers in Linguistics* 5, 193-210.

Ifantidou, Elly (2001), *Evidentials and Relevance*, Amsterdam: John Benjamins.

Itani, Reiko (1998), A relevance-based analysis of hearsay particles: With special reference to Japanese sentence-final particle *tte*, in R. Carston, and S. Uchida (eds.), (1998), *Relevance Theory. Applications and Implications*, Amsterdam: John Benjamins, 47-68.

Moeschler, Jacques (2016), Where is procedural meaning located? Evidence from discourse connectives and tenses, *Lingua* 175-176, 122-138.

Padilla Cruz, Manuel (2018), Expressive APs and expletive NPs revisited: Refining the extant relevance-theoretic procedural account, *Lingua* 205, 54-70.

Padilla Cruz, Manuel (in press), Qualifying insults, offensive epithets, slurs and expressive expletives: A relevance-theoretic approach, *Journal of Language Aggression and Conflict*.

Scott, Kate (2013), This and that: A procedural analysis, *Lingua* 131, 49-65.

Scott, Kate (2016), Pronouns and procedures: Reference and beyond, *Lingua* 175-176, 69-82.

Sperber, Dan, and Deirdre Wilson (1995), *Relevance. Communication and Cognition*, 2nd edition, Oxford: Blackwell.

Wharton, Tim (2003), Interjection, language, and the 'showing/saying' continuum, *Pragmatics and Cognition* 11, 39-91.

Wharton, Tim (2009), *Pragmatics and Non-verbal Communication*, Cambridge: Cambridge

University Press.
Wilson, Deirdre (2012), Modality and the conceptual-procedural distinction, in E. Wałaszewska, and A. Piskorska (eds.), (2012), *Relevance Theory. More than Understanding*, Newcastle: Cambridge Scholars Publishing, 23-43.
Wilson, Deirdre (2016), Reassessing the conceptual-procedural distinction, *Lingua* 175-176, 5-19.
Wilson, Deirdre, and Robyn Carston (2007), A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts, in N. Burton-Roberts (ed.), (2007), *Pragmatics*, Basingstoke: Palgrave, 230-259.
Wilson, Deirdre, and Dan Sperber (1993), Linguistic form and relevance, *Lingua* 90, 1-25.
Wilson, Deirdre, and Tim Wharton (2006), Relevance and prosody, *Journal of Pragmatics* 38, 1559-1579.

Unpacking Polish teachers' sexuality discourses: risk and concealment

Łukasz Pakuła & Joanna Chojnicka
(Adam Mickiewicz University in Poznań & University of Bremen)

Keywords: Poland; school; teacher; sexuality; discourse

Schedule: We ~~12.30~~ 12.00 Room 2

Polish school is "nationalistic, conservative, and unifying; it only recognizes one 'difference' – that of gender" (Dzierzgowska and Rutkowska 2008: 5). It thus serves to socialize young people to (re)produce and solidify heteronormative and heterosexist discourses. This observation has been corroborated by numerous studies concerned with inspecting gender and sexuality representations in textbooks (Kochanowski et al. 2013; Chmura-Rutkowska et al. 2015) or required readings (Rient et al. 2014), and also those triangulating textbook analyses with classroom interaction research (Pakuła et al. 2015; Pawelczyk and Pakuła 2015). Furthermore, sexual non-normativity, next to poverty, has surfaced as the axis of identity which triggers most bullying within the schooling environment (Gawlicz et al. 2015).

Our research project offers an unprecedented insight into Polish LGBT teachers' perspectives on the intersections of sexuality and education. Following Nelson (2009), we see teacher-mediated discourse as a powerful tool enabling the linguistic construction of sexuality discourses – with the potential to (co-)form students' perceptions and attitudes – instead of merely topicalizing LGBT themes. For this reason, LGBT teachers' perspectives were elicited in a focus group and follow-up interviews whose analysis, drawing on Feminist Discourse Analysis (Lazar 2014), Queer Linguistics (Motschenbacher 2011), and Conversation Analysis (Wilkinson and Kitzinger 2014), serves to unpack teachers' constructions of sexuality in educational sites.

Our data show the discursive struggle between the need for an inclusive classroom and the jeopardy that executing this need entails for their professional identity construction, and – as a consequence – their careers within the state-run educational system.

References

Chmura-Rutkowska, Iwona, Maciej Duda, Marta Mazurek and Aleksandra Sołtysiak-Łuczak (eds.), (2015), *Gender w podręcznikach. Projekt badawczy. Raport, T. 1-3.* Warszawa: Feminoteka.

Dzierzgowska, Anna and Ewa Rutkowska (2008), *Ślepa na płeć. Edukacja równościowa po polsku. Raport krytyczny*. Warszawa: Fundacja Feminoteka.

Gawlicz, Katarzyna, Paweł Rudnicki and Marcin Starnawski (eds.), (2015), *Dyskryminacja w szkole – obecność nieusprawiedliwiona. O budowaniu edukacji antydyskryminacyjnej w systemie edukacji formalnej w Polsce.* Warszawa: Towarzystwo Edukacji Antydyskryminacyjnej.

Kochanowski, Jacek, Zbigniew Lew-Starowicz, Robert Kowalczyk and Krzysztof Wąż (2013), *Szkoła milczenia: przegląd treści szkolnych podręczników do biologii, WOS i WDŻR pod kątem przedstawienia w nich problematyki LGBTQ i treści homofobicznych.* Toruń: Stowarzyszenie Na Rzecz Lesbijek, Gejów, Osób Biseksualnych, Osób Transpłciowych Oraz Osób Queer "Pracownia Różnorodności."

Lazar, Michelle M. (2014), Feminist Critical Discourse Analysis: Relevance for Current Gender and Language Research, in: S. Ehrlich, M. Meyerhoff and J. Holmes (eds.), (2014), *The Handbook of Language, Gender, and Sexuality*. Oxford: Blackwell Publishing, 180–199.

Motschenbacher, Heiko (2011), Taking Queer Linguistics further: sociolinguistics and critical heteronormativity research, *International Journal of the Sociology of Language* 212, 149–179.

Pakuła, Łukasz, Joanna Pawelczyk and Jane Sunderland (2015), *Gender and sexuality in English language education: Focus on Poland*. London: British Council.

Pawelczyk, Joanna and Łukasz Pakuła (2015), Constructing gender and sexuality in the EFL classroom in Poland: Textbook construction and classroom negotiation?, in: S. Mills and A. S. Mustapha (eds.), (2015), *Gender Representations in Learning Materials in an International Context*. London: Routledge, 193–212.

Rient, Robert, Ewelina Seklecka, Mateusz Walczak, Anna Walicka and Edyta Zierkiewicz (eds.), (2014), *Męskośc i kobiecość w lekturach szkolnych. Analiza treści lektur w szkole podstawowej i gimnazjum z perspektywy równości płci. Raport z badań.* Wrocław: Fundacja Punkt Widzenia.

Wilkinson, Sue and Celia Kitzinger (2014), Conversation Analysis in Language and Gender Studies, in: S. Ehrlich, M. Meyerhoff and J. Holmes (eds.), (2014), *The Handbook of Language, Gender, and Sexuality*. Oxford: Blackwell Publishing, 141–160.

A multifactorial approach to possession constructions: Corpus-Driven study of internal and external possession in Brazilian Portuguese

Dafne Palú & Olaf Mikkelsen
(Universidade Católica Portuguesa & Université Paris 8)

Keywords: multifactorial feature analysis, construction grammar, possession construction, alternations, Portuguese

Schedule: Thu 9.00 Room 2

Possession as a semantic relationship can be expressed in different ways. In internal possession constructions (IPCs), the possessor (PR) and the possessum (PM) form a single NP constituent (of which the PM is the head), while in external possession constructions (EPCs) the PR appears as a verbal argument (Van Peteghem 2001, Heine 1997), usually encoded by a dative pronoun in European languages (Haspelmath 1999). While the alternation between these two types of constructions is very common cross-linguistically (Payne & Barshi 1999, Haspelmath 1999), EPCs are usually more restricted (Martin 1999), and the exact conceptual-functional motivations behind the choice of

construction remain unclear. The present study contrasts IPCs and EPCs in Brazilian Portuguese (BP) in which the former (1) is expressed with a possessive pronoun and the latter (2) with a dative clitic. Remarkably, dative EPCs are still available in European Portuguese (EP), but they have nearly disappeared from BP, except in idioms or in written literary speech (Lamiroy 2001, Morais & Ribeiro 2014).

(1) Toquei a sua mão
touch-PST.1 her-DEF hand
'I touched his/her hand'

(2) Toquei-lhe a mão
touch-PST.1 she-DAT hand-DEF
'I touched his/her hand'

A series of complex factors governing this alternation have been proposed in the literature. For instance, it is said that EPCs tend to map a relationship between an "animate, human and ego" PR (Payne & Barshi, 1999: 14) and an inalienable PM (e.g. body part or kin term). Haspelmath (1999: 111) explains that external PRs "are only possible if the possessor is thought of as being mentally affected by the described situation", while PMs encoded by lower level grammatical relations (e.g. oblique objects vs. transitive subjects) favor EPCs (1999: 113).

This study employs this study employs multifactorial feature analysis (MFA) (Dirven *et al*. 1982, Geeraerts *et al*. 1994, Gries 2003) to confirm quantitatively the role of these factors in BP. Thanks to MFA, we can operationalize hypotheses in terms of usage-features, which allows us to annotate and test them for their descriptive and predictive accuracy. Another advantage is that the set of features employed can be formalized in terms of an attribute-value matrix permitting direct integration of the results into any phrase-structured grammar, in this study, Construction Grammar (Goldberg 1995, 2006).

The data are drawn from a corpus of literary texts from the late 19th century, contrasting the narration's monitored speech and the imitation of colloquial and oralizing patterns in the dialogues. The analysis manually annotates every example of IPC and EPC in the 3rd person within the corpus. The feature set is based on the results of earlier introspective studies, mentioned above, and mixed-effects binary logistic regression is used to model the results of the feature analysis. We hope to identify the syntactic, semantic and pragmatic factors at play, and quantify their relative contributions to the choice of construction.

References

Dirven, R., L. Goossens, Y. Putsey & E. Vorlat, E. (1982). *The scene of linguistic action and its perspectivization by SPEAK, TALK, SAY, and TELL*. Amsterdam: Benjamins.

Geeraerts, D., Grondelaers, S. & Bakema, P. (1994). *Structure of Lexical Variation. Meaning, naming and context*. Berlin: Mouton.

Goldberg, Adele. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele. (2006). *Constructions at work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gries, St. Th. (2003). *Multifactorial Analysis in Corpus Linguistics: A study of particle placement.* Continuum: London.

Haspelmath, M. (1999). External Possession in a European Areal Perspective. In D. L. Payne, & I. Barshi (Eds.), *External Possession* (pp. 109-135). Amsterdam/Philadelphia: John Benjamins.

Heine, B. (1997). *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.

Lamiroy, B. (2001). La construction à possesseur externe dans les langues romanes. Dans *The*

Expression of Possession in Romance and Germanic Languages (pp. 101-114). S.l.: Clusium.

Martin, J. B. (1999). External Possession in Creek. In D. L. Payne, & I. Barshi, *External Possession* (pp. 229-250). Amsterdam/Philadelphia: John Benjamins.

Morais, M., & Ribeiro, I. (2014). Possessivos de 3ª pessoa: o português arcaico e o português brasileiro contemporâneo. *Filologia e Linguística Portuguesa*, 16 (spe.), pp. 15-51. doi:http://dx.doi.org/10.11606/issn.2176- 9419.v16ispep15-51

Payne, D. L., & Barshi, I. (1999). External Possession: What, Where, How, and Why. In D. L. Payne, & I. Barshi (Eds.), *External possession* (pp. 3-29). Amsterdam/Philadelphia: John Benjamins.

Van Peteghem, M. (2001). Datif possessif et inaliénabilité en français, en roumain et en russe. In L. Tasmowski (Ed.), *The Expression of Possession in Romance and Germanic Languages* (pp. 149-161). S.l.: Clusium.

A case of morphologically bound complementation in Abaza

Anastasia Panova
(National Research University Higher School of Economics, Moscow)

Keywords: Northwest Caucasian languages, Abaza, complementation, grammaticalization, morphologization

Schedule: We 17.00 Room 6

One of the well-known and cross-linguistically common varieties of clause reduction is clause union where the matrix and complement predicates share one set of grammatical relations (Noonan 2007: 84) but the predicates are still morphologically independent. However, the mirror-image of clause union, i.e. the situation when two predicates become a single verb morphologically but retain their syntactic independence, is also attested. Among the known examples are the so-called verificative in Agul (Maisak 2016) and certain complement types in Yaqui (Guerrero 2008). In the present study I discuss another example of this strategy: the construction with the element *ʒəš'a* 'seem' in Abaza, a polysynthetic Northwest Caucasian language. The study is based on fieldwork data collected in the village Inzhich-Chukun (Abazinsky district, Karachay-Cherkes Republic, Russia) in 2017-2018.

In Abaza grammars (e.g. Tabulova 1976: 209-210) *ʒəš'a* is described as a verbal suffix. However, wordforms with *ʒəš'a* in fact contain two morphologically complex predicates: the matrix verb and its sentential complement (1).

(1) [*awəj* *d-ʕa-j-wa-š*]-***ʒə***-*w*-***š'***-*əj-ṭ*
DIST 3SG.H.ABS-DIR-go-IPF-FUT-**LOC**-2SG.M.IO-**seem**-PRS-DCL
'You think he will come.'

That the construction with *ʒəš'a* is a single wordform is proven by negation, which in finite forms is marked jointly by the affixes *g'(ə)-* and *-m* (2).

(2) *sara* [*d-ʕa-j-ta*] ***g'***-*qa-s-ç-wə*-***m***
I 3SG.H.ABS-DIR-go-ADV **NEG**-LOC-1SG.ERG-believe-IPF-**NEG**
'I don't believe he came.'

When negation applies to a construction with *ʒəš'a* (3), the prefix *g'-* appears in the prefixal part of the whole construction, even when only its second part (the main clause) is negated.

(3) [*awəj* *d-**g'**-ʕa-j*]-*ʒə-s-š'-wə-**m***
DIST 3SG.H.ABS-**NEG**-DIR-go-LOC-1SG.IO-seem-IPF-**NEG**
'I don't think he came.'

The second diagnostic is the formation of temporal subordinate clauses. When a dependent clause contains a complementation construction, the temporal prefix *an(ə)-* appears on the matrix predicate (4).

(4) *d-š-psə-z* ***anə**-l-ba*
3SG.H.ABS-REL.MNR-die-PST.NFIN **TMP**-3SG.F.ERG-see
d-ĉ̣əwa *d-a-la-ga-ṭ*
3SG.H.ABS-cry.IPF 3SG.H.ABS-3SG.N.IO-LOC-begin-DCL
'When she saw that he had died, she started crying.'

In the construction with *ʒəš'a* (5) the prefix *an(ə)-* appears to the left of the dependent verb stem, even though it modifies the matrix verb.

(5) [*d-**an**-psə*]-*ʒə-l-š'a*
3SG.H.ABS-**TMP**-die-LOC-3SG.F.IO-seem
d-ĉ̣əwa *d-a-la-ga-ṭ*
3SG.H.ABS-cry.IPF 3SG.H.ABS-3SG.N.IO-LOC-begin-DCL
'When she thought he had died, she started crying.'

Although the two predicates in the construction with *ʒəš'a* behave as a single wordform, they retain their syntactic independence: each of the predicates has its own tense markers, argument structure, and can be separately modified by adverbs (6).

(6) *sara* ***pasata*** [*wara* ***ŝabəžta*** *w-ʕʷ-**wa***]-*ʒ-s-š'-**wə-n***
I **earlier** you **fast** 2SG.M.ABS-run-**IPF**-LOC-1SG.IO-seem-**IPF-PST**
'Before, I thought you run fast.'

From a diachronic perspective, morphologically bound complementation, like clause union, seems to be an intermediate stage of grammaticalization of a matrix verb into a verbal affix. But while in case of clause union syntactic integration of clauses comes first and morphological integration is the second stage of development, in the case of morphologically bound complementation the order is reversed: first morphological union, then (if at all) syntactic integration. Thus, the most interesting questions are what makes morphology outrun syntax, where else such morphologically and syntactically peculiar development can be expected and which properties of linguistic systems make it possible.

The study was supported by the Russian Science Foundation, grant No. 18-78-10128.

References

Guerrero, Lilián (2008), Yaqui complementation, an example of a syntax-semantic mismatch, in R. J. Edwards et al. (eds), (2008), *Proceedings from the Panels of the 41 Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, 175–188.

Maisak, Timur (2016), Morphological fusion without syntactic fusion: The case of the "verificative" in Agul, *Linguistics* 54(4), 815–870.

Noonan, Michael (2007), Complementation, in T. Shopen (ed), (2007), *Language typology and syntactic description* (2nd ed.), vol. II, *Complex constructions*, Cambridge: Cambridge University Press, 52–150.

Tabulova, Nurja T. (1976), *Grammatika abazinskogo iazyka. Fonetika i morfologiia* [A grammar of Abaza. Phonetics and morphology], Cherkessk: Karachaevo-Cherkesskoe otdelenie Stavropol'skogo knizhnogo izdatel'stva.

Assessing cross-linguistic similarity with *Sound Comparisons*

Ludger Paschen & Paul Heggarty
(Leibniz-Zentrum Allgemeine Sprachwissenschaft, MPI-SHH Jena)

Keywords: phonetics, database, comparative linguistics, acoustic, voice onset time

Schedule: Thu 14.00 Room 9

With over 90,000 single word recordings and c. 50,000 narrow phonetic transcriptions, the *Sound Comparisons* online database offers a unique way of exploring phonetic diversity in a diverse range of language families from across the globe, from Quechua and Mapudungun to Oceanic and several branches of Indo-European. At soundcomparisons.com, users can instantaneously hear and compare word-by-word recordings, and customise sub-sets of languages and phonetic data using phonologically-informed filters for a range of features. Any custom selection of sound files and transcriptions can be downloaded, shared and cited by a unique URL. Transcriptions follow new initiatives for standardising usage of phonetic symbols in Unicode (Anderson *et al*. forthcoming). The intuitive (and multilingual) interface turns *Sound Comparisons* into an accessible and powerful resource for investigating phonetic variation across languages.

After briefly presenting this functionality, we illustrate applications in several fields. *Sound Comparisons* has particular value for researching cross-linguistic contrasts on the phonetic dimension, to a high level of resolution. This is what distinguishes *Sound Comparisons* from the best-known analyses of large-scale cross-linguistic databases, from lexicostatistics (Swadesh 1952) to the latest phylogenetic models (e.g. Chang *et al*. 2015). Those compare languages only in cognate status, a simple binary (or at most, multi-state) variable, irrespective of phonological form. Other algorithms perform pair-wise comparisons of segments, and assign similarity scores on phonological resemblance (e.g. List *et al*. 2018). But there is little consensus on how to define such scores, and most studies use only broad phonemic transcriptions, ignoring even major phonetic differences.

In *Sound Comparisons*, however, most data-sets target not meaning lists but known cognates (irrespective of meaning shifts), to support high-resolution comparison of net differences in phonetics, arisen since a known common ancestral stage. Early publications focused on phonetic distances in English dialectology (McMahon *et al*. 2007, Maguire *et al*. 2010) and across Germanic languages (Heggarty *et al*. 2010). *Sound Comparisons* has no bias towards traditional isoglosses, and invites instead multifactorial approaches to dialect continua, migrations and mixing of speech communities, illustrated here within German dialectology. Fine phonetic transcriptions have also contributed to the classification debate on Huilliche *vs*. Mapudungun (Sadowsky *et al*. 2015).

Most recently, Labov *et al*. (2013) and Pigoli *et al*. (2018), among others, have advocated factoring in acoustic detail for large-scale comparative studies. Our final application illustrates how novel quantitative comparisons can exploit not just high-resolution transcriptions, but also the acoustic information in the recordings in *Sound Comparisons*. We explore the acoustic reality of a primary diagnostic criterion in the internal classification of West Germanic: stop voicing. Quantifying differences in voice onset time (VOT) between dialect clusters yields surprising results. Instead of a

clear-cut voiced/voiceless split, we find a relatively smooth gradient in VOT, from negative though short to long. Voiced reflexes of the Proto-Germanic *D series are in Low German actually more similar to some voiceless stops in High German varieties than to some heavily pre-voiced stops in varieties of English. We close by discussing what qualification this might entail for this key criterion in West Germanic.

References

Anderson, Cormac, Tiago Tresoldi, Thiago Chacon, Anne-Maria Fehn, Mary Walworth, Robert Forkel & Johann-Mattis List, (forthcoming), CLTS: A Cross-linguistic database of phonetic transcription systems. *Yearbook of the Poznań Linguistic Meeting 4 (2018),* 1-27. http://doi.org/10.2478/yplm-2018-0001

Chang, Will, Chundra Cathcart, David Hall & Andrew Garrett. 2015. Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91(1), 194–244. http://doi.org/10.1353/lan.2015.0005

Heggarty, Paul, Warren Maguire & April McMahon. 2010. Splits or waves? Tree or webs? How divergence measures and network analysis can unravel language histories. *Philosophical Transactions of the Royal Society B* 365, 3829–3843. http://doi.org/10.1098/rstb.2010.0099

Labov, William, Ingrid Rosenfelder & Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: linear incrementation, reversal, and reanalysis. *Language* 89(1), 30–65. http://doi.org/10.1353/lan.2013.0015

List, Johann-Mattis, Mary Walworth, Simon J. Greenhill, Tiago Tresoldi & Robert Forkel (2018), Sequence comparison in computational historical linguistics, *Journal of Language Evolution* 3(2), 130–144. http://doi.org/10.1093/jole/lzy006

Maguire, Warren, April M. S. McMahon, Paul Heggarty & Dan Dediu. 2010. The past, present and future of English dialects: Quantifying convergence, divergence and dynamic equilibrium. *Language Variation and Change* 22 (1): p.69–104. http://doi.org/10.1017/S0954394510000013

McMahon, April M.S., Paul Heggarty, Robert McMahon & Warren Maguire. 2007. The sound patterns of Englishes: representing phonetic similarity. *English Language and Linguistics* 11 (01): p.113–142. https://www.cambridge.org/core/journals/english-language-and-linguistics/ell-20th-anniversary-collection

Pigoli, Davide, Pantelis Z. Hadjipantelis, John S. Coleman & John A. D. Aston. 2018. The statistical analysis of acoustic phonetic data: exploring differences between spoken Romance languages. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 67(5), 1103–1145. http://doi.org/10.1111/rssc.12258

Sadowsky, Scott, María José Aninao, María Isabel Cayunao & Paul Heggarty. 2015. Huilliche: ¿geolecto del mapudungun o lengua propia? Una mirada desde la fonética y la fonología de las consonantes. In: Fernández Garay, Ana & María Alejandra Regúnaga (eds) *Lingüística indígena sudamericana: aspectos descriptivos, comparativos y areales*. Buenos Aires: Editorial de la Facultad de Filosofía y Letras, Universidad de Buenos Aires.

Swadesh, Morris. 1952. Lexico-statistic dating of prehistoric ethnic contacts: with special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96(4): p.452–463. http://www.jstor.org/stable/3143802

Non-verbal predication and copular constructions in Maa (Eastern Nilotic)

Doris L. Payne
(University of Oregon and SIL International)

Keywords: syntax, nonverbal predication, informationstructure, tense-aspect-mood

Schedule: Sat 11.00 Room 1

In non-verbal predication, the lexical portion of a predicate is not a verb (Hengevald 1992; Overall, Vallejos & Gildea 2018). Non-verbal predication involving copulas is important for understanding the origin of cleft and certain focus constructions (Declerck 1988), given how copulas relate distinct portions of a predication. Their relational function also motivates their development as conjunctions (Amfo 2010), though their development into inter-clausal conjunctions is not well studied.

Non-verbal predication has not been systematically investigated in Eastern Nilotic languages, though Tucker & Mpaayei (1955: 90-93) list paradigms of Maa 'be', and Dimmendaal (1982: 74-78) gives a somewhat abstract analysis for Turkana. Maa (Kenya, Tanzania) has a rich inventory of copulas for expressing identity, set-inclusion, and attribution. This paper examines meaning contrasts among the copulas *ra, taa, naa* and *akʊ* in non-verbal predication, and motivates their differential development into discourse-related cleft elements, intensifiers, and interclausal connections. The study is based on corpus analysis of over 120 texts, supplemented by elicitation.

Ra inflects for subject and indicates non-perfect(ive) aspect and realis modality (1). It can take NP or adjective complements. *Ra* does not appear to function outside of non-verbal predication.

(1) á-**rá** ɔl=áikɔ́shúaani.
1SG-be MSG=glutton
'I am a glutton.'

As a lexical verb, non-perfect(ive) *akʊ(n)* 'become' inflects for person (2).

(2) n-í-**ˬákʊ́** íyˬié ɛn=dáa ɛnyɛ̂.
CN-2-become you.SG.NOM FSG=food 3PL.POSS
'You will become their food.'

However, the dominant use of *akʊ(n)* is as a discourse conjunction in the frozen 3rd-person form *nɛ́ˬákʊ́* 'and it becomes', roughly meaning 'and then, so, therefore'. The inchoative sense has been metaphorically extended to express that a situation "comes into being" or is (weakly) consequent to a temporally or logically prior situation.

The copula *taa*, formed on *aa* 'be', is inflected for perfect(ive) aspect (3) and/or subjunctive mood (4). This copula also inflects for subject and has inchoative meaning (which always arises when the morphological perfect(ive) combines with stative verbs).

(3) Ɛ́-**táá** ɛn=kíne in=kírí.
3-be.PF FSG=goat.NOM FPL=meats
'The goat has become meat.'

(4) ɛ-**taá** ɪl=tʊŋaná sapúkī
3-be.SBJV MPL=people.NOM big.PL
'When the people (can/should/might) become grown…'

As a discourse particle, *taa* does not inflect for person. It communicates exclusive or specificational focus in cleft-like constructions (5). It may co-occur with other verbs or the *ra* copula (6), where its modality sense gives rise to an emphatic intensifier.

(5) en=títō **taá** ınâ n-ɛıtó apá erikorí
FSG=girl FOCUS that.NOMREL.F-NEG.PF before 3-lead-ITIVE-PASS
'That is a girl who has not been led.' (compare: 'That girl has not been led.')

(6) Tɛ n-í-r↓á **taá** ol=ó-ítīēū...
if CN-2-be EMPHATIC MSG=MSG.REL.NOM-dare
'If you are really the one who can dare…'

Finally, a related form *naa*, also containing *aa*, also functions as a copula in non-verbal predication, but is especially prominent in cleft-like constructions and as an inter-clausal conjunction akin to 'and'. The study deepens typological knowledge of the relationship between non-verbal predication, clefts, and interclausal conjunctions, and demonstrates the unique profiles that copular forms, potentially involving the same historical verb root, may come to have.

References

Amfo, Nana Aba A. (2010) Noun phrase conjunction in Akan: The grammaticalization path, *Pragmatics* 20, 27-41.

Declerck, Renaa (1988), *Studies on Copular Sentences, Clefts, and Pseudo-clefts*, Leuven University Press/Foris.

Dimmendaal, Gerrit (1982), *Turkana*, Dordrecht: Foris.

Hengeveld, Kees (1992), *Non-verbal Predication: Theory, Typology, Diachrony*, Berlin: Mouton de Gruyter.

Overall, Simon, Rosa Vallejos & Spike Gildea (eds). (2018), *Nonverbal Predication in Amazonian Languages*, Amsterdam: John Benjamins.

Tucker, A. N. & J. T. O. Mpaayei (1955), *A Maasai Grammar, with Vocabulary*, London: Longmans, Green & Co.

Transitivity and Modality in Waray

Thomas E. Payne & Voltaire Q. Oyzon
(SIL International and University of Oregon, USA & Leyte Normal University, Tacloban City, Philippines)

Keywords: transitivity, modality, voice, inflection, verbs, grammatical relations, Philippine languages

Schedule: Fri 11.30 Room 6

Philippine languages are well-known for complex verb morphology. Recent approaches have treated verbal affixation as largely a matter of "voice", with modal and temporal notions being secondary nuances of voice affixes, plus lexical subcategorization of roots (see, e.g., Himmelmann 2008). In this paper we provide a fresh approach to verb morphology in Waray, the mother tongue and language of wider communication for most inhabitants of Samar, Eastern Leyte, and parts of Biliran islands in the Eastern Visayas region of the Philippines.

In this paper we argue that, counter to all other approaches, verbal inflection in Waray directly expresses exactly two dimensions which we identify as grammatical Transitivity and Modality. Relative time ("tense"), and temporal structure ("aspect") may be inferred based on various features of the grammatical construction and the context but are not directly coded in the inflectional paradigm. There are exactly twelve inflectional categories in Waray. These are strictly paradigmatic, in that at

least one and no more than one occurs with every verb functioning as a predicate. Grammatical transitivity (transitive or intransitive) is reflected in eleven of the twelve categories.

The "voice system" of Waray arises as a consequence of the interaction between this inflectional paradigm, and two stem-forming (or "derivational") applicative categories. The two applicative stem-forming categories (expressed with the suffix *-an* and prefix *i-*) create transitive stems with locative, instrumental, or benefactive absolutive arguments.

Another discussion that this study contributes to is whether or not roots in Philippine languages are "precategorial," i.e., not inherently categorized as nouns, verbs or modifiers (Foley 1998). While it is the case that the vast majority of roots may be used as Predicates or Referring expressions, this paper shows that in Predicative constructions, transitivity is a robust and pervasive distinction. In this sense, roots are not precategorial -- underived roots as well as derived stems have "basic" transitivity values when used predicatively.

We document and give evidence for all claims from a large corpus of Waray written and conversational data (3NS Corpora project). We argue that this approach significantly refines and simplifies the general understanding of how predicates in Philippine languages are constructed, including the famous Philippine "voice" system.

References

3NS Corpora project. (https://corporaproject.org/). National Network of Normal Schools, Philippines.

Foley, William A. (1998), Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the Workshop on Voice and Grammatical Functions in Austronesian Languages, LFG98 Conference, Brisbane, Australia.

Himmelmann, Nikolaus (2000), Lexical categories and voice in Tagalog. In Austin, Peter K. & Simon Musgrave, Voice and grammatical relations in Austronesian languages. Stanford: Center for the Study of Language and Information.

How sentential and nominal anaphora differ: Evidence from Russian

Olga Pekelis
(Russian State University for the Humanities)

Keywords: anaphora, proposition, Russian, pronoun, antecedent

Schedule: Fri 14.30 Room 7

Based on the Russian corpus data, this study investigates the features that distinguish the sentential anaphora (reference to a proposition) from the nominal one (reference to an individual). The former has received noticeably less attention in the literature than the latter (Ascher 1993, Gast&König 2008, Needham 2012 being among the few examples), and ultimately, sentential anaphora has been claimed to show “many (if not all) of the same behaviours as anaphora in other domains” (Snider 2017: 56). The following suggests, however, that this is not exactly the case.

The relationship between a pronoun and its antecedent is *more flexible* (i.e. *less restrictive* and *less clear-cut*) for the sentential anaphora than for the nominal one.

The propositional referent does not usually coincide syntactically with the antecedent (hence, the term *antecedent-trigger*, coined by Cornish 1992). In (1), the pronoun *ètogo* refers to a proposition that is in the scope of the imperative, but not to the imperative meaning itself:

(1) *[Ostavajsja zdes']*$_i$, *esli ty tak ètogo*$_i$ *hočeš.*
stay.IMP here if you.NOM so this.GEN want.PRS.2SG
'Stay here, if you wish it so much.' (RNC)

Contrary to nominal antecedents, sentential antecedents are *syntactically manifold*: they can be independent or dependent clauses, finite or non-finite clauses etc.

It is not always clear which exactly syntactic constituent should be treated as an antecedent(-trigger). Thus, in (2a) and (2b) the proposition denoted by the pronoun *èto* is the same; the antecedents, however, seem at first sight to be syntactically different.

(2) a. [*Sem'ja bolšaja*]$_i$ — *èto*$_i$ *horošo.*
family.NOM big.NOM this good
'The family is big, this is good.' (RNC)
b. [*Čto sem'ja bolšaja*]$_i$, *èto*$_i$ *horošo.*
that family.NOM big.NOM this good
'That the family is big, that is good.'

Sentential anaphora is sensitive to the sentence boundary in a way nominal anaphora does not seem to be. On the one hand, for instance, conditional and restrictive relative clauses provide a propositional discourse referent only in case the pronoun is part of the same sentence (Snider 2017). On the other hand, there are adverbial clauses in Russian that are available as antecedents only from outside the sentence they are part of.

Sentential pronouns differ from the nominal ones both in grammatical categories for which they are marked and in parameters by which they differ from each other. The basic sentential pronouns in Russian have default number and gender values, and the opposition between personal and reflexive pronouns, fundamental for the nominal anaphora, has no parallels in the sentential domain. Instead, Russian sentential pronouns differ in terms of the referential properties of the situation they denote (Letuchij 2011) and with respect to the available syntactic types of their antecedents.

I argue, within the functional-typological framework, that the phenomena listed in 1-3 are essentially due to more general distinctions between a proposition and an individual, and provide an account for 1-3 that is in line with this assumption. I also elaborate on a separate definition of a sentential antecedent that takes into consideration its peculiar traits.

This research has been supported by the Russian Foundation for Humanitarian Research, grant # 17-04-00517(a).

Abbreviations

2 = second person; GEN = genitive; IMP = imperative; NOM = nominative; PRS = present; SG = singular.

References

Asher, Nicholas (1993), *Reference to abstract objects in English*, Dordrecht: Kluwer Academic Publishers.

Cornish, Francis (1992), So be it: The discourse–semantic roles of *so* and *it*, *Journal of Semantics* 9(2), 163–178.

Gast, Volker & König, Ekkehard (2008), Sentence anaphora in English and German. *Handout presented at the First meeting of the International Society for the Linguistics of English*, University of Freiburg.

Letučij, Alexander B. (2011), Pronominalizacija sentencial'nogo argumenta v russkom jazyke (Pronominalization of sentential arguments in Russian), in *Proceedings of the conference «Dialog»*, Moscow.

Needham, Stephanie Maureen (2012), Propositional anaphora in English: The relationship between *so* and discourse, *Carleton University dissertation*.

RNC - Russian National Corpus. [URL: www.ruscorpora.ru]

Snider, Todd Nathaniel (2017), Anaphoric reference to propositions, *Cornell University dissertation*.

Diachronic paths of future anterior in Slavic from a typologic perspective

Yana Penkova

(V.V. Vinogradov Institute of Russian language, Russian Academy of Sciences)

Keywords: future anterior, grammaticalization, aspect specification, epistemic modality, Slavic languages

Schedule: Thu 9.30 Room 1

Acknowledgments: This study is supported by the Russian Foundation for Basic Research (research project no. 17-34-01061-OGN “Slavic future anterior in the typological perspective”)

The paper is dedicated to the Slavic future anterior (FA), paths of its diachronic development and possible explanations of them regarding their interconnection with the aspect, syntactic behavior, modality, and evidentiality. Under the label of FA fall the constructions, which are simultaneously marked for future tense and perfect or anterior: e.g., *bǫdǫ* + *l*-PTCP in Old Church Slavonic, Old Check, Old Russian, Old and modern BCMS (and others), *budu* + *n*-PTCP in Northern Russian, *biću* + *l*-PTCP in Čakavian, *ще съм* + *l*-PTCP in Bulgarian, *ќе имам* + *н*-PTCP in Macedonian. The data for the research is driven from modern and diachronic corpora of Slavic languages (diachronic corpora are available only for a part of Slavic languages), from editions of Slavic monuments (for historical data) and through a special translation questionnaire, compiled by Dmitri Sichinava and me (for modern Slavic).

FA has never been among separate TAM categories believed to deserve a particular interest. According to some influential authors (Bybee et al. 1994; Thieroff 2000), the meaning of FA (as well as the one of pluperfect – its counterpart in the domain of past) is supposed to be fully compositional. Beginning with Ö. Dahl (1985) and further to some recent approaches to pluperfect (Sichinava 2013) it has been shown that pluperfect can develop a variety of idiomatic secondary meanings. A similar approach to FA can allow us to identify a range of its non-compositional uses, e.g., epistemic, or presumptive uses that are cross-linguistically universal (found far from the SAE; cf. in Agul, Huallaga Kechua). Applied to diachronic data, such kind of approach allows tracing diachronic evolution of FA.

Slavic languages among other European languages seem to be the most advanced in the way they grammaticalize FA: two main grammaticalization routes can be traced.

Following the first one, FA undergoes a functional shift: it loses its retrospective function but preserves FTR. Such an evolution combined with some changes in aspect took place in the history of some Slavic languages (Polish, Rusyn, Western Ukrainian, Slovenian, Kajkavian), but it is not limited to Slavic (cf. Latin, Oscan, Umbrian, Dalmatian, Wymysorys). The opportunity to occur in conditional protasis and to neutralize with present tense pushes such development, cf. in Kajkavian:

Kajkavian: *Jako ti **jozdravim***[heal.PRS.1SG], *moja luba bodeš.* 'If I **heal** you, you will become my wife';

Kajkavian: *Ti baba, ako ne **budeš**[be.FUT.2SG] moje brate nazaj **oživila***[revive.PTCP], *onda moj cucek bude tebe zaklal.* 'You, woman, if you do not **revive** my brother back, my dog will bite you.'

Following the second path, FA loses its FTR and develops towards an epistemic-evidential marker, which, according to corpora and questionnaires, covers personal-non visual uses. The source of such development is the opportunity to mark an epistemic distance and deferred realization. Such development took place in Čakavian, e.g.:

Čakavian: ***Bićemo*** [be.FUT.1PL] ***spuali*** [sleep.PTCP] *ka si ti doša.*'We [probably] **were sleeping**, when you came'.

But the tendency towards such evolution can be observed in the history of East Slavic and Balkan Slavic.

References

Bybee, Joan L., Perkins, Revere & William Pagliuca (1994), *The Evolution of Grammar: Tense, aspect and modality in the languages of the world*, Chicago / London: The University of Chicago Press.

Dahl, Östen (1985), *Tense and aspect systems*, Oxford: Blackwell.

Sichinava, Dmitri V. (2013), *Tipologiya plyuskvamperfekta. Slavyanskij plyuskvamperfekt*, Moscow: AST-press.

Thieroff, Rolf (2000), On the areal distribution of the tense-aspect categories in Europe, In Ö. Dahl (ed.), *Tense and Aspect in the Languages of Europe*, Berlin / New York: Mouton de Gruyter, 265–308.

Insubordination and discourse-variation: a multi-genre analysis of quotative insubordination in Spanish

Sofía Pérez Fernández, Pedro Gras & Frank Brisard
(University of Antwerp)

Keywords: insubordination, discursive variation, quotatives, corpus analysis, interactional linguistics

Schedule: We 17.30 Room 2

This paper offers a quantitative and qualitative corpus-based analysis of sentence-initial *que* 'that' as a quotative marker, as in examples (1)-(3). Quotative *que* can receive three different interpretations (Gras 2011, 2013, 2016, Gras & Sansiñena 2015): echoic (1), copy (2) and indirect reported speech (3).

(1) A: Yo no fumo 'I don't smoke'
B: ¿*Que* tú no fumas?
COMP you no smoke-2.SG.PRS.-IND
'[did you say] that you don´t smoke?'

(2) A: Ven a cenar mañana
'come for dinner tomorrow'
B: ¿Qué?
'what?'

A: *Que* vengas a cenar mañana
COMP come-2SG.PRS.SBJ to have-dinner-INF tomorrow
'[I've said] you should come tomorrow for dinner'

(3) Ha llamado tu madre. *Que* vayas a verla.
COMP go-2SG.PRS.SBJ to see-INF er.OBJ.SG
'Your mum has called. [She said] that you should go see her'

It has been stressed that quotative constructions do not only have grammatical restrictions, but also discursive ones, as they only receive a quotative interpretation in highly specific discourse contexts (Gras 2011, 2013, 2016; Sansiñena 2015; Gras and Sansiñena 2015). However, what has not been noticed is the fact that these constructions also serve interpersonal functions, indicating information about the relation between participants in the communicative situation, which explains their discourse- and register-related restrictions. The aim of this paper is to fill this gap and examine whether these uses are sensitive to register (variation produced by the level of formality) or channel (spoken vs written language), as well as explore the kind of interpersonal information they encode.

For this purpose, we have conducted a quantitative analysis of a set of corpora that consists of five different genres: two oral (conversation and interview) and three written genres (social media, novels and news reports). The results show that quotative constructions are more productive in informal interactive genres (mostly spontaneous conversation and its representation in literary texts) than in formal registers (interviews and news reports), which suggests that they may perform interpersonal as well as discourse-connective functions simultaneously.

References

Gras, Pedro (2011), Gramática de construcciones en interacción. Propuesta de un modelo y aplicación al análisis de estructuras independientes con marcas de subordinación en español. Doctoral dissertation, University of Barcelona.

Gras, Pedro (2013), Entre la gramática y el discurso: valores conectivos de *que* inicial átono en español. In D. Jacob and K. Ploog (eds), (2013), *Autour de* que, Frankfurt am Main: Peter Lang, 89-112.

Gras, Pedro (2016), Revisiting the functional typology of insubordination: *Que*-initial sentences in Spanish. In N. Evans and H. Watanabe (eds). *Dynamics of Insubordination* (Typological Studies in Language), Amsterdam: John Benjamins, 113-143.

Gras, Pedro. and Sansiñena, María Sol (2015), An interactional account of discourse-connective *que*-constructions in Spanish, *Text and Talk* 35, 505-529.

Sansiñena, S. 2015. The multiple functional load of *que*. Doctoral dissertation. University of Leuven.

The interaction between extravagance and competition in syntactic change

Peter Petre

Schedule: We 17.30 Room 14

Grammatical material is typically rather unobtrusive. Yet sometimes we want to draw attention to a grammatical aspect of our message, and give it more expressive power or extravagance, an effect that may be a key motivation in (unidirectional) syntactic change (Meillet 1912, Detges & Waltereit 2002, Haspelmath 2000). This paper scrutinizes this notion by addressing how different types of extravagant

usage play different roles in bringing about such syntactic change (either or not as part of a grammaticalization process).
Two major types are distinguished, and will be discussed on the basis of existing as well as new data (from EMMA and PPCME2).

1. **Unexpected selection**: using an expression in a novel way, like *be going to* to express near future outside the context of motion. **Unexpected extension** (cf. Himmelman's 2004 host-class expansion), as in the first use of future *be going to* with inanimate subjects (Petré & Van de Velde 2018), is a milder variant, where an existing form-meaning pair is used in a context where another form is conventional.
2. **Redundancy**: adding an extra component to an already complete message. The redundant element can be:
    - Non-tautological and rich in meaning. A well-known example is the use of *pas* 'step' to reinforce negation (e.g., van der Auwera 2009);
    - Non-tautological but shallow in meaning, as *to be* in *grow to be so great* as an alternative to *grow so great* (own data). Here the addition, giving more phonetic weight to the expression, may be primarily iconic (Haiman 1980);
    - Tautological (multiple negation);

Each type triggers a different kind of competition. Unexpected selection triggers **paradigmatic competition** with a conventionalized alternative, potentially leading to the kind of long-term variation that is central to variationist sociolinguistics (e.g. *be going to* competing with *will*, Tagliamonte et al. 2009). Subsequent unexpected extensions of the incoming form imply that extravagance does is not just a first trigger of grammaticalization, but continues to be relevant in subsequent stages.
Redundancy may involve a more diverse **syntagmatic competition**. *Ne ... pas* does arguably not involve real competition, as the ousting of *ne* by *pas* presupposes a stage where the phrase was holistically interpreted. Meaningless *to be* did not really compete with *grow* and disappears again after its extravagant effect had worn down. True competition is however involved in the redundant combination of *when*- and *happen*-constructions, as in (1), where either construction could have introduced the sentence (cf. Brinton 1996).

(1) *Then fel hit, when scho was ded, fendys comen and fattyn hur soule to hell*. (*a*1500)
'Then it happened, when she was dead, devils came and fetched her soul to hell.'

Interestingly, their combination peaks early in their development as episode markers, suggesting that they both benefited from it. More generally, extravagant redundancy appears to be limited in time. Its effect wears down, leading to erosion, and *de facto*, syntactic change. As such, it is closely related to what are known as critical contexts in grammaticalization research (Diewald 2006).

References

Detges, Ulrich & Richard Waltereit. 2002. "Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar." *Zeitschrift für Sprachwissenschaft* 21. 151-195.
Diewald, Gabriele. 2006. Context types in grammaticalization as constructions. *Constructions* 1(9).
Haiman, John. 1980. The Iconicity of Grammar: Isomorphism and Motivation. *Language*. 56 (3): 515–540.
Haspelmath, Martin. 1999. "Why is grammaticalization irreversible?" *Linguistics* 37: 1043-1068.
Keller, Rudi. 1994. *On language change: The invisible hand in language*. London & New York: Routledge.
Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12: 130-148.

Petré, Peter & Freek Van de Velde. The real-time dynamics of the individual and the community in grammaticalization. *Language* 94(4). 867–901.

Tagliamonte, Sali A.; Mercedes Durham; and Jennifer Smith. 2014. Grammaticalization at an early stage: future be going to in conservative British dialects. *English Language & Linguistics* 18.75–108.

Van der Auwera, Johan. 2009. The Jespersen cycles. In Elly van Gelderen (ed.), *Cyclical change*, 35-71. Amsterdam: Benjamins.

The quantitative fingerprint of exaptation

Jozefien Piersoul, Freek Van de Velde & Karlien Franco
(University of Leuven, University of Leuven & University of Toronto)

Schedule: Fri 9.30 Room 14

Predicting when a particular change will emerge is next to impossible (Weinreich et al.'s 1968 problem of actuation), but the process of propagation is not haphazard. Like other social phenomena – which also suffer from the problem of actuation – propagation in language change is known to often follow an S-curve, with a slow onset, a rapid spread of an innovation in the middle, and a slow end phase. Historical linguists have shown that the logistic curve is a good approximation of many of the changes they study (Weinreich et al. 1968: 113, Kroch 1989, Denison 2003, Pintzuk 2003, and Blythe & Croft 2012), but they are generally not really interested in whether there is additional information in the residuals. The gold standard is to have an optimal fit between the curve and the data.

In this talk, we will see whether a deviating curve can actually reveal something. It will be investigated to what extent three recent changes in Dutch follow the parametrized logistic S-curve. More specifically, the regression parameter settings estimated on observed values from the period 1850-1949 in a register-stable acrolectal text corpus (120 million tokens) will be used to draw the future trajectory of the S-curve from 1950 to 1999. Subsequently, observed frequency proportions per year from 1950 to 1999 will be compared to the predicted proportions for those 'blinded' fifty years.

The three changes at issue are:

(a) the hortative alternation (*laat ons* infinitive vs. *laten we* infinitive) (Van de Velde 2017)
(b) the shift from demonstrative stems (*th*-paradigm) to interrogative stems (*wh*-paradigm) in relative pronouns with a superlative antecedent (*het mooiste dat ik ooit gelezen heb* vs. *het mooiste wat ik ooit gelezen heb*) (Van der Horst & Van der Horst 1999: 169-174)
(c) the demise of the inflection on the prenominal 1SG possessive (*mijne ouders* vs. *mijn ouders*) (De Vooys 1953: 92-93)

For each of these changes, we have collected datasets to reliably calculate the projected S-trajectory. We expect that compared to (a) and (b), the change under (c) will be 'undershot' by the estimation on the earlier part of the corpus. The change is not just a late instance of the slow overall schwa apocope of Modern Dutch (Van Loon 2014: 258-264). Rather, it is part of a currently ongoing exaptation of the adjectival schwa, which comes to be used as a marker for the watershed between the determiner and the adjective in the prenominal zone (see Van de Velde & Weerman 2014). We expect that this exaptation will speed up the change during its course, as it is catalysed by the reorganising system: the schwa apocope gets 'support' from other changes in the adjective/determiner system (see also Van Loon 2005 on 'synergy'). To check whether the deviance of the curve is not an artefact of the data, we

will look for consistency in the pattern. The distance between the projected trajectory and the observed proportions should increase in a monotonic fashion.

References

Blythe, R.A. & W. Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2): 269-304.

Denison, D. 2003. Log(ist)ic and simplistic S-curves. In: R. Hickey (ed.), *Motives for language change*. Cambridge: Cambridge University Press. 54-70.

De Vooys, C.G.N. 1953. *Nederlandse spraakkunst*. 3rd edn. Groningen: Wolters.

Kroch, A.S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199-244.

Pintzuk, S. 2003. Variationist approaches to syntactic change. In: B.D. Joseph & R.D. Janda (eds.), *The handbook of historical linguistics*. Oxford: Blackwell. 509-528.

Van de Velde, Freek. 2017. 'Understanding grammar at the community level requires a diachronic perspective. Evidence from four case studies'. *Nederlandse Taalkunde* 22(1): 47-74.

Van de Velde, Freek & Fred Weerman. 2014. 'The resilient nature of adjectival inflection in Dutch'. In: Petra Sleeman, Freek Van de Velde & Harry Perridon (eds.), *Adjectives in Germanic and Romance*. Amsterdam: John Benjamins. 113-145.

Van der Horst, J. & K. van der Horst. 1999. *Geschiedenis van het Nederlands in de twintigste eeuw*. 's-Gravenhage: SDU.

Van Loon, J. 2005. *Principles of historical morphology*. Heidelberg: Winter.

Van Loon, J. 2014. *Historische fonologie van het Nederlands*. Schoten: Uitgeverij Universitas.

Weinreich, U., W. Labov & M. Herzog. 1968. *Empirical foundations for a theory of language change*. In: W.P. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press. 95-188.

Towards a grammar of *Aggressive Expressions*

Cecilia Poletto & Alessandra Giorgi
(University of Padua/ Goethe University, Frankfurt & University of Venice)

Schedule: Thu 16.00 Room 7

The aim of the work: In this work we consider the properties of *Aggressive Expressions* – henceforth AEx – in Italian (see Obenauer 1994, Den Dikken and Giannakidou 2002). These expressions mostly appear in special questions, for instance surprise-disapproval ones, and exclamatives. They convey a strong emotional content, usually a negative evaluation, by the speaker and are associated with a special intonation. Our aim is to provide a syntactic account for their distribution in order to achieve a better understanding of the grammar of non-declarative sentences and on the relation between the clausal architecture and the internal DP structure.

The data: These expressions are often names for the male organ – identified either with vulgar term *cazzo*, or with euphemisms such as *cavolo* (lit: cabbage), *tubo* (lit: pipe), *piffero* (lit: fife) etc. – terms such as *diavolo* (lit: devil) and *accidenti* (a damn).
We studied various properties of these items, and observed that they exhibit slightly different distributions. We analyzed their occurrence in *wh*- phrases, as in the following example:

(1) Cosa *cazzo/ diavolo/ *tubo* fai?
What the fuck are you doing?

Cazzo and *diavolo* (devil), but not *tubo* (pipe) can lexicalize (part of) a wh-phrase. We also considered their occurrence as Negative Polarity Items:

(2) Non ho capito *un cazzo*/ **un diavolo*/ *un tubo*
I didn't understand shit

In this case, *cazzo* and *tubo* (pipe) function as NPIs, but not *diavolo* (devil). Furthermore, we considered their agreement properties:

(3) Quest**a** *cavolo* di finestr**a** non si apre
this-F cabbage-M of window-F does not open
(4) Quest**o** *cavol**o*** di finestra non si apre
this-M cabbage-M of window-F does not open
'this damned window doesn't open'
(5) *Quest**a*** *tubo* di finestr**a** non si apre
this-F pipe-M of window-F does not open
(6) *Quest**o*** *tubo* di finestra non si apre
this-M pipe-M of window-F does not open
(7) *Quest**a*** *diavolo* di donn**a** non arriva
this-F devil-M of woman-F doesn't arrive
(8) Quest**o** *diavol**o*** di donna non arriva
this-M devil-M of window-F doesn't arrive
'this damned woman doesn't arrive'

Cavolo (cabbage), like *cazzo,* can be transparent to agreement, admitting the demonstrative to agree directly with the noun *finestra* (window), as in example (3). *Tubo* (pipe) does not enter the construction at all, independently of agreement. Finally, *diavolo* (devil) only agrees with the demonstrative.

Generalizations: As a common property, these terms all express an evaluative feature. Generalizing on the basis of the observations above, we can say that they are located along a continuum from lexical to functional. *Diavolo* appears to be the most lexical, in that it retains its morphological properties not allowing the agreement pattern in (7). On the other end of the spectrum, *cazzo* and *cavolo* seem to be quite functionalized, in that they can be transparent to agreement, as shown in (3) and (4).

Diavolo however, does not work as an NPI. This is expected on the basis of the consideration that it does not express only a negative evaluation, but also a positive one, i.e. *quel diavolo di donna* (that devil of a woman) can also mean a very vivacious and active woman.

Tubo, on the other hand, expresses a negative evaluation – hence, it appears as NPI – but retains its nominal features to the point that it cannot be part of a larger nominal expression, hence it does not appear in *wh*- phrases, or in constructions such as (5) and (6).

Furthermore, we believe that a typology of AEx can shed new light on the parallel between the DP and the CP: since these elements are evaluative, they occupy positions in the DP that correspond to the ones in the CP, so that they contribute to the semantic and syntactic composition of different sentence types thereby showing that the features internal to DPs are visible to the clausal spine and that they must match.

References

Obenauer, Hans-Georg. 1994. Aspects de la syntaxe A-barre: Effets d'intervention et mouvements des quantifeurs. *Thése de doctorat d'état*, Université de Paris VIII.

Den Dikken, Marcel, Giannakidou, Anastasia, 2002, From Hell to Polarity: ''Aggressively Non-D-Linked'' Wh-Phrases as Polarity Items, *Linguistic Inquiry*, 33, 1, 31–61.

On the distribution and discourse pragmatic functions of the morpheme *=taŋ* in Chepang

Marie-Caroline Pons
(University of Oregon)

Keywords: Trans-Himalayan, Chepang, Discourse, Information Structure, Focus of assertion

Schedule: Fri 9.30 Room 6

This paper describes the morphosyntactic distributions and functions of the enclitic *=taŋ* in Chepang, a Trans-Himalayan language of Nepal. I show that its primary function is pragmatic, marking the focus of assertion (Lambrecht 1994:213), i.e. the "pragmatically non-recoverable" (Lambrecht 1994:218) piece of information that the addressee is expected to take for granted as a result of hearing it in an utterance (1). The focus of assertion function of *=taŋ* is non-contrastive. In (1), Speaker B does not understand what Speaker A said. Speaker B asks Speaker A to essentially repeat what s/he said through the use of *do* 'what' marked with the morpheme *=taŋ*. Speaker B's question seeks the informational content in Speaker A's utterance. Speaker A felicitously answers Speaker B's overt interrogation that reflects a presupposition such as 'the goat *x*-ed.' or '*x* happened.'

(1) A: *metɕʰja waŋ=a*
goat come.arrive=PST
'The goat arrived.'
B: *do=taŋ ?*
what=FOC
'What?' 'What's that?' 'What happened?'
A: *metɕʰja waŋ=taŋ=a !*
goat come.arrive=FOC=PST
'The goat arrived!' (CH_BMB_BAN_082218_2)

In addition to marking the focus of assertion, *=taŋ* shows three extended functions. In traditional narratives, *=taŋ* helps structure the sequence of events. Its use correlates with events that move the timeline of the narrative forward (Hopper 1979, Labov and Waletzky 1967), i.e. Main Event Line (MEL). The addressee's attention is dragged along a series of events, result of the speaker's manipulation of presuppositions and assertions (2). With modifiers and negated verbs, *=taŋ* shows a contrastive focus interpretation and reflects the speaker's "epistemological stance" (Mushin 2001) with regard to the information expressed. The speaker's presuppositions regarding the addressee's knowledge or beliefs trigger assertions with the epistemic effect of adding authority to the informational content (3). Finally, under language contact with Nepali, *=taŋ* has developed the function of evidential hearsay/quotative marker when used with reported information that was previously actually "said" (4).

(2) A. *saŋ-saŋʌ* *wah=ti,ʌni* *ʤa* *ɾʌ* *biɾalo,*
together walk=SEQ so tiger and cat
ʤa *ɾʌ* *biɾalo=i,* *sja=taŋ* *sat=a=kan* *dej !*
tiger and cat=ERG cow=FOC kill=PST=2/3.A IS
'They were walking together, and the cat and the tiger, the cat and the tiger, they killed a cow, hey!'

B. *ʌni* *"ŋa=pa,* *sjak=pa* *ʤe=ŋʌ=lʌ !"*
so 1SG=IS raw=IS eat=1SG=NEG
'Then, "But I don't eat raw (meat)!" (the tiger said).'

C. *daj=ti,* *biɾalo=kʌj* *mʰe=taŋ*
say=SEQ cat=DAT fire=FOC
boŋ=laŋ *lʰok=a=kan* *dej !*
look.for=PUR send=PST=2/3.A IS
'Having said that, the tiger sent the cat to look for fire.'

(3) *"ʤʌjk=taŋ=lʌ* *ŋa=pa* *al=na=ŋ !* *al=na=ŋ !"*
bite=FOC=NEG 1SG=IS go=IMPF=1SG go=IMPF=1SG
'"They won't bite (me), I'm gonna go, I go!" (the daughter-in-law said).'

(4) *Ameɾika-aŋ* *mʌj-sej* *dʌbej=i* *um=taŋ=na=i.*
America-LOC banana-CL medicine=INST ripe=IS=IMPF=PL
'They say that in America they ripen bananas with chemicals.'

These three extended functions arose diachronically from reanalyses of the function of focus of assertion. Synchronically, these functions do not seem to be fully grammaticalized, since the primary function of focus of assertion remains, i.e. not yet semantically bleached, they may be described as "semantic-pragmatic expansions" (Tabor and Traugott 1998, Himmelmann 2004) of the function of non-contrastive focus of assertion.

References

Himmelmann, Nikolaus. P. 2004. Lexicalization and grammaticization: Opposite or orthogonal. In: Bisang Walter; Himmelmann Nikolaus P.; and Björn Wiemer (ed.), *What makes grammaticalization?: a look from its fringes and its components*, Vol. 158: 21-42. Walter de Gruyter.

Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In *Syntax and Semantics*, Vol 12, pp. 213-241. Academic Press.

Labov, William, and Joshua Waletzky. 1967. Narrative analysis. In Helm June (ed.), *Essays on the Verbal and Visual Arts,* 12-44. Seattle: U. of Washington Press.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Mushin, Ilana. 2001. *Evidentiality and epistemological stance*: *Narrative retelling*. Amsterdam & Philadelphia: John Benjamins.

Tabor, Whitney, and Elizabeth Closs Traugott. 1998. Structural scope expansion and grammaticalization. *Typological Studies in Language*, 37, 229-272.

Pre-causative detransitivization and the restriction of causatives to intransitives or to non-agentive intransitives

Roland A. Pooth
(Ghent University/Cologne)

Keywords: Coding asymmetry, causativization, detransitivization, prototypical causation, frequency, cognitive effort

Schedule: We 11.00 Room 6

Many languages other than English show a coding asymmetry in form of a restriction of causative markers to intransitives or to non-agentive intransitive verbs (Nedjalkov & Sil'nitsky 1973: 7). In other languages, "a transitive verb must be made intransitive before a causative can be added" (Dixon 2012: 252). In such languages, the respective causative construction involves overt detransitivizing or intransitivizing morphology; an illustrative example is provided from Paumarí (Arawá family, Brazil; Chapman & Derbyshire 1991: 185–6; Dixon 2012: 252). The detransitivizing marker on the verb is *-a-*

(1) Paumarí

a. *bi-noki-hi* *ida* *gora*
3sg.A-see-theme art.f.O house.f.O
He saw the house.

b. *noki-a-hi* *ida gora*
see-detransitivizer-theme (see a.)
The house is visible.

c. *bi-na-noki-a-hi* *ida gora*
3sg.A-caus-see-detransitivizer-theme (see a.)
He made the house become visible.

d. *ho-ra* *na-noki-a-hi-vini* *hi-hi* [*ida gora*]$_{2ndO}$
1sg-acc caus-see-detransitivizer-appl-dep aux-theme (see a.)
He showed me the house, he made the house become visible to me.

In our talk, we will initially discuss a sample of illustrating data, as in (1), and we will show that both restrictions are two sides of the same coin. We will make a distinction between two subtypes of this restriction to capture both their similarity and difference: Such languages not allowing transitive verbs to undergo causativization fall back on the application of causativization to intransitive or non-agentive intransitive counterparts of transitive verbs on the level of their lexicon, whereas languages with pre-causative detransitivization show the same type of restriction on a non-lexical level. We will discuss the ultimate rationale for this restriction. The two comparable restrictions will be explained as parallel outcomes of the cognitive semantics of causativization from the background of prototypical causation in semantic terms (Lakoff & Johnson 1980: 70) and the cognitive effect it must have. We will identify an underlying semantic detransitivizing component of causativization: when applied to verbs that are prototypically transitive in semantic terms, causativization must have a detransitivizing semantic component in that causatives typically marginalize the agent by introducing the causer. Languages with pre-causative detransitivization/intransitivization thus provide evidence for the view that causativization of transitive verbs can be split up into a semantically detransitivizing operation plus adding a causer by transitivization, whereas causativization of intransitive verbs simply involves introducing a causer. We will finally argue that this coding asymmetry can generally be understood as an effect of the general structure of human cognition and conceptualization in terms of prototypical causation in semantic terms, implying that causativization of transitives involves a doubled deviation

from prototypical causation, whereas causativization of intransitives is cognitively more basic than that, and more frequent. We will provide evidence for this hypothesis, thereby strengthening an explanation of grammatical coding asymmetries in terms of degrees of predictability, frequency, and cognitive effort. We can thus add more evidence to the hypothesis that when two such grammatical constructions occur with significantly different frequencies, the less frequent construction tends to be coded with more segments (and more effort), while the more frequent construction tends to be coded with less effort (and less segments) (Haspelmath 2008, 2018ab).

References

Chapman, S. & D. C. Derbyshire (1991): Paumarí. In: Derbyshire, D. C. & G. K. Pullum (eds.): *Handbook of Amazonian Languages.* Volume 3. Berlin: Mouton De Gruyter, 161– 332.

Dixon, R. M. W. (2012): *Basic Linguistic Theory. Volume 3. Further Grammatical Topics.* Oxford U.P.

Haspelmath, M. (2008): Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19 (1), 1–33.

Haspelmath, M. (2016): Universals of causative and anticausative verb formation and the spontaneity scale. *Lingua Posnaniensis* 58 (2), 33–63, doi:10.1515/linpo-2016-0009

Haspelmath, M. (2018a, manuscript): Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability, https://www.academia.edu/37510601/Explaining_grammatical_coding_asymmetries_Form-frequency_correspondences_and_predictability

Haspelmath Martin (2018b, blogpost): Asymmetric coding in grammars and frequency-induced predictability, https://dlc.hypotheses.org/1056

Lakoff, G. & M. Johnson (1980): *Metaphors We Live By.* Chicago U.P.

Nedjalkov, V. P. & G. G. Sil'nickij (1969): Tipologija morfologičeskogo i leksičeskogo kauzativa. In: Xolodovič (ed.) A. A. (ed.): *Tipologija kauzativnyx konstrukcij*. Leningrad = (1973): The typology of morphological and lexical Causatives, in: F. Kiefer (ed.): *Trends in Soviet theoretical linguistics.* Foundation of Language supplementary series. Vol. 18. Dordrecht & Boston.

Human impersonals in Western Romance and South Slavic: A speech corpus study of *man*-impersonals

Pekka Posio & Max Wahlström
(University of Helsinki)

Keywords: impersonality, referentiality, speech corpora, Western Romance, South Slavic

Schedule: We 11.30 Room 8

The grammaticalization of nouns meaning 'man' or 'person' into impersonal or indefinite pronouns (so-called *man*-impersonals) has been proposed to be an areal feature of the core area of Standard Average European (SAE) languages like French (on) and German (man), atypical of null subject languages (Siewierska 2011, Giacalone Ramat & Sansò 2007). In this paper, we provide evidence on the grammaticalization of nouns meaning 'person' into impersonalization devices in two null subject languages, European Portuguese and Bulgarian. Our research questions are (1) to what extent the constructions being studied have a grammaticalized status and (2) why they are found in Portuguese

and Bulgarian, but not in other languages closely related to them. The study is carried out within a usage-based, contrastive corpus linguistic framework.

While sporadically mentioned in earlier research (see, e.g., Nunes 1919, Dimova 1981, Posio 2017), the potentially grammaticalized uses of the lexeme 'person', illustrated in examples (1) and (2), have not been studied in a systematic way in naturally occurring speech.

(1) European Portuguese (CORDIAL-SIN, VPA-30)

mas	*esse*	*peixe,*	*já*	*uma*	*pessoa*	*às vezes*	*não*	*o*	*conhece*
but	this	fish,	even	INDEF	person	sometimes	not	it	recognise.3SG

'but this fish, one doesn't even recognize it sometimes'

(2) Bulgarian (Corpus of Spoken Bulgarian)

ama	*vsjaka*	*duma*	*pri*	*neja*	*e*	*da*	*se*	*previva*	*čovek*	*ot*	*smjah*
but	every	word	by	her	is	COMP	REFL	fold.3SG	person	with	laughter

'but her every word makes one roll with laughter'

In order to analyze the use of man-impersonals in Portuguese and Bulgarian, we contrast the use of nouns denoting 'person' in these languages with two other closely related varieties, Spanish and Croatian, respectively, in which *man*-impersonals are either marginal (Croatian) or not found in the contemporary language (Spanish). In addition, we examine another impersonalization strategy, the second person singular (2sg), in all four languages. The data come from five speech corpora (Posio [in progress], Fernández Juncal 2005, Martins 2000–, Aleksova 1994, Kuvač Kraljević & Hržica 2016), consisting of sociolinguistic and dialectal interviews as well as spontaneous conversations among friends and family members, comprising altogether more than 1 million tokens.

Based on our comparisons between the language pairs (Portuguese/Spanish and Bulgarian/Croatian), we establish that both Portuguese and Bulgarian display grammaticalized uses of nouns denoting 'person' in expressing impersonality and generic reference. This is supported by evidence of phonetic and morphological reduction and frequency of use of the grammaticalizing NPs. In contrast, this strategy is non-existent or marginal in Spanish and Croatian, which in turn employ 2sg as an impersonalization strategy more frequently and productively. We hypothesize that this inverse correlation between the two impersonal strategies is explained partly by the general tendency to express overt subjects: in both language pairs, the language with more overt subject marking prefers the noun-based impersonal strategy, whereas the language with less expressed subjects favors the use of 2sg. We further discuss the role of pragmatic factors, such as the general use of address forms in discourse, in the competition between the two impersonal strategies.

References

Aleksova, Krasimira. 1994. *Corpus of spoken Bulgarian.* Sofia. http://folk.uio.no/kjetilrh/bulg/Aleksova/.

Posio, Pekka. [in progress]. *Português falado no Porto. A sociolinguistic interview corpus of Portuguese spoken in Porto.*

Dimova, Anna. 1981. Die Polysemie des deutschen Pronomens „man" unter Berücksichtigung seiner Äquivalente im Bulgarischen. *Deutsch als Fremdsprache* 18(1). 38–44.

Fernández Juncal, Carmen. 2005. *Corpus de habla culta de Salamanca.* Burgos: Instituto Castellano y Leonés de la Lengua.

Giacalone Ramat, Anna & Andrea Sansò. 2007. The spread and decline of indefinite man-constructions in European languages. In Paolo Ramat & Elisa Roma (eds.), *Europe and the Mediterranean as linguistic areas: Convergencies from a historical and typological perspective*, 95–132. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Kuvač Kraljević, Jelena & Gordana Hržica. 2016. *Croatian Adult Spoken Language Corpus* (HrAL). Zagreb. https://ca.talkbank.org/access/Croatian.html.

Martins, Ana Maria [coord.] 2000–. CORDIAL-SIN: *Corpus Dialectal para o Estudo da Sintaxe / Syntax-oriented Corpus of Portuguese Dialects.* Lissabon. http://www.clul.ulisboa.pt/ en/10-research/314-cordial-sin-corpus.

Nunes, José Joaquim. 1919. *Compêndio de gramática histórica portuguesa (fonética e morfologia).* Lisbon: Livraria Clássica Editora.

Posio, Pekka. 2017. Entre lo impersonal y lo individual: Estrategias de impersonalización individualizadoras en el español y portugués europeos. *Spanish in Context* 14(2). 209–229.

Siewierska, Anna. 2011. Overlap and complementarity in reference impersonals: man-constructions vs. third person plural-impersonals in the languages of europe. In Andrej Malchukov & Anna Siewierska (eds.), *Impersonal constructions: A cross-linguistic perspective.* 57–90. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Elevation as deictic category in Khinalug

Monika Rind-Pawlowski

Schedule: Thu 9.00 Room 12

Khinalug is a Nakh-Dagestanian language with approximately 2.300 speakers, located in Khinalug village in Northern Azerbaijan. The language has preserved petrified relicts of a directional preverb system that marks elevation with reference to the movement target. This system, even though typical for Nakh-Dagestanian languages, is no longer productive in Khinalug, the speakers do not recognize the directional meanings of the morphemes any more. The respective preverbs combine with distinct verbal cores only, and create lexicalized semantics.

č- downwards into contact position
(e.g. *čiviri* “sit down”)

g- upwards into contact position
(e.g. *giviri* “sit onto a riding animal”)

t- away from contact position
(e.g. *tačʰkui* “stand up”)

n- movement along
(e.g. *našxiri* “sharpen”)

χ- remaining in contact position / movement together with the object
(e.g. *χalkʰuri* “touch”; *χɯkuri* “run after”)

š- immersion / movement within / around sth.
(e.g. *šalkʰuri* “sink into water, get soaked”; *šälfiri* “sew a thread around a patch”)

Additionally, Khinalug has developed a new system that marks elevation with reference to the speaker’s position, which is assumably unique among Nakh-Dagestanian languages. This system is very productive: the preverbs can be interchanged freely, according to the respective direction.

a- upwards, away from speaker
(e.g. *akui* “go upwards, away from speaker”)

za- downwards, away from speaker

(e.g. *zakui* "go downwards, away from speaker")
la- at same level, away from speaker
(e.g. *lakui* "go away from speaker at same level")
kʰa- from any level, towards speaker
(e.g. *kʰakui* "come towards speaker from any level")
al- from above, towards speaker
(e.g. *alkui* "come towards speaker from above")
qal- from below, towards speaker
(e.g. *qalkui* "come towards speaker from below")
tʰal- from same level, towards speaker
(e.g. *tʰalkui* "come towards speaker at same level")

The preverbs of both systems can cluster to express complex movements. The speaker-oriented morpheme takes the initial position (*a-ttɯ-kui*, *za-ttɯ-kui* etc., "go away from/towards speaker (see above) from specific point of departure"). In these clusters, *č-* can express movement into contact position irrespective of direction, replacing the functions of *g-* (*a-ččɯ-kui*, *za-ččɯ-kui* etc. "go away from/towards speaker to a specific target). Clusters can contain several elevation morphemes (*a-tʰɯ-š-kui*, *za-tʰɯ-š-kui* etc. where speaker-oriented preverb + *t- (~ tt ~ tʰ-)* "departure" + *š-* "immersion" + *kui* "go" yield the meaning "leave an area (towards/away from speaker)". Sometimes the downward meaning of *č-* is maintained (*e-n-č-kui*, *ze-n-č-kui* etc. "go along a path, then downwards to target (towards/away from speaker)".
Also the locative-directive case *-r* combines with directional preverbs (e.g. *a-r-kiri*, *za-r-kiri* etc. "send").
Another Khinalug peculiarity is the usage of the copulas

oa above
tʰoa far/samelevel
qoa near/below
attä neutral

as morphemes in the finite verb system. Combining with the imperfective participle, they form the present, and combining with the perfective participle, the past tense. Besides position, these copulas can also express evidentiality: *tʰoa* is regularly used for non-witnessed information, and the present on *tʰoa* serves as historical present in narrations; *qoa* is used for witnessed information, *attä* is preferred when the action is performed by the the speakers themselves.

A computer-assisted quantitative typology of spatial deictic coding strategies

Maja Robbers & Harald Hammarström
(Uppsala University)

Keywords: Information extraction, Semantic parsing, Quantitative typology, Spatial relations, Deixis

Schedule: Thu 14.30 Room 4

Studies on spatial orientation (e.g. Levinson 1996, 2003) and spatial deixis (since Bühler 1934) have mostly been focussed on small-scale comparisons (e.g. Denny 1978). With a rich descriptive tradition

in mind, a large-scale comparative approach shall identify and statistically evaluate coding patterns and areal trends in the world's languages. A comprehensive typology including static and dynamic spatial deictic relations, i.e. coding of deictic Place/Goal/Source, ought be based on the distinction between systems that prefer morphological strategies to code Place/Goal/Source relations directly on the Ground as opposed to predominantly verb-internal coding strategies.

[1] shows a typical European solution involving the spatial adverbs *dort* 'there' and *hier* 'here' that overtly code Ground. Goal is coded by the clitic *=hin* 'to' and Source by the preposition *von* 'from'. In the bipartite verb stem in [2], Ground is coencoded with Source by the verb in the initial slot and combined with an allatival Goal verb in the second slot. [3] shows a typical Mesoamerican construction including a spatial adverb that codes the distance level compoment of the Ground notion and a dynamic spatial deictic component coded by the accompanying verb.

[1] German (Indo-European)
Von ***dort*** *kam* *er* ***hier=hin***.
from there come.3SG.PST 3SG.MASC here=ALL
'From there he came (to) here.'

[2] Arapaho (Algic)
héétnéénino*o* *héétnehnóuúhcehíno*o.
eti-neeni-noo eti-**eh-nouuhcehi**-noo
IC.FUT-be(AI)-1S IC.FUT-from.here-run.outside(AI)-1S
'I will be the one to go out there.' (Cowell & Moss 2008: 224)

[3] Guerrero Nahuatl (Uto-Aztecan)
Jesús ***ompa*** *o-quis.*
Jesus there PERF-leave
'(As) Jesus went on from there.' (NGU, Mat 9:27)

[1-3] represent a peek into the diversity of coding types found in spatial deictic systems. Further important criteria to develop a typology are e.g. the number of distance levels (with a minimum of two, i.e. proximal and distal) and possible specialized Grounds (e.g. elevational stages). Information on these typological features and on the parts of speech that are deemed relevant for the encoding of spatial deixis is usually found in descriptive grammars. Since the manual extraction of these linguistic facts has proven to be too time-consuming, it is tested which facts can be assessed systematically by the aid of automatic extraction of pre-defined typological features from digitalized language descriptions (Virk et al. 2017). Precisely, on the basis of a PDF collection that spans around 4500 languages, simple keyword and pattern recognition is employed to extract facts that are essential to the typology. This method serves to detect information on relevant parts of speech and predicts frequencies of types of coding systems.

Moreover, the distribution of tasks between human and machine shall be discussed, considering the possibility to parallel the computer-based study by a manual evaluation of a smaller sample, which is expected to have mutual improving effects on both methodologies as it opens up several possibilities for cross-checking and verification. The study on spatial deictic coding strategies in progress represents an informative test case for exploratory data analysis in the emerging field of classical linguistic typology collaborating with computer-based methods.

References

Bühler, Karl (1934), *Sprachtheorie. Die Darstellungsfunktion der Sprache*, Jena: Gustav Fischer.

Cowell, Andrew & Alonzo Moss Sr. (2008), *The Arapaho language*, Boulder: University Press of Colorado.

Denny, Peter (1978), Locating the universals in lexical systems for spatial deixis, in D. Farkas, W. M. Jacobsen, and K. W. Todrys (eds.), *Papers from the Parasession on the lexicon, Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, 71-84.

Levinson, Stephen C. (1996), Frames of reference and Molyneux's Question: crosslinguistic evidence, in P. Bloom, M. A. Peterson, L. Nadel, and M. F. Garrett (eds.), *Language and Space.* Cambridge: MIT press, 109-169.

Levinson, Stephen C. (2003), *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Wycliffe Bible Translators, Inc. (1987), *In yencuic iyectlajtoltzin Dios itech ica toTeco Jesucristo: New Testament in Nahuatl, Guerrero* (NGU).

Virk, Shafqat Mumtaz, Lars Borin, Anju Saxena & Harald Hammarström (2017), Automatic Extraction of Typological Linguistic Features from Descriptive Grammars, in K. Ekštein, and V. Matoušek (eds.), *TSD 2017: Text, Speech, and Dialogue (Proceedings).* Berlin: Springer, 111-119.

A universal pathway in the grammaticalization of discourse markers? Zooming in on Spanish *en plan (de)* and English *like* and *kind of/kinda*

Paula Rodríguez-Abruñeiras
(Universitat de València)

Keywords: *en plan (de)*; *like*; *kind of/kinda*; grammaticalization; cross-linguistic pragmatics

Schedule: Thu 14.00 Room 14

Grammaticalization has been traditionally described as the change from more lexical to more grammatical (cf. Meillet 1912: 131, Kuryłowicz 1975 [1965]: 52, Langacker 1977: 106, Hopper and Traugott 2003: 1, Fanego 2010, Hansen 2017, Narrog and Heine 2018: 1, among many others). Such a change may affect different areas of the language, and its manifestations can be seen at the semantic-pragmatic, morpho-syntactic, phonological and functional levels (see McMahon 1994: 161). This paper aims at shedding new light on the grammaticalization of DMs cross-linguistically, ultimately trying to figure out whether different formulas in different languages may follow the same transitional steps or *clines*. In Hopper and Traugott's (2003: 6) words, a cline "is a metaphor for the empirical observation that cross-linguistically forms tend to undergo the same kinds of changes or have similar sets of relationships, in similar orders". More specifically, the corpus-based study investigates the various uses of the Spanish formula *en plan (de)*, ranging from propositional to expressive meanings. All these functions coexist in present-day Spanish, but they did not emerge at once. Thus, evidence shows that the original adverbial meanings of manner (cf. 1) and purpose have gradually given rise to a wide variety of pragmatic functions, such as approximator, marker of exemplification or filler (cf. 2), among others.

(1) A lo mejor te puedes ir a Egipto con ella ***en plan de*** amigos. (CREA, 1989, *Pares y Nines*)
'Maybe you can go to Egypt with her *as* friends.'

(2) ¡Qué va! No, eso Era ***en plan*** no sé, ***en plan*** simplemente pues eso, divertirse. (CREA, year not provided. Interview CSC004)
'No way! No, that was *like* I don't know *like* just so, to have fun.

Previous research on the current uses of *en plan (de)* (cf. Rodríguez-Abruñeiras forthcoming in 2020) has revealed a strikingly similar set of functions when compared to English *like* and *kind of/kinda* as described in Fleischman and Yaguello (2004) and D'Arcy (2017) (*like*) and Margerie (2010) (*kind of/kinda*). Therefore, the aim of this presentation is threefold: (i) to describe the various uses of *en plan (de)* in present-day Spanish; (ii) to track the diachronic changes undergone by the formula in its process of grammaticalization, first from a prepositional phrase to an adverbial, and then from an adverbial to a DM; and (iii) to compare the grammaticalization cline of *en plan (de)* with that of English *like* and *kind of/kinda* so as to bring to the fore the potential existence of a universal pathway of grammaticalization in the emergence of DMs. For our purposes, the data on *en plan (de)* are mostly taken from two Spanish reference corpora, namely CORDE (*Corpus Diacrónico del español* 'Diachronic Corpus of Spanish') and CREA (*Corpus de Referencia del Español Actual* 'Reference Corpus of Present-day Spanish'). Two additional sources of information are occasionally consulted to supplement the data from these corpora, that is, the *Corpus oral de lenguaje adolescente (de Madrid)* ('Oral Corpus of Teen Talk (from Madrid)') or COLAm and the Web. The total number of examples retrieved from the three corpora consulted amounts to 1245. The results will be compared to Fleischman and Yaguello's (2004) and Margerie's (2010) research on *like* and *kind of/kinda*, respectively.

References

Primary sources

Real Academia Española. Banco de datos (*CORDE*) [en línea]. *Corpus diacrónico del español*. <http://www.rae.es>

Real Academia Española. Banco de datos (*CREA*) [en línea]. *Corpus de referencia del español actual*. <http://www.rae.es>

Universidad de Bergen. 2001. Departamento de español e inglés. *COLAm* [en línea]: *Corpus oral de lenguaje adolescente de Madrid*. <http://www.colam.org>.

Secondary sources

D'Arcy, Alexandra (2017), *Discourse-pragmatic variation in context: Eight hundred years of* like, Amsterdam/Philadelphia: John Benjamins.

Fanego, Teresa (2010), Paths in the development of elaborative discourse markers: Evidence from Spanish, in K. Davidse, L. Vandelanotte and H. Cuyckens (eds), *Subjectification, Intersubjectification and Grammaticalization*, Berlin: Mouton de Gruyter, 197-237.

Fleischman, Suzanne and Marina Yaguello (2004), Discourse markers across languages. Evidence from English and French, in C.L. Moder and A. Martinovic-Zic (eds), *Discourse Analysis across Languages and Cultures*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 129-147.

Hansen, Björn (2017), "What happens after grammaticalization? Post-grammaticalization processes in the area of modality, in D.V. Olmen, H. Cuyckens and L. Ghesquière (eds), *Aspects of Grammaticalization*, Berlin/Boston: Walter de Gruyter, 257-280.

Hopper, Paul J. and Elizabeth Closs Traugott (2003) [1993], *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.

Kuryłowicz, Jerzy (1975) [1965], *Esquisses Linguistiques II*, München: Wilhelm Fink.

Langacker, Ronald Wayne (1977), Syntactic reanalysis, in C.N. Li (ed), *Mechanisms of Syntactic Change*, Austin: University of Texas Press: 57-139.

Margerie, Hélène (2010), On the rise of (inter)subjective meaning in the grammaticalization of *kind of/kinda*, in K. Davidse, L. Vandelanotte and H. Cuyckens (eds), *Subjectification, Intersubjectification and Grammaticalization*, Berlin: Mouton de Gruyter, 315-346.

McMahon, April M.S. (1994), *Understanding Language Change*, Cambridge: Cambridge University Press.

Meillet, Antoine (1912), L'évolution des formes grammaticales, *Scientia* 6(12), 384-400.

Narrog, Heiko and Bernd Heine (2018), Introduction: Typology and grammaticalization, in H. Narrog and B. Heine (eds.), *Grammaticalization from a Typological Perspective*, Oxford: OUP, 1-15.

Rodríguez-Abruñeiras, Paula (forthcoming in 2020), 'Me vais a permitir que me ponga en plan profesor': Las funciones de *en plan de* en un estudio de corpus, *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*.

Style-shifting in Flemish TV and radio commercials: A diachronic study

Laura Rosseel, Jacoba Waumans & Dirk Geeraerts
(KU Leuven, Università di Bologna & KU Leuven)

Keywords: social meaning of language variation, Dutch language variation, advertising language, standardization, sociolinguistics

Schedule: We 14.30 Room 2

The aim of advertising is to persuade consumers to buy a product or service. In this process of persuasion, marketers make avid use of the stylistic values of language variation (Van Gijsel et al. 2004). Studying this deliberate exploitation of linguistic styles in the genre of advertising offers the possibility to uncover the social meanings attached to linguistic features and language varieties (Geeraerts & Van de Velde 2013). Against this background, we investigate style-shifts between two varieties of Dutch in Belgium, Standard Belgian Dutch (SBD) and Colloquial Belgian Dutch (CBD), in a corpus of 711 radio and TV commercials broadcast in 2018. Furthermore, we compare our results to those from an older study based on a similar corpus of commercials sampled in Flanders, the Dutch speaking area of Belgium, in 2001 (Van Gijsel et al. 2008). This comparison will allow us to assess whether the social meanings of the varieties under study have changed.

The study focuses on style-shifts *within* commercials. In order to do so, a colloquiality index was calculated per spot element (i.e. a subsection of a commercial). This index was based on 13 linguistic features characteristic of CBD (e.g. *gij* vs. *jij* forms in the pronominal system or t-deletion in short function words). Regression analyses on the colloquiality index in the 2018 data show results that are remarkably similar to those reported in the 2001 study. We note that CBD is mainly used in 'minidramas'. This is a type of spot element that portrays a situation from daily life. Additionally, the colloquiality index is significantly higher in humoristic ads, as well as in stylish telecom ads which are predominantly aimed at a younger audience. This use of CBD indicates that the variety is associated with entertainment and informality and potentially carries a more youthful indexicality. SBD, by contrast, is more strongly represented in informative spot elements, such as product presentations. The variety is also particularly well-represented in the payoff of ads, i.e. a brief statement at the end of the ad containing practical information like the company name, a slogan or temporary promotions. This usage suggests that SBD carries social meanings related to formality, seriousness and authority. Interestingly, these diachronically stable results echo findings from experimental language attitudes studies on SBD and CBD, which report a similar division of labour in the social meanings of SBD and CBD (e.g. Impe & Speelman 2007; Rosseel 2017). In our paper, we will discuss these findings in the

light of recent debates about the evolution of Belgian Dutch and claims regarding the alleged standardization of CBD.

References

Geeraerts, Dirk, & Van de Velde, Hans (2013), Supra-regional characteristics of colloquial Dutch, in F. Hinskens and J. Taeldeman (eds), (2013), *Language and Space: Dutch*, Berlin: de Gruyter, 532–556.

Impe, Leen, & Speelman, Dirk (2007), Vlamingen en hun (tussen)taal - een attitudineel mixed guiseonderzoek, *Handelingen van de Koninklijke Zuid-Nederlandse Maatschappij Voor Taal-En Letterkunde*, *61*, 109–128.

Rosseel, Laura (2017), *New approaches to measuring the social meaning of language variation. Exploring the Personalized Implicit Association Test and the Relational Responding Task*, KU Leuven PhD dissertation.

Van Gijsel, Sofie, Geeraerts, Dirk, & Speelman, Dirk (2004), A functional analysis of the linguistic variation in Flemish spoken commercials, in G. Purnelle, C. Fairon and A. Dister (eds), *Le poids des mots. Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles*, Louvain-la-Neuve: Presses Universitaires de Louvain, 1136–1143.

Van Gijsel, Sofie, Speelman, Dirk, & Geeraerts, Dirk (2008), Style shifting in commercials, *Journal of Pragmatics*, *40*(2), 205–226.

The use of ethnolectal features by pre-adolescent children in German-speaking Switzerland

Melanie Röthlisberger & Laura Rosseel
(University of Zurich, University of Leuven)

Keywords: Swiss-German, children, language variation, ethnolects, pre-adolescence

Schedule: We 17.00 Room 2

Features of ethnolectal Swiss-German have often been associated and discussed in connection with adolescents with a migrant background (see e.g. Tissot, Schmid & Galliker 2011). While these features have also been used by non-migrant adolescents for stylistic and indexical purposes (Auer 2002, Schmid 2017), little is known about their usage and the social meaning attached to them within the youngest speech group, i.e. pre-adolescent children. In this study, we investigate the extent to which such ethnolectal features (see, for instance, examples 1 and 2) have spread to the spoken vernacular of pre-adolescent Swiss-German children below the age of 12.

(1) Pragmatic expressions, e.g. *altä, diga* (as a pragmatic markers)
(2) Syntactic expressions, e.g. omission of prepositions, articles, pronouns or auxiliaries, as in *Chani bleistift?* <Can I pen?> for *Chani en bleistift ha?* <Can I have a pen?>

The present study seeks to empirically investigate earlier anecdotal claims about the usage of these features by tapping into spoken data collected through the diapix task (Baker & Hazan 2011) and via story telling. Diapix is a technique designed to elicit speech from two or more speakers who are asked to identify the differences between two (or more) nearly identical pictures without seeing the other persons' picture. Ethnolectal and other youth linguistic features are identified and extracted from the data on the basis of Auer's (2002) list of seven ethnolectal features (see also Tissot, Schmid &

Galliker 2011) and on the basis of perceptual data provided by the caregivers, peers and the children themselves. The participants are all primary school children aged between 8 and 12 years old who live in and the surroundings of the city of Winterthur, an urban area in Switzerland with roughly 100,000 inhabitants. Two schools were chosen on the basis of their locality, with one school situated in an ethnically more diverse area and the other in a more rural area with a lower number of migrant citizens. In that way, we aimed to balance our data regarding the number of children with a migrant background.

Preliminary analysis of the data indicates that children younger than 12 make use of ethnolectal features dependent on locality and their gender but seemingly regardless of their ethnic background. The most frequently attested ethnolectal features seem to be change of genus of nouns and omission of definite articles. These findings are consonant with similar overgeneralization effects observed in other languages (e.g. Quist 2008 on Danish, Cornips 2008 on Dutch, and Kotsinas 2001 on Swedish) where immigrant speakers overgeneralize the use of one grammatical gender to other contexts. Taken together, our study contributes to growing research on variation and change in the speech of multilingual speakers. What is more, by focusing on children we specifically advance current understanding of children's acquisition of language variation (see De Vogelaer & Katerbow 2017).

References

Auer, Peter (2002), "Türkenslang": Ein jugendsprachlicher Ethnolekt des Deutschen und seine Transformationen, in A. Häcki Buhofer (ed), *Spracherwerb und Lebensalter*, Tübingen/Basel: Francke, 255–264.

Baker, Rachel and Valerie Hazan (2011), DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs, *Behavior Research Methods* 43(3), 761–770.

Cornips, Leonie (2008), Loosing grammatical gender in Dutch: The result of bilingual acquisition and/or an act of identity?, *International Journal of Bilingualism* 12(1–2), 105–124.

De Vogelaer, Gunter and Matthias Katerbow (2017), *Acquiring sociolinguistic variation*. Amsterdam: John Benjamins.

Kotsinas, Ulla-Britt (2001), Pidginization, creolization and creoloids in Stockholm, Sweden, in N. Smith, and T. Veenstra (eds), *Creolization and contact*, Amsterdam: John Benjamins Publishing Company, 125–156.

Quist, Pia (2008), Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice, *International Journal of Bilingualism* 12(1–2), 43–61.

Schmid, Stephan (2017), Differenzierungsprozesse im Sprachgebrauch von Jugendlichen in der Deutschschweiz: zur sozialen Interpretation von ethnolektalen Sprechweisen in Schweizer Medien, *Bulletin Suisse de Linguistique Appliquée (Vals-Asla)* Nspécial(t. 1), 105–116.

Tissot, Fabienne, Stephan Schmid and Esther Galliker (2011), Ethnolektales Schweizerdeutsch: Soziophonetische und morphosyntaktische Merkmale sowie ihre dynamische Verwendung in ethnolektalen Sprechweisen, in E. Glaser & N. Frey (eds), *Dynamic des Wandels - Wandel und Variation: Akten des 3. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*, Stuttgart: Steiner, 319–344.

ø

On Some Remarkable Differences Between Wh-phrases and Left-Peripheral Foci

Vieri Samek-Lodovici

Schedule: Fri 9.30 Room 7

<pdf>

𝒟

Argument role alignment affects oscillatory activity during sentence planning in the brain

Sebastian Sauppe, Kamal K. Choudhary, Nathalie Giroud, Damián E. Blasi, Shikha Bhattamishra, Mahima Gulati, Aitor M. Egurtzegi, Ina Bornkessel-Schlesewsky, Martin Meyer & Balthasar Bickel

(University of Zurich, Indian Institute of Technology Ropar, Concordia University, University of Zurich, Indian Institute of Technology Ropar, Indian Institute of Technology Ropar, University of Zurich, University of South Australia, University of Zurich & University of Zurich)

Keywords: psycholinguistics, cognitive neuroscience of language, syntax, Hindi

Schedule: We ~~17.30~~ 17.00 Room 15

Virtually all theories of grammar assume that linguistic expressions share overlapping syntactic configurations, modeled by derivation (Adger, 2003; Chomsky, 1995) or inheritance (Hilpert, 2014; Pollard & Sag, 1994) mechanisms. One of the most prominent overlaps is known as argument role alignment and manifests itself for example in the frequent alignment of intransitive S ('sole') arguments with transitive A (agentive) arguments in the form of nominative marked "subjects". It remains open, however, whether role alignments are are only computational patterns (in Marr's sense) or whether they are also relevant at the neural level of language processing (Embick & Poeppel, 2015; Marr, 1982/2010). To explore whether role alignments are neurophysiologically detectable during sentence production, we conducted a combined eye tracking and EEG picture description experiment (Griffin & Bock, 2000) in Hindi. Here, we focus on the dynamics of event-related desynchronization (ERD) in the EEG α band during the structural encoding phase of sentence planning (Ferreira, 2010). α ERD is a pattern of desynchronization of neural oscillations in the frequencies of 8 to 13 Hz and is associated with a wide range of functions, including sentence-level processing during comprehension (Kielar, Panamsky, Links, & Meltzer, 2015; Meyer, 2018) and more general memory and attentional processes (Hanslmayr, Staudigl, & Fellner, 2012; Klimesch, 2012). Hindi is especially suited to explore the effect of role alignment because this language exhibits a split-ergative case marking system (Bickel & Nichols, 2009). S and A arguments align in an unmarked nominative case in the imperfective aspect (A=S) while in the perfective, A arguments carry ergative case marking (-ne), distinct from S (A≠S). We hypothesized that α ERD during structure planning should be sensitive to this difference in alignments if these are neurally relevant. Specifically, we predicted that the planning of transitive sentences with nominative As leads to larger α ERD responses because speakers need to plan a structure that overlaps with intransitives (A=S), whereas ergative As show no such overlap. Fifty Hindi speakers described pictures of events using intransitive SV and transitive APV sentences with nominative or ergative subjects, while EEG was recorded (between subjects, N = 25 per group). Analyses of the time course of α band activity during structure planning phases (400-1000 ms after stimulus onset (Sauppe, 2017); speech onset was always > 1500 ms) revealed larger α ERD for nominative A sentences and nominative S sentences than for ergative A sentences. The effect was distributed broadly over frontal, central and parietal electrode sites (based on a combination of growth curve regression and decision trees (Fokkema, Smits, Zeileis, Hothorn, & Kelderman, 2018; Mirman, 2016), statistically controlling for nuisance variables, including speech onset and NP length). Thus, planning sentences with role alignment (A=S) engenders more α ERD responses than planning uniquely specified configurations (A≠S). Our findings suggest that role alignment is relevant for the

neural processes subserving sentence planning and it increases attentional and selectional demands during sentence planning (Sadaghiani & Kleinschmidt, 2016).

References

Adger, D. (2003). Core Syntax: A Minimalist Approach. Oxford University Press.

Bickel, B., & Nichols, J. (2009). Case Marking and Alignment. In A. L. Malchukov & A. Spencer (Eds.), The Oxford Handbook of Case (pp. 304–321). Oxford: Oxford University Press.

Chomsky, N. (1995). The Minimalist Program. MIT Press.

Embick, D., & Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: Correlational, integrated, and explanatory neurolinguistics. Language, Cognition and Neuroscience, 30(4), 357–366.

Ferreira, V. S. (2010). Language production. Wiley Interdisciplinary Reviews. Cognitive Science, 1(6), 834–844.

Fokkema, M., Smits, N., Zeileis, A., Hothorn, T., & Kelderman, H. (2018). Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. Behavior Research Methods, 50(5), 2016–2034.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. Psychological Science, 11(4), 274–279.

Hanslmayr, S., Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. Frontiers in Human Neuroscience, 6, 74.

Hilpert, M. (2014). Construction Grammar and Its Application to English. Edinburgh University Press.

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. {NeuroImage}, 105, 507–524.

Klimesch, W. (2012). α-band oscillations, attention, and controlled access to stored information. Trends in Cognitive Sciences, 16(12), 606–617.

Marr, D. (2010). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Cambridge: MIT Press. (Original work published 1982)

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. The European Journal of Neuroscience, 48(7), 2609–2621.

Mirman, D. (2016). Growth Curve Analysis and Visualization Using R. CRC Press.

Pollard, C., & Sag, I. A. (1994). Head-Driven Phrase Structure Grammar. University of Chicago Press.

Sadaghiani, S., & Kleinschmidt, A. (2016). Brain Networks and α-Oscillations: Structural and Functional Foundations of Cognitive Control. Trends in Cognitive Sciences, 20(11), 805–817.

Sauppe, S. (2017). Word Order and Voice Influence the Timing of Verb Planning in German Sentence Production. Frontiers in Psychology, 8, 1648.

What do Indo-Aryan elements in Kanashi and Kinnauri say about their past?

Anju Saxena, Bernard Comrie, Lars Borin & Padam Sagar
(Uppsala University, University of California Santa Barbara, University of Gothenburg & Nako)

Keywords: Sino-Tibetan; Indo-Aryan; language contact; borrowing; prehistory

Schedule: We 17.00 Room 8

Kanashi (xns) is a Sino-Tibetan (ST) language spoken by a population of 1,500–2,000 in only one village – Malana (Kullu district, Himachal Pradesh, India). Malana is surrounded by Indo-Aryan (IA) speaking villages. Unfortunately there is very little published information about Kanashi: some short word lists (Harcourt 1871; Diack 1896; Grierson 1928; Tobdan 2010); a grammatical sketch (6 pages) and two short texts in Grierson (1909); and brief descriptions by Sharma (1989) and Saxena and Borin (2013).

The focus of this presentation is on borrowing in Kanashi. Based on our own fieldwork data, we will show that a substantial portion of the Kanashi lexicon is borrowed from IA.

Many di- and polysyllabic Kanashi nouns recognizably of IA origin exhibit a final part *-aŋ*, *-iŋ* or *-as* which is added to a borrowed IA noun stem. There are no apparent morphophonological factors determining the distribution of *-aŋ*, *-iŋ* or *-as*.

A similar adaptation process of IA borrowings is also observed in verbs, for example, there is a transitivizing/causative formation, where the transitive verbs contain *-j(aː)*.

Interestingly these same two processes are also found in Kinnauri (kfk), a closely related ST language spoken some 200 km away in lower/middle Kinnaur (Saxena 2017), but not in other ST languages of this region – even though they also have many IA borrowings. In Kinnauri, too, some IA noun stems receive an added *-aŋ*, *-iŋ* or *-as*. Whenever comparative data is available, we find that in almost all cases, the same IA item takes the same ending in Kanashi and Kinnauri (*-aŋ*, *-iŋ* or *-as*).

The origin of this final part is not clear; it does not appear in indigenous words and is not present in the present-day surrounding locally dominant IA languages (Kullu Pahari and Kotgarhi respectively) or in Hindi.

This raises some interesting questions. Firstly, what does this say about the older stages of Kanashi and Kinnauri? One possible scenario could be that the process of linguistic adaptations of IA borrowings took place or began when these two languges were part of a common earlier variety – before they separated from each other and moved to their present locations. Secondly, what is the source of these endings? In this presentation we will describe the two adaptation strategies in Kanashi and Kinnauri and discuss possible origins for them.

Acknowledgements

The work presented here has been financed in part by the Swedish Research Council, projects *Documentation of an endangered language: Kunashi* (contract 2014-560) and *South Asia as a linguistic area? Exploring big-data methods in areal and genetic linguistics* (contract 2014-969). We express our gratitude to our language consultants for their encouragement and their invaluable help.

References

Bailey, T. Grahame (1908), *The Languages of the Northern Himalayas, Being Studied in the Grammar of Twenty-Six Himalayan Dialects*, London: Royal Asiatic Society.

Diack, Alexander H. (1896), *The Kulu Dialect of Hindi: Some Notes on its Grammatical Structure, with Specimens of the Songs and Sayings Current Amongst the People, and a Glossary,* Lahore: The Civil and Military Gazette.

Grierson, George A. (1909), *Linguistic Survey of India, vol 3: Tibeto-Burman Family. Part I: General Introduction, Specimens of the Tibetan Dialects, the Himalayan Dialects, and the North Assam Group*, Calcutta: Government of India, Central Publication Branch.

Grierson, George A. (1928), *Linguistic Survey of India, Vol 1, Part II: Comparative Vocabulary,* Calcutta: Government of India, Central Publication Branch.

Harcourt, Alfred F. P. (1871), *The Himalayan Districts of Kooloo, Lahoul, and Spiti,* London: Wm H. Allen & Co.

Saxena, Anju (2017), Sangla Kinnauri. G. Thurgood and R. J. LaPolla (eds), *The Sino-Tibetan Languages*. Second edition, London: Routledge, 756–772.

Saxena, Anju and Lars Borin (2013), Carving Tibeto-Kanauri by its joints: Using basic vocabulary lists for genetic grouping of languages, in L. Borin and A. Saxena (eds), (2013), *Approaches to Measuring Linguistic Differences*, Berlin: Mouton de Gruyter, 175–198.
Sharma, Devidatta (1989), *Tribal Languages of Himachal Pradesh*, Delhi: Mittal Publications.
Tobdan (2010), *Exploring Malana. An Ancient Culture Hidden in the Himalayas*, New Delhi: Indus Publishing Company.

Quotes and their function: An empirical analysis

Marcel Schlechtweg
(University of Kassel)

Keywords: Quotation marks, acoustic correlates, quotation, name informing, duration

Schedule: Thu 11.30 Room 9

Quotation marks are a well-known metalinguistic tool. Their primary purposes are to (a) introduce directly reported speech (e.g., *Tom said: "I'll be there!")*, (b) signal reference to linguistic shapes (e.g., *"Leipzig" has seven letters.*), (c) indicate an ironic interpretation of an expression (e.g., *The "bed" was in reality a saggy mattress on the floor.*), or (d) inform about the name of a concept (e.g., *The "mitt" is a special type of glove for protecting a person's hand.*) (for overviews, see, e.g., Brendel et al., 2011, Saka & Johnson, 2017). While the semantics and pragmatics of quotation marks have been well explored from a theoretical perspective over the last decades (see, e.g., Davidson, 1979, Recanati, 2001, Saka, 1998), empirical evidence on the role quotation marks play for language users has been rare.

In a recent production study, sixteen native speakers of German read eight nouns (*Kaper* 'caper', *Pappel* 'poplar', *Kutte* 'robe', *Kippa* 'kippah', *Koppel* 'paddock', *Kuppe* 'tip/peak', *Pita* 'pita', *Pauke* 'timpani') embedded in the four conditions presented in (1) in order to test the role quotation marks play in spoken language.

(1) a. *Viele Mönche tragen die **sogenannte Kutte** täglich von morgens bis abends.*
'Many monks wear the **so-called robe** everyday from morning to night.'
b. *Viele Mönche tragen die **sogenannte „Kutte"** täglich von morgens bis abends.*
'Many monks wear the **so-called "robe"** everyday from morning to night.'
c. *Viele Mönche tragen die **wohlbekannte Kutte** täglich von morgens bis abends.*
'Many monks wear the **well-known robe** everyday from morning to night.'
d. *Viele Mönche tragen die **wohlbekannte „Kutte"** täglich von morgens bis abends.*
'Many monks wear the **well-known "robe"** everyday from morning to night.'

The analyses showed that quoted nouns (see 1b/d) were pronounced with a significantly longer first syllable, plosive of this syllable, constriction of this plosive, VOT of this plosive, and a significantly higher fundamental frequency in the vowel of this syllable than non-quoted nouns (see 1a/c). Based on this main effect, it is argued that the presence/absence of quotation marks has an influence on the acoustic realization of an item. The implications of the results for theories of quotation as well as the interface between semantics, pragmatics, and phonetics will be discussed.

References

Brendel, Elke, Jörg Meibauer & Markus Steinbach (Eds.), (2011), *Understanding Quotation* (Mouton Series in Pragmatics 7), Berlin: De Gruyter Mouton.
Davidson, Donald (1979), Quotation, *Theory and Decision, 11(1)*, 27–40.
Recanati, François (2001), Open quotation, *Mind, 110(439)*, 637–687.
Saka, Paul (1998), Quotation and the use-mention distinction, *Mind* 107(425), 113–135.
Saka, Paul & Michael Johnson, (Eds.) (2017), *The semantics and pragmatics of quotation*, Berlin, Germany: Springer.

Typological causative prominence and causal verb text frequencies: Refinements based on language documentation corpora

Frank Seifart, Stefan Schnell, Maria Bardají i Farré, Christian Döhler, Sonja Gipper, Katja Hannß, Katharina Haude, Birgit Hellwig, Anna Margetts, Sonja Riesberg, Harriet Sheppard, Claudia Wegener & Nikolaus P. Himmelmann
(CNRS-DDL, Leibniz-ZAS and University of Cologne; University of Bamberg; University of Cologne; University of Cologne; University of Cologne; University of Cologne; CNRS-SeDyL and University of Cologne; University of Cologne; Monash University; University of Cologne and CoEDL-Australian National University; Monash University; University of Cologne; & University of Cologne)

Keywords: Form-frequency correspondences, causative marking, causal verb frequencies, language documentation, language sampling

Schedule: We 11.30 Room 6

Verbs with meanings such as BOIL or BREAK can be ranked typologically according to *causative prominence*, i.e. their likelihood to receive overt (anti-)causative marking when expressing a (non-)causal event: compare, for instance, Spanish *hervir* '(come to) boil [intrans.]' vs. *hacer hervir* '(make) boil [trans.]' with *romper-se* 'break [intrans.]' vs. *romper* 'break [trans.]' (Haspelmath 1993). Haspelmath et al. (2014) argue that such overt (anti-)causative marking tends to correlate with the text frequencies of corresponding causal and noncausal event expressions, so that *romper-se* and meaning equivalents in other languages should be less frequent in texts than *romper* and respective equivalents, contributing to general coding efficiency in language. However, both data on typological causative prominence and data on text frequencies of corresponding verb forms remain typologically unrepresentative, being restricted largely to better-studied languages of Europe, North America and North-East Asia, and are hence unreliable so far, leaving the strength of such conclusions doubtful.

Here, we present refined results for both sides of the putative correlation. We selected six verb concepts with different causative prominence: high (DRY, BOIL), mid (TURN, BURN), and low (OPEN, BREAK). For each of them, we first establish refined typological causative prominence scores, based on a genealogically and areally balanced sample of 100 languages, combining results from our own coding with previously published results (Haspelmath 1993, Nichols, Peterson & Barnes 2004, and The WATP Consortium 2014). Secondly, we establish text frequencies of causal vs. noncausal verb uses in spoken language corpora of over 20,000 words each from 14 languages (Bora, Chipaya, Komnzo, Nafsan, Movima, Qaqet, Saliba, Savosavo, Sudest, Tetun Dili, Totoli, Vera'a, Waima'a, and Yurakaré), representing various genres, as collected in language documentation projects.

Our results on typological causative prominence are largely consistent with previously reported tendencies; yet, our data lead to refinements of scores and suggest re-ordering of verbs in causative ranking in some cases. Our results on text frequencies of causal vs. noncausal events in language documentation corpora are overall surprisingly consistent with previously reported frequency counts in large, written corpora (Haspelmath et al. 2014) (Figure 1). However, our results also deviate from previously reported ones in some respects, some of which may be due to specificities of corpus contents and sizes. Moreover, although we find a good correlation of causal event frequencies with typological causative prominence ranks, linear correlations with causative prominence scores show a less good fit. Additionally, we find that for individual verb pairs in individual languages, there is indeed an overwhelming tendency of overt marking on the less frequent verb form.

In sum, our study confirms the existence of form-frequency correspondences in morphological coding asymmetries, but suggests some refinements. Unlike previous studies, our investigation is based on text frequencies from spoken language corpora from a wide variety of languages, a much more appropriate type of data for studying wide-reaching long-term processes of language change than written corpora of mostly western languages. In general, our study demonstrates the potential contribution of recently compiled language documentation corpora to current theorizing in usage-based linguistics.

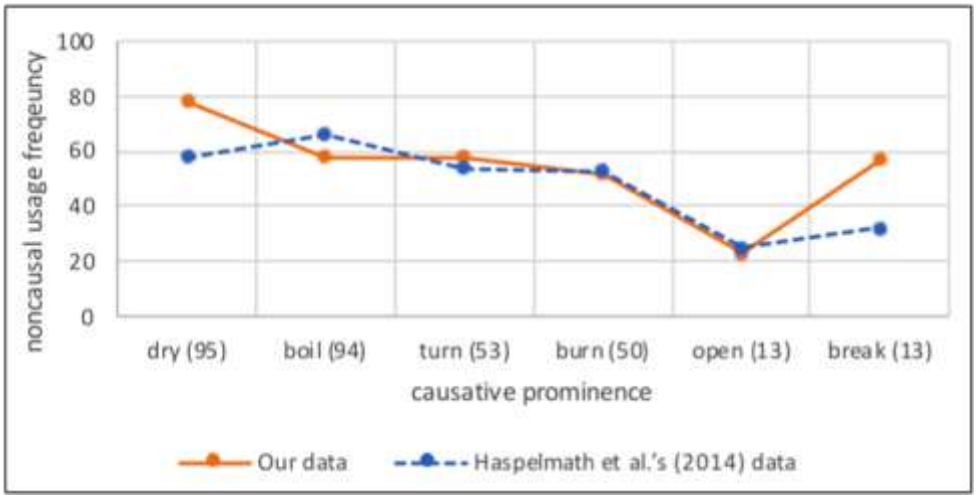


Figure 1: Preliminary results compared to Haspelmath et al. (2014)

References

Haspelmath, Martin (1993), More on the typology of inchoative/causative verb alternations. In B. Comrie, and M. Polinsky (eds), (1993), *Causatives and Transitivity*, Amsterdam / Philadelphia: John Benjamins, 87-120.

Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog & Elif Bamyaci (2014), Coding causal–noncausal verb alternations: A form–frequency correspondence explanation, *Journal of Linguistics* 50(3), 587-625.

Nichols, Johanna, David A. Peterson & Jonathan Barnes (2004), Transitivizing and detransitivizing languages, *Linguistic Typology* 8(2), 149–211.

The WATP Consortium (2014), *The World Atlas of Transitivity Pairs*, Tokyo: National Institute for Japanese Language and Linguistics. http://watp.ninjal.ac.jp/.

Aspect in loan verbal neologisms in Czech

Magda Ševčíková
(Charles University, Prague)

Keywords: verb, neologism, aspect, suffix, prefix

Schedule: Thu 9.00 Room 1

Borrowing of verbs in Czech necessarily implies adoption of both a thematic suffix and an ending and thus a full integration into the conjugation system, ex. (1). To accommodate loan verbs, Czech uses native suffixes (Mravinacová 2005, Bláha 2010) and matches the "indirect insertion pattern" in the typological classification of loan verbs by Wichmann & Wohlgemuth (2008).

(1) *klik-a-t*
click-IPFV-INF
'to click.IPFV'

The present paper focuses on how loan verbs form their aspectual counterpart in Czech, either by changing the thematic suffix in a simplex verb (2a), or by deriving a prefixed verb (3a). The main hypothesis to discuss is that the formation of aspectual pairs is determined by whether the loan verb is interpreted as having either a verbal base, or a noun base. Verbs with verbal bases prefer suffixes to change aspect while verbs with noun bases derive aspectual counterparts by prefixation. Loan verbs are thus assumed to follow the same patterns that were recently proposed for native verbs, cf. (2b) and (3b) (Ševčíková, 2018).

(2)(a) *klik-a-t* – *klik-nou-t*
click-IPFV-INF click-PFV-INF
'to click.IPFV' 'to click.PFV'

(b) *pad-a-t* – *pad-nou-t*
fall-IPFV-INF fall-PFV-INF
'to fall.IPFV' 'to fall.PFV'

(3)(a) *reset-ova-t* – *z-reset-ova-t*
reset-IPFV-INF PFV-reset-IPFV-INF
'to reset.IPFV' 'to reset.PFV'

(b) *cukr-ova-t* – *po-cukr-ova-t*
sugar-IPFV-INF PFV-sugar-IPFV-INF
'to sugar.IPFV' 'to sugar.PFV'

The study is based on language data stored in a neologism database (Neomat), namely on a sample of 100 verbs with different loan bases that have a direct verbal counterpart in English, which is the main donor language for Czech loanwords (Warmbrunn 1994, Martincová 1998, 2004, 2005). For each verb, verbs sharing the same root morpheme are extracted from a representative corpus of Czech (Křen et al. 2015).

A preliminary analysis suggests that if a suffixed counterpart is not attested for a verb in the analysed data, it is not found even in larger and/or more recent corpus data (e.g. Křen et al. 2018) while the number of prefixed verbs for a particular base is usually growing when increasing the dataset. In some groups of verbs, a base combines with different suffixes without changing the aspect (4a,b), the prefixation pattern is then assumed to apply (4c, d).

(4)(a) *googl-i-t* (b) *googl-ova-t*
google-IPFV- INF google-IPFV- INF

	‘to google.IPFV’		‘to google.IPFV’
(c)	*vy-googl-i-t*	(d)	*vy-googl-ova-t*
	PFV- google-IPFV- INF		PFV- google-IPFV- INF
	‘to google.PFV’		‘to google.PFV’

The hypothesis that formation of aspectual pairs with loan verbs is not a matter of chance but is rooted in structural features of Czech will be supported by an analysis of the inner structure of individual groups of verbs, their relations to non-verbal derivatives, corpus frequency, word class of the source verb in the donor language, and other features. The results will be compared to Moravscik’s hypothesis (1975; revised by Wichmann & Wohlgemuth 2008) that loan verbs are treated mostly as nouns in the borrowing languages.

Acknowledgements
This work was supported by the Grant No. GA19-14534S of the Czech Science Foundation.

References

Bláha, Ondřej (2010), Slovesné neologismy v češtině, *Bohemica Olomucensia* 2(3), 11–16.
Martincová, Olga (1998), *Nová slova v češtině. Slovník neologizmů*, vol. 1, Praha: Acaedmia.
Martincová, Olga (2004), *Nová slova v češtině. Slovník neologizmů*, vol. 2, Praha: Acaedmia.
Martincová, Olga (2005), Nová slovesná pojmenování, in *Neologizmy v dnešní češtině*, Praha: ÚJČ AV ČR, 119–133.
Moravcsik, Edith (1975), Borrowed verbs, *Wiener Linguistische Gazette* 8, 3–30.
Mravinacová, Jitka (2005), Přejímání cizích lexémů. In *Neologizmy v dnešní češtině*, Praha: ÚJČ AV ČR, 187–211.
Ševčíková, Magda (2018), Discerning action nouns formed by a zero suffix from unmotivated nouns, presented at the “Word-formation Theories III & Typology and Universals in Word-formation IV” Conference in Košice, June 2018.
Warmbrunn, Jürgen (1994), *Englische lexikalische Entlehnungen im Wortschatz der tschechischen Gegenwartssprache*, Münster: LIT Verlag.
Wichmann, Søren & Wohlgemuth, Jan (2008), Loan verbs in a typological perspective, in: T. Stolz et al. (eds), (2008), *Aspects of Language Contact: New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*, Berlin: de Gruyter, 89–121.

Language data resources

Křen, Michal et al. (2015): *SYN2015: A representative corpus of written Czech*. Prague: Institute of Czech National Corpus, Faculty of Arts, Charles University. Available online http://www.korpus.cz. Accessed 2019-01-04.
Křen, Michal et al. (2018): *Corpus SYN, version 7 from 29. 11. 2018*. Prague: Institute of Czech National Corpus, Faculty of Arts, Charles University. Available online http://www.korpus.cz. Accessed 2019-01-04.
Database of excerption material Neomat (2015). Version 3.0. Praha: Oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR. Available online http://neologismy.cz. Accessed 2019-01-04.

Sentence processing and the that/zero complementizer alternation: Investigating grammatical complexity and predictive processing

Christopher Shank, Anouschka Foltz & James Bragg
(Bangor University; University of Graz; & Bangor University)

Keywords: on-line sentence processing, that/zero complementizer alternation, wrap-up processing, grammatical simplification / complexity, liner mixed effects model

Schedule: We 16.30 Room 15

This paper examines the distribution of *that*/zero complementation alternation patterns in terms of cognitive processing. Two structural factors, which have been shown to be the most reliable predictors of the presence of the *that* form are examined: the absence of extra elements in the matrix clause (Tagliamonte and Smith 2005; Torres Cacoullos and Walker 2009; Shank, Plevoets & Cuyckens, 2014) and the length of the complement clause subject (Elsness 1984; Kearns 2007a, 2007b; Shank & Plevoets, 2018). We attempt to answer the following research questions: (1) Is the use of the zero form, versus the complementizer *that*, a case of general grammatical simplification or increased complexity? Alternatively, does increased complexity depend on whether readers are expecting *that* in the sentence? (2) Do readers predict the length of the complement clause subject in sentences that use the zero form versus the complementizer *that*? To answer these questions, two self-paced reading tasks were designed around the mental state complement-taking verbs *think*, *guess*, *know* and *believe*. Participants read sentences one word at a time and answered comprehension questions about the sentences.

Task 1 contained the following sentences.
He (immediately) thinks (3a) *that* this dog is going to run away
(3b) this dog is going to run away
(3c) that dog is going to run away

If *that* causes a general decrease or increase in complexity, the presence of *that* should decrease or increase reading times across the board. But if increased complexity depends on whether readers are expecting *that* in the sentence, then participants should read *this/that dog* more quickly in (3a) compared to (3b) and (3c) without extra material in the matrix clause (i.e. without *immediately*) and more quickly in (3b) and (3c) compared to (3a) with *immediately* in the matrix clause.

Task 2 contained the following sentences.
I think (4a) the mother is polite.
(4b) *that* the mother is polite.
(4c) the mother of the taxi driver is polite.
(4d) *that* mother of the taxi driver is polite.

If the presence of *that* leads readers to predict a long complement clause subject, whereas the zero form leads them to expect a short complement clause subject, then readers may expect the subject phrase to end at *mother* for the zero form, but not following *that*. In this case, we would expect additional clause-final wrap-up processing (Rayner et al., 2000) at *mother* for zero sentences, but not for *that* sentences.

Sixty native-English speaking UK undergraduate students completed either task 1 or task 2. Task 1 had a 2 matrix clause type x 3 *that*/zero/ambiguous x 4 verb design, and task 2 had a 2 complement clause subject length x 2 *that*/zero x 4 verb design. Both experiments had 48 target and 96 filler sentences. The results were analysed using linear mixed effects models with participant and items as

random effects. The results and subsequent discussion provide further insight on how distributional properties of the *that*/zero alternation relate to online sentence processing vis-à-vis grammatical complexity and/or simplification.

References

Elsness, J. (1984). *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65. 519–533.

Kaltenböck, G. (2006) . '. . . That is the question': Complementizer omission in extraposed that-clauses. *English Language and Linguistics* 10. 371–396.

Kearns, K. (2007a). Epistemic verbs and zero complementizer. *English Language and Linguistics* 11. 475–505

Kearns, K. (2007b). Regional variation in the syntactic distribution of null finite complementizer. *Language Variation and Change* 19. 295–336.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. The Quarterly Journal of Experimental Psychology Section A, 53(4), 1061-1080.

Rissanen, M. (1991). On the history of that zero in object clause links in English. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 272–289. London: Longman.

Rohdenburg, G. (1996). Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7. 149–182.

Shank, C., Plevoets, K., & Cuyckens, H. 2014. A Diachronic Corpus based Multivariate Analysis of "I think that" versus "I think zero*" Polysemy and Synonymy. Corpus Methods and Applications in Cognitive Linguistics*. Glynn, D & Robinson, J. (Eds). John Benjamins Publishing Co

Shank, C & Plevoets, K. 2018. Investigating the impact of structural factors upon that/zero complementizer alternation patterns in verbs of cognition: A diachronic corpus-based multifactorial analysis. *Research in Corpus Linguistics* 6: 83–112

Tagliamonte, S & Smith, S. (2005). *No momentary fancy!* The *zero* 'complementizer' in English dialects. *English Language and Linguistics* 9. 289–309.

Thompson, S.A. & Mulac, A. (1991a). The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15. 237–251.

Thompson, S.A. & Mulac, A. (1991b). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, Volume II, 313–329. Amsterdam: John Benjamins.

Torres Cacoullos, R & Walker, J. A. (2009). On the persistence of grammar in discourse formulas: A variationist study of *that*. *Linguistics, 47*, 1–43.

Non-finite bias: Loanword accommodation of French-origin loan verbs in Middle English

Marlieke Shaw
(University of Leuven)

Keywords: historical linguistics, loanword accommodation biases, inflectional categories, codeswitching vs. borrowing, contact-induced change

Schedule: We 17.30 Room 8

According to Poplack, Sankoff and Miller (2009), loanwords need to accommodate to the grammatical structure of their recipient language. Loan verbs, for instance, "must be inflected for tense, mood, and person" (Poplack et al., 2009, p. 62). The dominant strategy for achieving this is *direct insertion*, which adds native inflections directly onto borrowed stems (Wohlgemuth, 2009). However, there is reason to suspect that even under direct insertion, loan verbs may favour some native inflectional categories over others (De Smet, 2014).

In the present paper, we focus on the influx of Anglo-French (AF) loanwords into Middle English (ME) during the second half of the fourteenth century (Dekeyser, 1986, and Jespersen, 1905). We endeavour to assess how French-origin loan verbs were accommodated to late ME by investigating whether AF verbs favoured any specific forms of the ME verbal paradigm.

Extracting a sample (1350-1420) from PPCME2 (Kroch & Taylor, 2000), a diachronic corpus consisting of ME texts, we ensure to cover a wide variety of genres. We code each attestation for its lemma, origin (Germanic or French) and inflectional form. In the following example (1), Germanic-origin *wenen* appears in its finite verb form, while French-origin *scornen* is a non-finite form:

(1)	*The*	*kyng*	*here*	*fader*	***wende***	*þat*	*she*	***hade***	***scorned****him*
	The	king	her	father	believed	that	she	had	mocked him

(1400, PPCME2, *The Brut or the Chronicles of England*, emphasis added)

From the manual corpus analysis, the main finding is that non-finite structures are significantly more prevalent in French-origin loan verbs (e.g. *scornen*) than in their Germanic equivalents (e.g. *wenen*) (which is consistent with earlier findings for English loan verbs in Dutch). The results are particularly striking for the inflectional categories of the bare nominal gerund, the passive construction, the past participle and the verbal gerund. At the same time, the effect proves to be author- and even text-specific. We presume that it manifests itself in texts characterised by codeswitching (e.g. Chaucer's *The Wife of Bath's Tale*), written by bilingual authors, where loan verbs have not really been integrated in their recipient language yet. In contrast, non-finite bias is less pronounced in texts hosting a lower number of French-origin verbs and in translations from Romance languages (e.g. Chaucer's *Boethius*), which presumably avoid codeswitches but not borrowings. We further test this explanation by comparing the behaviour of high-frequency and low-frequency AF loan verbs – the latter being more likely to be codeswitches.

Finally, we speculate that non-finite bias may have contributed to indirect contact-induced syntactic change. Green (2017) observes an overall increase in non-finite structures in English. We suggest that the AF loan verbs may have consistently favoured non-finite forms over finite forms, and that those loanword accommodation biases may be held partly responsible for the increased reliance on non-finite verb forms in fourteenth-century ME. As such, this study throws light on strategies for loanword accommodation (cf. Wohlgemuth, 2009) and on the possible indirect effects of language contact (i.e. contact-induced change).

This research was funded by grant G0D1418N from Fonds voor Wetenschappelijk Onderzoek (FWO) – Vlaanderen.

References

De Smet, Hendrik (2014), De integratie van Engelse leenwerkwoorden in het Nederlands, in F. Van de Velde, H. Smessaert, F. Van Eynde, and S. Verbrugge (eds), (2014), *Patroon en argument: Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, Leuven: Leuven University Press, 75-87.

Dekeyser, Xavier (1986), Romance loans in Middle English: a re-assessment, in D. Kastovsky, and A. Swedek (eds), (1986), *Linguistics across historical and geographical boundaries*, Berlin: de Gruyter Mouton, 253-265.

Green, Clarence (2017), *Patterns and Development in the English Clause System: A Corpus-Based Grammatical Overview*, Singapore: Springer.
Jespersen, Otto (1905), *Growth and Structure of the English Language*, Leipzig: Teubner Publisher.
Kroch, Anthony, and Taylor, Ann (2000), The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Retrieved 11 January 2019, from https://www.ling.upenn.edu/histcorpora/PPCME2-RELEASE-4/index.html
Poplack, Shana, Sankoff, David, and Miller, Christopher (2009), The social correlates and linguistic processes of lexical borrowing and assimilation, *Linguistics* 26(1), 47-104. https://doi.org/10.1515/ling.1988.26.1.47
Wohlgemuth, Jan (2009), *A typology of verbal borrowings*, Berlin: de Gruyter Mouton.

Towards a typology of semantic domains of CLEAN, DIRTY, TRANSPARENT and TURBID

Maria Sidorova
(Lomonosov Moscow State University and Institute of Linguistics, RAS)

Keywords: lexical typology; semantics; frame-based approach; grammaticalization; borrowing

Schedule: We 15.00 Room 14

This research has been supported by Russian Foundation for Basic Research, grant № 19-012-00627a. I am grateful to Egor Kashkin & participants of INFUSE 2016 workshop in Szeged, Daria Mordashova & participants of International Research School 2018 in Yakutia, Vera Maltseva, Ksenia Patrikeeva, and Diana Zaripova for their help with collecting data.

This paper undertakes a typological research on the domains of CLEAN, DIRTY, TRANSPARENT and TURBID, which are still underdescribed (some initial contributions were made by Arkhangelskij et al. (2011), Gotlan (2014)). The study of these domains is interesting for the typology of semantic caritives (see e.g. Tolstaya 2008; Koptjevskaja-Tamm, Miestamo 2015), as well as for the general typology of how different semantic domains can be related (see e.g. Zalizniak et al. 2012; Juvonen, Koptjevskaja-Tamm (eds.) 2016). My sample includes 29 languages: Indo-European (Danish, English, German, Italian, Hindi, Persian, Polish, Russian, Serbian, Slovenian, Spanish), Uralic (Estonian, Finnish, Hill Mari, Komi, Moksha, Udmurt, Western Khanty, Hungarian), Altaic (Altai, Buryat, Khakas, Tatar, Yakut), Kra–Dai (Thai), Austronesian (Malay, Tagalog), Sino-Tibetan (Chinese), Dravidian (Tamil). I adopt the frame-based approach to lexical typology, see Rakhilina & Reznikova (2013, 2016). The data were carried out mainly by elicitation. The questionnaire was aimed at studying collocational properties of lexemes.

There are three main points to be considered in this paper.

Firstly, I examine the semantic oppositions within these domains (more detail will be provided in the talk; the oppositions will be visualized on the semantic map), for example:

CLEAN: no dirt / no admixtures (Spanish *puro* ‘pure’ vs. *limpio* ‘clean’), original / created cleanness (Moksha *čistaj* ‘clean’ vs. *ur'andaj* ‘tidy, neat’);

TRANSPARENT: completely transparent (Slovenian *prozoren* ‘transparent’) / transparent only to some extent (Slovenian *prosojen* ‘translucent’); aggregate state (Italian *trasparente* ‘transparent (solid

objects & liquids & rarely air/sky)' vs. *limpido* 'clear (liquids & air/sky)' vs. *terso* 'clear (air/sky)');

DIRTY: full / partial coverage of surface (Buryat *budagdahan* about partial coverage vs. *mu:xai, xirətəi* and other lexemes about full coverage); degree of dirtiness (Komi *s'əd* about a lower degree of dirtiness vs. *n'as't'i* about a higher degree of dirtiness);

TURBID: an external physical impact (Moksha *šor'af* – only if there is such impact, otherwise *mutnaj*; degree of opacity – not completely penetrable / impenetrable to the light (English *turbid* vs. *opaque*).

Secondly, I deal with the origin of the lexemes concerned, including semantic shifts lying behind them. I will show that most lexemes in these domains can be considered "secondary" in various ways, being transferred there in four main ways:

Borrowing from another language. Proto-Altaic **ĕ́ro* 'clean' (StarLing) was borrowed into several Uralic languages (cf. Moksha *aru* 'clean (water, air)', see Paasonen (1990-1999: 66)), as well as some earlier Indo-European lexemes: Finnish *puhdas* 'clean' and its Estonian cognate *puhas* originating from **puHtos* 'clean' (Häkkinen (2004: 962)).

Shift from an adjacent semantic domain. Lexemes meaning CLEAN and DIRTY often develop from colour terms: WHITE for CLEAN / TRANSPARENT and BLACK for DIRTY (Khanty, Komi, Udmurt, Estonian). Less frequent in my sample are shifts from BRIGHT to TRANSPARENT (Moksha) and from BROWN to TURBID (Khakas).

Derivation from a noun. Terms for DIRTY are often derived from a noun 'dirt' with attributivizers (Khanty, Moksha, Khakas, Buryat, Altai). Some languages use a juxtaposed noun 'dirt'. Terms for CLEAN may be morphological caritives from a noun 'dirt' (Komi, Buryat).

Compounding based on a periphrastic construction. This pattern is especially frequent for the domains of TRANSPARENT (often lit. 'see-through') and TURBID (often lit. 'not-see-through').

Thirdly, lexemes from CLEAN / TRANSPARENT domains tend to grammaticalize into focus particles (cf. English *purely*, Russian *чисто* 'purely') and develop quantificational properties (Hill Mari *cišti* 'purely; all' < Russian *чистый* 'clean'). This topic has been elaborated mainly in Uralic and partially in Indo-European languages so far, and the tendency is as follows: the lexeme undergoes grammaticalization only if it can mean 'pure, without admixture'. Note that Heine & Kuteva (2004) do not deal with grammaticalization of adjectives.

References

Arkhangelskij, T. A., M. G. Tagabileva & L. S. Kholkina (2011). KACHESTVENNYE PRIZNAKI 'CHISTYJ', 'GRYAZNYJ', 'PROZRACHNYJ', 'MUTNYJ': k postroeniyu semanticheskoj tipologii [Towards a semantic typology of qualities 'CLEAN', 'DIRTY', 'TRANSPARENT', 'TURBID'], in: *ACTA LINGUISTICA PETROPOLITANA. Transactions of the Institute of Linguistic Studies of the RAS*, Vol. VII. P 3 – SPb: Nauka. P. 257–260.

Gotlan Yu. A. (2014). NICHT NUR SAUBER, SONDERN REIN, in: *Bulletin of Tomsk State University*, Vol. 387, P. 5–7.

Häkkinen, Kaisa (2004). *Nykysuomen etymologinen sanakirja*. WSOY, 2004.

Heine, Bernd & Tania Kuteva (2004). *World Lexicon of Grammaticalization.* Cambridge University Press, Cambridge.

Juvonen, Päivi & Maria Koptjevskaja-Tamm (eds.) (2016). *The lexical typology of semantic shifts.* Berlin: De Gruyter Mouton.

Koptjevskaja-Tamm, Maria & Matti Miestamo (2015). Antonyms and derivational negation: a pilot study of cross-linguistic variation. In *ALT 2015* (11th Conference of the Association for linguistic typology). Albuquerque, New Mexico, USA. Abstract booklet. P. 85–86.

Paasonen, Heikki (1990-1999). *Mordwinisches Wörterbuch*. Helsinki: Lexica Societatis Fenno-Ugricae, XXIII, I-VI.

Rakhilina, E. V. & T. I. Reznikova (2013). Frejmovyj podkhod k leksicheskoj tipologii [Frame-based approach to lexical typology], in: *Voprosy yazykoznaniya* [Topics in the study of language] 2. P. 3–31.

Rakhilina, E. V. & T. I. Reznikova (2016). Frame-based methodology for lexical typology, in: *The Lexical Typology of Semantic Shifts*. Berlin, Boston: Mouton de Gruyter. P. 95–129.

StarLing. *The Tower of Babel. An International Etymological Database Project.* (Available online at: http://starling.rinet.ru/intrab.php?lan=en, Accessed on 2017-04-13).

Zalizniak, Anna A., M. Bulakh, D. Ganenkov, I. Gruntov, T. Maisak & Russo M. (2012). The Catalogue of Semantic Shifts as a Database for Lexical Semantic Typology, in: *Linguistics*. 50 (3). P. 634–670.

Combining morphological and syntactic annotation

Ivan Šimko and Teodora Vuković

Schedule: Thu 16.00 Room 4

<pdf>

The second best word order: How ordinals restrict superlatives and why it's not reciprocated

Ruby Sleeman
(Goethe University Frankfurt am Main)

Keywords: ordinal numerals, superlatives, Dutch, German, generative syntax

Schedule: Thu 14.30 Room 7

Ordinals are said to occur high in the DP-spine, before adjectives (see a.o. Cinque 2010, Svenonius 2008). The same is said of superlatives (Dixon 1982 as cited by Scontras et al. 2017). Dutch data from the Corpora from the Web (COW; Schäfer 2015, Schäfer & Bildhauer 2012) show that ordinals and superlatives can co-occur prenominally, in either order:

(1)	a.	*De*	*tweede*	*hoogste*	*berg* (7,180 occurrences)
		the	second	highest	mountain
	b.	*De*	*hoogste*	*tweede*	*berg* (1,807 occurrences)
		the	highest	second	mountain

How can we account for the co-occurrence of both these orders from a cartographic perspective? On the basis of Dutch corpus data and German compound distribution, I propose that further refinements are needed in the cartography of DP, for three reasons: (i) Superlatives and ordinals can occur in either order in Dutch; (ii) This is not the result of focus movement but an effect of scopal interaction; (iii)

Ordinals can modify not only nouns but also superlatives, and can be situated inside the superlative's extended projection. (2a,b) represent the readings for (1a,b) respectively:

(2) a. [second] [highest mountain]
b. [highest] [second mountain]

An analysis for (2b) could be (3): *highest* originates lower but moves over *second* into a landing site for focused adjectives, call it Spec of KindP (see Svenonius 2008):

(3) $[_{KiP}$ highest [Ki^0 $[_{SortP}$ second [$Sort^0$ $[_{nP}$ ~~highest~~ [n^0 [√P table]]]]]]]

However, there is no evidence to assume that either order is the result of focus movement: (i) No special focus intonation is required; and (ii) Focus movement would predict ambiguities to arise due to reconstruction possibilities. No such ambiguities arise.

There is a second reading for (1a), which is unavailable for (1b):

(4) [second highest] mountain

In this reading, mountains are ranked according to their highness. The superlative *highest* is thus directly modified by the ordinal. Along the lines of Corver (2005), for this reading I propose the structure in (5):

(5) $[_{DP}$ de $[_{FP}$ $[_{SupP}$ **tweede** $[_{Sup'}$ $[_{Sup}$ **hoog-ste**] $[_{AP}$ ~~hoog~~]]] $[_{F'}$ F $[_{NP}$ berg]]]]

In (5), the superlative morpheme is the head of the SuperlativeP. The ordinal is in the specifier of this projection. The ordinal has no such slot available. The German equivalent of (1a), *der zweite höchste Berg*, can **only** have the reading in (2a). For (4), German requires compounding (see 6a), corroborating the close relation between the ordinal and the superlative in this reading. The reverse is not possible (6b).

(6)	a.	der	zweit-höchste	Berg
		the	second-highest	mountain
	b.	*der	höchst-zweite	Berg
		the	highest-second	mountain

To conclude, any cartographic attempt at defining the predetermined locations for prenominal modifiers needs to take into account the following facts: (i) Superlatives and ordinals can occur in either order in Dutch; (ii) This should not be analyzed in terms of focus movement but as the effect of scopal interaction; (iii) Not only can ordinals modify nouns, they can additionally occur as the modifier of a superlative and be located directly inside its extended projection - but not vice versa - as corroborated by German compounds.

References

Cinque, Guglielmo (2010), *The Syntax of Adjectives. A Comparative Study*, Cambridge, MA: The MIT Press.

Corver, Norbert (2005), Double comparatives and the Comparative Criterion, *Revue Linguistique de Vincennes* 34: 165-190.

Schäfer, Roland (2015), Processing and Querying Large Web Corpora with the COW14 Architecture, in *Proceedings of Challenges in the Management of Large Corpora (CMLC-3)* (IDS publication server), 28–34.

Schäfer, Roland and Felix Bildhauer (2012), Building Large Corpora from the Web Using a New Efficient Tool Chain, in *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493.

Scontras, Gregory, Judith Degen and Noah D. Goodman (2017), Subjectivity predicts adjective

ordering preferences, in *Open Mind: Discoveries in Cognitive Science, 1*(1), 53-65.
Svenonius, Peter (2008), The position of adjectives and other phrasal modifiers in the decomposition of DP, in L. McNally and C. Kennedy (eds), (2008), *Adjectives and adverbs: Syntax, semantics, and discourse*, Oxford, UK: Oxford University Press, 16–42.

Change and variation in prepositional relative constructions in Portuguese: A corpus-based multifactorial and cognitive analysis of pied-piping, chopping and resumptive relatives

Augusto Soares da Silva, Susana Afonso & Dafne Palú
(Universidade Católica Portuguesa & University of Exeter)

Keywords: relative constructions, constructional variation, profile-based analysis, entrenchment, Portuguese

Schedule: Thu 15.00 Room 7

In Portuguese, there are three types of prepositional relative constructions: pied-piping (1), chopping (2) and resumptive (3) relatives:

(1) *Ô Ney, o homem é produto do meio em que vive.* (BP, NURC-RJ)
'Ney, man is a product of the environment in which he lives'
(2) a. *porque há uma canção que eu gosto muito.* (EP, C-Oral)
'because there is a song that I like very much'
b. *O primeiro restaurante que comemos foi horrível.* (BP, Google)
'The first restaurant [in] which we ate was horrible'
c. *Dieta que morre de fome não rola.* (BP, Google)
'A diet that you die of hunger doesn't work'
(3) *[devido] a umas algas que agora esquece-me o nome delas*. (EP, CRPC)
'due to some algae whose name I forgot it'

The pied-piping relative, the only standard type of prepositional relative constructions in Portuguese, is usually indicative of formal, educated speech, and occurs more frequently in European Portuguese (EP) than in Brazilian Portuguese (BP). Chopping and resumptive relatives are thus regarded as non-canonical constructions, mostly in EP (Raposo et al. 2013). The former has long been thought to be an innovation of BP (Tarallo 1983) that EP has only recently developed (Arim et al. 2005), while the latter is the least productive and the most marked construction of the two, both in EP and BP (Alexandre 2000). Existing studies point out the differences of register and the stage at which these constructions are in both national varieties, or describe it in formal terms, such as wh-movement and raising (Tarallo 1983, Brito 1991, Kato 1993, Kenedy 2017).

We argue that the alternation between these constructions is not determined by structural factors alone, but mostly by cognitive and social factors (i.e. semantic, pragmatic-discursive and lectal). This cognitive-functional study provides a qualitative and multifactorial usage-feature/profile-based analysis of corpus data, complete with multivariate statistics modeling (multiple correspondence analysis and logistic regression). The corpus includes written and spoken data from EP and BP from 1970s and 2000s. Approximately 800 occurrences were annotated for factors such as animacy,

specificity and syntactic/semantic role of the head noun, type of preposition and valence and semantics of the relative clause predicate.

We show that the depronominalization and grammaticalization of the relative pronoun into a complementizer facilitate non-canonical strategies and open way to more flexible semantic and pragmatic relationships between nominal head and relative clause. Relative clauses consist of a reference-point relationship (Langacker 1993) between a noun/topic and a proposition, which is interpreted in the conceptual frame evoked by the noun/topic. Accordingly, the interpretation of chopping relatives depends on the ability to attribute a unique semantic role to the head noun and establish a semantic connection with the relative predicate inside the conceptual frame. This happens when the head noun is coreferential with a complement of the relative clause verb (2a) or with an adjunct (2b), but also when there is no syntactic relationship between the head noun and the relative clause predicate (2c), thanks to conceptual integration (Fauconnier & Turner 2002, Nikiforidou 2005). Finally, we argue that resumptives answer to the cognitive principle of referent accessibility (Ariel 1991) and that the greater the level of cognitive entrenchment (Wiechmann 2015), the greater the probability for chopping and resumptive strategies.

References

Alexandre, Nélia (2000), *A estratégia resumptiva em relativas restritivas do Português Europeu*, Dissertação de Mestrado, Lisboa: Universidade de Lisboa.

Ariel, Mira (1991), The function of accessibility in a theory of grammar, *Journal of Pragmatics* 16, 443-463.

Arim, Eva, Maria Celeste Ramiro, and Tiago Freitas (2005), Mudança em curso e os média: o caso das relativas, in M.H. Mateus, and M.F. Nascimento (eds.), *A Língua Portuguesa em Mudança*, Lisboa: Caminho.

Brito, Ana Maria (1991), *A Sintaxe das Orações Relativas em Português*, Lisboa: INIC.

Fauconnier, Gilles, and Mark Turner (2002), *The Way we Think: Conceptual Blending and the Mind's Hidden Complexities*, New York: Basic Books.

Kato, Mary (1993), Recontando a história das relativas em uma perspectiva paramétrica, in I. Roberts, and M. Kato (eds.), *Português Brasileiro: uma viagem diacrônica: homenagem a Fernando Tarallo*, Campinas: Editora da Unicamp, 223-261.

Kenedy, Eduardo (2017), *A língua Portuguesa no Brasil e em Portugal. O caso das orações relativas*, Niterói: Editora da Universidade Federal Fluminense.

Langacker, Ronald W. (1993). Reference point constructions, *Cognitive Linguistics* 4 (1), 1-38.

Nikiforidou, Kiki (2005), Conceptual blending and the interpretation of relatives: A case study from Greek, *Cognitive Linguistics* 16 (1), 169-206.

Raposo, Eduardo P., M. Fernanda Nascimento, M. Antónia Mota, Luísa Segura, and Amália Mendes (2013), *Gramática do Português. Vol. II*, Lisboa: Fundação Calouste Gulbenkian.

Tarallo, Fernando (1983), *Relativization strategies in Brazilian Portuguese*, Ph.D. dissertation, Philadelphia: University of Pennsylvania.

Wiechmann, Daniel (2015), *Understanding Relative Clauses: A Usage-Based View on the Processing of Complex Constructions*, Berlin/Boston: Mouton De Gruyter.

Why do non-deontic modal, conditional and habitual meanings cluster? Answers from Polish

Anna Socka & Björn Wiemer
(University of Gdańsk, University of Mainz)

Keywords: Polish, aspect, modality, non-past tense use, telicity

Schedule: Sat 11.30 Room 1

Slavic languages do not generally belong to understudied languages; nonetheless, against a European background certain properties in the TAM domain remain poorly understood. Concomitantly, they are approached in disparate ways by linguists of generative vs functional-typological orientation. In particular, this concerns the distribution of perfective (pfv.) and imperfective (ipfv.) aspect over modal meanings and the relation of the latter to other meanings in which propositional (= truth-conditional) content is lacking (e.g., imperatives) or suspended (conditionals, habituals, future). Modern Polish data helps understand how a paradigmatic pfv.:ipfv. contrast yields meaning clusters in the aforementioned domains, provided two recurrent shortcomings in the analysis of its distribution are avoided.

Thus, according to the 'Aspect-Modality-Link' (AML) (Abraham 2005, Abraham/Leiss 2008, and Abraham et al. 2011), modals combining with lexically 'pfv.' infinitives trigger a deontic meaning, whereas modals combining with 'ipfv.' infinitives yield an epistemic meaning. This is motivated by assumedly different event structures (biphasic for 'pfv.' predicates and deontic modals vs monophasic for 'ipfv.' predicates and epistemic modals). Inherent to this reasoning are two shortcomings. First, (im)perfectivity is confused with the lexical feature of (a)telicity (*q*-boundedness in Croft 2012) and associated semantic effects on clause level. For grammatical aspect (a)telicity is not a defining feature; instead, *t*-boundedness (Croft 2012) is crucial (Wiemer 2014, Wiemer/Seržant 2017). Second, state-changing (= telic) events are conflated with correlated resultant states. However, pfv. verbs (in active voice) focus on events, and resultant states only follow from telic events via entailment.

An account of these distinctions solves most facts that contradict the AML, and explains why the conversational background (Kratzer 1981) varies depending on contextual clues. For instance, clausal complements [finite modal+pfv. infinitive] of epistemic predicates partake in modal concord (Ligara 1997: 100), and non-embedded *musieć* 'must' + pfv. infinitive can read epistemically, regardless of tense:

(1) *Sekretarka musi*[PRS] */ musiała*[PST] *(zaraz)* ***wrócić***[PFV.INF],
bo zostawiła kosmetyczkę na biurku
'The secretary must / was supposed to **return** (in a moment),
for she has/had left her cosmetic bag on the desk.' (Weiss 1987: 135)

Moreover, the feature crucial for aspect choice is not [± epistemic], but rather [± deontic]. Pfv. verbs pattern with circumstantial and dispositional (= participant-internal) modality (Wiemer 2001, 2014, and Divjak 2009, 2011), and these form a cluster with aspect choice in conditional and habitual contexts. Corroboration comes from present tense forms of pfv. verbs; in many environments these meanings overlap:

(2) *A bywa, że z żalu po człowieku i pies zaraz* ***umrze***[PFV.PRS.3SG].
'It happens that in grief over its late owner a/the dog immediately **dies** as well.'
(NKJP) habitual—conditional—dispositional

Similar patterns are also found outside of Europe (Narrog 2016: 96f., De Wit 2017). We propose an explanation of this clustering pattern which avoids the aforementioned conceptual conflations and is

compatible with recent findings on the aspect–future interface (Copley 2008, Mucha 2013, and Błaszczak et al. 2014). The explanation combines the strict pfv.:ipfv. contrast with usage patterns, as shown in a corpus-based pilot study on the distribution of the meanings of pfv. present-tense stems in Polish.

Acknowledgments: Research has been funded by German Science Foundation (WI 1286/19-1) and Polish Science Foundation (2016/23/G/HS2/00922), programme *Beethoven II*.

References

Abraham, Werner (2005), Event arguments and modal verbs, in C. Maienborn, and A. Wöllstein (eds), (2005), *Event arguments: Foundations and applications*. Tübingen: Niemeyer, 243–276.

Abraham, Werner, and Elisabeth Leiss (2008), Introduction, in W. Abraham, and E. Leiss (eds), (2008), *Modality-Aspect Interfaces: Implications and Typological Solutions*, Amsterdam / Philadelphia: John Benjamins, xi–xxiv.

Abraham, Werner, Peter Meihsner, Jadwiga Piskorz, and Kinga Piskorz (2011), Zurparadigmatischen Ausbildung von Modalverben, von Deontik und Epistemik im jüngeren gesprochenen Polnisch, in A. Kątny, and K. Lukas (eds), (2011), *Germanistik in Polen. Geschichte – Perspektiven – interdisziplinärer Dialog*, Frankfurt/M. etc.: Peter Lang, 157–190.

Błaszczak, Joanna, Patrycja Jabłońska, Dorota Klimek-Jankowska, and Krzysztof Migdalski (2014), The riddle of 'future tense' in Polish, in Ph. De Brabanter, M. Kissine, and S. Sharifzadeh (eds), (2014), *Future Times, Future Tenses*, Oxford: Oxford University Press, 165–204.

Copley, Bridget (2008), *The Semantics of the Future*, New York: Routledge.

Croft, William (2012), *Verbs: Aspect and Causal Structure*, Oxford: Oxford University Press.

De Wit, Astrid (2017), *The Present Perfective Paradox across Languages*, Oxford: OxfordUniversity Press.

Divjak, Dagmar (2009), Mapping between domains. The aspect–modality interaction in Russian, *Russian Linguistics* 33, 249–269.

Divjak, Dagmar (2011), Predicting aspectual choice in modal constructions: a quest for the Holy Grail?, in M. Grygiel, and L. A. Janda (eds), (2011), *Slavic Linguistics in a Cognitive Framework*, Frankfurt/M. etc.: Peter Lang, 67–85.

Kratzer, Angelika (1981), The notional category of modality, in H.-J. Eikmeyer, and H. Rieser (eds), (1981), *Words, Worlds, and Contexts*, Berlin, New York: De Gruyter Mouton, 38–74.

Ligara, Bronisława (1997), *Polskie czasowniki modalne i ich francuskie ekwiwalenty tłumaczeniowe*. Cracow: Universitas.

Mucha, Anne (2013), Temporal interpretation in Hausa, *Linguistics and Philosophy* 36, 371–415.

Narrog, Heiko (2016), The expression of non-epistemic modal categories, in J. Nuyts, and J. van der Auwera (eds), (2016), *The Oxford Handbook of Modality and Mood*, Oxford: Oxford University Press, 89–116.

Weiss, Daniel (1987), Polsko-niemieckie paralele w zakresie czasowników modalnych (na tleinnych języków zachodniosłowiańskich), in G. Hentschel, G. Ineichen, and A. Pohl (eds), (1987), *Sprach- und Kulturkontakte im Polnischen*, Munich: Sagner, 131–156.

Wiemer, Björn (2001), Aspect choice in non-declarative and modalized utterances as extensions from assertive domains (Lexical semantics, scopes, and categorial distinctions in Russian and Polish), in H. Bartels, N. Störmer, and E. Walusiak (eds), (2001), *Untersuchungen zur Morphologie und Syntax im Slavischen*, Oldenburg: BIS, 195–221.

Wiemer, Björn (2014), *Mora da* as a marker of modal meanings in Macedonian: on correlations between categorial restrictions and morphosyntactic behaviour, in W. Abraham, and E. Leiss (eds), (2014), *Modes of Modality. Modality, Typology, and Universal Grammar*, Amsterdam / Philadelphia: John Benjamins, 127–166.

Wiemer, Björn, and Ilja A. Seržant (2017), Diachrony and typology of Slavic aspect: What does morphology tell us?, in W. Bisang, and A. Malchukov (eds), *Unity and diversity in grammaticalization scenarios*, Berlin: Language Science Press, 230–307.

Corpus

Polish National Corpus:www.nkjp.pl

Noun vs. verb inflectional synthesis: A complexity trade-off?

Elena Sokur & Johanna Nichols
(National Research University Higher School of Economics, Moscow & University of California, Berkeley)

Keywords: Typology, inflectional categories, noun synthesis, noun phrase

Schedule: Sat 11.30 Room 6

Inflectional synthesis of the verb has been surveyed in a medium-sized worldwide sample (Bickel & Nichols 2005/2013, expanded in Bickel et al. 2017) and has an interesting distribution, with high synthesis peaking in the Pacific Rim population and nearby. This paper proposes a typology and questionnaire for inflectional synthesis of the noun, complementary to inflectional synthesis of the verb: a total of weighted points for the inflectional categories borne by the noun in the individual language. By inflectional categories we understand morphological categories that are sensitive to the grammatical environment in which they are expressed (Anderson 1992, Bickel & Nichols 2007). Here we report a pilot study surveying all inflectional categories, binned into the broad categories of case, number, gender, animacy, possession, predicative possession, and construct state (possession marking without indexation of possessor categories). We include categories regardless of whether their expression is affix, clitic, or isolating formative (term from Bickel & Nichols 2007). We survey phrasal as well as strictly nominal affixes. We also include definiteness, etc. (binned as "article", a term of convenience) and classifiers, all of which are nominal and affixal in some languages though mostly phrasal. Our decision is to include these in the default synthesis count but also track them separately. Wordhood and positioning in the phrase are not always accurately described in earlier grammars, but the categories are usually described, and our goal is to survey categories and formatives, not particular form classes.

We hypothesized that there would be complementarity, or a complexity trade-off, between verb inflectional synthesis and noun inflectional synthesis (both of which are measures of complexity). In fact, so far our survey of 105 languages (planned total 120+) finds the opposite. The correlation coefficient between noun and verb inflection was calculated for three macrocontinents (Africa, Eurasia, North America) and for Africa-Eurasia, all Americas, and worldwide. Results for each group and the total were similar: There is no correlation between noun and verb synthesis. They vary independently of each other, there is no trade-off, and inflectional morphological complexity varies considerably from language to language. (Local areas and individual families of course have strongly individual profiles.)

Noun synthesis levels in our sample range from 0 to 8; those in the mid part of the range are most frequent. The continental means are similar, all around 3.9; for northern Eurasia it is 3.94, for North America 3.93; this is striking in view of the strong preference for head marking in North America vs. dependent marking in Eurasia.

Inflectional categories of the noun are well described for particular languages and particular categories, and NP structure is described for some families and areas, but we still know little about the cross-linguistic distribution of noun inflection. This paper will fill the gap and present a first worldwide distributional typology of noun synthesis.

Acknowledgments
The research was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

References

Anderson, Stephen R. 1992 *A-morphous Morphology*. Cambridge: Cambridge University Press.

Bickel, Balthasar, and Johanna Nichols. 2005. Inflectional synthesis of the verb. Martin Haspelmath, Matthew Dryer, Bernard Comrie and David Gil, ed., *World Atlas of Language Structures*, 94-97. Oxford: Oxford University Press.

Bickel, Balthasar and Johanna Nichols. 2013. Inflectional synthesis of the verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Bickel, Balthasar, Johanna Nichols, Taras Zakharko, Alena Witzlack-Makarevich, Kristine Hildebrandt, Michael Riessler, Lennart Bierkandt, Fernando Zúñiga and John B. Lowe. 2017. The Autotyp database, release version 0.1.0.

Bickel, Balthasar, and Johanna Nichols. 2007. Inflectional Morphology. *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon*, 2nd ed., ed. Timothy Shopen. Cambridge University Press, 2007.

Maltese Evaluative Morphology

Michael Spagnol
(University of Malta)

Keywords: Semitic, Maltese, Morphology, Diminutive, Augmentative

Schedule: We 12.00 Room 8

This paper is concerned with morphological ways of marking evaluation in Maltese, in particular diminution (*ktieb* 'book' > *ktejjeb* 'booklet', *ċikkulat-ina* 'chocolate-DIM, small chocolate'), augmentation (*ġibj-un* 'reservoir-AUG, large reservoir', *bu-ras* 'AUG-head, large-headed'), attenuation (*isfar* 'yellow' > *safrani* 'yellowish'), intensification (*żgur-issmu* 'certain-SUP, most certain'), endearment (*żgħir-u* 'little-DIM, little one'), contempt (*sakran-azz* 'drunk-PEJ, drunkard'), and hypocorism (*Mikiel* > *Kelinu*).

In previous literature, only some of the Maltese evaluative word formation strategies mentioned above are described without going into much detail. Emphasis is usually placed on the root-and-pattern means of expressing evaluation in Maltese, paying little or no attention to affixation (cf. Sutcliffe 1936, Aquilina 1959, 1979, and Borg & Azzopardi-Alexander 1998). The topic is also occasionally discussed en passant in a typological context (cf. Bauer 1997, Grandi 2002a, 2002b, Körtvélyessy 2015, and Grandi & Körtvélyessy 2015), and recently it was the focus of psycholinguistic experiments (cf. Drake 2015, 2018).

The structure of the paper is twofold. First, I discuss templatic and concatenative evaluative constructions attested in Maltese on the basis of a database of over 1,000 evaluative forms compiled using Aquilina's (1987-1991) and Serracino-Inglott's (1975-1989) dictionaries. Secondly, the position of Maltese is considered in a wider crosslinguistic perspective, mainly within the Mediterranean area.

References

Aquilina, Joseph (1959), *The Structure of Maltese. A Study in Mixed Grammar and Vocabulary*, Malta: The Royal University of Malta.

Aquilina, Joseph (1979) *Maltese Arabic Comparative Grammar*, Malta: University of Malta.

Aquilina, Joseph (1987-1991), *Maltese-English Dictionary*, Malta.

Bauer, Laurie (1997), Evaluative morphology: in search of universals. *Studies in Language* 21, 533-577.

Drake, Shiloh (2015), The Form and Productivity of the Maltese Diminutive. Paper presented at GĦILM 5: The Fifth Annual Conference on Maltese Linguistics. Torino, Italy.

Drake, Shiloh, (2018), The Form and Productivity of the Maltese Morphological Diminutive. *Morphology*, 28, 3, 297-323.

Grandi, Nicola, (2002a), *Morfologie in contatto. Le costruzioni valutative nelle lingue del Mediterraneo*. Milano: Franco Angeli.

Grandi, Nicola, (2002b), Development and Spread of Augmentative Suffixes in the Mediterranean Area, in Paolo Ramat & Thomas Stolz (eds), *Mediterranean Languages*, Bochum: Dr. Brockmeyer University Press, 171-190.

Grandi, Nicola & Livia Körtvélyessy (eds), (2015), *Edinburgh Handbook of Evaluative Morphology*, Edinburgh: Edinburgh University Press.

Körtvélyessy, Livia, (2015), *Evaluative Morphology from a cross-linguistic perspective*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Serracino-Inglott, Erin, (1975-1989), *Il-Miklem Malti*. Malta.

Sutcliffe, Edmund F, (1936), *A Grammar of the Maltese language with chrestomathy and vocabulary*, London: Oxford University Press.

Information Structural variation in 17th Century It-Cleft Sentences: Relics and Rhetoric in Schemata

William Standing

Schedule: Fri 12.30 Room 14

While grammatical changes have primarily been examined for language as an abstract system, they are essentially the result of changes in the mental grammars of actual language users. In constructionist approaches to change, grammatical items are modelled as schematic nodes in a hierarchical and dynamic associative network, which implies inter alia that constructions and the links between them are subject to change through analogical attraction, both in terms of formal and semantic make-up, but also in terms of frequency (Hilpert, 2014). Change is necessarily a function of individual grammars: the variation and change in the grammar at large is essentially the result of substantive changes in the mental grammars of actual language users. The systematic pressures of analogy and frequency are also primarily applicable to the individual grammar: the connections and relationships these pressures represent can only be conceived of as existing in the language system of a single speaker.

Influential schematic changes have been attested at the community level, however it is still unclear to what extent similar processes are at work in language change in individuals. We tackle this key problem through analysis of individual-level change in the case of *it*-clefts, and the ramifications for related constructions, information structure and communicative function at both the aggregate and individual level. Data come from a sample of the EMMA-corpus (*Early-Modern Multiloquent Authors*), a new large-scale longitudinal corpus comprising the writings of 50 individuals from the 17th-century London-based elite, prolific over a prolonged period of time during their adult lives.

The *it*-cleft is a fixed syntactic structure primarily utilized for the unambiguous marking of focus (Lambrecht, 2001). This structural formula is, historically, mirrored by an information structural pattern, consistently consisting of new foci and given topics within the frame of the construction (Patten, 2012). Substantive changes to the *it*-cleft occur during the period in question, a widening diversity of focus type, and reorganisation of the aforementioned information structural formula, with more new topics and given foci. The fundamental changes to the information status and function of constituent elements of the construction coincides with the emergence of a number new functional types, including the rise of rhetorical and informative presuppositional uses of the cleft (Ball, 1999).
This paper assesses the potential connection between the collapse of the formulaic information structural constraints, the increasing non-compositionality and the emergence of new communicative strategies in the cleft. This increasing diversity of type and function is presumed to be linked to individuals' increasing rates of cleft usage throughout the 17th century and the establishment of a high-level cleft schema (Schmid 2015), the construction becoming dissociated from related structures at the aggregate level and in individual usage. Establishing the manner in which individuals motivate, incorporate and propagate the changing configuration of this construction is of vital importance to formulating a framework of language change.

References

Ball, C. 1994. *The Origin of the Informative Presupposition It-Cleft*. Journal of Pragmatics 22. 603-628.

Hilpert, M. 2014. *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.

Lambrecht, K. 2001. *A Framework for the Analysis of Cleft Constructions*. Linguistics 39.3. 463-516.

Patten, A. 2012. *The English It-Cleft: A Constructional Account and Diachronic Investigation.* Berlin: De Gruyter.

Schmid, H. 2015. *A Blueprint of the Entrenchment-and-Conventionalization Model*. Yearbook of the German Cognitive Linguistics Association 3. 1-27.

Aktionsart of non-agentive perception verbs in Polish

Przemysław Staniewski & Adam Gołębiowski
(University of Wrocław)

Keywords: aspect, *Aktionsart*, perception, Polish, semantics-syntax interfaces

Schedule: We 16.30 Room 1

In our paper, we analyse the impact of aspectual morphology and semantics of the complements on the *Aktionsart* of non-agentive perception verbs in Polish. We also propose solutions regarding the *Aktionsart* of the source-based verbs since this issue has not been dealt with in depth so far. Depending

on the context and perspective (source vs. experiencer) they can be analysed, e.g. as states, activities or inchoatives.

Based on broadly accepted Viberg's (1984, 2001) classification of perception verbs and employing Croft's (2012) model for analysis of verbal aspect and event structure (regarding aspect especially in Slavic languages see Breu 1997, 2000, 2007, Lehmann 2005, 2009, 2010, Łaziński & Wiemer 1997, Łaziński 2011 among many others) we focus in our analysis on (a) experiencer-based verbs which take the experiencer as grammatical subject, (e.g. *Jane saw the picture*) and (b) source-based verbs which occur with source of sensation in the subject position (e.g. *Jane looked nice*). Although there is a vast literature concerning the semantics and syntax of perception verbs in various languages the issue of the *Aktionsart* of the source-based verbs still remains unexplored. The *Aktionsart* of the experiencer-based verbs, on the other hand, has not been fully addressed yet. They are viewed in the literature as states and/or (inchoative) achievements (Rogers 1974; Viberg 1984, 2001). Compare the following examples:

(1) Piotr słysz-ał-ø muzyk-ę.
Piotr.NOM.SG hear-ed-ø.3SG.PST.M music-ACC.SG
'Peter heard the music.' – state (durative; imperfective)

(2) Piotr u-słysz-ał-ø muzyk-ę.
Piotr.NOM.SG PFV-hear-ed-ø.3SG.PST.M music-ACC.SG
'Peter heard the music.' – inchoative achievement (punctual; perfective)

(3) Piotr u-słysz-ał-ø strzał-ø.
Piotr.NOM.SG PFV-hear-ed-ø.3SG.PST.M shot-ø.ACC.SG
'Peter heard a shot.' – achievement (punctual; perfective)

(4) Piotr słysz-ał-ø strzał-ø.
Piotr.NOM.SG hear-ed-ø.3SG.PST.M shot-ø.ACC.SG
'Peter heard a shot.' – achievement (punctual; imperfective)

(5) Piotr czuł-ø zapach-ø kwiat-ów.
Piotr.NOM.SG feel-ø.3SG. PST.M smell-ø.ACC.SG flowers-GEN.PL
'Peter smelled the flowers.' – state (durative; imperfective)

(6) Piotr po-czuł-ø zapach-ø
Piotr.NOM.SG PFV-feel-ø.3SG. PST.M smell-ø.ACC.SG
kwiat-ów.
flowers-GEN.PL
'Peter smelled the flowers.' – inchoative achievement (punctual; perfective)

Based on the introspective and corpus data, we first argue that the *Aktionsart* in the domain of non-agentive perception verbs can be analysed in a more fine-grained fashion than assumed by Rogers (1974) and Viberg (1984, 2001); see examples (1) – (6). Secondly, we demonstrate to what extent the intra-linguistic factors such as grammatical aspect in Polish and semantics of the complements affect the *Aktionsart* as shown e.g. in sentences (1), (2) and (3). Thirdly, we claim that the *Aktionsart* can be constrained by extra-linguistic factors, i.e. by the nature of the perceptual modality. In this respect compare the difference between the experiencer-based verbs in the auditory and olfactory domains. While the auditory domain allows for four semantic variants which change due to the verbal aspect and the lexical semantics of embedded arguments ('music' vs. 'shot') the possibilities in the olfactory domain are limited to state (5) and inchoative achievement (6). "Pure" achievement reading is in our view excluded in the olfactory domain due to the fact that we perceive a smell as long as it dissolves in the air or we get adapted to it which implies that the sensation lasts for some time.

References

Breu, Walter (1998), Komplexe aktionale Verbklassen, insbesondere Inchoativa, in T. Berger and J. Raecke (eds), (1998), *Slavistische Linguistik 1997*, München: Sagner, 55-80.

Breu, Walter (2000), Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und lexikalische Interaktion), in: W. Breu (ed), (2000), *Probleme der Interaktion von Lexik und Aspekt*, Tübingen: Niemeyer, 21-54.

Breu, Walter (2007), Der Verbalaspekt im Spannungsfeld zwischen Grammatik und Lexik, *Sprachwissenschaft* 32(2), 123-166.

Croft, William (2012), *Verbs. Aspect and Causal Structure*, Oxford: Oxford University Press.

Łaziński Marek (2011), Polish aspectual prefixes, their order and functions: A study based on the National Corpus of Polish, *Word Structure* 4(2), 231-243.

Łaziński, Marek and Björn Wiemer (1997), Terminatywność jako kategoria stopniowalna, *Prace Filologiczne* 15, 99-126.

Lehmann, Volkmar (2005), Zur Beschreibung von Aspektfunktionen: Beispiel Sachverhalts-koinzidenz, in S. Kempgen (ed), (2005), *Slavistische Linguistik 2003*. München: Sagner, 145-165.

Lehmann, Volkmar (2009): Aspekt und Tempus, in: T. Berger, K. Gutschmidt, S. Kempgen and P. Kosta (eds) (2009), *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*, vol. 1, Berlin / Boston: De Gruyter Mouton, 526-556.

Lehmann, Volkmar (2010), Der Aspekt in den slavischen Sprachen im Licht der kognitiven Linguistik, in T. Anstatt and B. Norman (eds), (2010), *Die slavischen Sprachen im Licht der kognitiven Linguistik*, Wiesbaden: Harrassowitz, 77-99.

Rogers, Andy (1974), *Physical Perception Verbs in English: A Study in Lexical Relatedness*, Ph.D. dissertation, University of California, Los Angeles.

Viberg, Åke (1984), The Verbs of Perception: A Typological Study, *Linguistics* 21(1), 123-162.

Viberg, Åke (2001), Verbs of Perception, in M. Haspelmath, E. König, W. Oesterreicher and W. Raible (eds), (2001), *Language Typology and Language Universals: An International Handbook*, Berlin: de Gruyter, 1294–1309.

On the contact-induced emergence of grammatical gender

Thomas Stolz & Nataliya Levkovych
(University of Bremen)

Keywords: grammatical gender, language contact, agreement, human nouns

Schedule: Thu 12.00 Room 2

There is as yet no comprehensive account of the behavior of grammar gender (GG) under the conditions of language contact. A particularly intriguing question is whether an originally GG-less language can acquire GG via language contact. Field (2002: 192) counts the genesis of GG among the cross-linguistic rarities whereas Matras (2009: 174) argues that "[g]ender may also be introduced into a language along with borrowed forms." Gardani (2012: 77) even considers GG to be relatively easy to copy. GG as defined by Corbett (1991) manifests itself in agreement of noun (= controller) and attribute (= target). In Thomason's (2001: 71) borrowing scale the copying of agreement patterns is possible only in prolonged and intensive language contact situations. Aikhenvald (2000: 388) assumes that "[b]orrowing of an agreement system is extremely rare" but mentions two cases herself

(Ayacucho Quechua and Ilokano). Stolz (2012: 94–104) discusses (potential) evidence of examples of GG-less languages copying GG from their partner in a given contact situation. The data are often inconclusive since the illustration of the phenomenon is restricted to isolated words or NPs so that it is not always possible to rule out codeswitching.

What is needed is a cross-linguistic inventory of all those cases which (presumably) attest to the genesis of GG in contact situations involving a donor language with GG and a replica language without GG. Especially telling are instances like (1)–(4). In these sentences, feminine GG is marked overtly on the target (bold): adjectives in (1)–(3) and definite article in (4). The controller (underlined: pronoun in (1), nouns in (2)–(4)) is neither a copy nor marked for gender. For Karaim (Csató 2001: 18), Yucatec Maya, Tetun Dili (Hajek 2006: 170–171), and (Correntinean) Guaraní alike, these feminine forms contrast with (originally) masculine forms to yield a binary paradigm.

(1) Karaim (Éva Csató p.c.) [donor: Russian]
Ol *e-d'i* *inteligentn-**a**.*
3SG COP-PAST.3SG intelligent-**F**
'She was intelligent.'

(2) Yucatec Maya (Chamoreau 2012: 84) [donor: Spanish]
*Bek'ech-it-**a*** *u* *y-íits'in.*
thin-DIM-**F** A3SG POS-younger_sister
'His/Her younger sister is slim.'

(3) Tetun Dili (Bible Tetun Dili, Hahuu / Jénesis 12:14) [donor: Portuguese]

ema	*sira*	*haree*	*feto*	*nee*	*bonit-**a***	*lahalimar.*
person	PL	see	woman	DEM.PROX	pretty-**F**	really

'[…] the people saw this really pretty woman'.

(4) (Correntinean) Guaraní (Cerno 2010: 26) [donor: Spanish]

Ani	*na*	*re-'u*	*mamíta*	***la***	*so'o!*
NEG.IMP	PAR	2-eat	mammy	**DEF.F**	meat

'Don't eat the meat, mammy!'

In the talk further evidence from e.g. (Lekeitio) Basque, (Erzya) Mordvin, Chamorro, etc. is analyzed. Three questions are paid special attention:

(a) Is GG-agreement possible without the participation of lexical copies in a given construction (i.e. do copied gender-markers attach to autochthonous stems)?
(b) Is GG-agreement copied only if massive borrowing of nouns and adjectives occurs?
(c) Does GG-agreement start with human nouns?

The talk marks the point of departure of a project which investigates the fate of GG in language contact situations. It contributes not only to the research program dedicated to GG but also to language contact studies in general.

Abbreviations: A = a-set, COP = copula, DEF = definite (article), DEM = demonstrative, DIM = diminutive, F = feminine, IMP = imperative, NEG = negative, PAR = modal particle, PL = plural, POS = possessive, PROX = proximal, SG = singular

References:

Aikhenvald, Alexandra Y. (2000), *Classifiers. A typology of noun categorization devices*. Oxford: Oxford University Press.

Bible – Tetun Dili Bible, *Genesis e o Novo Testamento em língua tétum Dili* [Genesis and the New testament in Tetum Dili]. Wycliffe Bible translations. https://bible.cloud/ebooks/pdf/TDTUBB/TDTUBB.pdf (accessed 11 June 2018).

Cerno, Leonardo (2010), Spanish Articles in Correntinean Guaraní. A Comparison with Paraguayan Guaraní, *STUF – Language Typology and Universals* 63(1), 20–38.
Chamoreau, Claudine (2012), Spanish diminutive markers -ito/-ita in Mesoamerican languages. A challenge for acceptance of gender distinction, in Martine Vanhove et al. (eds.), *Morphologies in contact.* Berlin: Akademie, 71–90.
Corbett, Greville G. (1991), *Gender*. Cambridge, New York: Cambridge University Press.
Csató, Éva (2001), Karaim, in Thomas Stolz (ed.), *Minor languages of Europe. A series of lectures at the University of Bremen, April–July 2000*. Bochum: Brockmeyer, 1–24.
Field, Fredric W. (2002), *Linguistic Borrowing in bilingual contexts*. Amsterdam, Philadelphia: John Benjamins.
Gardani, Francesco (2012), Plural across infection and derivation, fusion and agglutination, in Lars Johanson & Martine Robbeets (eds.), *Copies versus Cognates in Bound Morphology*. Leiden, Boston: Brill, 71–98.
Hajek, John (2006), Language contact and convergence in East Timor: The case of Tetun Dili, in Alexandra Y. Aikhenvald & R.M.W. Dixon (eds.), *Grammars in contact*. Oxford: Oxford University press, 163–178.
Matras, Yaron (2009), *Language contact*. Cambridge: Cambridge University Press.
Stolz, Thomas (2012), Survival in a niche. On gender-copy in Chamorro (and sundry languages), in Martine Vanhove et al. (eds.), *Morphologies in Contact*. Berlin: Akademie, 93–140.
Thomason, Sarah G. (2001), *Language contact. An introduction.* Washington/DC: Georgetown University Press.

Reassessing morphological autonomy

Adam J.R. Tallman and Sandra Auderset

Schedule: Fri 16.00 Room 6

<pdf>

English placeholders: Their role in social interaction

Jarmila Tárnyiková
(Palacky University Olomouc)

Keywords: placeholders, social interaction, corpus data, context retrievable functions, social spaces

Schedule: We 11.30 Room 2

There are various communicative situations in which we are intentionally vague. The studies on vagueness (Channell 1994, Cuttings 2007, and Amiridze et al. 2010) confirm that languages offer a range of devices for accomplishing the intended goal. In this paper the focus is on relatively peripheral, yet communicatively relevant means of vague language, i.e. *placeholders* (PHs), with restriction to Noun PHs, such as *Mr Thingy*, *John Whatsisname*, *whatchamacallit* or *whatsit*, their

forms, functions and distribution in British and American English. Within the theoretical framework of a functional and systemic grammar, the PHs are approached as systemic parts of vague language network, as pro-forms referring to yet-to-be-specified referents, delayed due to word-formulating difficulties, which are caused by temporarily forgotten, difficult-to-pronounce, or deliberately withhold naming units. Many of the PHs are of uncertain spelling because of their main distribution in spoken discourse (Enfield 2003).

Using the BNC and the COCA data, I present the results of quantitative and qualitative analyses, assuming that irrespective of the genetic relationship between the two variants of English, there are differences in context retrievable strategies typical of different 'social spaces' (Evison et al. 2007).

As a pre-requisite to the analytical part, I first enumerate the diversity of terms used to reflect various roles of PHs in interaction, then advocate the preference of the term *placeholder* over such labels as *tongue-tippers*, or *mouth-fillers*, to finally survey the already identified roles of PHs (Amiridze et al. 2010) and compare them with my own findings.

In the analytical part, two types of relations will be activated to taxonomize the results: the paradigmatic relation of alternations (*Thingy/Whatsisname/So and so*), and the syntagmatic relation of co-occurrence. These will be used to project the PHs into the surrounding contexts in order to verify the following research tasks: Do the PHs represent a close set or are they open to innovations? Are the corpus data sufficient for grasping the spectrum of strategies underlying PHs use? Are there significant differences between the British and American usage?

The results confirm three basic strategies emergent from the data, all associated with facework: (i) the H(earer) is a recipient of a 'promise' (signalled by the PH) that after the word-formulating problems the PH will be substituted by a proper referent and the utterance will continue in its projection (*Have you seen the whatchamacalit, the suspension bridge?*); (ii) the S(peaker) uses the PH to intentionally suppress the proper referent leaving its identification to the H (*Take the whatsit. – What whatsit, you mean the virginity testing procedure?*) ; (iii) the S uses the PH to acknowledge due-ness and gives the H explicit signals to paticipate in identifying the referent (*Tom whatsisname, you know him. – Tom Abercrombie? – Him.*).

The analysis has proved that placeholders are integrated parts of our talking habits, endowed with a whole spectrum of context retrievable functions, and that the British-American interface reveals comparable functions but differences in the spectrum and hierarchy of PHs manifestations.

References

Amiridze, Nino, Boyd H. Davis and Margaret Maclagan (eds), (2010), *Fillers, Pauses and Placeholders. Typological Studies in Language*, vol. 93, Amsterdam / Philadelphia: John Benjamins.

Amiridze, Nino (2010), Placeholder Verbs in Modern Georgia, in N. Amiridze et al (eds), (2010), 75-102.

Channell, Joanna (1994), *Vague Language*, Oxford: Oxford University Press.

Cutting, Joan (2007), Introduction to vague language explored, in J. Cutting (ed), (2007), *Vague language explored*, Basingstoke: Palgrave Macmillan, 3-20.

Enfield, Nicolas James (2003), The definition of WHAT-d'you-call-it: Semantics and pragmatics of recognition deixis, *Journal of Pragmatics* 35, 101-117.

Evison, J., M. McCarthy, and A. O'Keeffe, (2007), Looking out for love and all the rest of it: Vague category markers as shared social space, in J. Cutting (ed.), (2007), *Vague language explored*, Basingstoke: Palgrave Macmillan,138-157.

The Thirteen (or More) Reflexive Pronouns in Khwarshi

Yakov Testelets

Schedule: Sat 12.30 Room 12

<pdf>

Questioning Samuels's Incipient Standards

Jacob Thaisen

Schedule: Thu 10.00 Room 2

<pdf>

Diversity and integration of Directional Verb Constructions in Mandarin Chinese

Mimi Tian
(Johannes Gutenberg-Universität Mainz)

Keywords: motion, directional verb, word formation, deixis, intonation

Schedule: We 17.00 Room 7

This study is a work in progress (Tian in prep.) on three-part Directional Verb Expressions (DVEs) in Mandarin Chinese. As in (1), three-part DVEs contain a verbal expression, followed by a directional expression, followed by a deictic expression. Such DVEs are a well-known phenomenon in many languages (e.g. Bisang 1992) but not yet fully understood in all their aspects. Basic questions, such as whether they are phrases or compound words and which part is the head, are still in dispute. This research explores the wide spectrum of diversity in DVEs with a corpus study, a questionnaire survey and audio recordings, and it illustrates how different levels of the grammatical system are integrated in the analyses.

(1)	*pǎo*	*jin*	*lai*
	Run	enter/in	towards-the-sphere-of-interest-of-the-speaker

Hundreds of verbs can appear as the first part. The second and third part can occur in full tones or in the neutral tone. Less obviously, there is a variety of relations between the first part and the second and third parts. Discussion around whether Mandarin is verb-framed or satellite-framed in Talmy's (2000) typology or even equipollently-framed (Slobin 2004) typically labels them as Manner and Path (e.g. Chen & Guo 2009), but there are more possibilities, as in Table 1.

role of the first part	role of the directional
• manner ***pǎo*** *shang qu* "**run** up" • state after motion	• path *pǎo* ***shang qu*** "run **up**" • endpoint of motion

zuò *guo lai* "**sit** (after coming) here" • cause ***shuō*** *jin qu* "**speak** and cause to go in" • purpose ***duǒ*** *guo lai* "**hide** by moving here"	*luò* ***shang lai*** "fall **on**" • result *wéi* ***jin lai*** "surround (it) so that (it) is **in**" • direction *kàn* ***guo lai*** "look **in this direction**"
scope of the directional	temporal relation
• $event_v$ only *kàn guo lai* "look in this direction" • participant only *shuō jin qu* "speak and cause to go in" • $event_v$ + participant *tuī guo lai* "push (it) here"	• simultaneous *pǎo shang qu* "run up" • $event_v$ before $event_d$ *zhāi hui lai* "pick up and bring back" • $event_v$ after $event_d$ *zuò guo lai* "sit (after coming) here"

Table 1

The general label of DVE is a cover notion for several different phenomena. The data are examined using the framework of Integrational Linguistics (Lieb 1983, 2013). Cases of word formation are distinguished from syntactic complementation and modification: (a) the directional is a verbal complement of the first part, e.g. *shuō jìn qu* "speak and cause to go in"; (b) the first part is a modifier of the directional verb, e.g. *chǎo guò lai* "come here quarrelling"; (c) the three parts form compound words in which different specific relations hold between the first part and the directional part, e.g. *dǎo xia lai* "fall down". While some first parts may licence only one structure, example (1) has more than one analysis, which may explain the dispute about its word-status and head-status in the literature. The analysis of an expression containing (1) depends on the occurrence and position of aspect markers and other constituents, and the intonation structure of the unit.

References

Bisang, Walter (1992), *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer. Vergleichende Grammatik im Rahmen der Verbserialisierung, der Grammatikalisierung und der Attraktorpositionen* [The verb in Chinese, Hmong, Vietnamese, Thai and Khmer. Comparative grammar concerning verb serialization, grammaticalization and attractor positions], Tübingen: Gunter Narr.

Chen, Liang, & Guo, Jiansheng (2009), Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of pragmatics*, 41(9), 1749-1766.

Lieb, Hans-Heinrich (1983), Integrational Linguistics. Vol. I.: General Ofutline. Amsterdam; Philadelphia: Benjamins. (= Current Issues in Linguistic Theory 17).

Lieb, Hans-Heinrich (2013), Towards a general theory of word formation: The process model.

Slobin, Dan I. (2004), The many ways to search for a frog: linguistic typology and the expression of motion events, in Strömqvist, S., Verhoeven, L. (eds), (2004), *Relating Events in Narrative: Typological and Contextual Perspectives*. Lawrence Erlbaum Associates, Mahwah, NJ. 219–257.

Talmy, Leonard (2000), *Toward a Cognitive Semantics*. Cambridge, Mass.: MIT Press.

Tian, Mimi (in prep), *Directional Verb Expressions in Mandarin Chinese: a tonal phonological, morphological, syntactic, lexical-semantic and syntactic-semantic analysis*. PhD thesis at University of Mainz.

Microdiachronic Evolution of Negative Concord: The Case of the Russian *ni*

Daniel Tiskin
(Saint Petersburg State University)

Keywords: Negative Concord, Russian, scalar particles, conjunction, negation

Schedule: We 11.30 Room 14

The Russian negative concord morpheme *ni-* has a range of uses, including prefixal, particle- and conjunction-like. The latter's distribution is notoriously complicated, combining the functions divided between *ni* and *niti* in Serbo-Croatian (Gajić 2018). For instance, it can scope above the negation *ne* (1) and be used without negation when it conjoins VPs (2) or nominal predicates (3) with no material preposed to the head (Tiskin 2018).

There are different accounts of the distribution of *ni*: e.g. Abels (2005) takes *ni*-items to be PPIs licensed in Spec,NegP, whereas Zeijlstra (2004) and Gajić assume agreement with a null Neg higher in the structure. Assuming that a successful analysis should envisage the possibility that the data involves (possibly ongoing) diachronic change, the present paper investigates the recent developments in the grammar of *ni-* using data from the Russian National Corpus (~1700—present, over 280 M tokens).

As for **conjunctive** uses of *ni… ni*, the proportion of negation-free predicate conjunction significantly decreases over time, for verbal (2) and nominal (3) predication alike. The proportion of *ni… ni* outscoping negation (1) decreases as well. This presents *ni… ni* as turning into an element licensed relatively low in the syntax, although the process is not complete yet: (2) and (3) are still acceptable.

Another piece of evidence comes from **particle** uses. Formerly, *ni* could serve as an *even*-type scalar particle associable with any salient lowest point on a scale, e.g. 'five minutes' in (4); this use is no longer available (except for lexically specified lowest points, e.g. *ni* (*odnoj*) *minuty* 'not a (single) minute', lit. 'NI one minute'; **prefixal** *ni-* is also etymologically scalar, Haspelmath 1997). Nowadays, (4) would require *i* instead of *ni* (cf. Chen 2018: ch. 4 for a similar alternation in Mandarin, *dou*/*ye*). Assuming the movement approach to 'even' under negation (Erlewine 2018 notwithstanding), we conclude that *ni* has lost its ability to be licensed as high as in (4).

Taken together, the data suggest that throughout the last centuries of the history of Russian, *ni-*elements have drifted towards low licensee status, although this process is not complete and has different pace in different fragments of grammar. Further work will have both to clarify the exact licensor position and to deal with "mixed" cases such as those where one conjunct lacks preposed material, as in (2), but the other has it and must therefore contain *ne*.

(1) Ni ja ničego ne znaju, ni ty ničego ne znaeš'.
NI I.NOM nothing.ACC NEG know.PRS.1SG NI thou.NOM nothing.ACC NEG know.PRS.2SG
'Neither I know anything nor do you'

(2) Ja ni zapreščaju, ni razrešaju…
I.NOM NI prohibit.PRS.1SG NI permit.PRS.1SG
'I neither prohibit nor permit'

(3) Soboju on byl ni duren, ni xoroš.
self.INS he.NOM be.PST.SG.M NI ugly.BREV.SG.M NI handsome.BREV.SG.M
'He looked neither ugly nor handsome'

(4) Divizija же Golovina ne ustojala **ni** pjati minut…
division.NOM.SG PRT Golovin.GEN.SG NEG stand.PST.SG.F NI five.GEN minute.GEN.PL
'As for Golovin's division, it did not stand even for five minutes'
(all examples: RNC)

References

Abels, Klaus (2005): "Expletive negation" in Russian: A conspiracy theory. *Journal of Slavic Linguistics* 13(1), 5–74.

Chen, I-Hsuan (2018): *Diachronic Changes Underlying Synchronic Distribution Scalar Inferences and Word Order*. Singapore: Springer.

Erlewine, Michael Yoshitaka (2018): *Even* doesn't move but associates into traces: A reply to Nakanishi 2012. *Natural Language Semantics* 26(3–4), 167–191.

Gajić, Jovana (2018): *Niti*—real negative coordination inside a strict Negative Concord language? *Proceedings of ConSOLE* XXVI, 1–21.

Haspelmath, Martin (1997): *Indefinite Pronouns.* OUP.

Tiskin, Daniel (2018): K obobščeniju semantiki russkogo *ni… ni* [Towards a unified semantics for the Russian coordinator *ni… ni*]. In: *The 15th Conference on Typology and Grammar for Young Scholars*, 167–170. St. Petersburg: ILS RAS.

Zeijlstra, Hedde (2004): *Sentential Negation and Negative Concord.* PhD dissertation, University of Amsterdam.

Dedicated possessive reflexives in languages with head marking of a possessor

Svetlana Toldova

Schedule: Thu 12.00 Room 12

This paper is devoted to a special type of reflexive pronouns, namely, dedicated possessive reflexives. According to Reuland (2011) and Despić (2011, 2015), these pronouns are available only in languages without preposed definite articles. There is a special class of languages without articles where there are at least two modes of bound possessor encoding. These are languages with overt morphological marking of possessor on the head of DP/NP. In these languages bound possessor can be expressed with possessive reflexives or via possessive affixes (1):

Moksha

(1) *er'* *loman'-s'* *kel'ksi* ***(es')*** *d'ɛd'ɛ-**nc***
every man_i-NOM.SG.DEF loves(PRAES.3SG) $self_i$ mother-ACC.3SG.$POSS_I$
'Every man loves his (own) mother'

3.POSS is triggered by overt possessor, the Subject bounds the possessor, the latter triggers the possessive marker on the noun, the index on 3.POSS is coindexed with the possessor NP/DP .

I use the data of some Finno-Ugric (Mari, Komi, Udmurt, Moksha (Finno-Ugric)) and Altaic languages (Bashkir, Tatar, Khakas, Kazakh (Turkic), Nanai, Ulchi, Udihe, Uilta (Tungusic). The work is based primarily on the fieldwork elicitation data as well as corpus data.

I claim that the languages under consideration have a bulk of properties in common as far as encoding reflexivity is concerned.

1. These languages have bare reflexive stems (lacking φ-features) as possessive reflexives. They are used in possessive constructions and postposition constructions;
2. These stems are the source for φ-featured compound reflexive pronouns (inflected in case and number); the compound pronouns can be reduplicated forms: bare form + inflected form; they are bound in co-argument binding domain, (2).

Hill Mari

(2) *tədə* ***[škəm-žə-m*** ***ške]*** *pəčk-ən*
that REFL-POSS.3SG-ACC REFL cut-PF.3SG
'He cut himself.'

3. the main claim is that although possessive reflexives manifest syntactic features of reflexives (they can be bound in their local domain, see Chomsky 1981), their primary function in the investigated languages is **emphatic**. They are associated with focus as markers of strong/inalienable possession (cf. English *own*, Spathas 2014).
The claim is based on the following observations.

(i). It is possible to express bound anaphora without these pronouns just using a possessive marker as in (1) or (3).
Uilta

(3) *Peta (mənə) iŋda-va-i ŋələ-či-ni.*
Peter_i own.refl_i dog-ACC-REFL.POSS_i frighten-PAST-3SG
'Peter is afraid of his (own) dog'

In (1a) it is a 3rd sg. marker. It can be bound by the Subject of the clause or can function as pronominal non-bound possessive marker. Tungusic languages, unlike Finno-Ugric ones, distinguish "pronominal"-like markers and dedicated reflexive possessive markers. Despite the difference, in both language types possessive reflexives are optional in the majority of contexts.

(ii) The possessive reflexives trigger an obligatory possessive marker on the head (*mənə iŋda-va 'own.refl dog-acc').

Taking into consideration (i) and (ii), one can make a conclusion that these pronouns are redundant in constructions with bound possessor.
(iii) They can function as intensifiers (see König 2000):
Uilta

(3) *tari-dda* *čī* *asimuna* *mānə* *əudu-ɣə-či* *nā-takk-ēri* .
that-ENCL so spouse self descend-PREAT-3PL earth-ALL-PL.REFL
lit. 'At once the husband and the wife went down themselves to their land'

To sum up, the investigated data show that the referent of the possessor is encoded primarily via possessive affixes (pronominal or reflexive), the usage of a dedicated pronoun is optional. It is used as an emphatic mean or as a focus of empathy holder.

References

Chomsky, Naom (1981), Lectures on government and binding, Dordrecht: Foris.
Despic, Miloje. 2011. Syntax in the Absence of Determiner Phrase. Doctoral dissertation. University of Connecticut.
Despic, Miloje. 2015. Phases, Reflexives, and Definiteness. Syntax. doi: 10.1111/synt.12031
Despić (2011, 2015)

König, Ekkehard, and Siemund, Peter (2000), Intensifiers and reflexives: a typological perspective, in Z. Frajzyngier, and T. Curl (eds.), (2000), Reciprocals: Forms and Functions, Amsterdam: John Benjamins Publishing Company, 41-74.
Reinhart, Tanya (1983), Anaphora and Semantic Interpretation, London: Croom Helm.
Reuland, Eric. 2011. Anaphora and Language Design. Cambridge, MA: MIT Press.
Spathas, G. (2014, August). Disentangling own: evidence from association with focus. In Semantics and Linguistic Theory(Vol. 24, pp. 271-289).

Demonstrative proadjectives in spoken Livonian: Morphosyntactic and semantic use

Marili Tomingas
(University of Tartu)

Keywords: Livonian, demonstratives, proadjectives, morphosyntax, semantics

Schedule: Thu 12.00 Room 7

Demonstrative proadjectives (e.g. Livonian *seļļi* 'such, like this', *tūoļi* 'like that') are a type of demonstrative words that has not been that much researched compared to the other demonstratives like demonstrative pronouns (e.g. Livonian *se* 'this') and demonstrative proadverbs (e.g. Livonian *täsā* 'here'). The reason can be that demonstrative proadjectives often tend to occur as attributes for other substantives and adjectives, so it can be more difficult to research proadjectives independently. In addition, demonstrative proadjectives are semantically very dependent on the text around them, so they can have different semantic functions depending on the context. Lately there have been more researches on demonstrative words, their use and meaning (e.g. Keevallik 2011, König & Umbach 2016), but morphosyntactic and semantic use of demonstrative proadjectives could be still researched more precisely to describe their diverse meanings, different forms and positions in sentences. An example of a sentence with a demonstrative proadjective that is non-attributive, but its meaning depends on following subordinate clause:

(01)

Ažā	*u'm*	***seļļi***,	*ku*	*ma*	*mūpõ*
thing	be.3SG	**such**	that	I	tomorrow
ä'b lī	*kuo'nnõ.* (LELS)				
NEG.1SG	will_be	at_home			

'The thing is like this that I am not going to be at home tomorrow.'

Demonstrative proadjectives are widely used in Finnic languages. Most of the Finnic languages have several different proadjectives (in Karelian there are even up to 20 different proadjectives), often distinguished by spatial oppositions (Larjavaara 1986). In Livonian there are two different proadjectives *seļļi* and *tūoļi* and in addition one demonstrative proadjective compound *seļļi-tūoļi* ('like this, like that') combining them both, showing different types of characteristics occurring at the same time. Livonian as an endangered Southern Finnic language which also has a literary language is an interesting and important source for research, as there are recordings of spoken Livonian where demonstrative words are used more often than in the literary language and their different uses and meanings could be spotted.

This research paper is based on the material of recordings in Archives of Estonian Dialects and Kindred Languages of University of Tartu (TÜ EMSA) with native Livonian speakers. For the analysis, 30 recordings with four different Livonian native speakers are used, recordings are dialogues between researchers and native Livonians which were recorded in 1986–2012. The research focuses on the question which demonstrative proadjectives and their forms are used in spoken Livonian the most. After that, the morphosyntactic and semantic use of demonstrative proadjectives is analyzed. Morphosyntactic analysis focuses on attributiveness or non-attributiveness of demonstrative proadjectives, their syntactic position and morphological use in the sentence. Semantic analysis describes semantic functions of demonstrative proadjectives in different types of sentences. For example, demonstrative proadjective could have a physical pointing function, a referring function to a previous or a following part of the text, a function referring to a well-known type of qualities and a function referring to a general type of qualities, depending on the structure of a sentence. The results of morphosyntactic and semantic analysis are presented descriptively.

References

Keevallik, Leelo 2011. Pro-forms as projective devices in interaction. *Discourse Processes* 48 (6), 404–431.

König, Ekkehard & Carla Umbach 2016. Demonstratives of manner, of quality and of degree: A neglected subclass. In E. Boef, M. Coniglio, E. Schlachter & T. Veenstra (eds.). *Demonstratives: syntax, semantics and typology*. Berlin: De Gruyter Mouton.

Larjavaara, Matti (1986) *Itämerensuomen demonstratiivit. Karjala, aunus, lyydi ja vepsä*. Mikkeli: Suomalaisen Kirjallisuuden Seura.

LELS = Viitso, Tiit-Rein and Valts Ernštreits (2012) *Līvõkīel-ēstikīel-leļkīel sõnārõntõz. Liivi-eesti-läti sõnaraamat. Lībiešu-igauņu-latviešu vārdnīca*. Tartu and Rīga: Tartu Ülikool and Latviešu valodas aģentūra. Available online at <http://www.murre.ut.ee/liivi/>. Accessed on 20.05.2019.

TÜ EMSA = Tartu Ülikooli eesti murrete ja sugulaskeelte arhiiv. Available online at <http://www.murre.ut.ee/arhiiv/>. Accessed on 20.05.2019.

The semantics of the left periphery in German V1/V2 interrogatives and V2 declaratives

Antonios Tsiknakis
(University of Wuppertal)

Keywords: verb second, topicalization, sentence mood, semantics/pragmatics interface

Schedule: Thu 15.00 Room 11

The talk deals with the contribution of finite verb movement and topicalization of a non-relative [–wh]-phrase to the meaning *(sentence mood)* of the German clause types exemplified in (1):

(1) a. Küsst$_1$ Maria den Postboten t$_1$? *(verb first y/n-interrogative)*
kisses Maria the postman
'Does Maria kiss the postman?'

b. Wen küsst$_1$ Maria t$_1$? *(verb second wh-interrogative)*
who.ACC kisses Maria
'Who does Maria kiss?'

c. [Den Postboten]$_2$ küsst$_1$ Maria t_2 t_1. *(verb second declarative)*
the postman.ACC kisses Maria
'Maria kisses the postman.'

The first part of the talk problematizes two competing approaches to sentence mood: firstly, the traditional approach according to which sentence moods include an illocutionary predicate encoding the speaker's attitude towards the descriptive content of the clause (Altmann 1993) and secondly, the approach presented in Portner (2005) according to which the illocutionary attitude of the speaker is derived from the semantic type of the descriptive content.

Against this background, the second part of the talk outlines a theory about finite verb movement and topicalization in German which stands in the tradition of previous works on the verb second phenomenon like Lohnstein (2000), Gärtner (2002), and Truckenbrodt (2006). The outlined theory is based on a comparison of the use potential of the clause types in (1) with the use potential of their corresponding clause types in (2) where neither verb movement nor topicalization takes place (Tsiknakis 2016, 2017, to appear):

(2) a. ob Maria den Postboten küsst *(verb final y/n-interrogative)*
whether Maria the postman kisses
'whether Maria kisses the postman'
b. wen Maria küsst *(verb final y/n-interrogative)*
who.ACC Maria kisses
'who Maria kisses'
c. dass Maria den Postboten küsst *(verb final that-clause)*
that Maria the postman kisses
'that Maria kisses the postman'

It will be proposed that finite verb movement is triggered by an interpretable clause type feature [F1] and topicalization by a feature [F2] which modifies the interpreation of [F1]. The interpretation of the proposed features will be modeled in a framework of dynamic semantics which represents the common information of the discourse participants in *(embedded) context sets* (Stalnaker 1978, Farkas 2003, Roberts 2012). I will argue that that [F1] restricts the use of clause to environments which in prototypical cases involve an update of a doxastic context set and that [F2] regulates whether the update operation is supposed to be completed directly in the context which results through the anchoring of the clause in a given context or inderectly in a contex which follows the output context of the anchoring.

Maintaining a balanced position between Altmann's maximalist and Portner's minimalist approach to sentence mood, the presented theory illustrates how morphosyntactic characteristics in the left periphery of German clauses can be interpreted at the semantics/pragmatics interface in a compositional manner.

References

Altmann, Hans (1993), Satzmodus, in J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*, Berlin / New York: de Gruyter, 1006-1029.

Farkas, Donka (2003), *Assertion, Belief and Mood Choice*, Paper presented at ESSLLI, Conditional and Unconditional Modality Workshop, Vienna, url: http://people. ucsc.edu/~farkas/papers/mood.pdf (visited on 31/10/2016).

Gärtner, Hans-Martin (2002), On the force of V2 declaratives, in *Theoretical Linguistics* 28, 33-42.

Lohnstein, Horst (2000), *Satzmodus – kompositionell. Zur Parametrisierung der Modus-phrase im Deutschen*, Berlin: Akademie-Verlag.

Portner, Paul (2005), The Semantics of Imperatives within a Theory of Clause Types, in K. Watanabe, and R. Young (eds), *Proceedings of Semantics and Linguistic Theory 14*, Ithaca, NY: CLC Publications, 235-252.

Roberts, Craige (2012), Information structure in discourse: Towards an integrated formal theory of pragmatics, in *Semantics and Pragmatics* 5, 1-69.

Stalnaker, Robert (1978), Assertion, in P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*, New York: Academic Press, 315-332.

Truckenbrodt, Hubert (2006), On semantic motivation of syntactic verb movement to C in German, in *Theoretical Linguistics* 32, 257-306.

Tsiknakis, Antonios (2016), *Die (a)symmetrische Distribution von [±wh]-Phrasen. Zur Syntax und Interpretation deklarativer und interrogativer Satztypen*, Phd thesis, Bergische Universität Wuppertal.

Tsiknakis, Antonios (2017), Zur Kategorie von Verbzweit- und Verbletztsätzen, in S. Nefedov, L. Gregorieva, and B. Bock (eds), *Deutsch als Bindeglied zwischen Inlands- und Auslandsgermanistik. Beiträge zu den 23. GeSuS-Linguistik-Tagen in St. Petersburg, 22.-24. Juni 2015*, Hamburg: Kovác Verlag.

Tsiknakis, Antonios (to appear), Verb movement and topicalization in German, in H. Lohnstein, and A. Tsiknakis (eds), *Verb Second – Grammar Internal and Grammar External Interfaces*, Berlin: de Gruyter.

"Your friend should keep her suggestions to herself *jor*": Pragmatic borrowing at left and right peripheries in Nigerian English

Foluke O. Unuabonah & Rotimi O. Oladipupo
(Redeemer's University)

Keywords: bilingual pragmatic markers, Global Web-based English Corpus, left periphery, right periphery, pragmatic borrowing

Schedule: We 17.30 Room 11

Pragmatic borrowing deals with the process and results of including discourse-pragmatic features such as interjections, pragmatic markers and tags from a source language into a recipient language (Andersen 2014, and Balteiro 2018). Such pragmatic items have been found in second language varieties of English such as Indian (Lange 2009) and Singapore (Leimgruber 2016) Englishes. These pragmatic items have also been found to occur both at left and right peripheries (Beeching and Detges 2014). While it has been confirmed that there is paucity of studies that examine pragmatic borrowing in Nigerian English (e.g. Unuabonah and Oladipupo 2018), studies that distinguish between borrowed pragmatic markers at the left and right peripheries in Nigerian English are almost non-existent. Indigenous (bilingual) pragmatic markers in Nigerian English such as *o*, *sha* and *abi*, which have been studied (see Unuabonah and Oladipupo 2018), have been found to occur mainly at the right periphery. However, other indigenous markers (such as *jor* and *shey*), which are yet to be studied, do occur both at the right and left peripheries. Thus, this study investigates the inclusion of four pragmatic markers: *jor, jare, shebi and shey* from Yoruba (a Nigerian local language) in Nigerian English, with a view to exploring their meanings, frequency, grammatical constraints and discourse-pragmatic functions. The

study further investigates the functional differences that occur when borrowed pragmatic markers occur at different peripheries.

The data for the study are taken from the Nigerian component (42,646,098) of the Global Web-based English Corpus (Davies and Fuchs 2015). These are analysed qualitatively and quantitatively, using the grammatical-pragmatic approach to the study of pragmatic markers, systemic-functional linguistics as well as the concept of pragmatic borrowing. The study indicates that while *jare* is found at the right periphery *jor*, *shey* and *shebi* are located both at the right and left peripheries. While *jor* shares similar meanings with the English pragmatic marker, *please*, and may also be used to indicate exasperation, *jare* shares similar meanings with the English pragmatic marker, *indeed*. *Shebi* and its modified form, *shey*, share similar meanings with the pragmatic marker, *right*, and may serve as interrogative markers in the left-periphery and question tags in the right periphery. The study, thus, contributes to the research on bilingual pragmatic markers and the impact of left versus right periphery on their forms and functions.

Acknowledgements
The authors appreciate the support of the Alexander von Humboldt Foundation.

References

Andersen, Gisle (2014), Pragmatic Borrowing, *Journal of Pragmatics* 67, 17-33.

Balteiro, Isabel (2018), Oh Wait: English Pragmatic Markers in Spanish Football Chatspeak, *Journal of Pragmatics* 133, 123-133.

Beeching, Kate, and Ulrich Detges (eds), (2014), *Discourse Functions at the Left and Right Periphery: Cross-Linguistic Investigations of Language Use and Language Change*, Leiden: Brill.

Davies, Mark, and Fuchs, Robert (2015), Expanding Horizons in the Study of World Englishes with the 1.9 Billion Word Global Web-Based English Corpus (GloWbE), *English World-Wide*, 36(1), 1–28.

Lange, Claudia (2009), 'Where's the party yaar!' Discourse particles in Indian English, in T. Hoffmann, and L. Siebers (eds.), (2009), *World Englishes: Problems, Properties and Prospects*, Amsterdam: John Benjamins, 207-226.

Leimgruber, Jakob (2016), Bah in Singapore English, *World Englishes* 35(1), 78-97.

Unuabonah, Foluke, and Oladipupo, Rotimi (2018): "You're not Staying in Island *Sha O*": *O*, S*ha* and *Abi* as Pragmatic Markers in Nigerian English. *Journal of Pragmatics* 135, 8-23.

The diachrony of the constructional networks of 'wonder' nouns

An Van Linden & Lieselotte Brems
(University of Liège and research fellow KU Leuven; University of Liège and research fellow KU Leuven)

Keywords: mirativity, grammaticalization, constructional network, semiotic nouns

Schedule: We 12.00 Room 11

This paper examines the diachronic development of constructions with nouns that denote something unexpected or astonishing. Specifically, it will compare constructions with *marvel* (1) and *surprise* (2) with those with *wonder* (3)-(4), which have been described in Van linden et al. (2016) and Gentens et al. (2016). The paper concentrates on constructions with complement clauses, cf. (1)-(4), and

distinguishes between lexical and grammatical uses of complement-taking predicate (CTP) clauses (cf. Boye & Harder 2007).

(1) So drastic a series of atrocities worried even Lord Burleigh, who compared them with the much-condemned Spanish activities in the Low Countries: "as things be altered it is *no marvel* the people have rebellions here, for the Flemings had not so much cause to rebel by the oppression of the Spaniards, as is reported to the Irish people" (WB)
(2) Tyson soon integrated into that environment and the authorities misguidedly believed he was being rehabilitated thanks to boxing. And it was *no surprise* when Tyson was released early to go to live with D'Amato permanently. (WB)
(3) It is *a wonder* to me that no one is laughing at the silly boots, but I suppose they have other worries at the moment, and so do I. (WB)
(4) After all the scaremongering since September 11 regarding good versus evil, with us or with the terrorists, it is *no wonder* Arab-phobia has hit new heights. (WB)

In (4), *it is no wonder* functions as a mirative qualifier, commenting on the complement proposition in terms of its (un)expectedness (DeLancey 2001). It can be paraphrased by an expectation adverb such as *of course* (Simon-Vandenbergen & Aijmer 2007: 172), and thus serves a grammatical function (Gentens et al. 2016). The speaker's lack of surprise about the complement proposition is justified by the *after*-PP. By contrast, in (3) *it is a wonder to me* conveys that the speaker is very surprised, with the *that*-clause containing the presupposed factive proposition that s/he is surprised about. The CTP-clause in (2) expresses the speaker's lack of surprise about the proposition in the *when*-complement. Note that – unlike in (4) – the expressions of (lack of) surprise in (3) and (2) are discourse-primary (e.g. they can be 'addressed' by 'how much wonder/(of a) surprise is/was it?') and thus show lexical use (cf. Boye & Harder 2007). In (1) *it is no marvel* functions as a mirative qualifier like in (4).

This paper aims to inventory the constructions with 'wonder' nouns in Present-day English, as well as to study their diachronic development, investigating how *marvel* and *surprise* came to compete with native *wonder* when borrowed into English after the Norman conquest. It will also examine to what extent grammatical CTP-clauses have been in paradigmatic contrast with juxtaposed and parenthetical clausal structures or adverbial structures (cf. Gentens et al. 2016), with adverbials of the form negation + noun (e.g. *no way, no doubt)* already entrenched in Early Modern English as a constructional template (cf. Davidse et al. 2015). Data will be drawn from the Penn-Helsinki Historical Corpora, the Corpus of Early Modern English Texts, the Corpus of Late Modern English Texts and WordBanksOnline (WB).

References

Boye, Kasper and Peter Harder (2007), Complement-taking Predicates: Usage and Linguistic Structure, *Studies in Language* 31, 569–606.

Davidse, Kristin, Simon De Wolf and An Van linden (2015), The development of (*there/it is / I have*) *no doubt* expressing modal and interactional meaning, *Journal of Historical Pragmatics* 16(1), 25–58.

DeLancey, Scott (2001), The mirative and evidentiality, *Journal of Pragmatics* 33(3), 369–382.

Gentens, Caroline, Ditte Kimps, Kristin Davidse, Gilles Jacobs, An Van linden and Lieselotte Brems (2016), Mirativity and rhetorical structure: The development and prosody of disjunct and anaphoric adverbials with '*no' wonder*, in G. Kaltenböck, E. Keizer and A. Lohmann (eds), (2016), *Outside the Clause. Form and function of extra-clausal constituents* [Studies in Language Companion Series 178], Amsterdam: John Benjamins, 125–156.

Simon-Vandenbergen and Karin Aijmer (2007), *The Semantic Field of Modal Certainty: A Corpus-based Study of English Adverbs*, Berlin: Mouton de Gruyter.

Van linden, An, Kristin Davidse and Lennart Matthijs (2016), Miracles and mirativity: From lexical *it's a wonder* to grammaticalised *it's no wonder* in Old English, *Leuvense Bijdragen – Leuven Contributions in Linguistics and Philology* 99-100, 385–409.

Corpora

WordbanksOnline Corpus https://wordbanks.harpercollins.co.uk/

YCOE: Taylor, A., A. Warner, S. Pintzuk and F. Beths (2003), The York-Toronto-Helsinki Parsed Corpus of Old English Prose, University of York.

PPCME2: Kroch, A. & Taylor, A. (2000), The Penn-Helsinki Parsed Corpus of Middle English (PPCME2), Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, release 4.

PPCEME: Kroch, A., Santorini, B. & Delfs, L. (2004), The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME), Department of Linguistics, University of Pennsylvania. CD-ROM, first edition, release 3.

CEMET & CLMET: De Smet, Hendrik (2013), *Spreading Patterns: Diffusional Change in the English System of Complementation*, Oxford: Oxford University Press, pp. 13–17.

Testing Greenberg's universals on a global scale

Annemarie Verkerk, Hannah Haynie, Russell Gray & Hedvig Skirgård

(Saarland University, Colorado State University, MPI for the Science of Human History & Australian National University)

Keywords: linguistic typology, universals, phylogenetic comparative methods, noun phrase, word order

Schedule: ~~We 11.00 Room 1~~ New schedule Fri 17.00 Room 6

Linguistic typology has a rich history of proposing universals; the Konstanz Universals Archive (Plank and Filimonova 2006) contains 2,029 entries (297 of them absolute). Most, if not all, have at some point been demonstrated as false given a certain language. As Bickel (2018) proposes, we should focus less on absoluteness, exceptions and rara, but instead recognize that all universals are probably probabilistic. Linguistics has evolved as a discipline, and we have found more nuanced and sophisticated ways of testing the power of a certain universal. One of the most important advances in the testing of rules like universals involves controlling for language history.

In recent years, several of Greenberg's word order universals have been claimed to be lineage-specific (Dunn et al. 2011) or alternatively to be divided between lineage-specific patterns and true universals (Jäger 2018a). Other universals have not yet been investigated using quantitative methods that control for language history. In this paper we investigate a large set of universals with a new large morphosyntactic dataset and control not for individual families but for an abstract global tree. The typological data are taken from Grambank (Skirgård et al. in review), a new database containing information on more than 1400 languages. The universals are gathered from Greenberg (1996) and Plank and Filimonova's (2006) massive collection. We select universals that dictate the presence or absence of two specific features (such as gender and number; or tensedness and nouny adjectives) and investigate whether these features are likely to co-evolve or not. Rather than investigating individual families (as done by Dunn et al. 2011 and Jäger 2018a), we make use of global language trees (Jäger

2018b, other global trees are in preparation). This allow us to incorporate small families and isolates, as well as quantitatively appreciate the fact that many separate families that are in a certain area probably share a past, even if we cannot currently with the comparative method confirm that they are indeed of one family.

Preliminary results suggest that some features indeed change together while others do not, supporting Jäger's (2018a) conclusion that universality is a matter of degree. We interpret universals in a diachronic sense: in testing for instance Greenberg's universal no. 36, "If a language has the category of gender, it always has the category of number", we compare a model in which gender is allowed to be acquired without the presence of number, to a model where such a change is not allowed. We find that indeed Greenberg's universal no. 36 is close to being universal, while Stassen's (1997) and Wetzer's (1996) universal "If a language is tensed, it will have nouny adjectives" is closer to the non-universal end of the spectrum. We propose an explanation for the attested spectrum of universality in terms of interconnectivity: some features are relatively stand-alone (such as numeral order, which doesn't pattern with other NP orders) while others are highly dependent on other features probably due to morphology (such as having nouny or verby adjectives).

References

Bickel, Balthasar (2018), Estimating diachronic biases of typological traits in unknown phylogenies, paper presented at the 12th Conference of the Association for Linguistic Typology, Canberra.

Dunn, Michael, Greenhill, Simon J., Levinson, Stephen C. & Gray, Russell D. (2011), Evolved structure of language shows lineage-specific trends in word-order universals, *Nature*, 473, 79-82.

Greenberg, Joseph H. (1966), Some universals of grammar with particular reference to the order of meaningful elements, in Joseph H. Greenberg (ed.), *Universals of grammar*, Cambridge: The MIT Press.

Jäger, Gerhard (2018a), A Bayesian test of the lineage-specificity of word order correlations, paper presented at the 12th International Conference on Language Evolution (Evolang XII), Torun.

Jäger, Gerhard (2018b), Global-scale phylogenetics linguistic inference from lexical resources, S*cientific Data*, 5, 180189. doi:doi.org/10.1038/sdata.2018.189

Plank, Frans & Filimonova, Elena (2006), The Universals Archive: A brief introduction for prospective users, *STUF - Language Typology and Universals* 53(1), 109-123.

Skirgård, Hedvig et al. (in review), Grambank, Jena: Max Planck Institute for the Science of Human History.

Stassen, Leon (1997), *Intransitive predication*, Oxford: Oxford University Press.

Wetzer, Harrie (1996), *The typology of adjectival predication*, Berlin: Mouton de Gruyter.

The place of spatial particles in the typological profile of Swedish

Åke Viberg
(Uppsala University)

Keywords: contrastive study, corpus linguistics, spatial semantics, spatial metaphor

Schedule: We 11.00 Room 14

In Swedish, spatial verbal particles with meanings such IN/OUT and UP/DOWN play a prominent role in the general structure of the language and its typological profile (Viberg 2006). Not only are they used as free words (*komma upp* 'come up'), but in addition they are frequently used as bound forms with an abstract meaning in verbs (*uppkomma*: *upp-* 'up' + *komma* 'come' meaning 'come into

existence') and in nouns (*uppkomst* 'coming into existence'). Several studies describe Swedish particles from a monolingual perspective (Teleman et al. 1999, Hellberg 2007, Johannison 1954, Norén 1996, Strzelecka 2003) but few studies such as Andersson (2005) have a contrastive perspective, except for studies related to Talmy's motion typology (Fagard et al. 2013, Olofsson 2018, Viberg 1992). The present paper is part of a general study of all the Swedish spatial particles from a contrastive/typological perspective (Viberg 2015, 2017). Data are taken from English and Swedish original texts and their translations in the English-Swedish Parallel Corpus (ESPC, Altenberg & Aimer 2000) and on Swedish original texts (fiction, around 600,000 words) and their translations into English, German, French and Finnish in the Multilingual Parallel Corpus (MPC, Viberg 2013). The present paper will focus on *upp*, whose core meaning is realized across a wide range of languages but in various forms, and on the language-specific *fram*. The investigation of *upp* is based on 3000 occurrences of free and bound *upp* in the Swedish original texts in MPC and their translations into the other languages. For each occurrence where *upp* is combined with a verb, the verb and its semantic field are coded. In fiction, motion verbs are dominant, but several fields are frequent (Perception 'look up', Possession 'give up' etc). The meaning of the resulting combination is also coded. In motion events, UP is realized in the verb not only in French but also to various degrees in the languages that have a rich system of particles. (There is a continuum on this point.) The use of case in German and Finnish also affects the choice of translations in certain contexts (often Zero, when *upp* is combined with a locative preposition). Bound *upp* primarily fills lexical gaps in Swedish in abstract domains and has very variable correspondences. *Fram* basically is conceptually related to the FRONT/BACK schema. This is clearly reflected in forms such as the compound preposition *framför* 'in front of'. The particle *fram* has many extended uses based on other schemas and is very language-specific which is reflected in its in English translations in the ESPC. Out of a total of 772 occurrences, the major translations are *out* (114 occurrences), *up* (57) and *forward* (36). In addition, a single verb is used as a translation (170) (e.g. *arrive*, *emerge*, *present*, *produce*). The analysis will be based on cognitive linguistic notions such as image schema (Ruddy 2007) and metaphor (Sullivan 2017) and related to work on the languages compared with Swedish (for example: Dewell 2011, Huumo & Ojutkangas 2006, Lindner 1982, 1983, Stolova 2015).

References

Altenberg, Bengt and Karin Aijmer. 2000. The English-Swedish Parallel Corpus: A resource for contrastive research and translation studies. In *Corpus Linguistics and Linguistic Theory*, edited by Christian Mair and Marianne Hundt, 15-23, Amsterdam – Atlanta/GA: Rodopi.

Andersson, Carina. 2005. Équivalence et saillance dans l'expression de la localization frontale dynamique en suédois et en français. Étude comparative et contrastive de *fram* et de (*s'*)*avancer*/*en avant*. [ACTA UNIVERSITATIS UPSALIENSIS. Studia Romanica Upsaliensia 73]

Dewell, Robert B. 2011. *The Meaning of Particle/Prefix Constructions in German*. Amsterdam: Benjamins.

Fagard, B., Zlatev, J., Kopecka, A., Cerruti, M. & J. Blomberg. 2013. The Expression of Motion Events: A Quantitative Study of Six Typologically Varied Languages. Berkeley Linguistics Society 39, 364-379.

Hellberg, Staffan. 2007. Polysemy across image schemas: Swedish *fram*. *Studia Linguistica* 61(1), 20-58.

Huumo, Toumas, and Krista Ojutkangas. 2006. "An Introduction to Finnish Spatial Relations, Local Cases and Adpositions." In Grammar from the Human Perspective: Case, Space and Person in Finnish, edited by Marja-Liisa Helasvuo, and Lyle Campbell, 11-20. Amsterdam: Benjamins.

Johannisson, Ture. 1954. Om sammansatta verb i svenskan. [About compound verbs in Swedish] *Nysvenska Studier*.

Lindner, Susan Jean. 1982. What goes up doesn't necessarily come down: The Ins and Outs of Opposites in *Papers from the Eighteenth Regional Meeting, Chicago Linguistic Society*, April 15-16, 1982, ed. by Kevin Tuite, Robinson Schneider, Robert Chametsky.

Lindner, Susan Jean. 1983. A lexico-semantic analysis of English verb particle constructions with *out* and *up*. Bloomington, Ind.: Indiana univ. linguistics club.

Olofsson, Joel. 2018. *Förflyttning på svenska*. [Motion in Swedish – on productivity from a construction grammar perspective.] Göteborgsstudier i nordisk språkvetenskap 32. Gothenburg, University of Gothenburg.

Ruddy, Todd. 2007. Image schemas. In Dirk Geeraerts & Hubert Cuyckens (eds.) *The Oxford handbook of cognitive linguistics*. New York: Oxford University Press, 214-235.

Stolova, Natalya I. 2015. *Cognitive linguistics and lexical change. Motion verbs from Latin to Romance*. Amsterdam: Benjamins.

Strzelecka, Elżbieta. 2003. Svenska partikelverb med *in*, *ut*, *upp* och *ner*. En semantisk studie ur kognitivt perspektiv. (Swedish phrasal verbs with *in*, *ut*, *upp* and *ner*. A semantic study from a cognitive perspective.) [Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet 62] Uppsala.

Sullivan, Karen. 2017. Conceptual metaphor, In Barbara Dancygier (ed.). 2017. *The Cambridge handbook of cognitive linguistics*. Cambridge: Cambridge University Press, 385-406.

Teleman, Ulf, Hellberg, & Erik Andersson 1999. *Svenska Akademiens Grammatik* [The grammar of the Swedish Academy] vol. 3, 417-435. Stockholm: Svenska Akademien.

Viberg, Åke. 1992. Universellt och språkspecifikt i det svenska ordförrådets organisation. *Tijdschrift voor Skandinavistiek* 13:2, 17-58.

Viberg, Åke. 2006. The typological profile of Swedish: An introduction. *Sprachtypologie und Universalienforschung*. (STUF) 59(1), 3–11.

Viberg, Åke. 2013. Posture verbs. A multilingual contrastive study. *Languages in Contrast* 13-2, 139-169.

Viberg, Åke. 2015. Motion verb typology and the expression of the endpoint of motion in Swedish. In: *Concepts and Structures - Studies in Semantics and Morphology.* [Studies in Linguistics and Methodology, Vol. 8]. (edited by Maria Bloch-Trojnar, Anna Malicka-Kleparska and Karolina Drabikowska). Lublin: Wydawnictwo KUL, 209-229.

Viberg, Åke. 2017. Contrasts in morphology. The case of UP/DOWN and IN/OUT as bound morphemes in Swedish and their English correspondents. In *Contrasting English and Other Languages through Corpora,* Markéta Janebova, Ekaterina Lapshinova-Koltunski & Michaela Martinková (eds). Cambridge: Cambridge Scholars, 32-74.

Self-quotation and logophoricity (in West-Africa) - geared to the same end

Rebecca Voll
(CNRS - LLACAN)

Keywords: self-quotation, logophoricity, reported speech, logophoric pronoun, person-marked quotatives

Schedule: We 17.30 Room 7

There is a handful of languages in the world which require a speaker to use either a special syntactic construction or a specialized morphological marker when quoting themselves rather than someone else. Such dedicated marking of self-quotation has only very recently received the attention of linguists (Michael 2012, Lionnet 2015) and it is said to be extremely rare crosslinguistically.

In this paper I argue that dedicated self-quotation markers in West-Africa are better understood with reference to logophoricity, a phenomenon which is more widespread in the area and which has been described more extensively.

I interpret the term logophoricity here in a wider sense, based on Nikitina (2018), who points out that "the distinction based on the presence or absence of overt markers [i.e. specialized pronouns which are restricted to reported speech and which refer back to the reported speaker] is artificial and does not capture the extent of the phenomenon of logophoricity. In particular, a substantial number of languages lacking specialized logophoric pronouns nevertheless display patterns that are functionally equivalent to those encountered in logophoric languages.". I use the Bantoid language Mundabli (Cameroon) as a rare example of a language with dedicated self-quotation markers in order to show how the system of self-quotation can be better understood with reference to logophoricity (in the sense of Spronck and Nikitina in progress) as these seemingly unrelated phenomena share functional motivations and exhibit similar properties and restrictions on use.

In order to support my claim, I compare self-quotation with three other logophoricity-related phenomena. First, I compare whether the restrictions on the use of self-quotation markers are similar to common restrictions on the use of logophoric pronouns. Second, I confirm whether the typology of the cross-linguistic distribution of person-marked quotative paradigms shows that languages with a two-way person distinction always distinguish first person from non-first person, rather than, e.g. speech-act participant vs. other. This would show the functional relatedness to another logophoricity-related phenomenon. And third, I will look more closely into the following issue: when the semantics of logophoric prononuns extend diachronically, the distinction that is preserved is the one between self- and non-self quotation.[i] Logophoric pronouns are either restricted to the 3rd person or they cover both 2nd and 3rd person. They never include the 1st person. The fact that pure 2nd person logophoric languages are not attested suggests that the semantics of 3rd person logophoric pronouns extend diachronically to include the 2nd person (Nikitina 2012b, Dimmendahl 2001). And when the semantics of logophoric prononuns extend diachronically to include the 2nd person, the distinction that is preserved is exactly the one between self- and non-self quotation.

Thus, the final results are expected to show that self-quotation markers are better understood with regard to logophoricity, and that they are ultimately geared to the same end: they are a different ways of marking the same opposition.

This project has received funding from the European Research Council (ERC) as part of the project 'Discourse reporting in African storytelling' under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 758232).

References

Dimmendaal, G. J. 2001 *Logophoric Marking and Represented Speech in African Languages as Evidential Hedging Strategies.* Australian Journal of Linguistics 21(1): 131—57.

Güldemann, Tom (2008). *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Empirical Approaches to Language Typology 34. Mouton de Gruyter.

Idiatov, D. 2010. Person–number agreement on clause linking markers in Mande. *Studies in Language* 34(4). pp. 832–868.

Lionnet, F. 2015. *More than reported speech: quotative evidentiality in Laal.* 8th World Congress on African Linguistics (WOCAL 8), Kyoto University, 21-24 August 2015. (slides)

Michael, L. 2012. Nanti self-quotation: Implications for the pragmatics of reported speech and evidentiality. Pragmatics and Society, 3(2).
Nikitina, T. 2012a. Logophoric Discourse and First Person Reporting in Wan (Westafrica). In Anthropological Linguistics, Volume 54, Number 3, Fall 2012, pp. 280-301.
Nikitina, T. 2012b. Personal deixis and reported discourse: Towards a typology of person alignment. *Linguistic Typology* 16.
Nikitina, T. (2018). *Discourse reporting in African storytelling.* Unpublished project summary.
Spronck, S. and T. Nikitina. 2019. In press. Reported speech forms a dedicated syntactic domain: Typological arguments and observations. *Linguistic Typology*.

When two systems meet - Analytic and synthetic indirect object marking in Torlak dialect

Teodora Vukovic

Schedule: We 12.30 Room 7

<pdf>

Differential Object Marking in Spanish and Portuguese: towards a more differentiated analysis by comparing profiles of different varieties

Albert Wall
(University of Zurich)

Keywords: Differential object marking, clitic doubling, null objects, leísmo, Ibero-romance languages

Schedule: We 12.00 Room 7

Maybe due to its popularity, the notion of Differential Object Marking (DOM) (Bossong 1985) has become a broad cover term for very different phenomena. This has led to recent attempts of establishing entire typologies of DOM phenomena, as in Witzlack-Makarevich & Seržant (2018). In this talk, I first show that this also applies to DOM studies on Spanish and Portuguese. Then, I analyze the different DOM phenomena in the two languages according to the above-mentioned typology. On the one hand, this contributes to a better understanding of such new typologies, on the other, it allows for a more adequate treatment of the grammatical idiosyncrasies of the different DOM phenomena in the two languages.

The grammatical phenomena that are currently being discussed as instantiations of DOM systems in Spanish and Portuguese are splits in (i) flagging of direct objects (DOs), the so-called "prepositional accusative", (ii) indexing of DOs ("clitic doubling"), (iii) overt vs. null object pronouns and (iv) pronominal expression of DOs ("leísmo"), as discussed in Flores & Melis (2007), López (2012), Schwenter (2014), Melis (2017) and Cyrino (2017), among others. Both languages and many of their varieties share the four grammatical phenomena, but in each variety only a subset works as a DOM system. For instance, Standard Spanish has both flagging and indexing DOM, while DOM in Portuguese is mainly instantiated by null vs. overt DO pronouns. Furthermore, there are differences in the variational dynamics. Subsystems of DOM may be robustly entrenched or only marginal and they

also may be rigorously split, split-fluid or fluid, in the terminology of Witzlack-Makarevich & Seržant (2018).

In this talk, I compare the "DOM profiles" of twelve varieties of Spanish and Portuguese from Europe as well as the Americas and show that if we combine the descriptions that can be found in the literature and own empirical work from Argentina, Mexico, Peru and Spain, there are roughly as many profiles as there are varieties. Contrary to many previous accounts that, based on supporting evidence, identify ever more DOM systems in these two languages, I also consider contradicting evidence, namely incompatibilities with the concept of DOM, counterintuitive patterns of interaction between the phenomena, different diachronies and the lack of an overarching generalization beyond the fact that the data can be interpreted as showing some kind of split. Given this more complete picture, I argue that a unified account currently seems to be out of reach, if feasible at all.

Applying the split/fluid distinction mentioned above to each DOM subsystem and combining it with the observation from Klein & de Swart (2011) that inherent properties of DOs trigger DOM while for external properties DOM is the "result", it will show that all Spanish varieties have DOM trigger in DO flagging. For any other DOM system in Spanish and Portuguese this is not the case, where we have only fluid systems, which, following Klein & de Swart (2011), should be dealt with in an underspecification analysis. Such and other patterns, in return, can again be compared with the patterns of other further languages within a more differentiated typology.

References

Bossong, Georg (1985), *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Cyrino, Sonia (2017), Reflexões sobre a marcação morfológica do objeto direto por a em portguês brasileiro, *Estudos linguísticos e literários* 58. Salvador, 83–103.

Klein, Udo & de Swart, Peter (2011), Case and referential properties, *Lingua* 121, 3–19.

López, Luis, (2012), *Indefinite Objects. Scrambling, Choice Functions and Differential Marking*. MIT Press.

Melis, Chantal (2018), Spanish indexing DOM, topicality, and the case hierarchy, in I. Seržant & A. Witzlack-Makarevich (eds), *The Diachronic Typology of Differential Argument Marking*, Berlin: Language Science Press, 97–128.

Schwenter, Scott (2014), Two kinds of differential object marking in Portuguese and Spanish, in P. Amaral & A. M. Carvalho (eds), *Portuguese-Spanish Interfaces. Diachrony, Synchrony and Contact*, Amsterdam: John Benjamins, 237–260.

Witzlack-Makarevich, Alena & Seržant, Ilja (2018), Differential argument marking: Patterns of variation, in I. Seržant & A. Witzlack-Makarevich (eds), *The Diachronic Typology of Differential Argument Marking*, Berlin: Language Science Press, 1–40.

Functionality of *de* in the context of Chinese adjective-noun combinations

Xu Xu
(Shanghai Jiao Tong University)

Keywords: adjective-noun combination, compound, de, language representation, semantics

Schedule: We 15.30 Room 15

De is a high frequency character in the Chinese language. In fact, it has the highest frequency among all words in the Lancaster Corpus of Mandarin Chinese (McEnery and Xiao 2003). However, despite (or because of) its ubiquity, the nature of Chinese *de* remains a mystery. Larson (2009) remarks that "Chinese grammarians have achieved no consensus on the nature of *de*, despite intensive study", and that some even consider it a grammatically functionless element, inserted for purely phonological reasons.

The status of Chinese adjective-noun combinations ([A N]) has long been debated. Some argue that semantically transparent Chinese [A N] such as *hong hua* 'red flower', in which both *hong* 'red' and *hua* 'flower' contribute to the meaning of the whole expression, are phrases (e.g., Li and Thompson, 1981; Paul, 2005). However, many others consider Chinese [A N] words (e.g., Zhu 1956, Sproat and Shih 1991, McCawley 1992, Duanmu 2000).

A cause for this debate might be the fact that *de* is often inserted in-between (e.g., *hong de hua* 'red DE flower'), resulting in a different form of adjective-noun combination ([A *de* N]). Recently, Xu (2018) has presented extensive linguistic evidence to show that Chinese [A N] are words, while Chinese [A *de* N] are phrases. The debate on the status of [A N] appears a corollary of past uncertainty among linguists about the relationship between [A N] and [A *de* N]. That is, those who considered [A N] phrases might have viewed [A N] equivalent to [A *de* N], whereas those who considered [A N] words might not, which in turn appears attributable to uncertainty about the functionality of *de*.

This study took a psycholinguistic approach, aiming to examine potential syntactic and semantic differences between Chinese [A N] and [A *de* N]. We report two experiments. Experiment 1 (n=132) adapted the structural priming paradigm in order to investigate whether [A N] and [A *de* N] are structurally differentiable. Structural priming, in short, refers to the effect that exposure to expressions with a particular syntactic structure would enhance the level of processing for expressions with the same structure (Branigan and Pickering 2017). We tested whether reading a set of adjective-noun combinations in one form, [A N] versus [A *de* N], would modify comprehension about another set of combinations in the same form or in the opposite form. With semantic features being controlled, we observed priming effects on comprehension of expressions in the same form but not in the opposite form, demonstrating structural distinction between the two forms.

Experiment 2 (n=20) examined how the two forms might differ in meaning. A visual memory task indicated that, as evidenced by different meaning representations, the two forms appeared to place different levels of emphasis on the attribute denoted by the adjective. These syntactic and semantic differences naturally imply that the two forms also differ in function. Complementary to linguistic evidence, such systematic distinctions based on psycholinguistic evidence are instrumental to advancing theories about Chinese adjective-noun combinations and functionality of the mysterious Chinese character *de*.

References:

Branigan, Holly P. & Martin J. Pickering, (2017), Structural priming and the representation of language, *Behavioral and Brain Sciences* 40, Cambridge University Press: e313.

Duanmu, San, (2000), *The phonology of Standard Chinese*, Oxford: Oxford University Press.

Larson, Richard K., (2009), Chinese as a reverse Ezafe language, *Yuyanxue Luncong (Journal of Linguistics)* 39, 30-85, Beijing: Peking University.

Li, Charles N. & Sandra A. Thompson, (1981), *Mandarin Chinese: A functional reference grammar*, Berkeley & Los Angeles: University of California Press.

McCawley, James D., (1992), Justifying part-of-speech assignments in Mandarin Chinese. *Journal of Chinese Linguistics*, 20(2), 211–246.

McEnery, Anthony & Zhonghua Xiao, (2003), *Lancaster Corpus of Mandarin Chinese*, Released by the European Languages Resources Association (Catalogue No. W0039) and the Oxford Text

Archive (Catalogue No. 2474).
Paul, Waltraud, (2005), Adjectival modification in Mandarin Chinese and related issues. *Linguistics* 43(4), 757–793.
Sproat, Richard & Chilin Shih, (1991), The cross-linguistic distribution of adjective ordering restrictions, in C. Georgopoulos & R. Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 565–593, Dordrecht: Kluwer Academic Publishers.
Xu, Zheng, (2018), The word status of Chinese adjective-noun combinations. *Linguistics* 56(1), 207–256.
Zhu, Dexi, (1956), Xiandai hanyu xingrongci yanjiu [A study of modern Chinese adjectives], *Yuyan yanjiu* 1, 83–112.

Property concepts and adjectives: A typological perspective

Jingting Ye
(University of Leipzig)

Keywords: property concepts, adjective, the attributive-prominence hierarchy, coding discrepancy, coding asymmetry

Schedule: We 12.00 Room 6

The adjective is notorious for its internal heterogeneity and has been subject to long-standing disputes in the previous literature (e.g. Chomsky 1970, Dixon 1977 and Schachter 1985). In many languages, the adjective constitutes a closed category and the counterparts of many English adjectives are either coded as nouns (e.g. in many Australian languages) or as stative verbs (e.g. in many African languages), which are termed "nouny adjectives" and "verby adjectives" by Wetzer (1992, 1996).

This paper deals with the coding discrepancy and the coding asymmetry of property concepts across languages. The coding discrepancy refers to the phenomenon that different property concepts are encoded differently, as illustrated by Japanese examples in (1).

(1) Japanese
a. *nagai* *ressya*
long train
"long train" (Backhouse 2004: 56)
b. *rippa* ***na*** *setubi*
impressive **ATTR** facilities
"impressive facilities" (Backhouse 2004: 59)

Two types of coding asymmetry are identified: the predicative-only coding and the attributive-only coding. The predicative-only coding refers to the phenomenon that only the predicative use is marked by a predicativizer, while the attributive use remains unmarked, as illustrated by (2). The attributive-only coding refers to the phenomenon that only the attributive use is marked by an attributivizer, while the predicative use remains unmarked, as illustrated by (3).

(2) English
a. *the red flower*
b. *The flower* ***is*** *red.*
(3) Cavineña (Pano-Tacanan, South America)

a. *wika* *arida=**ke***
hook big=**ATTR**
"the big hook" (Guillaume 2008: 360)
b. *éna=tu* *arida*
water=3SG big
"The water is high (lit. big)." (Guillaume 2008: 44)

This study investigates the attributive and predicative use of 27 different property concepts in 30 languages with coding discrepancy. The sample consists of 47 languages from different language families and different geographical areas. The property concepts, shown in (4), are chosen based on Dixon (1977), excluding SPEED, which is identified by Hallonsten Halling (2018) as a core semantic type of adverbs.

(4) **Property concepts from 6 semantic types:**
a. DIMENSION: big, small, long, short
b. AGE: young, old, new
c. VALUE: good, bad, beautiful, ugly
d. COLOUR: white, black, red, green
e. PHYSICAL PROPERTY: soft, hot, cold, dry, heavy
f. HUMAN PROPENSITY: afraid, crazy, angry, happy, sad, clever

The main result is formulated as an attributive-prominence hierarchy in (5). The aim of this hierarchy is to make predictions about the formal encoding of property concepts in languages with coding discrepancy.

(5) **The attributive-prominence hierarchy**:
AGE > COLOUR > DIMENSION > PHYSICAL PROPERTY > VALUE > HUMAN PROPENSITY

From the left-most to the right-most point of the hierarchy, the attributive-prominence decreases and the predicative-prominence increases. If property concepts at one point of the hierarchy show the **attributive-only coding**, all property concepts **to the right** also tend to show the **attributive-only coding**. If property concepts at one point of the hierarchy show the **predicative-only coding**, all property concepts **to the left** also tend to show the **predicative-only coding**.

References

Backhouse, A. E. (2004), Inflected and Uninflected Adjectives in Japanese. In: R. M. Dixon & A. I. Aikhenvald (Eds.), (2004), *Adjective classes: A cross- linguistic typology*. Oxford: Oxford University Press, 55-73.

Chomsky, N. (1970), Remarks on nominalization. *Readings in English Transformational Grammar*. Ginn, Waltham, Massachusetts: Blaisdell, 184-222.

Dixon, R.M. (1977), Where have all the adjectives gone? *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 1(1), 19-80.

Guillaume, A. (2008), *A grammar of Cavineña.* Berlin: Mouton de Gruyter.

Hallonsten Halling, P. (2018), *Adverbs: A typological study of a disputed category*. Stockholm: Stockholm University.

Schachter, P. (1985), Parts-of-speech systems. In T. Shopen (Ed.), (1985), *Language typology and syntactic description*, vol.1, Cambridge: Cambridge University Press, 3-61.

Wetzer, H. (1992), "Nouny" and "verby" adjectivals: A typology of predicative adjectival constructions. In: *Meaning and grammar: Cross-linguistic Perspectives.* Berlin: Mouton de Gruyter, 223-262.

Wetzer, H. (1996), *The typology of adjectival predication*, Berlin: Walter de Gruyter.

Does smell speak our language? A preliminary study of smell terms in Estonian and German

Karin Zurbuchen & Mari Uusküla
(Tallinn University)

Keywords: prototype and categorization, perception, olfaction, odor naming, cultural factors

Schedule: We 14.30 Room 15

Humans can distinguish at least about one trillion different odors (Bushdid et al., 2014). A number of studies have confirmed that in urbanized Western societies smells are not particularly codable in languages (de Val, Huisman, Wnuk, and Majid, 2017; Majid, 2015; Maijd, Burenhult, 2014; Wnuk, Majid, 2014). It has been reported that in some non-Western cultures odors play a more significant role and the members of indigenous hunter-gatherer communities have a greater varietey of lexical items at their disposal to speak about smells (e.g. Burenhult, Majid, 2011; Maijd, Burenhult, 2014).

Our empirical study provides evidence that odors are difficult to express both in Estonian and in German. The data was collected from 20 Estonian native speakers (7 male, M=37.62) and 20 German native speakers (10 male, M=28.49). The method was adapted from a similar empirical study carried out on Estonian language two decades ago (Sutrop 2002) in order to compare the results of both studies. We report on commonalities and differences in odor naming in Estonian and German and identify whether there are prototypical odor terms in the two languages by applying basic colour criteria to the olfactory domain (Berlin & Kay 1969). We hypothesize that there is only one basic word in the olfactory domain for both languages — *lõhn* ‘odor’ for Estonian and *Geruch* ‘odor’ for German. In addition, native speakers of both Estonian and German were creative in describing odors more comprehensively than expected. This might be due to cultural similarities between Estonians and Germans, or other factors. In both languages, native speakers tended to use basic descriptive adjectives (e.g. in Estonian *meeldiv* ‘pleasant’, *ebameeldiv* ‘unpleasant’, in German *angenehm* ‘pleasant’, *unangenehm* ‘unpleasant’) and adjectives from the taste domain (e.g. in Estonian *magus* ‘sweet, *soolane* ‘salty’, in German *süß* ‘sweet’, *salzig* ‘salty’) and the weather/climate domain (e.g. in German *feucht* ‘humid’).

References

Berlin, B., Kay, P. (1969). Basic color terms: their universality and evolution. Berkeley: University of California Press.

Burenhult, N., Majid, A. (2011). Olfaction in Aslian ideology and language. – *The Senses & Society*, 6 (1), 19–29.

Bushdid, C., Magnasco, M. O., Vosshall, L. B., Keller, A. (2014) Humans Can Discriminate More than 1 Trillion Olfactory Stimuli. – *Science*, 343, 1370–1372

de Valk, J. M., Wnuk, E., Huisman, J. L.A., Majid, A. (2017). Odor-color associations differ with verbal descriptors for odors: A comparison of three linguistically diverse groups. *Psychonomic Bulletin & Review*, 24, 11711179

Majid, A. (2015). Cultural factors shape olfactory language. *Trends in Cognitive Sciences*, 19, 629–630

Majid, A., Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. – *Cognition*, 130, 2, 266–270.

Sutrop, U. (2002). The vocabulary of sense perception in Estonian: structure and history. Opuscula Fenno-Ugrica Gottingensia, 8. Frankfurt am Main: Peter Lang.

Wnuk, E., Majid, A. (2014). Revisiting the limits of language: The odor lexicon of Maniq. *Cognition*, 131, 125–138.

POSTER PRESENTATIONS

Schedule: Thu 10.30-11.30

Categorical Status of Adjectives in Russian Sign Language

Kirill Aksenov and Anna Klezovich

<pdf>

Realisation of Phonological Quantity in the Metrics of Vocal Music: Case of a Quantity Language

Heini Arjava
(University of Helsinki)

Keywords: Textsetting, quantity, prosody, sonority, metrics

Many researchers of both linguistics and musicology, perhaps most notably Patel (2003 & 2008) and Lerdahl & Jackendoff in their classical generative treatment in (1983), have noted the similar rhythmic premises of music and metrical language. Theoretical and corpus-based studies of textsetting have studied this music-linguistic relationship mainly from the perspective of metrical stress, focusing on Western European languages in which stress is an important and variable feature (e.g. Palmer & Kelly 1992, Halle & Lerdahl 1993, Kiparsky 2006, Proto & Dell 2013). Syllable and segment weight, on the other hand, have so far been more neglected in music linguistic study. As Western poetic and musical traditions differ conspicuously in their treatment of quantity, however, problematising the issues of phonological length are a highly relevant area in the study of textsetting.

The present study examines how phonological quantity is realised metrically in a Finnish song corpus of modern Western music. Finnish is a language where quantity is both phonologically and semantically important, and offers thus a good test environment for the study of music-linguistic quantity. The main questions include, 1) how simple one-to-one quantity alignment is manifested in songs, that is, what is the relative frequency of tolerated clashes between lengths of phonological elements and their respective musical notes; and 2) how prosodic features are used to realise hierarchical, heavily asymmetric length relations. The non-binary asymmetries are generally absent in spoken language (Ryan 2019), but abound in music (cf. the nPVI calculations in Patel & Daniele 2003, and see the 3:1 length ratio in the first word of Grid 1 below); and this complicates the selective pressure put on the prosody of the language of textsetting.

(1)

x								x		
x				x				x		
x		x		x		x		x		
x	x	x	x	x	x	x	x	x	[…]	
	x	x								
Hap.		py	birth.		day		to		you	(English)
Pal.		jon on.		ne.		a		vaan	(Finnish)	

I argue that prosodic unmarkedness and well-formedness re. quantity are aimed for (but often missed) by the combination of both phonological weight values of syllables, and the singability and prominence of segments in the nucleus-to-nucleus intervals, singability valued here by phonetic

features such as sonority, continuity, friction, and delay of release. The preliminary data indicates that in addition to general avoidance of short syllables on overlong notes (overpassing the neutral length corresponding to a linguistic semi-foot), asymmetric segment pairs also make use of a more complicated gradient sonority weight scale, as well as symmetric differences in the change in sonority. Overall, besides broadening the understanding of the under-studied field of linguistic textsetting, the study of these prosodic choices in song texts may also shed light on the general perception of prosody in spoken language.

Acknowledgments: This research has been supported by the Kone Foundation.

References

Halle, J. and F. Lerdahl (1993), A generative textsetting model, *Current Musicology* (55), 3.

Kiparsky, P. (2006), A Modular Metrics for Folk Verse, in B. E. Dresher and N. Friedberg (eds), (2006), *Formal Approaches to Poetry: Phonology and Phonetics*, Berlin: Mouton de Gruyter, 7–49.

Lerdahl, F. and R. S. Jackendoff (1983), *A generative theory of tonal music*. Cambridge, Massachusetts: The MIT Press.

Palmer, C. and M. H. Kelly (1992), Linguistic Prosody and Musical Meter in Song, *Journal of Memory and Language* 31, 525–542.

Patel, A. D. (2003), Rhythm in Language and Music: Parallels and Differences, *Annals of the New York Academy of Sciences*, 999, 140–143.

Patel, A. D. and J. R. Daniele (2003), An empirical comparison of rhythm in language and music, *Cognition* 87(1), B35–B45.

Patel, A. D. (2008), *Music, language, and the brain*, Oxford: New York: Oxford University Press.

Proto, T. and F. Dell (2013), The structure of metrical patterns in tunes and in literary verse. Evidence from discrepancies between musical and linguistic rhythm in Italian songs, *Probus*, 25(1), 105–138.

Ryan, K. M. (2019), *Prosodic Weight: Categories and Continua*, USA-OSO: Oxford University Press.

When a child's certainty turns colourful: An experimental view upon the acquisition of epistemic adverbs in child Romanian

Adina Camelia Bleotu
(University of Bucharest & ICUB)

Keywords: first language acquisition, epistemic adverbs, scalarity, colouring, truth value judgment task

Aim Relying on two experimental studies on the acquisition of epistemic modality, a colouring task and a TVJT on the modal adverbs *sigur* 'certainly' and *poate* 'maybe/ possibly' conducted on 2 groups of 25, respectively, 20 Romanian monolingual 5-year-old children (and 2 groups of 20 adults controls), the paper shows children's mastery of the meaning of epistemic modal adverbs as early as 5 years old, but uncovers the absence of scalar implicatures at this stage. The paper focuses on epistemic adverbs rather than modals, since epistemic adverbs are understudied in acquisition and also more frequent in Romanian.

Previous findings Previous experimental work on epistemic modals (Hirst & Weil 1982, Noveck 2001, Noveck, Ho & Sera 1996 a.o.) has shown that children are sensitive to the relative strength of modals from very early on, being aware a modal scale. However, 5-year-olds still have problems with understanding logical modals, and a full grasp of epistemic modality seems to be in place only at 7 years old. Similar results have been obtained experimentally for epistemic lexical items such as *(un)sure*: 4 to 6-year-old French children attribute certainty more easily than 8-year-olds (Bassano a.o. 1992, Bassano 1996). From a production perspective, corpus data shows a difference between epistemic adverbs and epistemic modals (possibly explained in terms of sentential embedding in English): children use epistemic adverbs (but not modals) before 3;00 (Cournane 2015).

Experiment 1 Experiment 1 is a colouring task where children are asked to colour certain drawn objects/ animals based on sentences with modal adverbs: *Sigur că pisica este portocalie* 'Certainly the cat is orange' (a), *Poate că pantalonii sunt roşii* 'Maybe the trousers are red' (b) and affirmative and negative control sentences. The results show that, similarly to adults, children master the meaning of the modal adverbs at stake and their strength (as also shown in Noveck, Ho & Sera 1996), often colouring the objects in a different colour for *poate* 'maybe' (46%), but never for *sigur* 'certainly' (see **Figure 1**).

Figure 1: Examples of answers in Experiment 1

Experiment 2 is a colour adaptation of an experiment on modals and scalarity Noveck (2001) previously conducted on 8-10 -year- olds. It is based on a scenario where children are told the Little Prince has magically made certain objects seem colourless but the objects actually have colour, the children are given hints and they have to say whether they agree with the sentences describing the colour of the objects. For example, children are shown a drawing with stars, as in **Figure 2,** and they are told that the third colourless star actually has the colour of the first star.

Figure 2: **Drawings used in Experiment 2**

They are asked to give truth value judgments for 6 different sentences presented in a randomized order, consisting of 1 test item (the underinformative sentence *Poate că steaua e galbenă* 'Maybe the star is yellow), 3 control items (*Poate că steaua e albastră.* 'Maybe the star is blue', *Sigur că steaua e galbenă/ albastră* 'Certainly the star is yellow/ blue') and 2 fillers (which are false). This procedure is

repeated 4 times, for each of the 4 drawings involving stars, flowers, ribbons, circles. The task reveals that children consider contextually underinformative sentences true (98.75%), unlike adults, who mostly consider them false. This suggests that children do not yet produce scalar implicatures with *poate* 'maybe/ possibly' and are significantly more logical than adults.

References

Bassano, Dominique, Hickmann, Maya & Champaud, Christian (1992), Epistemic modality in French children's discourse: *to be sure* or *not to be sure?*, *Journal of Child Language* 19(2), 389-413.

Bassano, Dominique (1996), Functional and formal constraints on the emergence of epistemic modality: A longitudinal study on French, *First Language* 16(46), 77-113.

Cournane, Ailis (2015), Revisiting the Epistemic Gap: Evidence for a Grammatical Source, *BUCLD 39 Proceedings*, Somerville, MA: Cascadilla Press.

Hirst, William, & Weil, Joyce (1982), Acquisition of the epistemic and deontic meaning of modals, *Journal of Child Language* 9, 659-666.

Noveck, Ira, Ho, Simon & Sera, Maria (1996), Children's understanding of epistemic modals, *Journal of Child Language*, 23, 3, 621-643.

Noveck, Ira (2001),When children are more logical than adults, *Cognition* 78, 165-188.

"Ok, qui d'autre na, nobody on the line right now?" A Diasystematic Construction Grammar approach to Discourse Markers in Bilingual Cajun Speech

Samuel Bourgeois
(University of Neuchâtel)

Keywords: discourse markers, bilingualism, Diasystematic Construction Grammar, Cajun French, English

Discourse markers (DMs) in bilingual speech have received much attention in language contact studies because their semantic and syntactic detachability makes them easy targets for code switching or borrowing (e.g. Muysken 2011; Neumann-Holzschuh 2008). Some of these past studies highlight the progressive adaptation of one DM system to the other as seen with English DMs taking over in American German dialects (Goss & Salmons 2000) as well as in a similar case with Chiac in New Brunswick (Chevalier 2007). Others emphasize that superstrate DMs are utilized to saliently strengthen discourse cohesion and coherence (Myers-Scotton 1993; de Rooij 2000). Though such studies of multilingual DM usage have provided rich insights, open questions remain with regard to non-salient uses and the emergence of mixed code DMs.

Using data from a bilingual radio program, this paper looks at the DM system of Cajun French (CF) and English speakers in southern Louisiana using the Diasystematic Construction Grammar (DCxG) approach as introduced in Höder (2012, 2014). This approach emphasizes the importance of analyzing contact phenomena from a multilingual standpoint (Höder 2012: 244). With the exception of only a few English DMs such as *you see*, *well* (see example 1) and to a lesser extent *of course* showing signs of being preferred over the CF equivalents, there is little indication that CF DMs are in the process of being supplanted by English ones overall.

(1) Host 2: Un tas de nouvelles ce matin?
'A bunch of news this morning?'

Host 1b: *Well*, uh, c'est comme on dit tout le temps, c'est pour ça qu'on est ici!
'Well, uh, it's like we say all the time, that's why we're here!'

(2) Host 1: Alright let's talk about Roy Motors; your Chevrolet Dealer in Opelousas for over eighty years. You know, *sont fiers de garder notre coutume, notre langage en français toujours vivant!*
'(they) are proud to keep our customs, our French language (still) alive!'

On the contrary, the radio show data in this study shows that the DM systems of Cajun bilinguals are simultaneously active to them. Along with evidence pointing to the development of a combined CF and English DM network, things are taken one step further with the development of mixed code DMs, which are used in both in CF and English. Such mixed code DMs include *na, yeah mais* and *mais yeah*. In addition to the usage of English DMs in CF and the usage the mixed-code DMs in both varieties, the radio data also shows that English and mixed-code DMs are used to mark the speaker's mid-utterance switching of languages from CF to English or English to CF (see example 2).

References

Chevalier, Gisèle (2007), Les marqueurs discursifs réactifs dans une variété de français en contact intense avec l'anglais, *Langue française* 154(2), 61-77.

de Rooij, Vincent A. (2000), French discourse markers in Shaba Swahili conversations, *International Journal of Bilingualism* 4(4), 447-467.

Goss, Emily L., & Joseph C. Salmons (2000), The evolution of a bilingual discourse marking system: Modal particles and English markers in German-American dialects, *The International Journal of Bilingualism* 4(4), 469-484.

Höder, Steffen (2012), Multilingual constructions: A diasystematic approach to common structures, in K. Braunmüller, & C. Gabriel (eds.), (2012), *Multilingual individuals and multilingual societies*, Amsterdam / Philadelphia: John Benjamins, 241-257.

Höder, Steffen. (2014), Constructing diasystems: Grammatical organisation in bilingual groups, in T. A. Åfarli, & B. Mæhlum (eds.), (2014), *The sociolinguistics of grammar*,Amsterdam/ Philadelphia: John Benjamins, 137-152.

Muysken, Pieter (2011), Code-switching, in R. Mesthrie (ed.), (2011), *The Cambridge Handbook of Sociolinguistics*, Cambridge: Cambridge University Press, 301-314.

Myers-Scotton, Carol (1993), *Social Motivations for Codeswitching: Evidence from Africa*.Oxford: Clarendon Press.

Neumann-Holzschuh, Ingrid (2008), Oui. YEAH! Zu Syntax und Pragmatik 'gedoppelter' Diskursmarker im Louisiana-Französischen, in E. Stark, R. Schmidt-Riese, & E. Stoll (eds.), (2008), *Romanische Syntax im Wandel*, Tübingen: Narr, 469-485.

The development of L2 phonetic skills following articulatory classroom-based phonetic training

Bartosz Brzoza
(Adam Mickiewicz University, Poznań)

Keywords: phonetic training, articulatory phonetics, speech perception, perceptual judgement task, L2

Laboratory-based phonetic training, especially involving high variability paradigm, has been evidenced to exert some positive influence on overall articulatory abilities (Insam and Schuppler 2015; Alshangiti and Evans 2015) and on speech perception (Hazan et al. 2005). However, the results of

experiments are mixed and studies with little or no developmental benefits are also observed (Aliaga-García and Mora 2009; Lively et al.1994). Considering various studies in the area, it seems that lab-based training generates overall better results than classroom-based training. The artificiality of the instruction and mostly immediate testing sessions might be why the inflated effect of phonetic training is reported. In real life, learners rarely learn the sounds of L2 in a lab, most of them being taught in formal classroom settings.

The current contribution is an attempt to evaluate the effectiveness of a classroom-based phonetic training. The systematic L2 articulatory pronunciation course consisted of 45 hours of training spreading over 8 months. It familiarized 24 participants with segmental phonetics and some suprasegmental elements of British English, Polish participants' L2. Participants' speech was recorded both pre- and post-training, and the speech samples were evaluated in a perceptual judgement task completed by 8 independent evaluators – practising phoneticians.

The results show an overall positive influence of instructed L2 phonetic training on the quality of selected vowels, word-final consonant voicing, and nativelikeness (statistically significant gains). As expected, the progress is less pronounced than this reported in L2 phonetic training studies employing lab methods of training. Such differences serve as points for comparison of these types of training procedures.

The results will be discussed in the light of the findings from online processing psycholinguistic tasks performed by the same group of participants (learners enrolled in a classroom-based training). Juxtaposing these types of data allows to observe how the development of phonetic skills transfers into language processing patterns during online spoken-word recognition tasks.

Acknowledgments: This research is supported by a grant from the Polish National Science Centre (Narodowe Centrum Nauki), grant no. UMO-2016/21/N/HS2/02605

References

Aliaga-García, Cristina and Mora, Joan C. (2009), Assessing the effects of phonetic training on L2 sound perception and production. In: M. Watkins, A. Rauber, and B. Baptista (eds), *Recent Research in Second Language Phonetics/Phonology: Perception and Production.* Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 2-31.

Alshangiti, Waafa and Evans, Bronwen (2015), Comparing the efficiency of vowel production training in immersion and non-immersion settings for Arabic learners of English. In: The Scottish Consortium for ICPhS 2015 (ed) *Proc. 18th ICPhS* Glasgow, online.

Hazan, Valerie, Sennema, Anke, Iba, Midori and Faulkner, Andrew (2005), Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication* 47: 360-378.

Insam, Milena and Schuppler, Barbara (2015), Evaluating the Effects of Pronunciation Training on Non-native Speech - A Case Study Report. In: A. Leemann, M.J. Kolly, S. Schmid and V. Dellwo (edss), *Trends in phonetics and phonology in German-speaking* Europe. Bern: Peter Lang, 317-330.

Leong, Christine, Price, Jessica, Pitchford, Nicola and Walter J.B. van Heuven (2018), High variability phonetic training in adaptive adverse conditions is rapid, effective, and sustained. *PLoS ONE* 13(10): e0204888.

Lively, Scott, Pisoni, David, Yamada, Reiko, Tohkura, Yoh'ichi and Yamada, Tsuneo. (1994), Training Japanese listeners to identify English /r/and/l/. III. Long-term retention of new phonetic categories. The Journal of the Acoustical Society of America 96(4):2076-2087.

Verbal Reduplication in Southern Italy. Subordination disguised as coordination?

Maria Vittoria D'Onghia
(Università per Stranieri di Siena)

Keywords: reduplication, romance syntax, unconditional clauses, verb movement, grammaticalization

Total reduplication (TR) is a widespread grammatical device consisting in reiteration of a so-called root or base used to express different functions. Although most of European languages show this mechanism, a dichotomy exists between Europe's core and periphery in relation to two main different type of reduplication. Reduplication may be, in fact, considered a grammatical device or, on the other side, a stylistic option (Stolz et al. 2011).

The aim of this paper is to investigate the syntax of grammatical verbal reduplications occurring in Apulo-Barese (AB), a dialectal variety spoken in the South of Italy, in Puglia. In AB, finite verbs can be reduplicated to express indefinite meaning.

1) Addo vè e vè, sə fescə i cumbagnə. [Bari]
   where. (he) goes. and. (he) goes RIFL. makes ART.pl. friends.
   "Wherever he/she goes, he/she makes friends»
2) [wh-] – $[VP]_{copy1}$ (-and) – $[VP]_{copy2}$ (DO)-(IO)-(SUBJ)

The reduplication structure (2) always occurs in a subordinate clause and consists in a wh- pronoun followed by a verbal reduplication. The two identical verbs always occur in indicative and can't be separate from the pronoun by subject or other items.
The function these constructions express is related to indefiniteness, since the pronoun can be alternatively translated by a wh-*ever* introducing a subordinate clause, also called Universal Concessive Conditionals (cfr. Haspelmath & König 1998; Caponigro 2003).

The reduplication sometimes appears like a verbal coordination, since you can find a conjunction in between. Furthermore, the nature of this conjunction is not clear: it disappears in specific syntactic contexts and, most importantly, it may appear like an [a] (AC > *a*) in certain dialects of the same Apulo-Barese area, whereas the traditional coordination conjunction in this area is ET > *e*.

The main scope of this paper is to show why verbal reduplication of Italian dialects could not be a case of coordination and how, by contrast, its origin could be investigated starting from a case of subordination. The following examples from Catalan and Spanish could represent the starting point to explain the reduplication occurring in Italian dialects.

3) Vagi on vagi, mai non la deixará. (Catalan)
4) Vaya adonde vaya, nunca la dejará. (Spanish)
   goes.SUBJ. where goes.SUBJ., never not her.OBJ. leaves.FUT.
   "Wherever she goes, he will never leave her"

AB verbal reduplication could be arisen from AdvP movement to the left periphery.

5) (cat.) Vagi [$_{AdvP}$ on vagi] → *On vagi vagi → (AB) Dove vai vai

References

Caponigro, Ivan (2003), *Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-linguistically*, Ph.D. dissertation. University of California, Los Angeles.

Haspelmath Martin, König Ekkehard (1998), Concessive Conditionals in the languages of Europe." In: van der Auwera, Johan (ed.), *Adverbial constructions in the languages of Europe*. Berlin: Mouton de Gruyter, 563-640.

Stolz Thomas., Stroh Cornelia & Urdze Aina (2011), *Total Reduplication: The Areal Linguistics of a Potential Universal*. (Studia Typologica 8). Berlin: Akademie Verlag.

Relativization strategies in Tomo Kan Dogon: why clausal architecture matters

Vadim Dyachkov
(Institute of Linguistics, Russian Academy of Science)

Keywords: syntax, relative clauses, information structure, Dogon languages, Tomo Kan

The paper will discuss an unusual relativization stategy in Tomo Kan language (< Dogon < Niger-Congo). The data were collected during my fieldwork in Mali and Burkina Faso in 2011 – 2018.

In relative clauses, the subject is always marked by a special particle *nɛ́*. Marking does not depend on which argument is relativized – for example, in (1) it is the object while in (2) it is the subject.

(1)	[á	nɛ́]	wɛ́rí	dòndí]	ɲàmbà-kɔ́	mà
	2SG	REL	money	hide.PFV	where-be.NONH	Q
	'Where is the money that you have hidden?'					

(2)	[à:^{n}	nɛ́	wɛ́rí	dòndí=^{n}]
	man	REL.L	money	hide.PFV=PTCP
	ɲàmbà-wɔ́		mà	
	where-be.HUM	Q		
	'Where is the man who have hidden the money?'			

In the first Tomo Kan grammar 1971] another particle is attested. However, my data suggest that using is not a common strategy in modern Tomo Kan. It is quite possible that is another instance of the Accusative marker which is used to mark prominent participants (see (Dyachkov ms, and Aissen 2003, Dimmendaal 2010) for the typology of the phenomenon). This strategy is exemplified in (3).

(3)	[àmbà	ní	ɔ̀jɔ̀	mɔ̀ndì]	càND-á:	ɲɛ̀ʔ^{n}è	yɛ̀
	god.L	REL	thing.L	create.PFV	middle-LOC	person	FOC
	gàrá:	gàʔàlé^{n}	sá				
	much	spirit	have				
	'Of all the God's creatures, the man is the most intelligent'.						

I will present a complex analysis of Tomo Kan relative clauses. Firstly, I will argue that *ɛ́* is topic marker and will test the predictions made by this analysis – for example, under certain conditions, constituents other than subject NPs can be marked by *nɛ́*. This will allow us to account for the cases when *nɛ́* is added to NPs which do not possess any subject properties as it can be shown by tests:

(4)	kò	jìnà	hùndá	nɛ́	ɲɛ̀ʔ^{n}è	wɔ́=^{n}
	DEM	house.L	yard	REL	person	be.HUM=PTCP

fú: w-è-l-è
all be.HUM-PL-NEG-PL
'People who were in the yard are't there anymore'.

Secondly, I will show that both the Accusative case and the subject role must be described in terms of information structure rather than verb argument structure. This approach will enable us to explain why the Accusative marker functions as a subject-adjoined particle like in (3) and as a Dative case marker in closely related Dogon languages. Thirdly, I will suggest a possible diachronic path for the development of *né* which might be originally a spatial adverbial. Finally, I will thoroughly describe the formal properties of relative clauses in terms of generative syntax.

Typologically, the paper will develop the ideas similar to those applied to other African (namely Mande) languages in order to account for unusual relativization strategies, cf. (Downing 1978, Keenan 1985, Lehmann 1986, Comrie & Kuteva 2011 and Nikitina 2012) for typological generalizations. I will compare my Tomo Kan data to the Mande data and discuss the differences.

The research has been supported by Russian Science Foundation, grant No. 17-78-20071.

Abbreviations

FOC – focus, HUM – human, L – low-tone contour overlay, LOC – locative, NONH – nonhuman, PFV – Perfective, PTCP – participle, Q – question particle, REL – relative clause particle

References

Aissen, Judith (2003), Differential Object Marking: Iconicity vs. Economy, *Natural Language and Linguistic Theory*, 21, 435 — 483.

Comrie, Bernard, Kuteva, Tania (2011), Relativization strategies, in Dryer, M., Haspelmath, M. (eds.), *The World Atlas of Language Structures Online*. Max Planck, 2011. Digital Library, Munich Available online at http://wals.info/supplement/8 (Supplement 8).

Dyachkov, Vadim (ms.), *A Grammar of Tomo Kan Dogon*. https://www.academia.edu/25632305/A_Grammar_of_Tomo_Kan_draft_in_English

Keenan, Edward (1985), Relative clauses, in Shopen, Timothy (ed.), *Language Typology and Syntactic Description, vol. II: Complex Constructions*. Cambridge University Press, Cambridge, 141–170.

Dimmendaal, Gerrit (2010), Differential object marking in Nilo-Saharan, *Journal of African Linguistics*, 31, 13–46.

Downing, Bruce (1978), Some universals of relative clause structure, in Greenberg, Joseph (ed.), *Universals of Human Language, vol. IV. Syntax*. Stanford, University Press, 375–418.

Léger, Jean (1971), *Grammaire dogon: Tomo-kan*. Bandiagara.

Lehmann, Christian (1986), On the typology of relative clauses, *Linguistics*, 24(4), 663–680.

Nikitina, Tatyana (2012), Clause-internal correlatives in Southeastern Mande: A case for the propagation of typological rara, *Lingua*, № 122, 319 — 334.

All (adult) Basques bilinguals? Natural consequences of early bilingualism in the phoneme inventory

Irantzu Epelde and Oroitz Jauregi

<pdf>

ꝺ

Cases denoting path in Komi: Semantic and typological perspectives

Riku Erkkilä & Niko Partanen
(University of Helsinki & University of Helsinki)

Keywords: Uralic languages, Komi, cases, path coding, semantics

In Komi-Zyrian (henceforth Komi), as is typical for Uralic languages, there is a rich system of spatial cases (Bartens 2000: 78–79). Typically there are also semantically very similar cases, with partial overlap of functions (i.e. illative, lative, and allative in Erzya; Erkkilä forthcoming). However, closely related cases also normally differ in some respects. Komi has two cases, traditionally called prolative and transitive, that primarily denote path (Bartens 2000: 107, Lytkin 1955: 145). It has been suggested that they are almost synonymous, the only difference being what kind of path they denote: prolative expresses an oblong path, whereas transitive denotes a path of any shape (Bartens 2000: 107).

The objective of this study is to investigate the parameters behind this variation in prolative marking in Komi. We look at the features of the landmark (e.g. shape, functional dimensions) and the semantics of the verb to determine what role they play in the selection of one prolative form over the other. The hypothesis is that either or both contribute to the variation. In the analysis, we make use of two concepts from cognitive linguistics, namely encyclopedic knowledge and conceptualization. Encyclopedic knowledge means that when used, a certain linguistic element activates all the knowledge the speaker has of said element, including, but not limited to, usual contexts, and speakers' own former experiences of using the element. This also applies to inflectional elements (Langacker 1991: 3–4, Tyler & Evans 2003: 14–18). Conceptualization refers to the way the speaker perceives the described situation and what linguistic means he uses to describe it (Langacker 2008: 30–33).

The preliminary view of our data points to the idea that there are significant differences between the use of prolative and transitive in Komi. Prolative is used almost exclusively to describe movement inside something (1) or on an elongated path (2). On the other hand, there is a strong tendency to use transitive when the movement occurs on a two-dimensional plane (3). The distribution of the cases also seems to vary across parts of speech.

(1) *bɨd* *kerka-ɔ va-sɔ* ***truba-jas-ɔd*** *nuɔd-ɔma*
all house-ILL water-3POSS.ACC **pipe-PL-PROL** bring-PTCP
'Water is brought to every house via pipes'

(2) *tajɔ* ***tuj-ɔd*** *ʒɔ* *muna-snɨ*
that **road-PROL** CNTR go-FUT.3PL
'They will go along that road (and not some other)'

(3) *vɔv-jas-lɨ* *jona bur* *lud* ***vɨv-ti*** *vetl-ɔm*
horse-PL-DAT much good meadow **surface-TRA** walk-PTCP
'Walking in the meadow is very good for the horses'

The study uses a written Komi corpus from Fenno-Ugrica collection (Fenno-Ugrica 2019). However, the data is supplemented with material from the Komi National Corpus (FU-Lab 2019) to gain additional depth in more marginal examples and to find comparable minimal pairs. The analysis is corpus-based and aims to explain the phenomena by evaluating all relevant instances in their naturally occurring contexts.

Path-expressing cases are often variously named (Haspelmath 2009: 515) and tend to have limited roles and frequency, so their systematic analysis, even within one language, contributes directly to the needs of general linguistics and typology.

References

Bartens, Raija (2000), *Permiläisten kielten rakenne ja kehitys*, Helsinki: Suomalais-ugrilainen seura.

Erkkilä, Riku (Forthcoming), *Ersän tulosijojen merkitykset [Semantics of Erzya directional cases]*, MA Thesis, University of Helsinki.

Fenno-Ugrica collection, The National Library of Finland, Browsed 15.1.2019, URL: http://fennougrica.kansalliskirjasto.fi/

FU-Lab Team (2019), Komi National Corpus, URL: http://komicorpora.ru/

Haspelmath, Martin (2009), Terminology of case, in Andrej Malchukov, and Andrew Spencer (eds), (2009), *The Oxford handbook of case*, Oxford: Oxford university press, 505–517.

Langacker, Ronald (1991), *Concept, image, and symbol. The cognitive basis of grammar*, Berlin: Mouton de Gruyter.

Langacker, Ronald (2008), *Cognitive grammar, a basic introduction*, Oxford: Oxford unversity press.

Lytkin 1955 = Лыткин, В. И. (ed), (1955), *Современный коми язык, учебник для высших учебных заведений, часть первая, фонетика, лексика, морфология*, Сыткывкар: Коми книжное издательство.

Tyler, Andrea, and Evans, Vyvyan (2003), *The semantics of English prepositions, spatial scenes, embodied meaning and cognition*, Cambridge: Cambridge university press.

Uncovering variation in Word Englishes: Finite vs non-finite complementation of REMEMBER with prospective meaning

Laura García-Castro
(University of Vigo)

Keywords: REMEMBER, clausal complementation, synchronic variation, corpus linguistics, GloWbE

REMEMBER is a polysemous verb that can govern finite and non-finite complement clauses (CCs), and it is precisely the meaning of the verb that constrains the distribution of CCs (Fanego 1996, Mair 2006, and Cuyckens, D'hoedt and Szmrecsanyi 2014). One of the three most frequently attested meanings is prospective, 'remember to do', which only allows *to*-infinitival CCs as a complementation option (*Cambridge Dictionary Online*, s.v. "remember", FrameNet, *Oxford Dictionaries Online*, s.v. "remember"), as illustrated in (1a).

(1a) *Please remember not to fill this empty space with some useless but "pretty" design elements*.

However, the analysis of a random sample of 12,000 hits retrieved from four varieties of English in the Corpus of Global Web-based English (GloWbE, Davies 2013; British, Indian, Bangladeshi and Sri Lankan English) reveals the existence of a new envelope of variation between finite and non-finite CCs. Compare (1b), an example from the Indian component of GloWbE, with (1a) above.

(1b) *Please remember that you shouldn't fill this empty space with some useless but "pretty" design elements.* (GloWbE IN)

Both clauses equate to definitions provided for the meaning 'remember to do', for example, 'do

something that one has undertaken to do or that is necessary or advisable' (*Oxford Dictionaries Online*, s.v. "remember") and "[a] cognizer thinks of and performs an action that is a self- or other-imposed task or some other kind of desirable behavior" (FrameNet). Specifically, in (1b) the modal *should* conveys the nuance that the task of *filling this empty space* is self- or other-imposed and/or that it is a desirable behaviour, which is traditionally expressed by means of a *to*-infinitival complement, as in (1a).

In this study, I analyse finite and non-finite CCs after REMEMBER with the meaning 'remember to do'; from a quantitative perspective I explore their distribution per variety, and qualitatively I survey a series of language internal variables (e.g. meaning of the CC verb and complexity measured in number of words) which might determine the choice of one CC-type over the other. The aim of these analyses is to answer the following research questions:

a. Is there variation in the distribution of finite and non-finite CCs after REMEMBER meaning 'remember to do' across sections of GloWbE?
b. Are there any language-internal variables determining the choice of CC after REMEMBER meaning 'remember to do'? If so, are such variables the same across sections of GloWbE?

Preliminary results show that the finite alternative is present in all the varieties of English surveyed (both L1 and L2 varieties), which lends support to the hypothesis that this is a new envelope of variation, and has been neglected in literature thus far. The qualitative analysis of the examples indicates that, as is also the case with the complementation of other meanings of REMEMBER (e.g. 'recall'), complexity factors are especially active in determining the choice of finite CCs, in line with the Complexity Principle (Rohdenburg 1996, 2006).

Acknowledgments: For their generous financial support, I am grateful to the Spanish Ministry of Economy, Industry and Competitiveness (Grants FFI2017-82162-P and BES-2015-071543) and the University of Vigo.

References:

Cambridge Dictionary Online (2018), S.v. "remember."Accessed June27, 2018. https://dictionary.cambridge.org/es/diccionario/ingles/remember

Cuyckens, Hubert, Frauke D'hoedt, & Benedikt Szmrecsanyi (2014), Variability in Verb Complementation in Late Modern English: Finite vs. Non-finite Patterns, in M. Hundt (ed), 2014, *Late Modern English Syntax,* Cambridge; New York: Cambridge University Press, 182-203.

Davies, Mark (2013), *Corpus of Global Web-Based English: 1.9 Billion Words from Speakers in 20 Countries (GloWbE).* http://corpus.byu.edu/glowbe.

Fanego, Teresa (1996), On the Historical Development of English Retrospective Verbs, *Neuphilologische Mitteilungen* 97(1), 71-79.

FrameNet (2016), "The Berkeley FrameNet Project." https://framenet.icsi.berkeley.edu/fndrupal/

Mair, Christian (2006), Nonfinite Complement Clauses in the Nineteenth Century: The Case of *remember*, in M: Hytö, M: Rydén, and E Smitterberg (eds.), (2006), *Nineteenth-century English: Stability and Change,* Cambridge: Cambridge University Press, 215-228.

Oxford Dictionaries Online (2018), S.v. "remember." Accessed June 27, 2018. https://en.oxforddictionaries.com/definition/remember

Rohdenburg, Günter, (1996), Cognitive Complexity and Increased Grammatical Explicitness in English, *Cognitive Linguistics* 7(2), 149-182.

Rohdenburg, Günter, (2006), The Role of Functional Constraints in the Evolution of the English Complementation System, in C. Dalton-Puffer, D. Kastovsky, N. Ritt, and H. Schendle (eds), (2006), *Syntax, Style and Grammatical Norms: English from 1500-2000*, Bern: Peter Lang, 143-66.

ꝺ

The quotativization of 'say' in Polish

Wojciech Guz
(The John Paul II Catholic University of Lublin)

Keywords: grammaticalization, quotatives, direct speech, verba dicendi, Polish

This corpus-based study investigates the use of Polish *mówić* 'say/speak' as a routinized strategy of marking direct quotation. The data come from Spokes – a corpus of conversational modern Polish available at http://spokes.clarin-pl.eu. The central claim is that, besides the lexical verb meaning, *mówić* is also used in casual speech as a quotative marker that shows signs of shifting to function word status. This shift, which may be termed 'quotativization', is rooted in several grammaticalization mechanisms: those of expansion, reduction, shift. Specifically, for *mówić* they are as follows:

- Limited morphological variation: historical present tense forms are essentially the only option; additionally, corpus attestations are restricted to 1st pers sg *mówię* and 3rd pers sg *mówi*, as in (1):

(1) (the Spokes corpus)
czasu mi szkoda | ostatnio sobie tak **myślę mówię** „kurde przepieprzam..."
'Such a waste of time. Recently I've been thinking MÓWIĘ "Damn, I'm letting (time) slip by"'

- Phonetic erosion and prosodic weakness (*mówię > m'wie*); the marker in (1) is barely audible/recognizable in the audio.
- Semantic expansion and bleaching: *mówi(ę)* may introduce non-actual hypothesized quotations, i.e. constructed discourse, when following non-speech verbs such as 'think' or 'wonder', as in (1).
- Category change: lexical > grammatical: *mówi(ę)* is no longer the report verb; it co-occurs with another verb – the main predicate verb – usually a verb of saying, as in (2), but also other report verbs capable of introducing quoted material, e.g. 'think', 'regret', 'write' and speech act verbs: 'apologize', 'shout', 'ask', 'apologize', etc.

(2) doktorowa mi **powiedziała mówi** „może tam by pani poszła? to może i do tego endokrynologa?"
'The doctor told me MÓWI "Maybe you could go there? And then maybe see the endocrinologist too?"'

- Syntactic flexibility: *mówi(ę)* is commonly embedded liberally in the quoted material – multiple occurrences in largely unrestricted positions. In longer stretches of discourse in particular, *mówi(ę)* has a discourse-organizational function in that it is used to re-signal the fact that the speaker is reporting the words of others.
- *Mówi(ę)* acts as an adapter particle accommodating the kind of syntactic material that could not otherwise be embedded directly under the lexical report verb; e.g. it accommodates imperative quotations next to desiderative predicates that require suitable past tense or infinitival complement clauses, as in (3):

(3)
prosił ją żeby **mówi** „weź | weź" **mówi** „ty zadzwoń"

'He was asking her to MÓWI "Go on, go on" MÓWI "You make the call"'

As has been demonstrated in the literature, 'say' is common source material for the development of quotative markers cross-linguistically (Heine and Kuteva 2002: 267–268; Güldemann 2008; Deutscher 2010). The quotativization of Polish 'say' is another case in point. So far, this path of development has not been studied closely in the Slavic domain, although similar uses of Russian *govorit'* 'say/speak' (truncated to *grit*/*ga'it*) are reported by Bolden (2004).

References

Bolden, Galina (2004), The Quote and Beyond: Defining Boundaries of Reported Speech in Conversational Russian. *Journal of Pragmatics* 36, 1071–1118.

Deutscher, Guy (2011), The grammaticalization of quotatives, in H. Narrog and B. Heine (eds), (2011), *The Oxford Handbook of Grammaticalization*, Oxford: Oxford University Press, 646–655.

Güldemann, Tom (2008), *Quotative indexes in African languages. A synchronic and diachronic survey*. [*Empirical Approaches to Language Typology* 34]. Berlin: Mouton de Gruyter.

Heine, Bernd and Tania Kuteva (2002), *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Revolutionary and mysterious: written Estonian through the 19th century

Helle Metslang & Külli Habicht
(University of Tartu)

Keywords: written language, morphosyntax, syntax, 19th century, Estonian

While in Western Europe the period from the French Revolution to World War I is regarded as the long 19th century − the time of the development of nationalism and the modernization of society (see e.g. Hobsbawm 1990), in the countries forming the western part of Czarist Russia similar developments emerged a bit later, from the beginning of the 19th century. The territory of modern Estonia entered the 19th century as a class society, where the upper class was formed by Baltic Germans and the lower class consisted of Estonian peasants. In the 19th century, the region became exposed to the ideas of the Enlightenment, important reforms of social life took place, and the economic situation and educational opportunities of the native population improved. The Estonian national awakening began, a national intelligentsia emerged, and the Estonian language gradually acquired a higher status in society.

The linguistic situation in Estonia was characterized by double diglossia (cf. Nordlund 2007). Both German and Estonian were used, but there were also distinct varieties of Estonian: standard written Estonian and spoken vernacular Estonian – the first being the prestigious variety (the H-variety, see e.g. Rutten 2016), the second associated with lower status (the L-variety). During the 19th century, the H-variety underwent substantial changes, which we chart in our study.

Our research question is: What variation can be observed in (morpho)syntactic features in Estonian texts over the course of the 19th century, and how does the variation reflect changes in the societal and sociolinguistic situation?

Our study is usage-based and relies on the framework of historical sociolinguistics. We compare the usage of characteristic indicator features in written Estonian, on the basis of influential authors

from different parts of the 19th century. For background, we also look at the usage of analogous features in German texts from the same time period.

Our research material comes from the Corpus of Old Written Estonian, the Corpus of Estonian Dialects (https://keeleressursid.ee/en/resources/corpora) and from published Estonian and German texts.

Our hypotheses are:
1) The written Estonian of the beginning of the 19th century, being an L2 of its (native German-speaking) users, was a type of collective interlanguage (cf. Makoni, Pennycook 2007);
2) The written Estonian in the middle of the 19th century can be seen as an amalgam (L3), where L1 speakers adopted interlanguage features used by L2 speakers (Thomason 2001);
3) From the last decades of the 19th century, the amalgam-like features of the written language became less prominent, and L1 features used by native Estonians gradually gained greater influence.

The indicator phenomena under study are: 1) a Finnic-type phenomenon which differs dramatically from German: DOM (see Ogren 2017), 2) complex structures which occur in both languages, but are presumably used to varying degrees: verb chains, complicated complex sentences; 3) presumably German-like word order (frame constructions).

The preliminary results support the hypotheses. In the middle of the 19th century, the overuse of the L1 features in the first group and of the interlanguage features in the second and third groups is observed. By the end of the 19th century, both types of overuse occur less frequently.

This research has been supported by the European Regional Development Fund (Centre of Excellence in Estonian Studies) and Estonian Research Council grant (PRG341).

References

Hobsbawm, Eric J. (1990), *Nations and nationalism since 1780: Programme, myth, reality*. Cambridge, Cambridge University Press.

Makoni, Sinfree, Alastair Pennycook (2007), Disinventing and Reconstituting Languages, in M. Sinfree and A. Pennycook (eds.), (2007), *Disinventing and Reconstituting Languages* (Bilingual Education and Bilingualism.), Clevedon / Buffalo / Toronto: Multilingual Matters LTD, 1–41.

Nordlund, Taru (2007), Double diglossia – lower class writing in 19th century Finland, *Multilingua* 26(2–3), 229–246.

Ogren, David (2015), Differential object marking in Estonian: prototypes, variation, and construction-specificity, *SKY Journal of Linguistics* 28, 277−312.

Rutten, Gijsbert (2016), Diaglossia, individual variation and the limits of standardization: Evidence from Dutch, in Russi, Cinzia (ed.) (2016), *Current Trends in Historical Sociolinguistics*, De Gruyter Open, 194–218.
https://www.degruyter.com/viewbooktoc/product/475491 (31.5.2019)

Thomason, Sarah Grey (2001), *Language Contact*, Edinburgh: Edinburgh University Press.

Information structure and syntactic choices in Northern Sarawak

Charlotte Hemmings
(University of Oxford)

Keywords: Austronesian; Information Structure; Voice; Word Order: Case

This poster explores the role of information structure in motivating syntactic choices in three closely related Western Austronesian languages spoken in Northern Sarawak: Kelabit, Lun Bawang and Sa'ban. The languages have symmetrical voice alternations or multiple transitive clause types that differ in the mapping of arguments to functions:

(1) Lun Bawang

a. **Actor Voice** (actor = subject, undergoer = object)

ne'	**nukat**	kelatih	uih	nalem
PFV.go	AV.dig	worms	1SG.NOM	yesterday

'I went to dig up worms yesterday'

b. **Undergoer Voice** (undergoer = subject, actor = object)

Tinukat	uih	kelatih	dih	feh
UV.PFV.dig	1SG.NOM	worms	DEM	PT

'I already dug up the worms'

They also have relatively free word order for the subject which can appear either pre-verbally or post-verbally:

(2) **Sa'ban**

a. **Súel** **nah** maan bii'
girl DEM AV.eat rice
'That girl eats rice'

b. Maan bii' **súel** **nah**
AV.eat rice girl DEM
'The girl eats rice'

Finally, in Kelabit and Lun Bawang pronominal arguments can be differentially-marked:

(3) **Kelabit**

a. Seni'er **kuh (/uih)** t=ieh
UV.see 1SG.GEN (/1SG.NOM) PT=3SG.NOM
'I saw him'

b. **Lun Bawang**

Pipag neh **ieh (/keneh)**
UV.PFV.slap 3SG.GEN 3SG.NOM (/3SG.OBL)
'He hit him'

Consequently, speakers make a number of syntactic choices when expressing a semantically transitive event – begging the question of what motivates these choices?

Cross-linguistically, word order and differential marking are known to be affected by information structure and in particular the roles of topic (what the sentence is about) and focus (the informative part of an utterance) (Lambrecht 1994, Dalrymple and Nikolaeva 2011). I argue on the basis of elicited grammaticality judgements and naturalistic corpus data that information structure also plays a role in motivating syntactic choices in Northern Sarawak. For example, word order can be used in Kelabit to place focus information initially (common in languages that follow the *principle of newsworthiness* (Mithun 1992)) as shown through negative contrast data where focus information precedes the particle *teh*:

(4) **Context: 'Did Andy hit John yesterday?' No…**

a. $[\text{Paul}]_{\text{focus}}$ teh suk nemupu' ieh
Paul PT REL PFV.AV.hit 3SG.NOM
'It was Paul who hit him (John)'

b.	[Paul]$_{focus}$	teh	suk	pinupu'	neh
	Paul	PT	REL	PFV.UV.hit	3SG.GEN
	'It was Paul that he (Andy) hit'				
c.	[nemepag	Paul]$_{focus}$	teh=ieh		
	AV.PFV.slap	Paul	PT=3SG.NOM		
	'He (Andy) slapped Paul'				

Similarly, OBL-marking in Lun Bawang may be linked to the status of the undergoer as topical (a common pattern in *differential object marking* (Iemmolo 2010)), since OBL undergoers cannot be realised clause-initially or clefted, unlike NOM undergoers:

(5) **Lun Bawang**

a.	***Keneh**	pipag	*ieh*	
	3SG.OBL	UV.PFV.slap	3SG.NOM	
	For: 'He was slapped by him'			
b.	***keneh** luk	pipag	*i=Yudan*	
	3SG.OBL	REL	UV.PFV.slap	NOM=Yudan
	For: 'He was the one Yudan slapped'			
c.	**Ieh**	luk	pipag	*i=Yudan*
	3SG.NOM	REL	UV.PFV.slap	NOM=Yudan
	'He was the one Yudan slapped'			

Consequently, information structure plays a role in motivating syntactic choices such as word order and pronominal form in the languages of Northern Sarawak. This is of interest as it shows that symmetrical voice languages can be affected by information structural considerations in similar ways to ergative and accusative languages.

Acknowledgements

I gratefully acknowledge the support of the Leverhulme Trust (ECF-2016-425) and the Kelabit, Lun Bawang and Sa'ban communities.

References

Dalrymple, Mary & Irina Nikolaeva (2011), *Objects and Information Structure*, Cambridge: Cambridge University Press.

Iemmolo, Giorgio (2010), Topicality and differential object marking: Evidence from Romance and beyond, *Studies in Language* 34(2), 239-272.

Lambrecht, Knud (1994), *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*, Cambridge: Cambridge University Press.

Mithun, Marianne (1992), Is Basic Word Order Universal, in Doris L. Payne (ed.), (2009), *Pragmatics of Word Order Flexibility*, Amsterdam: John Benjamins, 15-61.

ø

Mari kinship terms from the typological, genealogical and areal point of view

Sampsa Holopainen, Veronika Milanova, Jeremy Bradley & András Czentnár
(University of Helsinki, University of Vienna, University of Vienna & Eötvös Loránd University, Budapest)

Keywords: Mari, Uralic, Kinship terms, Language contact, Chuvash

The purpose of this presentation is to analyze the etymology and development of kinship terms in Mari, a Uralic language spoken in the Volga-Kama Region. We aim to ascertain how the current Mari system and terminology of collateral relatives evolved: which terms are inherited from Proto-Uralic, what their semantic development has been, and how areal influences have shaped the Mari kinship system. An equally important aim is the description of Mari collateral relative system from the point of view of typology of kinship terminologies.

The study of kinship terminology in Uralic linguistics is far from advanced. Despite complex and diverse kinship systems found among the Uralic languages (especially those of Volga-Kama Region and Siberia), there are few studies of this topic. Udmurt kinship terminology is well studied and described by Szíj (1998); there is some work on Mordvin kinship terms as well (Mészáros 2001), but there are no modern studies of Mari kinship terms.

The most conspicuous feature of Mari kinship terminology is relative age distinction for collateral relatives, i.e. specific terms for older and younger siblings (also older and younger siblings of parents and spouses). This is typical for both the Uralic and Turkic languages of the Volga-Kama Region but unusual from a European perspective. We argue that this peculiarity of Mari kinship terminology is most likely contact-induced from neighbouring Turkic languages, Chuvash and Tatar (which have also provided numerous loanwords into Mari), rather than an inherited Uralic feature. This is supported by the fact that some words denoting siblings in Mari are loanwords from Tatar, Chuvash or earlier stages of Chuvash-type (Bulgar) Turkic, e.g. *šol'o* 'younger brother' < Chuvash *šăллăм* 'my younger brother' (Räsänen 1920: 13).

Setälä (1900) argued on Turkic influence on this system in the Uralic languages of the Volga-Kama Region and against the reconstruction of the system to Proto-Uralic, but this question has not been properly studied since then. Relative age distinction is notably absent for Uralic languages (such as Komi) that have not been in intensive contact with Turkic. For this purpose we will compare Mari sibling terminology with sibling terminologies in Eurasia and in other parts of the world. Through a critical analysis of individual Mari etymologies and comparison to terms in the neighbouring Udmurt, Mordvin and Chuvash languages we also aim to shed light on more precise dating of the genesis of the relative age distinction in those languages.

Moreover, we will contrast the situation of Mari with sibling terminologies in modern Iranian and Indo-Aryan languages. Relative age distinction is not an originally Indo-European feature, but is found in some Iranian varieties that have been subject to strong contact influence. Most words denoting younger and especially older siblings are either of Turkic (or Mongolic) or of South Asian origin (see examples in Ansari et al. 2015, and Milanova et al. 2018: 7–9 with references). The Iranian and Indo-Aryan situation can thus be seen as a typologically similar pattern of linguistic and cultural transfer.

References

Ansari, Zakia R., Srishti Shabnam, and Brajesh Priyadarshi (2015), Kinship terms of Urdu in Mysore and Lucknow varieties: a comparative study, *Indian Linguistics* 76 (1–-2), 113–130.

Mészáros, Edit (2001), A mordvin rokonságnevek eredet szerinti vizsgálata, *Néprajz és Nyelvtudomány* 41, 169–184.

Milanova, Veronika, Sampsa Holopainen, and Jeremy Bradley (2018), Contact phenomena in IE kinship and social terms and beyond. [handout for the international colloquium *Loanwords and substrata in Indo-European languages*, Limoges, 4th–7th June 2018] [kinship.mari-language.com/limoges.pdf, kinship.mari-language.com/limoges-table.pdf]

Räsänen, Martti. (1920), *Die tschuwassischen Lehnwörter im Tscheremissischen*. Helsinki: Société Finno-Ougrienne.

Setälä, Eemil N. (1900), I. N. Smirnow's Untersuchung über die Ostfinnen, *Journal de la Société Finno-Ougrienne* 17, 17–28.
Szíj, Enikő. (1998), *Термины родства и свойства в удмуртском языке*. Budapest: MTA.

Expressing certainty and uncertainty in child and child-directed speech in three typologically different languages

Victoria V. Kazakovskaya, Reili Argus & Sigal Uziel-Karl
(Institute for Linguistic Studies (Russian Academy of Sciences), Tallinn University & Achva Academic College)

Key words: Epistemic modality, first language acquisition, Russian, Estonian, Hebrew

The study aims to describe the acquisition of certainty and uncertainty expressions (epistemic markers) in child speech, as compared to child-directed speech, in three typologically different languages – Estonian, Russian and Hebrew. Our goal is to examine to what extent language typology affects the means that children and their main caregivers use for expressing epistemic semantics.

Naturalistic speech samples of typically developing monolingual children (three per language) recorded from ages 1;3 to 6;2 were analysed. 280,5 hours of recordings were transcribed and morphologically coded using CHILDES (MacWhinney 2000). Then the certainty and uncertainty expressions in the data of both children and their caregivers were isolated and coded for part-of-speech, degree of confidence (high, middle, low), position in utterance (initial, middle, final) and pragmatic function. The age at which expressions of certainty and uncertainty emerge, their diversity (*lemmas*) and frequency (*tokens*) in spontaneous speech, as well as similarities and differences in epistemic development between the children and caregivers, were also taken into consideration.

A developmental analysis of the epistemic markers revealed that they emerge at the multi-word utterance stage, after the expressions of deontic and dynamic modality (Stephany, Akcu-Koç 2020). Regardless of typological features of the language acquired, all children under observation start with uncertainty markers, cf., however, (Hickmann, Bassano 2016). In all three languages epistemic markers used by children are mostly adverbs (e.g., *vist* and *äkki* 'maybe', *muidugi* 'of course' in Estonian; *betax* 'sure', *ulay* 'maybe', *kanir'e* 'apparently, it seems like', *batu'ax* 'certain' in Hebrew) and parenthetical modal words (e.g., *naverno(e)* 'probably', *možet (byt')* 'maybe', *konečno* 'of course' in Russian). Marking their utterance (proposition) epistemically, children start by evaluating objective situations, viz. by guessing the location or identity of objects. Later they begin to reflect on the cognitive (viz. psychological, mental) state of others and still later of themselves (see children's theory of mind). The functions of epistemic markers develop from pure epistemic ones (primary) towards secondary functions (e.g., conversational strategy and politeness). Differences between speakers of the languages under investigation lie in the frequency of usage of epistemic markers and the degree of confidence. Particularly, in Russian, epistemic marking begins with high-degree markers in both uncertainty and certainty domains, in Estonian, with middle-degree, in Hebrew, with middle uncertainty and high certainty.

The results suggest that both frequency and diversity of epistemic markers in the child-directed speech and cultural factors like politeness strategies may have greater effect on the acquisition of these markers across languages than the typological characteristics of the language.

References

MacWhinney, Brian (2000), *The CHILDES project. Tools for analyzing Talk. Electronic edition. Part 1: The CHAT transcription format*.

Hickmann, Maya and Bassano, Dominique (2016), Modality and Mood in First Language Acquisition, in J. Nuyts, and J. van der Auwera (eds), (2016), *The Oxford Handbook of Modality and Mood*, Oxford: Oxford University Press, 430-447.

Stephany, Ursula and Akcu-Koç, Ayhan (eds), (2020), *Development of Modality in First Language Acquisition: A Cross-linguistic Perspective*. Berlin: de Gruyter (in press).

What happens when the future disappears? A case study from Komi-Zyrian

Alexandra Kellner & Niko Partanen
(Unaffiliated & University of Helsinki)

Keywords: Komi, Uralic languages, TAM-systems, language change, dialectal variation

Although Komi-Zyrian has morphologically marked future tense in the third person, the Udora dialect lacks this opposition (Bartens 2000: 179, Sorvacheva et. al. 1990: 68). As a well described and fairly recent change in this dialect, the topic has received some earlier attention and analysis, e.g. by Ponaryadov (2004).

The loss of a tense distinction is arguably a large change for the TAM system, meaning that the Udora dialect of Komi, when compared to the other dialects, which all still maintain the distinction, offers an almost laboratorial context for studying the adaptations and restructurations initiated by the loss of this opposition. Among these changes are neutralization of the distinction, and when future is explicitly expressed, a complete shift to the auxiliary verbs.

To substantiate our study, we use a corpus of Udora dialect data that contains both archival and contemporary data. The earliest data in our corpus was collected in the late 19th century, which also offers a historical dimension. Comparative data is taken especially from Ižma (Blokland et al.), often seen as a direct sister dialect of Udora.

The example below shows temporally identical verbs with different conjugations, indicated in bold:

vot	*stav-ɨs*	*lokt-**əni***	*i*	*stav-ɨs*	*prəʃajts-**asni***	*gort*
well	all-3SG	come-3PL.PRES	and	all-3SG	praise-3PL.PRES	grave
gəgər-ɨs	*pukt-**asni***	*ɟeŋga*				
around-3SG	put-3PL.PRES	money				

"Well, everyone comes (pres. marked) and pays their respects (fut. marked) around the grave, places (fut. marked) money there" (Vazhgort 2013)

Our research shows that the choice of tense marker is connected to the derivational type of the verb, which raises the question of the degree to which the nullified present–future markers have shifted to indicate derivational functions. The following example shows two derivations of the same root verb, a causative and a frequentative, occurring with their respective conjugations:

Ivan	*tsarevitɕ*	*məd-asɨlas*	*suvt-**as**,*	*a*	*kar*	*sulal-**ə***
Ivan	Tsarevich	tomorrow	get_up-3SG.PRES	but	city	stand-3SG.PRES

"Ivan Tsarevich gets up (fut. marked, lit. stands up) next morning, and the city is (still) standing (pres. marked)." (Fokos-Fuchs 1916: 96)

There are also morphemes that are homonymous but have distinct origins and functions. One of these is the *-as* suffix used in Standard Komi in the third person possessive form of gerunds:

mun-ig-ən
go-GER-INSTR
'while going'
*mun-ig-**as***
go-GER-INSTR.3SG
'while (he is) going'

However, in Udora, only a new reanalysed form occurs, seemingly as an attempt to avoid the use of *-as* in such a function. This is an example where a part of Udora morphology that is not related to tense marking has anyway undergone a shift launched by the disappearance of future tense, resulting in the following form:

mun-ig-ən-is
go-GER-INSTR-3SG
'while (he is) going'

Since such a form is not attested in any other Komi varieties, it can be confidently seen as an Udora innovation. The innovations and adaptations such as these, although not drastically changing the language in itself, show that an individual change can ripple through various layers of morphosyntax, resulting in many ways in a distinct system with changes on multiple levels.

References

Bartens, Raija (2000), Permiläisten kielten rakenne ja kehitys, Helsinki: Suomalais-Ugrilainen Seura.

Blokland, Rogier; Chuprov, Vassily; Fedina, Maria; Fedina, Marina; Dmitriy Levchenko; Niko Partanen and Rießler, Michael (2019), Iźva Komi Documentation Project corpus, Funded by Kone Foundation, URL: https://hdl.handle.net/1839/00-0000-0000-001B-99BC-F@view

Fokos-Fuchs, Dávid Rafael (1916), Zürjén szövegek, Finnugor füzetek.

Ponaryadov = Понарядов, Вадим (2004), О глагольных окончаниях -ö и -ас в удорском диалекте, Вопросы коми диалектологии, Труды Института языка, литературы и истории Коми НЦ УрО РАН; Вып. 65, pp. 104–12.

Sorvacheva et. al. = Сорвачева, В. А., and Безносикова. Л. М. (1990), Удорский диалект коми языка, Издательство "Наука".

Proper names in Daghestanian languages

Zaira Khalilova
(The G. Tsadasa Institute of Language, Literature and Art of Dagestan Scientific Centre, RAS, Makhachkala)

Keywords: proper names, kinship terms, possessor, associative, Daghestanian languages

The class of nouns is typically divided into subclasses of common nouns and proper nouns (proper names). However there is no clear universal boundary between these two subclasses, as morphosyntactic tests and features that distinguish proper names from common nouns are language-specific. Proper names, more specifically personal names, in the Daghestanian languages Chamalal,

Avar, Bezhta, and Khwarshi show unique features (shared with kin terms) that distinguish them from common nouns, namely the presence of special morphology in marking possession and association.

Depending on the language in the adnominal possessive construction a possessor which is a personal name or a kinship term can be marked either with a special genitive or a combination of a basic plural marker and the ordinary genitive, or simply marked with the ordinary genitive. Such possessive constructions express common possession and association.

The use of a special genitive with personal names is found in the Andic languages. For example, in Chamalal the possessor is marked either with the ordinary genitive suffix *-λ* when it is a common noun, but by an agreement affix agreeing with the head (possessum) when it is a personal name, a typologically unusual possibility (e.g. *Aħmadi-b xaj* 'Axmed's dog', *Aħmadi-w Murad* 'Murad, son of Axmed').

In the Tsezic languages personal names as possessors can be marked with the combination of a basic plural marker and the ordinary genitive in order to express association. For example, Bezhta *Ramazan-la-s Muħamad* means 'Muhammad, son of the Ramazan family', also Avar *Ramazanil-az-ul jas* 'daughter of the Ramazan family'. When the possessor is marked with an ordinary genitive, the construction expresses only possession and associative meaning is lost, e.g. Avar *Ramazan-i-l jas* 'Ramazan's daughter' (Khalilov & Khalilova 2007).

Some Tsezic languages have gone further in marking specific semantic relations, e.g. possessor has a dedicated attributive suffix marking family relations of son and daughter, with the associative meaning being lost, e.g. *Karim-zo Raisat* 'Raisat, the daughter of Karim' (Khalilova 2009). One feature rare even in Daghestanian languages is that Bezhta uses special morphology, the suffix *-ko*, only with male proper names to denote 'daughter of X', e.g. *Ramazan-la-ko* 'the daughter of Ramazan'.

Special morphology used with personal names is equally possible with all kinship terms as possessors, e.g. Bezhta *isi-la-s öžö* 'son of the sister's family', Avar *was-az-ul jas* 'daughter of the son's family'.

In Daghestanian languages personal names and kinship terms go together in having special referential properties in discourse. Personal names and kin terms are always animate human and semantically definite. There is a strong cross-cultural tendency for kin terms and personal names to function as pragmatic alternants (Fleming & Slotta 2016). Personal names and kin terms share the main functions of referring to and addressing human individuals (Helmbrecht to appear).

The aim of this paper is to show that personal names and kin terms in the Daghestanian languages have specific semantic and pragmatic features that are reflected in distinctive morphology.

References

Fleming, L. & Slotta, J. (2016), The pragmatics of kin address: A sociolinguistic universal and its semantic affordances. *Proceedings of CLS 51 (2015),* Chicago Linguistic Society, 165-179.

Helmbrecht, Johannes. To appear. On the form and function of proper names – a typological perspective. In: *Sprachtypologie und Universalienforschung (Language Typology and Universals) STUF (Special issue on Referentiality)*.

Khalilov, M. & Khalilova, Z. (2007), Das awarische Personennamensystem. In Andrea Brendler, Silvio Brendler (eds.), *Europäische Personennamensysteme: ein Handbuch von Abasisch bis Zentralladinisch*. Hamburg: Baar, 80-90.

Khalilova, Z. (2009), *A Grammar of Khwarshi*. PhD dissertation, Leiden University.

Phonetics of the Aromanian Dialect of Selenica – A Preliminary Analysis

Anastasia Kharlamova
(Institute for Linguistic Studies of the Russian Academy of Sciences)

Keywords: phonetics, Aromanian language, Balkan linguistics, dialectology, consonants

This paper is dedicated to our study of the Aromanian dialect in Selenica, Albania. It is based on audio recordings gathered during our expedition to the town in June 2018. The recordings include spontaneous and replies to our questionnaire from four informants – three men (a 53-year-old priest, a 63-year-old folk music conductor and a 70-year-old retired teacher) and a woman who did not give her last name, age or occupation.

The Aromanian dialect of Selenica has never been studied from a purely linguistic point of view before, usually being mentioned in sociolinguistic papers and statistic tables. Our goal is to describe the phonetics and phonology of the dialect and to determine to which dialect group it belongs.

The materials we have now show an eclectic picture when it comes to dialectal features (for details on the dialectal structure of Aromanian, see (Narumov 2001)). There are some features characteristic to the Northern Aromanian dialects – such as, for example, the clear differentiation between dental and palatoalveolar affricates; however, the [mn] consonant group is retained and does not get replaced with [m:] (as could have been expected in the Northern Aromanian dialect zone). Furthermore, there are some cases of vowel nasalisation, which is, again, a Northern (Farşerot) feature, while we also believe we can speak of two different mid-row vowels in our data, which is characteristic to the Southern dialects.

There are also some phonetic features that are not usually described as common to Aromanian. A highly noticeable one is the presence of palatal affricates [cç] and [ɟʝ], which often occur in Albanian borrowings, such as [ʃocçata] < *shoqata* ‘society’. [cç] and [ɟʝ] have previously been found in the Albanian and Aromanian speech of Elbasan Aromanians (Kharlamova 2018).

Some wordforms demonstrate change of affricates into fricatives, such as [mesovo] ‘Metsovo (a town in Epirus and a major center of Aromanian culture)’, and reduction of consonant clusters, for example, [kosantsə] ‘Constanța’ and [fes’ival’] ‘festival’. The reduction of affricates’ plosive phase is highly prominent in the Turia dialect of Aromanian, spoken in Pindos, Greece (see the full description of the dialect and the text collection in (Bara et al. 2005), the analysis of affricates in (Kharlamova 2017)).

The research on Selenica Aromanian has only started, with this being the initial stage, and the combination of Northern and Southern Aromanian features already found in the materials shows that it can turn out to be part of a previously unmentioned intermediate dialect group between the two main dialect zones.

Acknowledgments
Funded by RNF grant 19-18-00244.
We would like to express the warmest thanks to our informants: Fr. Spiro Bulika, Pascal Kotta, Kozma Mitro and Gjina.

References

Bara, Maria, Kahl, Thede, and Sobolev, Andrey (2005), *Yuzhnoarumynskiy govor sela Turya (Pind)* [The South Aromanian dialect of the village Turia (Pindos)]. München: Biblion Verlag.

Kharlamova, Anastasia (2017), Affrikaty v arumynskoy spontannoy rechi [Affricates in Aromanian spontaneous speech] in: Monakhov, S., Vasilyeva, I., Khokhlova, M (eds), (2017), *Proceedings*

of the 45th International Philological Conference (IPC 2016), Paris; Amsterdam: Atlantis Press, 298–300.

Kharlamova, Anastasia (2018), *Affrikaty v spontannom rechevom potoke na materiale albanskogo i arumynskogo yazykov* [Affricates in spontaneous speech flow with data from the Albanian and Aromanian languages]. MA thesis presented at Saint-Petersburg State University, Department of General Linguistics.

Narumov, Boris (2001), Arumynskiy yazyk/dialekt [The Aromanian language/dialect] in: Zhdanova, T. et al. (eds), (2001), *Yazyki mira. Romanskie yazyki* [Languages of the world. Romance languages]. Moscow: Academia, 636–656.

Comitative constructions in Kazym Khanty

Irina Khomchenkova
(Lomonosov Moscow State University; Vinogradov Russian Language Institute, RAS)

Keywords: comitative, coordination, associative plural, Khanty, Uralic

This paper deals with a comitative marker *piλa* in Kazym Khanty (< Ob-Ugric < Uralic), which is used (inter alia) in two constructions – with a singular verb form and with a non-singular one (dual if one refers to two items and plural otherwise):

(1) *pet'a-jen* *wənt-a* [*maša-jen* *piλ-a*] *măn-əs* / *măn-s-əŋən*
P.-P.2SG forest-DAT M.-P.2SG with-DAT go-PST.3SG go-PST-3DU
'Peter went to the forest with Mary'.

I will analyze these constructions, comparing them with a coordinating construction. The data come from my fieldwork in the Kazym village (Khanty-Mansi Autonomous District, Russia) in 2018.

The first construction is a genuine comitative construction – a "morphosyntactic construction, used to express a non-obligatory participant set in a given situation, such that the predicate denoting it is not repeated more than once; the individual participants are expressed separately; the expressions denoting these participants differ in structural rank" (Arkhipov 2009: 224). The second construction is an instance of a coordinating comitative construction – a construction "exhibiting some (or all) features of coordination proper" (Arkhipov 2009: 234). To compare these constructions, I used standard tests for the analysis of comitatives (see, e.g. (Dyła, Feldman 2008)).

On the one hand, the construction with a dual verb form is different from coordination. For example, in both comitative constructions the ComP can be extracted in questions (2) contrary to coordination (3), which means that in the construction with a dual verb form the central NP and the ComP do not form a constituent.

(2) [*χʉj* *piλ-a*] *maša-jen* *wənt-a* *măn-əs* / *măn-s-əŋən*?
who with-DAT M.-P.2SG forest-DAT go-PST.3SG go-PST-3DU
'With whom Mary went to the forest?'

(3) **χʉj* *pa* *maša-jen* *wənt-a* *măn-əs* / *măn-s-əŋən?*
who ADD M.-P.2SG forest-DAT go-PST.3SG go-PST-3DU
Expected: 'With whom Mary went to forest?'

On the other hand, the construction with a dual verb form is different from the genuine comitative. For example, stative verbs cannot be used in genuine comitative constructions (4) unlike the constructions with a dual verb form (4) or coordination (5).

(4) *maša-jen* *tərəm* *aś-en-a* [*pet'a-jen* *piλ-a*] *ewəλ-λ-əŋən* / **ewəλ-əλ*
M.-P.2SG sky father-P.2SG-DAT P.-P.2SG with-DAT believe-NPST-3DU believe-NPST.3SG
'Mary and Peter believe in God'.

(5) [*maša-jen* *pa* *pet'a-jen*] *tərəm* *aś-en-a* *ewəλ-λ-əŋən*
M.-P.2SG ADD P.-P.2SG sky father-P.2SG-DAT believe-NPST.3DU
'Mary and Peter believe in God'.

To sum up, the construction with a dual verb form has features both of comitative and coordinating constructions, and the central NP does not form a constituent with the ComP. The question arises which element in this construction is responsible for licensing the dual verbal agreement. I claim that it is licensed by a zero associative plural marker on the central NP, which is supported by the fact that in Kazym Khanty associative plural can be expressed with a zero marker (Sokolova 2018). Similar analysis has been proposed for Tzotzil < Mayan (Aissen 1989), so the relation between associatives and comitatives is important from the typological point of view (see also (Vassilieva 2005)). In the talk, I will give a more thorough analysis and compare my data with the data of some other Uralic languages.

Acknowledgments
The research has been supported by RFBR, grant № 19-012-00627.

Abbreviations
2, 3 – 2, 3 person, ADD – additive, DAT – dative, DU – dual, NPST – non-past, P – possession, PST – past, SG – singular

References
Aissen, Judith L. (1989), Agreement Controllers and Tzotzil Comitatives, *Language* 65(3), 518–536.

Arkhipov, Alexandre (2009), Comitative as a cross-linguistically valid category, in A. Arkhipov and P. Epps (eds), (2009), *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, Berlin / New York: Mouton de Gruyter, 223-246.

Dyła, Stefan, Feldman, Anna (2008), On comitative constructions in Polish and Russian. In G. Zybatow, L. Szucsich, U. Junghanns & R. Meyer (eds), (2008), *Formal Description of Slavic Languages*. Bern, Switzerland: Peter Lang. Available from: https://msuweb.montclair.edu/~feldmana/publications/fdsl.pdf

Sokolova, Irina S. (2018), Associativy v Kazymskom dialekte khantyjskogo jazyka [Associatives in Kazym Khanty], in D. Mishchenko (ed), (2018), *XV Conference on typology and grammar for young scholars (St. Petersburg, 22–24 November 2018). Book of abstracts*, St. Petersburg: ILS RAS, 156-158.

Vassilieva, Maria B. (2005), *Associative and Pronominal Plurality*, PhD dissertation, Stony Brook University.

Huayiyiyu, a Chinese corpus to study Asian languages in the 15C – 16C

Ji-Eun Kim

Nutshell: This study introduces a way to study Historical Phonology of Asian languages by revisiting *Huayiyiyu,* Chinese-Barbarian dictionary published to train official interpreters in the 15th – 16th centuries.

Background: Historical Linguistics has been relying on comparative method and internal method (Shields, 2010), while they have considerable limitations at the same time. The comparative method has relative temporal limitations, sociohistorical limitations, linguistic domain limitations, and "delicacy" limitations (Harrison, 2008). The internal method shares the limitation of the comparative methods, when the gap between orthography and phonology could be another problem when adopting the data from a single language. It is because the orthography of a language is not the exact transcription, but the mnemonic of a language (Lass, 2015). However, there has been little study on new methods to overcome the limitations of these two primary methods.

Data: *Huayiyiyu* transcribes 12 different spoken languages in Asia in Chinese characters, including either disappeared or endangered languages such as Cham and Jurchen.

Method: *Huayiyiyu* is considered to have a system of phonetic notation just like IPA. The characters were used with different phonetic values from those used to mnemonically record Chinese language. Therefore, this study examines the phonetic value and system of each 'Chinese phoneme' used in *Huayiyiyu* by following four steps. **(1)** Digitize *Huayiyiyu* aligned with its *yinxi* information. **(2)** Select and append reliable reconstructions on each word list (language) into the dataset of **(1). (3)** Find out transcription principles only based on stable phonologies of each language. **(4)** Reconstruct the phonetic value of each phoneme based on the examined principles.

Results: (1) Reliable studies were selected and compiled, which will be published and shared. **(2)** a) The morphemes were transcribed phonetically instead of phonemically, b) each morpheme was transcribed in the same manner for each of the wordlist, c) elements which do not fit into the Chinese syllable structure can be the object of *chumgi*. d) elements satisfying the first principle should be voiced, constituent or released sound, and e) phonetic scope of each *yinxi* transcribed in the *Huayiyiyu* is as [1]. **(3)** For example, the phonetic value of 15th century Korean: There was evidence to support the existence of 'ㅱ[ɱ]' at the coda and possible phonetic values for each vowel were presented in a schematic diagram.

Discussion: This study **(1)** partially overcomes the limitations of existing methods, **(2)** enables languages with few historical resources to be studied in Historical Phonology, **(3)** creates synergy between historical studies on each language, and **(4)** can be applied to study the disappeared or endangered languages. However, we need to further work on (1) sharing the dataset to be edited by the historical linguists of each language and (2) confirming phonological stability of each language.

References

Harrison, S. P. (2003). On the limits of the comparative method. *The handbook of historical linguistics*, *213*, 243.

Lass, R. (2015). Interpreting Alphabetic Orthographies. In *The Oxford handbook of historical phonology*.

Shields, K. (2010). Linguistic Typology and Historical Linguistics. In *The Oxford Handbook of Linguistic Typology*.

ꝺ

Basque relational marker: a diachronic study

Dorota Krajewska
(University of the Basque Country, UPV/EHU)

Keywords: Basque, morphology, syntax, diachrony

Basque relational marker *-ko* derives adjectival modifiers from: (a) local-case-marked nominals (*etxera* 'to the house' > *etxerako bidea* 'road home'); (b) instrumental-marked nominals (*zurez* 'of wood' > *zurezko etxea* 'wooden house'); (c) adverbs and postpositional phrases (*beti* 'always' > *betiko aitzakia* 'the usual excuse' or *literaturari buruz* 'about literature' > *literaturari buruzko liburua* 'a book about literature'); (d) bare nouns (*etxe* 'house' > *etxeko lana* 'homework'); and (e) determiner-less noun phrases (*ospe handi* 'great fame' > *ospe handiko gizona* 'a man of great fame').

This paper examines the functions and morphosyntax of *-ko* in a corpus of 16th to 20th century texts (30 texts with 3800 examples of *-ko* were analyzed). I propose that changes observed show the diachronic dimension of possession-modification scale proposed by Nikolaeva & Spencer (2010): inalienable possession > alienable possession > nominal modification > adjectives. The typical function of Basque relational marker is to express nominal modification, but over time some phrases with *-ko* have gradually acquired several properties of canonical adjectives. This change is visible in morphology and syntax.

As regards morphology, in its most prototypical functions *-ko* can be considered contextual inflection (Booij 1996), but in some of its uses, the suffix has moved towards derivation on the inflection-derivation continuum. For instance, the semantics of some *-ko* modifiers is no longer fully predictable from their base. Also, some phrases with *-ko* start to behave as adjectives by acquiring comparative and superlative forms, which become common in the 19th century texts.

In syntax, even though in the modern standard language the modifier is placed before the noun, in old texts the opposite ordering is also attested. This has been usually put down to the influence of Romance languages. Here, however, I argue that the behavior of these modifiers can be seen as another example of them moving towards adjectives. The placement of a *-ko* phrase depended on its similarity with adjectives: intuitively, the more prototypical adjective-like it was, the more it was likely to appear after the noun, in the position where canonical qualitative adjectives go in Basque. To clarify this, I use the concept of anchoring and non-anchoring relations (Koptjevskaja-Tamm 2002) and show that only non-anchoring modifiers could be used postnominally. Nevertheless, not all did: only some frequently used *-ko* phrases have lexicalized and become more like adjectives. Thus, the word-order change with the relational marker was not generalized and happened in a piecemeal fashion, but following the possession-modification scale.

References

Booij, G., 1996, «Inherent versus contextual in-ection and the split morphology hypothesis», in *Yearbook of morphology* 1995, Springer, Dordrecht, 1–16.

Koptjevskaja-Tamm, M., 2002, «Adnominal possession in the European languages: form and function», *STUF-Language Typology and Universals*, 55(2), 141–172.

Nikolaeva, I. & A. Spencer, 2010, «The Possession-Modification Scale: A universal of nominal morphosyntax», URL http://citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.611.158&rep=rep1&type=pdf, unpublished manuscript.

Differential Marking on Locative Phrases in Mandarin Chinese

Pei-Jung Kuo
(National Chiayi University)

Keywords: preposition, locative PP, TP domain, VP domain, internal topic

Djamouri et al. (2013) and Paul (2015) observe that when locative PPs in Chinese appear in the TP internal topic position, the preposition *zai* ('at') is optional, as in (1a). However, if locative PPs emerge in the VP-adjoined position, *zai* is obligatory, as in (1b). Note that the negation *bu* ('not) and modal *neng* ('can') mark the boundary between TP and VP.

(1) a. Ni [(zai) woshi li] bu neng fang dianlu.
You at bedroom inside not can put electric.stove
'You cannot put an electric stove in the bedroom.'
b. Ni bu neng [*(zai) woshi (li)] fang dianlu.
You not can at bedroom inside put electric.stove

To explain this puzzle, I propose that the locative NPs/PPs in Mandarin Chinese can be base-generated in the TP and the VP domain respectively (Ramchand and Svenonius 2014 and Takamine 2017). This proposal is supported as follows. Firstly, Paris (1979) shows htat while the path PP in (2) or commitative PP in (3) are restricted to the VP domain, locative PPs can emerge in the TP domain as in (1a).

(2) Zhangsan (*cong yinhang) bu neng (cong yinhang) jie qian.
Zhangsan from bank not can from bank borrow money
'Zhangsan cannot borrow money from the bank.'
(3) Zhangsan (*gen Lisi) bu neng (gen Lisi) yiqi zuo che.
Zhangsan with Lisi not can with Lisi together take car
'Zhangsan cannot take the car with Lisi together.'

Secondly, two locative PPs in different domains imply the possible co-occurrence of both phrases. This prediction is borne out in (4).

(4) Xuesheng-men [zai jaoshi li] bu neng [zai zuozi shang] tuya.
student-PLU in classroom inside not can at table up scrawl
'Students cannot scrawl on the tables in the classroom.'

Moreover, I propose that the preposition *zai* is a kind of differential marking which is observable in the TP domain. The preposition *zai* is sensitive to the animacy and definiteness features of the following locative NP. The preposition *zai* becomes obligatory when the following locative NP is [+animate] as in (5). When the locative NP is an indefinite one as in (6), the sentence becomes ungrammatical. And the preposition is optional with a definite locative NP in (7) or a generic locative NP in (1a).

(5) Ni [*(zai) Zhangsan lianshang] bu neng zhi hua hua.
You at Zhangsan face.up not can only draw flower
'You cannot draw only flowers on Zhangsan's face.'
(6) *Ni [zai yi-jian woshi li] bu neng fang dianlu.

You at one-CL bedroom inside not can put electric.stove
'You cannot put an electric stove in a bedroom.'

(7) Ni [(zai) zhe-jianwoshi li] bu neng fang dianlu.
You at this-CL bedroom inside not can put electric.stove
'You cannot put an electric stove in this bedroom.'

The above patterns are reminiscent of differential object marking (Bossong 1985, Aissen 2003, Seržant and Witzlack-Makarevich 2018). In addition, since the TP-domain situated locative NP is in the internal topic position (Paul 2005), this further links the current differential locative marking to information structure as in Dalrymple & Nikolaeva (2011).

References

Aissen, Judith (2003), Differential Object Marking: Iconicity vs. Economy, *Natural Language and Linguistic Theory* 21, 435-483.

Bossong, Georg (1985), *Differentielle Objektmarkierung in den Neuiranischen Sprachen*, Gunter Narr.

Dalrymple, Mary & Irina Nikolaeva (2011), *Objects and information structure*, Cambridge: Cambridge University Press.

Djamouri, Redouane, Waltraud Paul and John Whitman (2013), Postpositions vs. prepositions in Mandarin Chinese: The articulation of disharmony in Theresa Biberauer and Michelle Sheehan (eds.), (2013), *Theoretical approaches to disharmonic word order*, Oxford: Oxford University Press, 74 - 105.

Paris, Marie-Claude. (1979), Some aspects of the syntax and semantics of the *lian…ye/dou* construction in Mandarin, *Cahiers de linguistique Asie orientale* 5(1), 47-70.

Paul, Waltraud (2005), Low IP area and left periphery in Mandarin Chinese, *Recherches linguistiques de Vincennes* 33, 111 -134.

Paul, Waltraud (2015), Postpositions: Double troubles in *New Perspectives on Chinese Syntax*, Berlin: De Gruyter.

Seržant, Ilja A. & Alena Witzlack-Makarevich (eds.) (2018), *Diachrony of differential argument marking*, Berlin: Language Science Press.

Takamine, Kaori (2017), *Putting Adpositions in Place: Sortal domains and Modifier PPs in Japanese*, John Benjamins Publishing Company.

Directive speech acts in Estonian child directed speech: SES factors and children's vocabulary

Andra Kütt
(Tallinn University)

Keywords: first language acquisition, sociolinguistics, directives, child directed speech, vocabulary, Estonian

Directives cover speech acts with different pragmatic force and cognitive complexity (Tomasello 2010: 84; Mauri, Sanso 2011: 3). The amount of directives in child directed speech has been considered to have a negative impact to children's speech development (Paavola-Ruotsalainen et al. 2017) and it has also been considered to serve as a predicting factor for child's academic success (Masur, Turner 2001). Not much attention has been paid to the question how can the amount of directives affect the child's language development. The correlations between socioeconomic factors like mothers' education, cognitive complexity of directives parents use, with child's language

development have been almost uncovered area in first language study. The present pilot study aims to identify **the correlations between the amount of directives and their cognitive complexity and the child's language development**.

The current preliminary research bases on recordings of spontaneous speech of six 6-year old children with their mothers (9 hours). All directive speech acts were coded according to their cognitive complexity (e. g. commands, requests where the addressee is not expressed, e.g. wishes, and the most complex requests as social norms where neither addressee nor the speaker is not explicitly expressed). Results revealed to the significant correlation between SES factors (e. g. mother's education), the amount and the types of directives used by mother, and child's language development.

The speech of the mothers with higher education included less directives. HighSES mothers do not use significantly more types of indirect requests but they use more encouragements, proposals, and questions for requests, while expressions of social norms are more frequents in the speech of LowSES mothers. In both groups, the correlation between the amount of directives and the size of child's vocabulary is significant: the child who hears more directives has smaller vocabulary and MLU. Still, the child's vocabulary is influenced not only by the general amount of directives, but also by their cognitive complexity.

References

Frank Masur, Elise and Turner, Margaret (2001), *Stability and Consistency in Mothers' and Infants' Interactive Styles*. *Merrill-Palmer Quarterly*: Vol. 47: Iss. 1. https://digitalcommons.wayne.edu/mpq/vol47/iss1/5

Paavola-Ruotsalainen, Leila, Lehtosaari, Jaana, Palomäki, Josefina, Tervo, Immi (2017), Maternal verbal responsiveness and directiveness: consistency, stability, and relations to child early linguistic development. *Journal of Child Language*, 1–21.

Tomasello, Michael (2010), *Origins of human communication*. Cambridge, MA: MIT Press.

Inalienable possession in Ṭuroyo: A corpus study

Nikita Kuzin
(Free University of Berlin)

Keywords: Neo-Aramaic, possession, inalienability, kinship terms

In our study, using the corpus of oral literature and dialogues (ca. 0.5 million tokens), we investigate the distribution of two sets of pronominal possessive suffixes in Ṭuroyo. The first set (general, frequent) suffixes can be attached to most nouns of the Ṭuroyo lexicon, while the second set (limited, less frequent) suffixes are used only for a closed set of nouns. In the focus of our study are those nouns of the closed set which can be used with both suffixes.

Table 2: Two sets of possessive suffixes for 3rd person possessor

	Set №1	Set №2
3 m.s.	-ayḏ-e	-e
3 f.s.	-ayḏ-a	-a
3 pl.	-atte	-ayye

In his grammar of Ṭuroyo, Jastrow (1967) adduces a list of 85 nouns which constitute the closed set. Most of these nouns can be split into three groups: body parts (*ʕayno* ‘eye’, *raġlo* ‘foot’), kinship terms (*emo* ‘mother’, *aḥuno* ‘brother’), and inherent properties (*ʕumro* ‘age’, *gawno* ‘colour’).

On the basis of the corpus data we summarize the general trend for all three groups: body part nouns are the most resistant group and only rarely and under special conditions can take the general set of suffixes. They are followed by nouns denoting inherent properties most of which also use the limited set. Finally, for kinship terms and other nouns the encroachment of the general set is much stronger.

For body parts the use of the series №1 is usually possible when a word is used in its non-basic, or historically secondary, meaning, like in (2):

(1) Lišon-i
tongue-POSS-1S
‘my tongue’

(2) Lišon-ayḏ-i
language-POSS-1S
‘my language’

The kinship terms are the most interesting group. As can be seen from table 2 below, using *emo* ‘mother’, the two sets are almost in complementary distribution: 1st and 2nd persons vs. 3rd person.

Table 3: Frequency of possessive suffixes for emo ‘mother’

	1 s.	1 pl.	2 m.s.	2 f.s.	2 pl.	3 m.s.	3 f.s.	3 pl.
Set №1	1	0	0	0	0	25	2	33
Set №2	122	8	42	13	5	2	0	1

Two terms within the kinship class behave differently: *abro* ‘son’ and *barṯo* ‘daughter’. The use of the general set is more common with these two nouns, and the forms with 1st person possessor, *abr-i* ‘my son’ and *barṯ-i* ‘my daughter’, are mostly used while addressing a person. This can be further compared with the frequencies for other animate nouns, and the hierarchy we have is the following: kinship terms of ascending generation → son, daughter → non-kin relations (friend, wife, husband).

To sum up, we argue that Ṭuroyo still has a strong opposition between inalieanble and alienable possession, and the choice of the suffix is lexically determined. Kinship terms and other nouns, such as *šuġlo* ‘work’ and *xabro* ‘word’, are the ones most penetrated by the innovative suffix series.The sharp split between 1st and 2nd person vs. 3rd person for kinship terms is probably due to the egocentric nature of 1st and 2nd pronouns and pronominal elements, especially in combination with kinship terminology (Dahl and Koptjevskaja-Tamm 2001).

References

Dahl, Östen and Maria Koptjevskaja-Tamm (2001), Kinship in grammar, in I. Baron, M. Herslund, and F. Sørensen (eds), (2001), *Dimensions of Possession*, Amsterdam / Philadelphia: John Benjamins, 201-226.

Jastrow, Otto (1967), *Laut- unf Formenlehre des neuaramäischen Dialekts von Miḏin im Ṭur ʕAbdin*, Bamberg.

Evidence for a tritonal default pitch accent L+M*+H in Valais German statements

Marius Zemp
(University of Berne)

Valais German (VG) is well known to sharply differ from other dialects of Swiss German (CHG) on the segmental and morphological levels. However, in Berne, where a notable number of VG speakers live, one regularly identifies these by overhearing their utterances from considerable distance, that is, without recognizing individual segments. A phonologically-contrastive study of intonation in VG and three dialects of the Swiss Midland (Berne, Zurich, and Lucerne) confirms the impression one gains from the mentioned observation: that VG sharply differs from other CHG dialects also in terms of intonation.

Arguably the most salient difference between VG and other CHG dialects is found in wide-focus statements starting with monosyllabic preverbal constituents (such as VG *ich* ~ Bernese German (BG) *i(g)* 'I'), where pitch rises onto the verb (VG/BG *bi* 'am') in VG but falls in the other investigated dialects, as illustrated for BG in Figure 1. Broad evidence (see Zemp 2018) will be presented in order to show that this contrast between VG and BG manifests itself in all wide-focus statements with monosyllabic sentence-initial constituents.

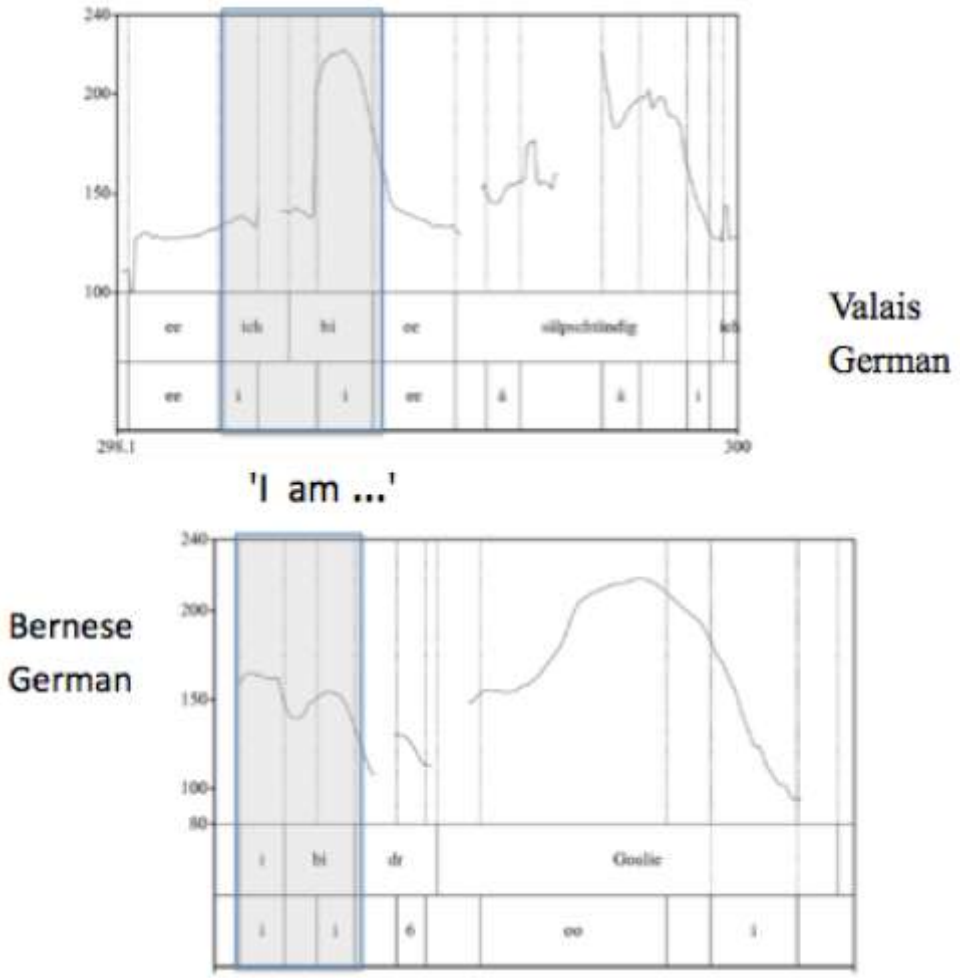


Figure 1: In wide-focus statements, after monosyllabic sentence-initial constituents, pitch rises onto the verb in VG but falls in BG (data are from the VG sitcom „Tschutter“ and the BG movie „Der Goalie bin ig“)

After plurisyllabic preverbal constituents, the contrast between VG and BG is regularly neutralized in that pitch drops onto the verb in both dialects. This indicates that the high pitch of the VG verb after monosyllabic constituents emanates not from an accent of the actual verb but from an accent of the preverbal element. In VG, this accent is thus characterized by a pitch rise whose peak lies within the preverbal element if it is plurisyllabic, but on the verb when the preverbal element is monosyllabic.

We get to the heart of the difference between VG and BG intonation by comparing wide-focus statements with plurisyllabic preverbal constituents that are stressed on their last syllable. As illustrated in Figure 2, in BG, pitch on the verb (*macht* ‘does’) is lower than on the stressed last syllable of the preverbal element (*ROUch* ‘smoke’).

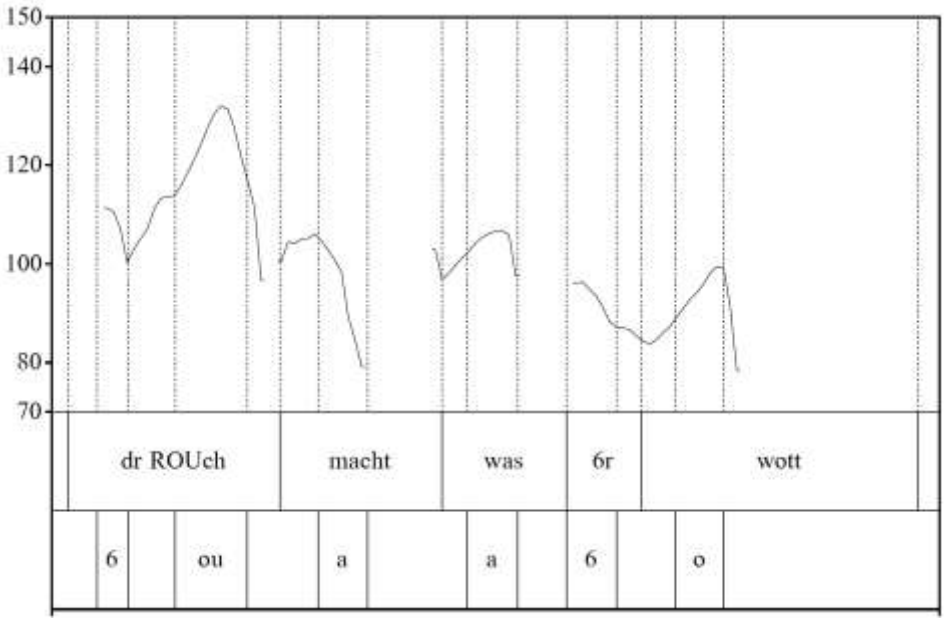


Figure 2: Neutral BG statement with pitch dropping from stressed second syllable of the preverbal element *dər Rouch* ‘the smoke’ onto the following verb *macht* ‘does’ (from „Der Goalie bin ig“, 33:24; speaker: Goalie)

By contrast, Figure 3 illustrates that in VG, the verb (*isch* ‘is’) receives a higher pitch than the stressed last syllable before it (*ab jEtz* ‘from nOw on’). At the same time, within the preverbal constituent, the stressed syllable exhibits the highest pitch. Thus, in VG, the stressed last syllable of a plurisyllabic statement-initial constituent is regularly both preceded and followed by a pitch rise. This intonation pattern is best represented as L+M*+H (where L means a low leading tone, M* a mid-high tone associated with the stressed syllable, and H a high trailing tone), and the consideration of further data suggests that this is in fact the default pitch accent of VG statements such as that in Figure 3, where it occurs four times in a row, downstepped each consecutive time (it appears that the syllable *er-* assigned with the leading tone of the fourth of these accents is too short for the pitch to actually drop below the following accented syllable *-Öff-*).

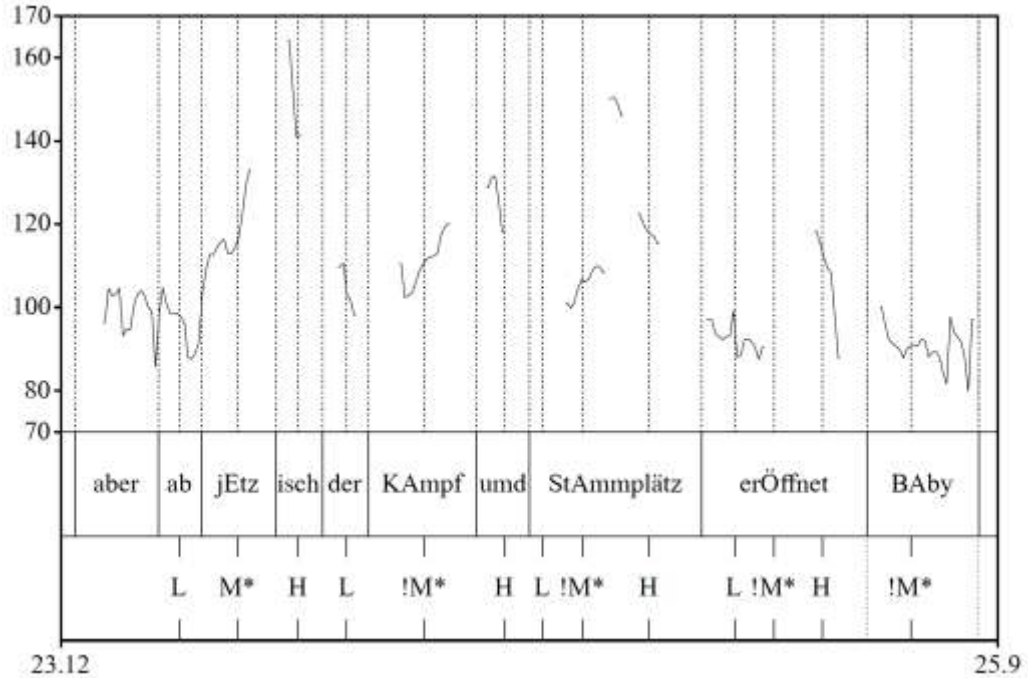


Figure 3: VG series of downstepped L+(!)M*+H, the default statement intonation (from „Tschutter“, Season 1, Episode 3; speaker: narrator)

Sources

Der Goalie bin ig. 2014. Ein Film von Sabine Boss. Nach dem Roman von Pedro Lenz. 92 Minuten. Michael Steiger & Anita Wasser – C-Films Theres Scherer-Kollbrunner – CARAC Film AG.

Tschutter. D neu Sitcom üsum Wallis. Season 1, episode 3: Propällär

URL: <https://www.youtube.com/watch?v=154h_eqhybo>, accessed on December 6 2018.

References

Zemp, Marius. 2018. Eine Eigenheit der Walliserdeutschen Intonation: default L+M*+H in Aussagesätzen . *Zeitschrift für Dialektologie und Linguistik* 85(2): 205–233.

ꕥ

Romance Purpose Prepositional Infinitives. A comparative and diachronic study

Laura Migliori
(Leiden University)

Keywords: Romance, Diachronic syntax, Purpose clause, Infinitive, Prepositions

Purpose Prepositional Infinitives (PREP + infinitive) are a widespread structure in Romance languages (Schulte 2007, a. o.):

(1) Mia figlia prende l’autobus per andare a scuola [Italian]
my daughter-F.3SG take-3SG the bus to go-INF. to school

(2) Ma fille prend le bus pour aller à l’école [French]
my daughter- F.3SG take- 3SG the bus to go-INF. to school
‘My daughter takes the bus to go to school’

Interestingly, this kind of complement clause is not attested in the ancestor language of Romance. Latin exhibits, in fact, other strategies for expressing a purpose sentence. Next to the inflected embedded clauses formed by *ut(i)/ne*/relative/*quo* + subjunctive (3), (4), (5), there were also indefinite purpose clauses, but never with an infinitive; see (6), (7) and (8):

(3) vere ut dicam [Cic. Pis. 68, 449]
thruthfully to speak-SUBJ.1SG
'to speak thruthfully'

(4) sed plerique perverse ne dicam inpudenter
but many-NOM. perverse-ADV. not to say-SUBJ.1SG insolent-ADV.
habere talem amicum volunt[…] [Cic. Lael., 82]
have-INF. such friend-ACC. desire-PRES.IND.3PL
'However, many people desire in a perverse, not to say in an insolent way, such a friend (that) […]'

(5) quinque legati ad regem missi qui res
five ambassadors-NOM. to king-ACC. sent-PP. who-NOM. thing-PL.ACC.
in Macedonia aspicerent [T. Liv. 42,6,4]
in Macedonia-ABL. control-SUBJ.IMPF.3PL
'Five ambassadors were sent to the king to examine the situation in Macedonia'

(6) ad pacem petendam [T. Liv. XXI, 13,1]
to peace-ACC. ask-gerundive-ACC.
'to ask for peace'

(7) Veientes pacem petitum
citizen of Veio-NOM.PL peace-ACC. ask-SUPINE
oratores Romam mittunt [T. Liv. 1, 15,5]
orator-ACC.PL. Rome-ACC. send-PRES.3PL
'The citizens of Veio sent orators to Rome to ask for peace'

(8) legati pacem petentes ad Scipionem venerunt [T.Liv. 22,20,8]
ambassador-NOM.PL peace-ACC. ask-PRES.PRT to Scipio-ACC. arrive-PF.3PL
'Some ambassadors arrived to Scipio to ask for peace (lit. asking for peace).

The main aims of this study will be: (i) to analyse the syntactic properties of this Romance complement clause, with special attention to synchronic variation; (ii) to formulate a hypothesis about its emergence during the Latin-Romance transition. The core properties of PPIs are subject coreference on the one hand, and the compulsoriness of a preposition for introducing the infinitive on the other hand. Moreover, synchronically, a distinction can be made between Western and Eastern Romance in that these two branches differ as far as the productivity and the characteristics of PPIs are concerned. As for their diachronic development, it will be proposed that the following factors, typical of the Latin-Romance transition, played a majour role in this syntactic change: (i) the change from a stative/active alignment to a nominative/accusative one (cf. La Fauci 1988; Ledgeway 2012, a. o.), with its consequences for subject encoding (cf. Migliori 2016, a.o); (ii) the extension of preposition functions (cf. Adams 2013). The emergence of PPIs can be, in this way, understood as a development which is coherent with the developments concerning the whole language system and not as an isolated development.

Language shift from Karelian to Russian in the latter half of the 20th century: demographics, domains and social networks

Niko Partanen and Janne Saarikivi

<pdf>

The pronoun *kes* ‘who’ and its referent’s animacy in Estonian dialects

Hanna Pook
(University of Tartu)

Keywords: dialect syntax, Estonian dialects, interrogative-relative pronouns, animacy

In many languages there seems to be a strong connection between the animacy of the referent and the choice between different ways of referring to it (Dahl & Fraurud 1996: 56). This is also true for the interrogative-relative pronouns *kes* ‘who’ and *mis* ‘what’ in Standard Estonian: *kes* is usually used to refer to animate entities, *mis* is used to refer to inanimate entities and the choice between those two pronouns is decided by the animacy hierarchy (see Silverstein 1976; Dixon 1979: 85) and the individuation scale (Erelt 2017: 743; see also Comrie 1989: 189; Yamamoto 1999: 131, 138–139). However, in certain Estonian dialects these pronouns can be used in a different way and *kes* can also refer to inanimate entities (examples 1–2).

(1) Mid dialect

see	*pidi*		*nii*	*pieenikke*	*ja*	*ilus*	*olema*
it	have_to:PST:3SG		so	fine	and	beautiful	be:SUP
see	***lõng***	***kellest***	*`koot´ti*				
that	thread	who:ELA	knit:IPS:PST				

‘It had to be so fine and beautiful, the thread that was used to knit.’

(2) Western dialect

raha	*põle*	*kaa*	***kellega*** *sa*	*`ostma*	*lähäd*	
money	NEG:be	also	who:COM	you	buy:SUP	go:2SG

‘There is also no money, what will you use to go buy [things with]?’

The aim of this paper is to give a thorough overview of this non-standard use of *kes*: which Estonian dialects and subdialects frequently use *kes* to refer to inanimate entities and which studied variables (sentence type, case, referential distance, abstractness, number of the referent and the position of the relative clause) contribute most to this feature.

The data for this paper comes from the Corpus of Estonian Dialects (http://www.murre.ut.ee/estonian-dialect-corpus/), which contains authentic spoken texts from all Estonian dialects. This analysis uses (manually) morphologically annotated dialect texts, from which all the uses of the pronoun *kes* were collected into the data: in total 2957 keywords accompanied by context and information about the informant. The approach to the data is mostly quantitative and corpus-based; the significance of variables has been tested with Chi-Square tests and modeled with conditional inference trees.

The results show that *kes* is typically used to refer to inanimate entities in the Northern-Estonia, mainly in the Coastal, Western, Eastern and Mid dialects; this use is most frequent in prenominal headed relative clauses and a majority of the pronouns in the elative or comitative case referred to an inanimate entity. However, the referential distance between the pronoun and the head of the relative clause (*e.g.* the referred entity) proved to not be a significant factor affecting the animacy of the referent.

There are many possible reasons for this non-standard use of *kes*, including grammatisation of the interrogative marker (see Heine & Kuteva 2006: 209, 227), language contact and general spread of SAE languages. Further research is currently being done on the counterpart of *kes* – the pronoun *mis* 'what' – that could give a more comprehensive picture of the reasons behind the interrogative-relative pronoun variation based on the referent's animacy.

Abbreviations

2, 3 – second, third person, COM – comitative, ELA – elative, IPS – impersonal voice, NEG – negation, PST – past tense, SG – singular, SUP – supine

References

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.

Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. (Ed.) Thorstein Fretheim & Jeanette K. Gundel. *Reference and Referent Accessibility* (Pragmatics and Beyond) 47–64.

Dixon, Robert M. W. 1979. Ergativity. *Language* 59–138.

Erelt, Mati. 2017. Komplekslause. Liitlause. *Eesti keele süntaks* (Eesti Keele Varamu III), 647–755. Tartu: Tartu Ülikooli Kirjastus.

Heine, Bernd & Tania Kuteva. 2006. *The changing languages of Europe* (Oxford Linguistics). Oxford; New York: Oxford University Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*, 112–171. Canberra: Australian National University.

Yamamoto, Mutsumi. 1999. *Animacy and reference: A cognitive approach to corpus linguistics*. Vol. 46. John Benjamins Publishing.

Comparing Hill Mari standard markers

Julia Sinitsyna
(Lomonosov Moscow State University)

Keywords: comparison of equality, standard marker, morphosyntax, typology, Uralic languages

This study deals with standard markers in equative and similative constructions in Hill Mari (<Uralic) – *gan*' and *-la (-lä)*. My data was collected in the villages of Kuznetsovo and Mikrjakovo (Mari El Republic, Russia) in 2016-2018.

Firstly, I will discuss semantic differences between these two markers. Mari grammars describe *-la* as 'comparative case' that compares one state or action with another one (Tuzharov 1987: 91). In terms of Haspelmath & Buchholz (1998), *-la* is used as a standard marker in equative constructions. Furthermore, *-la* has other functions, e.g. it marks names of games with the verb 'to play' (*futbol-la mad-aš* (football-SIM play-INF) 'to play football'). Traditionally *gan*' is considered as a comparative

postposition (Savatkova 2002: 260), but there are no detailed studies of its functions. Our data shows that *gan*’ is used in both equative and similative construction and *-la* is a special similative marker.

Secondly, I will analyze the morphosyntactic features of these markers. As regards *gan*’, its morphosyntactic properties are not typical of Hill Mari postpositions. It does not govern the case of a dependent noun (3b) unlike other Hill Mari postpositions which usually require nominative case. Constructions with *gan*’, unlike prototypical PPs, can be subjects and control predicate agreement (1).

(1) *g’itar gan’-vlä oxol-âštâ šalg-at.*
guitar EQU-PL corner-IN stand-NPST.3PL
‘There are guitar-like instruments in the corner’.

I consider two types of constructions shown in (3b) and (1). In (3b) *kreslоštâ gan*’ is a clausal standard of comparison (see Merchant (2009) for the differences between clausal and phrasal comparatives) containing a verbal element in its syntactic structure. This element is elided because it is identical to the compare VP.

The construction in (1) is a phrasal comparative containing only a NP as a standard. The head noun (e.g. *instrum*’*ent* ‘instrument’) is elided as it is known from the context. According to some research on ellipsis (e.g., Saab, Lipták 2016: 15), grammatical features of an elided noun are “obligatorily spelled out on the linearly *last remnant* preceding the elliptical gap”.

The existing studies of *-la* do not discuss sentences like (2) which are grammatical in the Mikrjakovo subdialect. Here *-la* appears on a finite verb form. This is not typical of Hill Mari case markers. In equative constructions, *-la* marks arguments connected only with SU and DO; locative arguments require an attributive suffix to become accessible to this standard marker (3a). On the other hand, *gan*’ marks standard NPs that derive their case assignment from the compare NP regardless of its syntactic position (3b).

(2) *mäläm tagačâ pet’a tol-eš-lä čuč-eš.*
me.DAT today P. come-NPST.3SG-SIM seem-NPST.3SG
‘It seems to me that Peter will come today’.

(3a) *pet’a pöken-äštə kreslo-štâ-*(šâ)-la šənz-ä.*
P. chair-IN armchair-IN-ATTR-SIM sit-NPST.3SG

(3b) *pet’a pöken-äštə kreslo-štâ gan’ šənz-ä.*
P. chair-INarmchair-IN EQU sit-NPST.3SG
‘Peter is sitting on the chair like in the armchair’.

As regards *-la*, I will argue that in the Mikrjakovo subdialect the marker in (2) and (3a) belongs to different morphosyntactic classes (subordinator of depended clauses and case marker respectively). This phenomenon is attested cross-linguistically, e.g see Chamoreau (2017). The grammaticalization path SIMILE > COMPLEMENTIZER is mentioned in Heine & Kuteva (2002: 273).

This research has been supported by RFBR, grant № 19-012-00627a.

References

Chamoreau, Claudine (2017), Multifunctionality and polysemy of the similative marker =kán in Pesh, in Y. Treis and M. Vanhove (eds), (2017), *Similative and Equative Constructions: A cross-linguistic perspective*, Amsterdam / Philadelphia: John Benjamins, 321–337.

Haspelmath, Martin. & Buchholz, Oda, (1998), Equative and similative constructions in the languages of Europe, in J. van der Auwera (ed.), (1998), *Adverbial constructions in the languages of Europe*, Berlin: Mouton de Gruyter, 277-334.

Heine, Bernard and Kuteva, Tania (2002), *World lexicon of grammaticalization*, Cambrige: CUP.

Merchant, Jason (2009), Phrasal and clausal comparatives in Greek and the abstractness of syntax, *Journal of Greek Linguistics* 9, 134–164.

Tuzharov, Gennadij (1987), *Grammaticheskije kategorii imeni suschestvitel'nogo v marijskom yazyke* [Grammatical categories of nouns in Mari], Yoshkar-Ola.

Saab A., Lipták A. 2016. Movement and deletion after syntax: Licensing by inflection reconsidered, *Studia Linguistica* 70(1), 66-108.

Savatkova, Anna (2002), *Gornoje narechije marijskogo jazyka* [Hill Mari dialect], Savariae: Berzsenyi Dániel Főiskola.

The Use of the Copula in Non-Copula Constructions in the Language of South Asia: An exploratory study

Anna Sjöberg & Anju Saxena
(Stockholm University; Uppsala University & Uppsala University)

Keywords: areal linguistics, Sprachbund, South Asia, copula, periphrastic constructions

At least since Emeneau (1956), many scholars have considered South Asia to be a linguistic area or *Sprachbund*. Several linguistic features have been investigated as possible candidates for South Asian areal features and in this presentation, I report the results of a study of the use of copulas in non-copula constructions and establish whether this is something that can be added to the list of things that the languages of South Asia do the same. The aim of the study is thus exploratory: to give an overview of the phenomenon and note potential areal and genetic patterns.

The material used in the study is the *Linguistic Survey of India* (Grierson 1903-1927). This gigantic work, compiled at the turn of the last century, contains descriptions of and materials from over 500 language varieties. This study looks at a subset of 206 varieties from the four major language families of South Asia: Indo-European (Indo-Aryan and Nuristani), Sino-Tibetan, Dravidian and Austroasiatic (Munda). The subset consists of those languages that are sufficiently described in the LSI to include information on the phenomenon. For each language, a set of questions concerning the use of the copula in non-copula constructions and the relationship to tense was answered.

The results of the study show that the usage of the copula in the non-copula constructions is common in the surveyed branches of all four families, with 171 of the languages using it in at least some construction. It is the most common in IA, Nuristani and Dravidian, followed by Munda where one language does not have it. In Sino-Tibetan, there is greater differences in the distribution between the subgroups.

In addition to the simple existence or non-existence of the phenomenon, the study looks at the distribution across tenses. In general, the usage is more common in past and present than in the future. Due to the nature of the material, and the time at which it was compiled, it is not possible to get any reliable results on the distribution across aspectual categories, which these forms are often used to encode.

The general geographical distribution of the features seems to support the copula in non-copula constructions as an areal phenomenon, based on the characteristics established by Masica (1976). He notes that it is common for a linguistic area not to be sharply cut off but rather to end in a zone where the features characteristic of the area are present in a somehow reduced manner. This can be observed in the results of this study, where languages to the extreme east and northwest deviate from the prototype found in the centre of the area.

To conclude, then, this exploratory survey of the phenomenon of copulas in non-copula constructions shows that they appear to be more common than the world-wide average in South Asia and that the distribution is such that it suggests contact-induced spread.

Acknowledgments: The work presented here has been financed in part by the Swedish Research Council, projects *South Asia as a linguistic area? Exploring big-data methods in areal and genetic linguistics* (contract 2014-969) and *Documentation of an endangered language: Kunashi* (contract 2014-560).

References

Emeneau, M.B. 1956. India as a linguistic area. *Language,* vol. 32: 3-16.

Grierson, G.A. 1903-1927. *The Linguistic Survey of India. Volumes 1-9*. Calcutta: Office of the Superintendent, Government Printing, India.

Masica, C.P. 1976. *Defining a linguistic area: South Asia*. Chicago: Univ. of Chicago P

A nonstandard affix reordering: The restrictive *kən* in Ulch

Natalia Stoynova
(RLI & IL RAS, HSE, Moscow)

Keywords: Tungusic languages, Ulch, restrictive markers, affix reordering, externalization of inflection.

The paper deals with the restrictive marker *kən* ‘only’ in Ulch (Southern Tungusic, Khabarovsk Kraj, Russia). This marker has nontrivial positional features: it can attach before inflectional suffixes, (1) or after them, (2).

(1) əpə-**kəm**-bə ǯapa-xam
bread-RSTR-ACC take-PST.1SG
‘I took only bread’ (lpd_170726_nst_BioPoshlaVShkolu)

(2) masa, masa gursəl-bə-**kən**
strong strong people-ACC-RSTR
‘only very strong people’ (lpd_170726_nst_BioRybalka)

One might postulate here the process of affix reordering "infection + morpheme X > morpheme X + inflection", described in Haspelmath (1993) as "externalization of inflection". M. Haspelmath explains such a shift by the general semantically based “inflection-outside-derivation principle”. One of his predictions implied by this principle is that the diachronic process in question is unidirectional (ibidem: 289), i.e. the reverse change is impossible.

However, there is an evidence, that external uses of the restrictive *kən* (after inflection) are innovative compared to internal ones (before inflection), not vice versa. So, this direction of diachronic development seems unexpected.

The aim of the paper is to show that exactly this direction of affix reordering takes place and to propose an explanation for this process.

The data come from texts (published ones, archive ones and our field records) and from elicitation tasks with 6 speakers.

The assumed shift from the internal position to the external one is supported by two empirical facts. 1) The frequency of the external *kən* is low, and it increases significantly in the texts recorded in the

2010s, compared to the texts of the 1960s (from 2% to 25%, 2-tailed exact Fisher test, p=0.0019). 2) In the elicited data from the older speakers both internal uses and external ones were attested, while the youngest one consistently prefers the external position in all contexts.

The following points are relevant for the discussion on this case of a nonstandard affix reordering:
a) The meaning 'only' is not predisposed to be expressed by derivational morphemes. It is potentially compatible with any part of speech. It has a very flexible semantic scope. Morphological restrictives are cross-linguistically rare, compared to restrictive particles and analytical constructions (cf. König 1991: 20).
b) The restrictive *-kān* is very probable to be diachronically related to the Tungusic derivational diminutive suffix -**kĀn* (cf. Whaley & Li 1998: 459–460, Fuente 2018: 130). If it is true, then *kān* gets an access to a new morphological position together with the new restrictive meaning which is less typical of derivation than the diminutive one. Cf. a similar motivation of affix reordering discussed in (Mithun 2000) on the data of Iroquoian languages.
c) Ulch is highly endangered (cf. Sumbatova & Gusev 2016). The external sociolinguistic factor may additionally force this rapid nonstandard change.

To conclude, the affix reordering process observed in Ulch, at first glance, seems to be opposite to the externalization of inflection. However, it follows in fact the same more general principle of morpho-semantic congruence. In contrast to typical derivational morphemes, which acquire a more appropriate internal position via the internalization of inflection, the marker *kān* acquires the position after inflection which is semantically more appropriate for restrictives.

References

Haspelmath, M. (1993), The Diachronic Externalization of Inflection // Linguistics, 31. P. 279–309.

Sumbatova, N. R. and V. Ju. Gusev (2016), Ulcjskij jazyk (The Ulch language) // Mikhaljchenko, V. Ju. (Ed.) Language and Society. An encyclopaedia [Jazyk I obšjestvo. Enciklopedija], Moscow: Azbukovnik.

Whaley, L. J. and F. Li. (1998), The suffix *-kan* in Oroqen // Studies in Language, 22. P. 447–471.

Mithun, M. (2000), The reordering of morphemes // Gildea, S. (Ed.) Reconstructing grammar, Comparative linguistics and gramamticalization, Amsterdam: John Benjamins. P. 231–258.

de la Fuente, J. A. A. (2018), Past Tenses, Diminutives and expressive palatalization: Typology and the limits of internal reconstruction in Tungusic // Apatoczky, A.B. and Ch. Atwood (Eds.) Philology of the Grasslands. Essays in Mongolic, Turkic and Tungusic Studies, Leiden–Boston: Brill. P. 112–137.

Grammaticalization and constructionalization of a 'TAKE' verb in Lan Hmyo

Yoshihisa Taguchi
(Chiba University)

Keywords: <Hmongic, Hmong-Mien, grammaticalization, constructionalization, TAKE>

Lan Hmyo (ISO 639-3: hml) is a Hmongic language spoken in Guizhou, China. It has a verb, *ʔmiA* (*A* designating tone), which means 'to give' when used as the main verb in a clause (see [1] below). However, it fulfills various functions as a first verb in some serial verb constructions (SVCs), including the following: A disposal construction (see [2]), an instrumental construction (see [3]), a

causative construction (see [4]), and a PUT construction (see [5]). In this presentation, the author reconstructs the original meaning of *ʔmiA* and the pathways of development for these constructions based on comparative data and crosslinguistic considerations. The comparative data suggest that *ʔmiA* was originally a TAKE verb and that the abovementioned constructions were derived from SVCs with *ʔmiA* as a first verb. The development of a disposal construction, instrumental construction, and causative construction from a construction in which a TAKE verb appears as a first verb is crosslinguistically well-known (Heine and Kuteva 2002). However, the constructionalization (Traugott and Trousdale 2013) of a PUT construction and GIVE construction from a TAKE construction is not as well-documented. This presentation describes how Lan Hmyo exploits the TAKE verb *ʔmiA* to supply a way of expressing three-participant events in addition to PUT and GIVE in the language. The author also positions the development of *ʔmiA* in Lan Hmyo within the phylogenetic stock. The data suggest that the abovementioned disposal construction and instrumental construction with a cognate word of *ʔmiA* are observed in other Hmongic languages and thus may indicate cases of shared innovation, while the other constructions appear to be relatively recent developments in Lan Hmyo.

[1]*jeB,kaɴB* ***ʔmiA****tsiA* *seɴA* *mbliA* *taA* *tsiA* *seɴA* *tshaɴB,* *ʁoɴA* *tsoC* *tiB* *muB, …*
FLL 1MI three CLF rice and three CLF millet 2 then take go
"I will give you three cups of rice and three cups of millet, so you take them home."

[2] *χoɴA* ***ʔmiA*** *ŋ̊aCmbloA* *luB* *ʨaɴA.*
then MI glutinous.rice come steam
"… next, we steam glutinous rice."

[3]*təBqweɴAnoɴBχoɴA-****ʔmiA*** *aAnteɴB* *l̥eA* *niB* *-moA* *taAmbuAseɴA,* *niB χoɴA-*
after.that then MI knife stab 3 PRT liver 3 then
ðaC *-loA.*
die PRT
"After that, he stabbed her in the liver with a knife; then she died."

[4] *jeB* *tsoC-* ***ʔmiA*** *niB* *muB* *ntaɴB* *nʥiBðiA.*
FLL then MI 3 go tread millstone
"Then they made him go and polish rice with a millstone."

[5] *niB* *tsoC-* ***ʔmiA*** *toɴBʐoɴA* *qoC* *həBnʈoɴA* *qloɴA …*
3 then MI stool GOAL inside tub
"Then he put a stool into the tub, …"

[Abbreviations]
CLF: Classifier FLL: Filler MI: ***ʔmiA*** PRT: Particle

References

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: New York: Cambridge University Press.

Traugott, Elizabeth Closs and Graeme Trousdale. 2014. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

A cross-linguistic study of body-part metaphors and their cognitive basis

Annika Tjuka
(Humboldt-Universität zu Berlin)

Keywords: Lexical Typology, Body-Part Metaphors, Meronymy, Cognitive Linguistics

Thus far, the cognitive foundation of lexicalized terms for object and landscape properties which are retrieved from the body domain have not been fully explored. Although approaches to lexical typology indicate the importance of this phenomenon for examining cross-linguistic variation in semantic categories (Kraska-Szlenk 2014, and Koptjevskaja-Tamm 2008), most explanatory claims are too general. Ullmann (1963) states: "when an analogy is obvious, it should give rise to the same metaphor in various languages; hence the wide currency of expressions like the 'foot of a hill' or the 'leg of a table'." Studies which investigate this claim show that the dimensions of *shape*, *spatial alignment*, and *function* play a crucial role in this process (Levinson 1994, Ibarretxe-Antuñano 2012, and Tilbe 2017). In Tzeltal, see (1), the shape information is added to the locative description of the object.

(1) waxal ta x-chikin mexa te p'ine
standing-of-vertical-cylinder at its-ear table the pot
'the pot is standing at the corner of the table' (Levinson 1994)

Body-part metaphors, such as *head of the mountain* 'the top of the mountain', illustrate the use of the dimension of *spatial alignment* to metaphorically map a body-part term to a landscape property. The *function* dimension is the basis for examples, such as *leg of the table*.

By comparing the utilization of the *shape* and *function* dimension, Tilbe (2017) reveals that i) speakers of Mesoamerican languages assign more body-part terms to objects than English speakers, and ii) languages differ in terms of which dimension is used more productively.

For my MA thesis, I conducted the first systematic typological study that investigates the following questions:

1. How productively do languages use body-part terms to express parts of objects and landscapes?
2. Is one of the three dimensions used more productively than the others?
3. How much variation do we find between languages with respect to 1) and 2)?

The elicitation study consisted of pictures depicting 92 body-part metaphors which were collected through observations of oral communications and literature review. 13 speakers of the following languages participated in my study: Czech, Marathi, Persian, Greek, Vietnamese, Wolof, Mandarin Chinese, Khoekhoe, Hungarian, Japanese, Hebrew, Turkish, and Bahasa Indonesia.

The cross-linguistic comparison shows that the languages Wolof and Mandarin Chinese use more body-part metaphors from the seed list compared to Japanese and Marathi. However, the occurrence of particular metaphors differs throughout the entire set of languages. The metaphors *leg of the chair/table* appear in all languages, whereas *heart of the apple* is not expressed in any of the investigated languages. The data implicates that a body-part metaphor is frequent in a wide range of languages if it is based on more than one dimension. The results also show that none of the dimensions is favored in the overall language set. In contrast, individual languages use one dimension more productively than others, e.g., Khoekhoe prefers the spatial alignment dimension.

The cognitive principles for conceptualizing body-part metaphors may be routed in perceptual properties of objects and landscapes, such as shape, spatial alignment, and function.

References

Ibarretxe-Antuñano, Iraide (2012), The importance of unveiling conceptual metaphors in a minority language, in A. Idström, E. Piirainen, and T. Falzett (eds), (2012), *Endangered metaphors*, Amsterdam: John Benjamins, 253-273.

Koptjevskaja-Tamm, Maria (2008), Approaching lexical typology, in Vanhove, M. (ed) (2008), *From polysemy to semantic change: a typology of lexical semantic associations*. Amsterdam: John Benjamins, 3-52.

Kraska-Szlenk, Iwona (2014), Semantic extensions of Body Part Terms: Common patterns and their interpretation, *Language Sciences* 44, 15-39.

Levinson, Stephen C. (1994), Vision, shape, and linguistic description: Tzeltal body-part terminology and object description, *Linguistics* 32(4-5), 791-856.

Tilbe, Timothy James, (2017), *Parts and Wholes in Mesoamerican Language and Cognition*, State University of New York at Buffalo dissertation.

Ullmann, Stephen (1963), Semantic universals, in J. Greenberg (ed), (1963), *Universals of language*, vol. 2, Massachusetts: The MIT Press, 373-396.

Nominal and verb derivatives in Russian child speech

Maria D. Voeikova & Victoria Kazakovskaya
(Institute for linguistic Studies, Russian Academy of Sciences – Saint Petersburg State University)

Keywords: morphemic derivation, L1 acquisition, Russian morphology, semantic categories

The report discusses the early phases of first language acquisition of nominal (noun and adjective) and verb derivatives in the morphologically rich Russian language approaching an ideal inflecting-fusional language type. The results are based on naturalistic longitudinal observations of two typically developing Russian-speaking monolingual children (1;5–3;0) from middle SES families (33 hours of recorded spontaneous speech transcribed and coded according to CHILDES MacWhinney (2000).

The Russian word-formation system is represented by morphemic and non-morphemic methods (Švedova 2005). The morphemic patterns are frequent and productive. They include affixation ('X+suffix', 'prefix+X', 'prefix+X+suffix'), compounding ('X+Y') and their combination (e.g. synthetic compounds 'X+V+suffix'). Within affixation, the 'X+suffix' pattern is the most productive for the nominals. Prefixation is the usual way of modifying the lexical meaning of verbs, whereas a restricted number of suffixes is mostly used to change their grammatical characteristics (aspect, voice). Within non-morphemic methods, which are scarce in the material under observation, conversion (viz. substantivation) is the most prominent in nominals. Verbs can mostly be derived by morphemic methods.

Our results show that noun derivatives are mostly formed according to the denominative ('N+affix') pattern with the biggest repertoire of suffixes (up to 60). The semantic categories of these derivatives are females (*medved-ic(a)* 'bear-FEM'), young animals (*kot-enok* 'kitten'), singulatives (*izjum-ink(a)* 'one raizin'), stylistic (viz. colloquial) modifications (*kartoš-k(a)* 'potato'), instruments (*gradus-nik* 'thermometer'), agents (*gitar-nik** 'guitar player'), diminutives (*ptič-k(a)* 'bird-DIM'). Less frequent deverbal nouns represent agentive (*gon-ščik* 'racer'), instrumental (*bri-tv(a)* 'razor'), locative (*ostanov-k(a)* 'stop'), processual or resultative (*lep-k(a)* 'molding') meanings. Rarely in child speech, deadjectival nouns are mostly built by suffixes expressing the semantics of carriers of the corresponding traits (*grjazn-ulj(a)* '~untidy person', *čern-ik(a)* 'blueberry'), localisers and abstract qualities (*glup-ost'* 'foolish'). Denominative adjectives express possession (*mam-in* 'mommy's'),

material (*beton-n(yj)* ‘concrete’), colour (*apel’sin-ov(yj)* ‘orange’), shape (*krug-l(yj)* ‘round’), evaluation (*protiv-n(yj)*‘ugly’), relation (*zvezd-n-yj* ‘star’). Deadjectival adjectives mostly express different grades of quality (*ser-enk(ij)* ‘grey-DIM’). Deverbal adjectives have the meaning of the resulting state or potential actions (*balova-nn(yj)* ‘indulged’). Most of derived verbs are formed by prefixation expressing different aktionsarts, e.g., the path of motion (*vy-jti* ‘out-go’), start (*po-jti* ‘start-go’) or end of the action. Suffixes are mostly used for imperfectivation (*zakr-yva-t’* ‘close-IPFV’) or semelfactive meaning (*pryg-nu-t’* ‘jump-once’). The reflexive suffix *-sja* has several semantic functions (like mutual or saturative aktionsarts) beyond its role in passive constructions. Thus, in children’s nominals, suffixation is the main method of word-formation, whereas prefixation (*pra-ded* ‘great-grandfather’) and mixed methods (*pod-osinov-ik* ‘orange-cap boletus’) are rare. Although nominals can be formed from almost any classes, most of them originate from nouns. Verb derivatives are usually built from verbs.

The preference for certain patterns is the same for both subjects in all three grammatical classes and mirrors the child- and adult-directed speech. Both children have the similar percentage of first-appearing derivatives in nominals and verbs. The number of occasionalistic novel derivatives [Gvozdev 2007; Ceitlin 2009) is scarce for all grammatical classes. The differences noted in child speech may be explained by input properties and special strategies of children and their caregivers.

References

Ceitlin, Stella N. (2009), *Očerki po slovoobrazovaniju i formoobrazovaniju v detskoj reči* [Essays on word and form building in child speech]. M.: Znak.

Gvozdev, Alexandr N. (2007 (1949)), *Voprosy izučenija detskoj reči*. [Problems of the investigation of child speech]. St. Petersburg: Detstvo-Press.

MacWhinney, Brian. 2000. *The CHILDES project. Tools for analyzing Talk. Electronic edition. Part 1: The CHAT transcription format*.

Švedova Natalia Ju. (ed.) (2005 (1980)). *Russkaja grammatika* [Russian grammar]. M.: Nauka.

Macrofunctionality. Genitives, adjectives and relative clauses in Gutob

Judith Voß
(University of Kiel)

Keywords: macrofunctionality, attribution, areal typology, genitives, relative clauses

A conflation of the morphological marking of two or all of the functions of genitives, adjectives and relative clauses into macrofunctional attributive or associative markers is cross-linguistically common, especially in East and South-East Asia (Gil 2001, e.g. Gil 2013 on Hokkien) but also in Northern Kurdish (Haig 2011). For South Asian languages information on the phenomenon is so far largely lacking. I present a preliminary analysis of macrofunctional attributive markers in Gutob, which I will then compare to findings in other Munda languages.

In Gutob, the enclitic *=nu* occurs in possessive constructions (1), in relative clause-type constructions (2), adjectivized adverbials (3) as well as in postpositional phrases (4). There is a large overlap in the functionality of *=nu* in all of these constructions and no clear boundary can be drawn between different functions. I will show that these seemingly different functions can all be reduced to attributivization/adjectivization. This is the case both with possessor marking and the adjectivization of adverbials. Neither nouns nor adverbials normally attributively modify a noun, but can do so if marked with *=nu*. Similarly, verbs or clauses cannot modify a noun unless they are marked by the

attributivizer that transforms the converb into a relative participle, i.e. a verbal adjective. Note that adjectives, which form a small, closed class in Gutob, attributively modify nouns without any special marking, as is shown in example (5). Therefore, the marker *=nu* has to be analysed as a single macrofunctional morpheme that marks the attributivization/adjectivization of non-adjectival elements in these constructions. The use of *=nu* in postpositional phrases can be explained by their diachronic development from relational nouns within possessive constructions. Similar patterns are found in other South Munda languages such as Gtaʔ (Anderson 2008), Gorum (Rau p.c.) and Kharia (Peterson 2011). The macrofunctionality in South Munda languages is especially interesting in an areal context. While the adjectivization of demonstrative adverbials by means of a genitive case marker is common in other South Asian languages, a conflation of genitive and relative clause marker has to my knowledge not been reported for either Dravidian or Indo-Aryan languages. Both in Sadri and Desia, the two major contact languages for South Munda languages, relative clauses and genitives are distinct (Peterson p.c., Mathews 2003).

(1) *niŋ=nu ɖaʔ kakor-gu ui-gi=be*
1sg=ATTR water cold-MID.CV LV(go)-MID.PST=HON
'My water got cold.' (Gutob-0444-20180216-2-K: 274)

(2) *tu riŋ-gu=nu lai maʔ iɖaʔ ɖiʔto som-oʔ beʔ-to*
DEM bring-MID.CV=ATTR cooked.rice curry gruel REP eat-ACT.CV LV(give)-HAB
'He eats that rice, curry and gruel he had brought.' (Gutob-0444-20170116-3: 0399)

(3) *aʔ=nu lok bana-gu=nen beɽ-oʔ*
now=ATTR people forget-MID.CV=PL LV(give)-ACT.PST
'The present-day people forgot (the old customs)' (Gutob-0444-20161125: 069)

(4) *basguɖa ji=nen=nu=boʔ*
Basguda aunt=PL=ATTR=LOC
'At the Basguda aunts' place' (Gutob-0444-20170119-4: 265)

(5) *pilej gisiŋ muiroʔ sitra gisiŋ muiroʔ*
white chicken one speckled chicken one
'One white chicken and one speckled chicken.' (Gutob-0444-20170112-1-S: 109)

References

Anderson, Gregory DS. 2008. Gtaʔ. *The Munda languages* 682–763.

Gil, David. 2001. Escaping Eurocentrism: Fieldwork as a process of unlearning. *Linguistic fieldwork.* 102-132.

Gil, David. 2013. Genitives, Adjectives and Relative Clauses. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. https: //wals.info/chapter/60.

Haig, Geoffrey. 2011. Linker, relativizer, nominalizer, tense-particle. In *Nominalization in Asian languages: Diachronic and typological perspectives.* Typological Studies in Language 96. Amsterdam: John Benjamins Publishing, 363-392.

Mathews, Susan. 2003. *Aspects of Desiya Grammar*. Lamtaput: Asha Kiran Society.

Peterson, John. 2011. *A grammar of Kharia: A South Munda language.* Brill's studies in South and Southwest Asian languages. Leiden: Brill.

Contact-induced innovations in Ladin: the case of gerund

Jelena Živojinović
(University of Verona; UiT The Arctic University of Norway)

Keywords: gerund, Ladin, Rhaeto-Romance, language contact, variation

This study aims to explore the use of gerund in Ladin. Although there is no standardized version of Ladin common to all the Valleys, each variety has published a grammar showing the presence of a synthetic gerund, such as *chantan*, *dormian*, etc. However, there are interesting cases showing variation in which alternative constructions are preferred with respect to the canonical gerund. I will illustrate the distribution of gerund patterns across the main varieties of Ladin and present some preliminary results of fieldwork conducted in Val Badia, Val di Fassa and Val Gardena.

What emerges from current literature is a highly restricted use of synthetic gerund, used exclusively in specific constructions. Namely, Casalicchio (2011) points out that gerundive clauses, particularly in Gardenese and Badiot (Northern Ladin varieties) with perception verbs require a synthetic form. However, this type of clause can only be introduced by three general verbs from a semantic point of view, *audì* (to hear), *sentì* (to feel – by touching) and *udëi* (to see), as in the following example:

e.g. *I ujins s'la cuinova* ***udan*** *Noé zumpran ch'la ercia coche na gran arca*. (Bib.pit. VT)
'Neighbours were laughing **by seeing** Noah working (shaping) that barn into a big ark'

This construction lacks the characteristics of durativity and the need of an explicit semantic subject. Therefore, Casalicchio (2011) concludes that the gerundive construction has lost a larger part of its semantic and aspectual features, by resembling the analogous infinitival structure of Italian, which has the role of superstrate/prestige language.

The test conducted in the Ladin Valleys provides some interesting preliminary results. Speakers were asked to translate a set of sentences, from Italian into Ladin and showed a strong preference for rephrasing the entire structure.

e.g. Ita. *Lavando i piatti ho rotto la mia tazza preferita.*
Lad. *Intratan che i lavâ i tais ai rot mia copa preferida.*
'I broke my favorite cup while washing the dishes'

What arises is a very clear resemblance of circumstantial Ladin clauses to the analogous German structure, lacking gerund. In fact, the Italian synthetic gerund is translated into a temporal clause, introduced by *intratan* and corresponding to a German *während.* Therefore, it is interesting to note that both German and Italian have an important impact on Ladin, influencing different aspects of it.
In my talk, I will claim that the category of gerund in Ladin is in the process of gradual change. I will present the results of fieldwork conducted in Val Badia, Val di Fassa and Val Garden and enrich the analysis by adding data from the corpus ALD-II, in order to postulate contact-induced change. The overall results should provide a larger picture on variation in the use of gerund across Ladin varieties.

References

Aikhenvald, Alexandra Y. & Dixon, Robert M. W. (2001), *Areal diffusion and genetic inheritance: problems in comparative linguistics*, Oxford: Oxford University Press.

ALD-II = Goebl, Hans et al. (eds.). (2012). Atlant linguistich dl ladin dolomitich y di dialec vejins, 2ª pert / Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 2ª parte / Sprachatlas des Dolomitenladinischen und angrenzender Gebiete", 2. Teil. Salzburg.

Benincà, Paola, & Haiman, John (2005), *The Rhaeto-Romance languages*, Abingdon-on-Thames, Routledge.

Casalicchio, Jan (2011), L'ˇuso del gerundio con i verbi di percezione gardenesi, *Ladinia XXXV*, 321-352.

Casalicchio, Jan (2012), The Syntax of Gerunds and Infinitives with Perception Verbs in Northern Italy. Poster presented at the first Italian Dialect Meeting, Leiden 24-26 May 2012.

De Smet Hendrik et al. (eds.) (2015), On Multiple Source Constructions in Language Change. Amsterdam: John Benjamins.

Chiocchetti, Nadia & Iori, Vigilio (2002), *Gramatica del ladin fascian*, Sèn Jan: Ist. Culturale Ladino.

Gasser, Tone (2000), *Gramatica ladina por les scores*. Bolzano: Istitut Pedagogich Ladin.

Hickey, Raymond (Ed.) (2012), *The handbook of language contact*, Hoboken: John Wiley & Sons.

Loporcaro, Michele (2009), *Profilo linguistico dei dialetti italiani* (Vol. 275). Bari: Laterza.

Ramat, Paolo & Da Milano, Federica (2011), Differenti usi di gerundi e forme affini nelle lingue romanze, *Vox Romanica 70*, 1-46.

Salvi, Giampaolo, & Renzi, Lorenzo (Eds.), (1991), *Grande grammatica italiana di consultazione: I sintagmi verbale, aggrettivale, avverbiale. La subordinazione*. Bologna: Il Mulino.

Salvi, Giampaolo (2011), *Morphosyntactic persistence from Latin into Romance*, The Cambridge History of the Romance Languages, Cambridge: Cambridge University Press.

Tekavčić, Pavao (1972), *Grammatica storica dell'italiano: volume II: Morfosintassi*, Bologna: Il Mulino.

Thomason, Sarah (2008), Social and linguistic factors as predictors of contact - induced change, *Journal of Language Contact 2.1*, 42-56.

WORKSHOP PRESENTATIONS

WORKSHOP 1

Schedule: Thu 14.00 – Fri 17.25 (Room 12)

A comprehensive perspective on reflexive constructions

Nicoletta Puddu & Katarzyna Janic

(University of Cagliari & University of Leipzig)

Keywords: grammaticalization, valency change, diachrony, reflexive, coreference/binding

The term reflexive refers to events wherein the patient object co-refers with the agent subject. Reflexivity has been extensively discussed from different angles. Everaert and van Riemsdijk (2005) investigate it from a syntactic perspective. Yet Huang (2000) incorporates a neo-Gricean pragmatic account, whereas Keenan (1988) delivers semantic expertise. Crosslinguistic studies also couch this topic in various theoretical frameworks. While on the one hand, there is a large body of literature nested in a generative perspective (Chomsky 1981, Everaert 1986, Reinhart and Reuland 2011), on the other hand, there are several functional-typological studies (Faltz 1985, Geniušienė 1988, Kemmer 1993, König 2001, König and Gast 2008).

The coding of reflexive interpretation demonstrates remarkable variation across the languages of the world that begs for explanations (cf. König and Siemund 2000, Déchaine & Wiltschko 2017). It extends from nominals through dedicated reflexive pronouns grammaticalized in some languages into verbal affixes, finally to verbal strategies, including a change in verbal paradigm. A reflexive meaning can also be available with possessive or personal pronouns. In the face of such immense variation, Volkova (2014) openly admits that reconciling such diversity specifically within the Binding Theory is hard to accomplish.

The classification of reflexive forms poses further problems (Puddu 2017). A common separation takes place along the morphological line, leading to verbal vs. nominal distinction. This dichotomy was traced by Faltz (1985) and recognized in both generative and functional traditions. But it runs into some difficulties when we consider, for instance, the cases wherein objects are coded on the verb and where the distinction between verbal vs. NP strategy relies merely on affix vs. clitic distinction. As subsequently argued by Faltz (1985), this distinction should rather be viewed as a continuum. The fact that nominal and verbal reflexives frequently provide evidence for a common etymology (Kazenin 2001) further supports the gradient approach.

This workshop seeks to investigate the multifaceted aspects of reflexivity from different theoretical perspectives. It is also interested in describing and classifying the general patterns that shape reflexivity in a language or language family(-ies) both in its current structure and historical development. Given that reflexivity is a well-charted territory in languages with a good record of data, we particularly encourage scholars working on less-documented languages to explore this empirical domain so that it could bring not only a new dimension to the theoretical linguistics but also serves as a solid research tool for typological studies.

Possible topics

- How is the reflexive interpretation coded in a language?
- What are crosslinguistic generalisations resulting from the work on reflexivity?
- What are other semantic effects associated with reflexive morphology and how can we explain their common formal source?
- What is the relationship of reflexive marker with the domain of valency?

- What are the possible paths of development of reflexive marker?
- How can reflexives be classified in a functional-typological perspective?

References

Chomsky, Naom (1981), *Lectures on government and binding*, Dordrecht: Foris.

Déchaine, Rose-Marie, and Wiltschko, Martina (2017), A formal typology of reflexives, *Studia Linguistica* 71(1- 2), 60-106.

Everaert, Martin (1986), *The Syntax of Reflexivization*, Dordrecht: Foris

Everaert, Martin, and Henk C. van Riemsdijk (eds.) (2005), The Blackwell companion to syntax, Malden, MA, USA: Blackwell Pub.

Faltz, Leonard (1985), *Reflexivization: a study in universal syntax*, New York: Garland Pub.

Frajzyngier, Zygmunt, and Curl, Tracy S. (eds.), (2000), *Reciprocals: Forms and Functions*, Amsterdam: John Benjamins Publishing Company.

Geniušienė, Emma (1987), *The typology of reflexives*, Berlin: De Gruyter Mouton.

Haspelmath, Martin (2008), A Frequentist Explanation of Some Universals of Reflexive Marking, *Linguistic Discovery* 6(1), 40-63.

Huang, Yan (2000), *Anaphora: a cross-linguistic approach*, Oxford: Oxford University Press.

Kazenin, Konstantin (2001), Verbal reflexives and the middle voice, in M. Haspelmath (et al.) (eds.), (2001), *Language Typology and Language Universals*, Berlin: Mouton De Gruyter, 916-927.

Keenan, Edward (1988), On semantics and the binding theory. In J. Hawkins (ed.), (1988), *Explaining Language Universals*, Oxford: Blackwell, 105-44.

Kemmer, Suzanne (1993), *The Middle Voice*, Amsterdam: John Benjamins Publishing Company.

König, Ekkehard (2001), Intensifiers and Reflexives, in M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.), (2001), *Language typology and language universals*, vol. 1, Berlin: Mouton de Gruyter, 747-60.

König, Ekkehard, and Gast, Volker (eds.), (2008), *Reciprocals and Reflexives: Theoretical and Typological Explorations*, Berlin: De Gruyter Mouton.

König, Ekkehard, and Siemund, Peter (2000), Intensifiers and reflexives: a typological perspective, in Z. Frajzyngier, and T. Curl (eds.), (2000), *Reciprocals: Forms and Functions*, Amsterdam: John Benjamins Publishing Company, 41-74.

Puddu, Nicoletta (2017). Nominal versus verbal reflexives: a categorial opposition?, *Presentation at the 50th Annual Meeting of the Societas Linguistica Europaea*, Zürich, 10-13 September 2017.

Reinhart, Tanya (1983), *Anaphora and Semantic Interpretation*, London: Croom Helm.

Reinhart Tanya, and Reuland, Eric (1993), Reflexivity, *Linguistic Inquiry* (24), 657-720.

Volkova, Anna (2014), *Licensing Reflexivity: Unity and variation among selected Uralic languages*. Doctoral dissertation, Utrecht University.

Reflexives in Ruuli (Bantu, JE103)

Alena Witzlack-Makarevich, Amos Atuhairwe and Erika Just

<pdf>

Reflexive systems in contact and the acquisition of binding: The case of Mano and Kpelle

Maria Khachaturyan & Anna Volkova
(University of Helsinki & National Research University Higher School of Economics)

Keywords: binding, locally bound pronominals, Mande, language acquisition, contact phenomena

The paper contrasts the anaphoric systems of two Mande languages: Mano (Southern Mande) and Kpelle (Southwestern Mande) and focuses on the contact between them. While Mano employs a dedicated 3rd person sg reflexive pronoun *ē* to express local binding, Kpelle lacks a dedicated reflexive pronoun. The speech of monolingual, early bilingual of different age and L2 Mano speakers show different patterns for expressing reflexivity.

The Mano reflexive pronoun *ē* (1a) contrasts with the non-reflexive pronoun *à* (1b). The reflexive pronoun can be used alone forming a SIMPLE reflexive marker (2) or be accompanied by the intensifier *dìè* forming a COMPLEX reflexive marker *ē dìè* (1a).

In Kpelle, a 3rd person sg pronominal prefix, which is normally ambiguous between a free and a bound reading (3b, 4), can be attached to an intensifier *kpɔ́yɔ́*, the combination always yielding a locally bound reading (3a).

Mano-speaking individuals with exposure to Kpelle show different patterns of variation in reflexive marking, allowing locally bound use of the pronominal *à*. A preliminary study of 20 individuals shows that the influence of Kpelle is the strongest in late L2 speakers with limited exposure to Mano and in young bilingual children. These individuals allow the *à* pronominal in all contexts tested. If the dedicated pronoun is ever used, it is used primarily in the direct object position. Approximately by the age of 12 early bilinguals perform in Mano almost like monolingual adults. However, there are two contexts where the variation persists even in the speech of adult L1 speakers of Mano: in the postpositional phrase (5) and in the possessor position of an inclusory construction, where the pronoun *à* is as frequent as the pronoun *ē* (6).

The late acquisition and the persistence of variation in the peripheral syntactic positions can be accounted for by the cross-linguistic tendency of dedicated reflexives to be used in core syntactic positions (Testelets and Toldova 1998). The direct object is the position where the variable needs most to be protected due to the IDI principle (Reuland 2011).

(1a) *mīī* *dò* *lèé* *ē* *dìè* *gɛ̀*
person INDEF 3SG.NEG 3SG.REFL REFL.INT see
'No-one sees himself.' (Mano)

(1b) *Kɔ̀ɔ́* *lèé* *à* *gɛ̀*
Ko 3SG.NEG 3SG see
'Ko does not see him/*herself.' (Mano)

(2) *ē* *ē* *zúlú*
3SG.PST 3SG.REFL wash
'He washed himself.' (Mano)

(3a) *nú* *tà* *hwó* *gbɔ̀yɔ̀* *káá-í*
person INDEF 3SG.BASE.NEG 3SG\REFL.INT see-PROG
'No-one sees himself.' (Kpelle)

(3b) *Kɔ̀lú* *hwó* *gàà-ì*
Ko 3SG.BASE.NEG 3SG\see-PROG
'Ko does not see him/herself.' (Kpelle)

(4) *ě* *lɔ́* *ɲèì* *bɛ̀lɛ́ŋ̀*
3SG.CONJ enter 3SG\hand.LOC DEF\house.LOC

'He entered his own/her house.' (Kpelle; Konoshenko, p.c.)

(5) *áà* *nù* *ē /* *à* *pàà*
3SG.JNT come:JNT 3SG.REFL 3SG at
'He came home (lit.: he came at him).' (Mano)

(6) *wà* *ē /* *à* *lòkóò*
3PL.IP 3SG.REFL 3SG mother
'he$_i$ and his$_i$ mother' (Mano)

References:

Khachaturyan, Maria. 2019. Inclusory pronouns in Mande: The emergence of a typological rarum. *Folia Linguistica* 53(1), 87–123.

Reuland, Eric. 2011. *Anaphora and language design*. MIT Press.

Testelets, Yakov G. and Svetlana Y. Toldova. 1998. Refleksivnye mestoimenija v dagestanskih jazykah i tipologija refleksiva. *Voprosy jazykoznanija* 4, 35–57.

Reflexivity in Khanty

Anna Volkova
(National Research University Higher School of Economics)

Keywords: binding, anaphora, locally bound pronominals, Uralic, variation

Khanty (Uralic) is a language that allows locally bound pronominals (Nikolaeva 1995). In my talk I contrast a variety of Khanty spoken in the village of Tegi (Volkova & Reuland 2014) with a variety spoken in the village of Kazym. Both varieties of Khanty have two types of verbal agreement: obligatory subject agreement and optional object agreement, as illustrated in (1).

(1) UtltiteXo poXlen'ki išək-s-əłłe / išək-s. [Tegi Khanty]
teacher boy praise-PST-SG.3SG / praise-PST.3SG *The teacher praised the boy.*

In Tegi Khanty a personal pronoun can be locally bound only if the verb carries object agreement (2), cf. the ill-formedness of (3).

(2) UtltiteXo$_i$ łuveł$_{i/k}$ išək-s-əłłe. [Tegi Khanty]
teacher he.ACC praise-PST-SG.3SG *The teacher praised himself / him.*

(3) *UtltiteXo$_i$ łuveł$_i$ išək-s. [Tegi Khanty]
teacher he.ACC praise-PST.3SG *The teacher praised him / *himself.*

In Kazym Khanty the inter-speaker variation is very high. While some speakers adhere to the pattern used in Tegi Khanty, others use one of the alternatives:

- a combination of a personal pronoun and object agreement is dedicated to expressing local binding, disjoint reading is unavailable (4). If the verb carries only subject agreement (5), the pronoun can take either a bound or a disjoint reading.

(4) Kašen Xö$_i$ łuwti$_{i/*j}$ išək-ł-əłłe. [Kazym Khanty, sp10]
every man he.ACC praise-NPST-SG.3SG *Every man praises himself /*him.*

(5) Kašen Xö$_i$ łuwti$_{i/j}$ išək-əł. [Kazym Khanty, sp10]
every man he.ACC praise-NPST[3SG] *Every man praises himself /him.*

- a combination of a personal pronoun and object agreement is dedicated to expressing disjoint reading, bound reading is unavailable (6). If the verb carries only subject agreement (7), the pronoun can take either a bound or a disjoint reading.

(6) Kašen učitel'$_i$ łuwti$_{j/*i}$ išək-s-əłłe. [Kazym Khanty, sp04]
every teacher he.ACC praise-PST-SG.3SG *The teacher praised him / *himself.*

(7) Kašen učitel'$_i$ łuwti$_{i/j}$ išək-əs. [Kazym Khanty, sp04]
every teacher he.ACC praise-PST[3SG] *The teacher praised him / himself.*

Some Kazym Khanty speakers also make use of a doubled pronoun which can be combined with both subject (8) and object (9) agreement on the verb:

(8) Učitel'-ət$_i$ łiwti łiw$_{i/*j}$ išək-s-ət. [Kazym Khanty, sp07]
teacher-PL they.ACC they praise-PST-3PL *The teachers praised themselves*.

(9) Učitel'-ət$_i$ łiwti łiw$_{i/*j}$ išək-s-əłłał. [Kazym Khanty, sp07]
teacher-PL they.ACC they praise-PST-PL.3PL *The teachers praised themselves*.

This level of variation poses a puzzle for any binding theory. Kazym Khanty creates an additional challenge, as speakers tend to avoid using negative or universal quantifiers in subject positions, hence the Heim test (Heim 1982) for binding is not available. The use of doubled pronouns in Kazym Khanty indicates that irrespective of the type of agreement on the verb, the variable needs to be protected by adding complexity (Reuland 2011). I discuss how this degree of variation can be addressed and what other syntactic properties we can expect to co-vary between speakers.

References

Heim, Irene (1982), The semantics of definite and indefinite noun phrases. Amherst, MA: University of Massachusetts PhD thesis.

Nikolaeva, Irina (1995), *Obdorskij dialekt xantyjskogo jazyka* (*The Obdorsk dialect of Khanty)* [in Russian]. Hamburg: Mitteilungen der Societas Uralo-Altaica.

Reuland, Eric (2011), *Anaphora and language design*. MIT Press.

Volkova, Anna and Eric Reuland (2014), Reflexivity without reflexives? *The Linguistic Review*, Vol. 31. No. 3-4. P. 587-633.

The reflexive pronoun *ži* in Northern Akhvakh: A corpus-based study

Denis Creissels
(University of Lyon)

Keywords: reflexivity, long-distance reflexives, logophoricity, Nakh-Daghestanian languages, Akhvakh

Northern Akhvakh, spoken in western Daghestan, belongs to the Andic sub-branch of the Avar-Andic(-Tsezic) branch of the Nakh-Daghestanian family. It is documented through Magomedbekova's (1967) monograph, Magomedova & Abdulaeva's (2007) dictionary, and a series of articles by the author of this presentation (Creissels 2008, 2009, 2010, 2012, 2013, 2014, 2016a, 2016b, 2017, 2018). The analysis of the reflexive pronoun **ži** of Northern Akhvakh is based on a corpus of oral texts collected mainly in Tadmagitl' and Logonitl' (Daghestan) including 625 occurrences of **ži**.

ži (whose case inflection involves a suppletive stem **ĩ-**) occurs in bare form and in a form enlarged by the addition of the intensifying particle **-da**. **ži** in its bare form is used as a long-distance reflexive, both in logophoric (1) and non-logophoric (2) contexts.

(1) **wac**:**oga** **eƛ’**:**awi,** **“ĩs**:**uƛa** **komokiɬ**:**a** **woq’a!”**
brother:ALL say:CPL:N ŽI:M:DAT help:DAT come:M:IMP
‘He$_i$ said to (his$_i$) brother “Come to my$_i$ aid!”’

(2) **bakala** **ox**:**ewi** [**ĩ-ɬ**:**i-ƛa** **komoki** **gweda**] **ak’**:**aɬ**:**iƛa** **šĩde.**
thanks give:CPL:N ŽI:N:DAT help do:PTCP woman:DAT bear:ERG
‘The bear$_i$ thanked the woman who had helped it$_i$.’

The addition of the intensifying particle **-da** to **ži** gives 3rd person intensive pronouns, and like the **da**-form of 1st and 2nd person pronouns, the **da** form of **ži** also has reflexive uses, and in strictly local configurations, the particle **-da** is required (3). The particle **-da** is also required in reflexive configurations involving a term in the construction of a verb and a genitive modifying another term in the construction of the same verb (4), and finally, **ži** can be used as a long-distance reflexive in non-logophoric configurations not only in its bare form, but also in the **da**-form (5).

(3) **χãs**:**ude** **žo**:**wudi** **ĩs**:**uƛ**:**ira-da** **ĩk’**:**was**:**e** **waša.**
king:ERG call:CPL:M ŽI:M:ALL-INT small:SEL boy
‘The king called (his) younger son to himself.’

(4) **qe** [**šĩ**] **maʔewidi** [[**ĩɬ**:**iƛ**:**i-da**] **rec’**:**aɬ**:**iq**:**a**].
then bear go:CPL:N ŽI:N:GEN-INT cave:ALL
‘Then the bear went to its cave.’

(5) **dibi** [[**ĩs**:**u-da** **č’iliɬ**:**ige** **k’onada**] **č’ari**] **bis**:**o**:**ruƛa** **w-oƛ**:**-ari.**
molla ŽI:M:GEN-INT house:LOC start:PTCP fire put.out:N:INF get.off:M:CPL
‘The molla got off to extinguish the fire that had started in his house.’

In this corpus study, particular attention is devoted to the syntactic functions fulfilled by **ži** and its antecedent. The three main conclusions are that (a) reflexivization in Northern Akhvakh supports the traditional distinction between ‘syntactic’ and ‘semantic’ cases (since the only NPs commonly involved in reflexivization are ergative arguments, nominative arguments, dative arguments, and adnominal genitives), (b) among syntactic cases, ergative arguments are particularly prone to acting as antecedents of **ži**, and (c) **ži** is particularly prone to occurring in adnominal genitive function.

Abbreviations

ALL allative, CPL completive, DAT dative, ERG ergative, GEN genitive, IMP imperative, INF infinitive, INT intensive, LOC locative, M masculine, N neuter, PTCP participle, SEL: selective.

References

Creissels, Denis. 2008. Person variations in Akhvakh verb morphology: functional motivation and origin of an uncommon pattern. *Sprachtypologie und Universalienforschung* 61(4). 309-325.

Creissels, Denis. 2009. Participles and finiteness: The case of Akhvakh. *Linguistic Discovery* 7.1. 106-130.

Creissels, Denis. 2010. Specialized converbs and adverbial subordination in Axaxdere Akhvakh. In Isabelle Bril (ed.) *Clause linking and clause hierarchy: syntax and pragmatics*. John Benjamins. 105-142.

Creissels, Denis. 2012. External agreement in the converbal construction of Northern Akhvakh. In Volker Gast and Holger Diessel (eds.) *Clause linkage in cross-linguistic perspective*. De Gruyter Mouton. 127-156.

Creissels, Denis. 2013. Floating genitives and possessive framing in Northern Akhvakh. In Anne Carlier and Jean-Christophe Verstraete (eds.) *The genitive*. Benjamins. 333-354.

Creissels, Denis. 2014. Functive-transformative marking in Akhvakh and other Caucasian languages. In V.A. Plungjan, M.A. Daniel, E.A. Ljutikova, S.G. Tatevosov & O.V. Fedorova (eds.) *Jazyk. Konstanty. Peremennye. Pamjati Aleksandra Evgen'eviča Kibrika*. Saint-Petersburg: Aleteja. 430-449.

Creissels, Denis. 2016a. Word-formation in Akhvakh. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, and Franz Rainer (eds.) *Word-Formation, An International Handbook of the Languages of Europe, volume 5*. De Gruyter Mouton.

Creissels, Denis. 2016b. Univerbation via liaison and the evolution of lexicon and grammar in Northern Akhvakh. *Languages of the Caucasus* 1(1). 107-125.

Creissels, Denis. 2017. Transitivity and valency in Northern Akhvakh. *Linguisticae Investigationes* 40(1). 82-117.

Creissels, Denis. 2018. Perfective tenses and epistemic modality in Northern Akhvakh. In Diana Forker & Timur Maisak (eds.), *The semantics of verbal categories in Nakh-Daghestanian languages*. Leiden / Boston: Brill. 166-187.

Magomedbekova, Z.M. 1967. *Axvaxskij jazyk (grammatičeskij analiz, teksty, slovar').* Tbilissi: Mecniereba.

Magomedova, P & I. Abdulaeva. 2007. *Axvaxsko-russkij slovar'.* Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Bantu reflexives: A comparative case-study of the Kikongo Language Cluster

Sebastian Dom & Leonid Kulikov
(Ghent University)

Keywords: reflexive, middle, Bantu, Kikongo Language Cluster, micro-variation

This paper presents a micro-variational study of reflexivity in Bantu languages belonging to the Kikongo Language Cluster, a continuum of 40-odd closely related language varieties spoken in Gabon, the Republic of the Congo (Congo-Brazzaville), the Democratic Republic of the Congo (Congo-Kinshasa), and Angola (including Cabinda) (de Schryver et al. 2015). The main focus of this paper is on a) formal encoding and microvariation, and b) reflexive-middle polysemy.

Reflexivity is encoded in all Kikongo varieties by means of a verbal morpheme that immediately precedes the verb root, which is also the most commonly found strategy in Bantu (Polak 1983). Within the morphological structure of the verb, this pre-root morpheme slot is also where object indexes occur. This is illustrated in (1) with examples from Kimbata.

(1) a. Reflexive
*tu-taa-**ki**-tal-a*
1PL.SUBJ-PRS.PROG-REFL-look-PRS
'We are looking at ourselves.'

b. Object indexing
*tu-taa-**ba**-tal-a*
1PL.SUBJ-PRS.PROG-3PL.G2.OBJ-look-PRS
'We are looking at them.'

The pre-root position of the reflexive morpheme is striking, given that all (other) valency-changing morphology occurs after the verb root (Schadeberg 2003). This is illustrated in (2) with the reciprocal suffix *-an* in Kimbata.

(2) Reciprocal
*tu-taa-tal-**an**-a*
1PL.SUBJ-PRS.PROG-look-RECP-PRS
'We are looking at each other.'

Formal variation of the reflexive prefix exists between different Kikongo varieties, as well as within one single variety between different tense/aspect constructions. A full inventory of the attested morphemes is given in (3).

(3) Formal variation of reflexive prefixes in the Kikongo Language Cluster
a. CV *ki-*, *ku-*, *ke-*, *di-*, *du-*, *li-*, *lu-*
b. (G)V (*y*)*i-*

Most CV forms are identical in shape to other verbal prefixes such as the infinitival prefix *ku-* or the object indexes of noun classes 5 *di-/li-* and 7 *ki-*. Hence, it is unlikely that these forms are reflexes of the Proto-Bantu reflexive prefix **ji-* (Polak 1983).

In many languages of the world, verbal reflexive forms are sources of middle voice morphology (Kemmer 1993, Kazenin 2001). Bantu languages provide valuable evidence for typology of middle voice, as multiple valency-decreasing affixes in various Bantu languages tend to develop middle voice polysemy patterns (Dom et al. 2016). Most interestingly, the Bantu derivational system with a variety of affixes encoding individual valency-decreasing categories, inherently delimits the evolution of these affixes into a full-fledged middle voice marker. This seems to be the case in the Kikongo varieties, where the only attested 'secondary' use of the reflexive prefix is the encoding of the autobenefactive meaning, in which it is obligatorily combined with the applicative suffix, as in the Kimanyanga example in (4).

(4) Auto-benefactive
yandi *u-a-**ki**-tung-**il**-a*
3SG.PRON 3SG.SUBJ-PST-REFL-build-APPL-PST
'He built a house for himself.'

In other words, we only observe the development 'reflexive > middle' at its initial stage. Obviously, this cross-linguistically very common grammaticalization path is not (yet) completed due to the inherent characteristics of the derivational system of the valency-changing categories and, especially, due to the fact that multiple affixes compete for the expression of middle voice.

References

de Schryver, Gilles-Maurice & Grollemund, Rebecca & Branford, Simon & Bostoen, Koen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87-162.

Dom, Sebastian & Kulikov, Leonid & Bostoen, Koen. 2016. The middle as a voice category in Bantu: Setting the stage for further research. *Lingua Posnaniensis* 58(2). 129-149.

Kazenin, Konstantin I. 2001. Verbal reflexives and the middle voice. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds), *Language typology and language universals: An international handbook*, vol. 2 (Handbooks of Linguistics and Communication Science 20). Berlin: de Gruyter.

Kemmer, Suzanne. 1993. *The middle voice* (Typological Studies in Language 23). Amsterdam: John Benjamins.
Polak, Louise. 1983. Le réfléchi en bantou. *Africana Linguistica* 9. 271-304.
Schadeberg, Thilo C. 2003. Derivation. In Nurse, Derek & Philippson, Gérard (eds), *The Bantu languages* (Routledge Language Family Series 4), 71-89. New York: Routledge.

Ø

The semantics of verbal reflexives in Kihehe

Lengson Ngwasi
(University of Gothenburg)

Keywords: reflexive, reciprocal, middle, passive, grammaticalization

The meanings associated with verbal reflexives have been one of the central areas of research cross-linguistically (cf. Geniušienė (1987); Kazenin (2001); Kemmer (1993)). The discussion has focused on the way verbal reflexives extend their functions to encode other meanings like reciprocal, middle, passive etc. It has been argued and shown by Geniušienė (1987); Kemmer (1993) and Heine (2000) that this extension in meaning is due to grammaticalization process leading from Nominal > Emphatic reflexive > Reflexive > Reciprocal > Middle > Passive, cross-linguistically. However, based on the fieldwork data from Kihehe (G62), I show that this grammaticalization cline is problematic for Bantu languages, for three reasons:

First, it has been established in Bantu that reflexive and reciprocal meanings are encoded by the reflexes of the reconstructed proto-Bantu verbal affixes **-(j)i-* and **-an-* respectively (cf. Meeussen (1967)). So far, there is no evidence suggesting that the sources of the verbal reflexives and reciprocals are nominals or emphatic reflexives in Bantu languages. The reconstructed proto-reflexive and reciprocal markers are just verbal affixes (see also Schadeberg (2003)).

Second, unlike in many Bantu languages, the reflexive prefix *-i-* (cf.1) has replaced the reciprocal suffix *-an-* and encodes reciprocal meaning in Kihehe (cf. 2a). Thus, the reflexive prefix *-i-* is ambiguous between reflexive and reciprocal meaning. The suffix *-an-* is found in a few constructions encoding inherent reciprocal situations (cf. 2b). In some cases, it is fossilized and the reflexive prefix is recruited for reciprocal meaning (cf. 2c).

1. *Juma a - ku - i - on - ag - a mu-ki-lole*
   Juma 1.SM-PRS-REFL-see-HAB-FV LOC-CL.7-mirror
   'Juma sees himself in the mirror'
2a. *Juma na Jeni va -ku - i - on - a mu - ki - lole*
   Juma Conj Jane 2.SM-PRS-REFL-see-FV LOC-CL.7-mirror
   'Juma and Jane see each other in the mirror'
b. *Juma na Jeni va - Ø - gav -an - a ma-tunda*
   Juma Conj Jane 2.SM-FUT-divide-REC-FV CL.6-fruit
   'Juma and Jane will share fruits'
c. *Juma na Jeni va - ku - i - taang'an-a neng'uni*
   Juma Conj Jane 2.SM-PRS-REFL-meet - FV today
   'Juma and Jane are meeting today'

Finally, the grammaticalization from Reflexive > Reciprocal > Middle > Passive is not universal for all Bantu languages with a polysemous reflexive marker. For instance, in Swahili (G42), the prefix *-ji-* encodes reflexive and middle, not reciprocal meaning. Based on Swahili and the fossilization of

the reflexive prefix *-i-* in some middle constructions in Kihehe (cf. 3b and c), I argue that the grammaticalization from the reflexive to other meanings does not necessarily abide to the cross-linguistic chain. In other words, it is not necessary for it to pass through the reciprocal stage to become a middle marker. Thus, the grammaticalization might be parallel or following this path: Reflexive > Middle > Reciprocal > Passive.

3a. *ku-i-chanula*	'to comb' (grooming)
b. *ku-ika*	'stepping down' (translational)
c. *ku-iwuka*	'to remember' (cognitive action)

In addition, I show that the reflexive prefix in Kihehe does not encode passive meaning. The passive meaning is encoded by the suffix *-w-*.

References

Geniušienė, E. (1987). *The typology of reflexives: Empirical Approaches to Language typology* (Vol. 2). Berlin; New York: Mouton de Gruyter.

Heine, B. (2000). Polysemy involving reflexive and reciprocal markers in African languages. In Z. Frajzyngier & T. Walker (Eds.), *Reciprocals: Forms and functions* (Vol. 2, pp. 1-30). Amsterdam; Philadephia.

Kazenin, K. I. (2001). Verbal reflexives and the middle voice. *Language typology and language universals, 2*, 916-927.

Kemmer, S. (1993). *The middle voice* (Vol. 23). Amsterdam; Philadephia: John Benjamins Publishing.

Meeussen, A. E. (1967). Bantu grammatical reconstructions. *Africana linguistica, 3*(1), 79-121.

Schadeberg, T. (2003). Derivation. In D. Nurse & G. Phillipson (Eds.), *The Bantu Languages* (pp. 71-89). London: Routledge.

The emergence of reflexive markers: the case of Hittite

Guglielmo Inglese
(Università degli Studi di Pavia)

Keywords: middle voice, reflexive, diachronic typology, Hittite

Typological research has shown that reflexive markers come about from different sources following selected paths of grammaticalization (cf. König & Siemund 2000, Schladt 2000, Heine & Kuteva 2002, Evseeva & Sakeberri 2018). In the case of polyfunctional markers that encode multiple valency reducing operations, already Kemmer (1993) suggested that it is reflexives that constitute the cross-linguistic dominant source of other functions, such as passive and anticausative. A textbook example of this development is the history of the reflexive marker in Romance languages (cf. Kemmer 1993, Puddu 2005). The unidirectionality of the REFLEXIVE > ANTICAUSATIVE (> PASSIVE) cline has remained virtually unchallenged in the literature (Haspelmath 1990, 2003, Heine 2002). However, this is not the entire picture. Even though with a much narrower distribution, reverse developments have been indeed documented (e.g. Dik 1983 on PASSIVE > REFLEXIVE in Uto-Aztecan). In this paper, in contrast with the received view concerning the unidirectionality of the development of reflexives, I argue that evidence for yet another possible source of reflexives, i.e. anticausatives, can be found in Hittite.

Hittite features a two-fold system of verbal voice opposition between the active and the middle inflection (Neu 1968, Hoffner & Melchert 2008, Luraghi 2012). Verbs that belong to the middle voice

can either be *media tantum*, i.e. they inflect in the middle voice only, e.g. *kiš-a(ri)* 'become', or they can be opposed to corresponding active verbs. Oppositional middles encode a range of intransitivizing functions, including reflexive, anticausative, passive, and reciprocals, as in (1). Based on the cross-linguistic behavior of reflexive markers, one expects the Hittite middle to originate as a reflexive marker and later expand its functional scope to the encoding of the other functions. However, a closer look at the distribution of middle verbs in the corpus of original Hittite texts shows that this hypothesis is untenable. In the earliest phase of the language, i.e. Old Hittite, most middle verbs belong to the *media tantum* (Tab. 1), and the few oppositional ones are predominantly anticausatives (Tab. 2), while reflexive middles remain rare throughout the history of the language. Moreover, since their earliest attestations, reflexive middles are often additionally reinforced by the reflexive particle *=za* (cf. Hoffner & Melchert 2008, Cotticelli Kurras & Rizza 2011, 2013). These facts suggest that reflexivity hardly constituted the original core of the middle inflection. As an alternative scenario, I argue that the Hittite middle was originally confined to the group of *media tantum*, which included verbs that denote spontaneous change-of-state events. These were liable to be opposed to active counterparts, giving rise to the anticausative alternation. The anticausative function then served as the starting point for the rise of the other oppositional functions, including the reflexive. In this shift, autocausatives events of self-induced motion, e.g. *ne-ari* 'turn (intr.)', constitute a key bridging context. This scenario is also fully compatible with recent models of reconstruction of verbal voice in Proto-Indo-European (Luraghi forthc.).

Examples

(1) *zinna-**tta(ri)*** 'finish (intr.)' (anticausative), *tamašš-**ta(ri)*** 'be oppressed' (passive), *šuppiyaḫḫ-**ta(ri)*** 'purify oneself' (reflexive) *zaḫḫ-**anda*** 'hit each other' (reciprocal).

References

Cotticelli Kurras, Paola & Alfredo Rizza. 2011. Die hethitische Partikel *-z(a)* im Licht neuer theoretischer Ansätze. In *Indogermanistik und Linguistik im Dilog. Akten der XIII. Fachtagung der Indogermanischen Gesselschaft vom 21. bis 27. September 2008 in Salzburg*, Thomas Krisch & Thomas Lindner (eds.), 12–30. Wiesbaden: Reichert.

Cotticelli Kurras, Paola & Alfredo Rizza. 2013. Reconstructing Proto-Indo-European Categories: The Reflexive and the Middle in Hittite and the Proto-language. *Journal of Historical Linguistics* 3 (1): 7–27.

Dik, Simon. 1983. The status of verbal reflexives. In *Problems in Syntax*, Liliane Tasmowski & Dominique Willems (eds.), 231-256. New York/London: Plenum Pres

Evseeva, Natalia & Iker Saleberri. 2018. Grammaticalization of nouns meaning 'head' into reflexive markers. A cross-linguistic study. *Linguistic Typology* 22 (3): 385-435.

Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Studies in Language* 14 (1): 25-71.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In *The new psychology of language: Cognitive and Functional Approaches to Language Structure, Vol. 2*, Michael Tomasello (ed.), 211-242. Mahwah, NJ: Lawrence Erlbaum.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization* [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds.), 83-101. Amsterdam/Philadelphia: John Benjamins.

Hoffner, Harry A. Jr. & H. Craig Melchert. 2008. *A Grammar of the Hittite Language. Part I: reference grammar*. Winona Lake, Indiana: Eisenbrauns.

Kemmer, Susanne. 1993. *The middle voice* [Typological Studies in Language 23]. Amsterdam/Philadelphia: John Benjamins.

König, Ekkehard & Peter Siemund. 2000. Intensifiers and reflexives: a typological perspective. In *Reflexives: Forms and functions* [Typological Studies in Language 41], Zygmunt Frajzyngier & Traci S. Curl (eds.), 41-74. Amsterdam/Philadelphia: John Benjamins.

Luraghi, Silvia. 2012. Basic valency orientation and the middle voice in Hittite. *Studies in Language* 36 (1): 1-32.

Luraghi, Silvia. Forthcoming. Basic valency orientation, the anticausative alternation, and voice in PIE. To appear in the *Proceedings of the 15. Fachtagung der Indogermanischen Gesellschaft.* Wien (Austria), 13-16 September 2016.

Neu, Erich. 1968. *Das hethitische Mediopassiv und seine indogermanischen Grundlagen* [Studien zu den Bogazköy-Texten 6]. Wiesbaden: Harrassowitz.

Puddu, Nicoletta. 2005. *Riflessivi e intensificatori: greco, latino e le altre lingue indoeuropee*. ETS Edizioni: Pisa.

Schladt, Mathias. 2000. The typology and grammaticalization of reflexives. In *Reflexives: Forms and functions* [Typological Studies in Language 41], Zygmunt Frajzyngier & Traci S. Curl (eds.), 103-124. Amsterdam/Philadelphia: John Benjamins.

Tables

Table 4: Distribution of *media tantum* in Old Hittite

	Media tantum	**Oppositional middles**
Old Hittite	27 (79%)	7 (21%)

Table 5: Diachronic distribution of oppositional middles in Hittite

	Anticausative	Passive	Anticaus./Pass.	Reciprocal	Reflexive
Old Hittite	3	2	0	1	1
Middle Hittite	6	7	6	0	1
New Hittite	9	13	8	0	0

Reconstructing the PIE reflexive function within an active alignment system

Verónica Orqueda

After a long time of debate, it has now been accepted the possibility that Proto- Indo-European displayed, at some stage, an active alignment system. Despite the wide range of active patterns that scholars have suggested for the reconstruction (Clakson, 2007), a common feature seems to be the semantic motivation in role marking. In particular, as Barðdal & Eythórsson (2009) and Pooth et al. (2018) claim, the opposition of agentive/anti-agentive case was determined by factors such as volitionality and self-motivation.

Although reflexivity is not usually a test for alignment, it would not be strange that particular reflexive strategies would be preferred or avoided according to the alignment system of each language. In particular this seems to be true if we accept syntactic and semantic constraints for reflexive clauses,

such as a high degree of agentivity (+volition and +control) of the Actor and the unacceptance of state predicates, which would locate reflexive constructions within the active sphere. As for typological evidence, data points towards a possible connection between accusative systems (or at least, head-dependent languages) and clitic/nominal strategies. Contrarily, semantic alignment systems seem to be more frequently associated to valency-reducing strategies, such as detransitivizing mechanisms.

In the light of such possible relations, the aim of this presentation is to analyze the expression of reflexivity in Proto-Indo-European under the hypothesis of an active alignment system. As I will try to show, the reconstruction of a nominal reflexive **swe-* is not necessary at an early stage of the language, and this is consistent with previous studies, such as the one by Puddu (2005) among others, who claim that **swe-* was not primarily reflexive. The same can be said of the middle voice, if we accept that the distinction between the active and the middle did not have any syntactic bases in reconstructed PIE (Clackson, 2007).

Following the hypothesis of polysemy of the antipassive and the reflexive constructions (Janic, 2016), I will support the plausible emergence of different syntactically-based reflexive constructions from a 'non-canonical antipassive' construction. Such a non-canonical antipassive construction, where the first participant was marked by the agentive case and the verb was marked for proto-middle (ditransitive) voice (*-o-*), is already proposed by Pooth et al. (2018).

By analyzing the fact that most daughter languages slowly changed from a semantically towards a syntactically-motivated system, I will also conclude that the reanalysis of the allative case as accusative explains well the need for disambiguating the reflexive constructions by filling a syntactic gap with different nominal strategies.

References

Barðdal, J. & Eythórsson, Th. (2009). "The origin of the oblique subject construction: An Indo-European comparison". In: Grammatical change in Indo-European languages. Ed. By Vit Bubenik, John Hewson & Sara Rose, 179-193. Amsterdam: John Benjamins.

Clackson, J. (2007). *Indo-European Linguistics*. Cambridge: CUP.

Janic, K. (2016). "On the reflexive-antipassive polysemy: typological convergence from unrelated languages". *Berkeley Linguistics Society* (36), 158–173.

Pooth, R., Kerkhof, P., Kulikov, L., Barðdal, J. (2018). "The Origin of Non-Canonical Case Marking of Subjects in Proto-Indo-European: Accusative, Ergative, or Semantic Alignment". Preprint. Available at https://www.researchgate.net/publication/307142782_On_the_Reflexive-Antipassive_Polysemy_Typological_Convergence_from_Unrelated_Languages.

Puddu, N. (2005). *Riflessivi e intensificatori: greco, latino e le altre lingue indoeuropee*. Pisa: ETS.

Reflexive constructions and areal relationships: A typological-diachronic analysis of replica grammaticalization

Anne C. Wolfsgruber & Iker Salaberri
(Humboldt-Universität zu Berlin & Public University of Navarre)

Keywords: body-part reflexives, areal relationships, replica grammaticalization, Romance languages, quantitative corpus data

Reflexive markers vary cross-linguistically, encompassing both nominal and verbal strategies (Faltz 1977: 15-109, König 2007: 107-124). In addition, reflexive constructions have been argued to be prone to spreading areally (Heine 1999: 9-10). Types of contact such as substrate and superstrate have

been related to this tendency: for example, the fact that Haitian Creole French presents body-part reflexives (BPRs) such as *kò* 'body' + pronoun and *tèt* 'head' + pronoun has been associated with a substratum of Kwa languages, in which such reflexive constructions are widespread (Lefebvre 1998: 170-171).

A problem with this view is, however, that Indo-European-based pidgins and creoles which do not have a demonstrable African substrate and are spoken outside the Atlantic area, such as Tok Pisin and Zamboanga Chavacano, also have BPRs (Michaelis et al. 2013). This suggests that the existence of such reflexives in Indo-European-based pidgins and creoles may be due not only to substrate influences, but also to the presence of BPRs in the late-Renaissance source languages (Chaudenson 1973: 360). Stated differently, the presence of BPRs in Tok Pisin and Zamboanga Chavacano suggests that many Indo-European-based pidgins and creoles may have undergone what has been labeled 'replica grammaticalization' in the literature (Heine & Kuteva 2003: 539, 2005: 92 and Ziegeler 2017: 316-319). This proposal has been rejected by some authors (Jourdain 1956: 139, Lefebvre 1998: 168-169), but never on the grounds of quantitative corpus data.

In view of this state of affairs, this study analyzes the occurrences of BPRs in the history of five Romance languages: Catalan, French, Occitan, Portuguese and Spanish. The data under investigation encompass texts written between 1400 and 1700 CE and have been extracted from digital corpora: *Frantext*, *Corpus diacrónico del español*, *Corpus do português*, *Corpus informatitzat del català antic* and *Concordance de l'occitan médiéval*. First results show that non-standard dialects and fixed expressions point toward the existence of BPRs in Romance languages at least until the late-17th century, which supports the view that the source languages may have passed on these reflexive constructions onto pidgins and creoles.

References (corpora)

Concordance de l'Occitan Médiéval (2001) (COM) 1. Turnhout: Brepols.

Concordance de l'Occitan Médiéval (2005) (COM) 2. Turnhout: Brepols.

Corpus Informatitzat del Català Antic (CICA). www.cica.cat (02/11/2018).

FRANTEXT (2018). https://www.frantext.fr/ (02/11/2018).

O Corpus do Português (CP). http://www.corpusdoportugues.org/ (02/11/2018).

REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) [en línea]. Corpus diacrónico del español. http://www.rae.es (02/11/2018).

References (literature)

Chaudenson, Robert (1973), Pour une étude comparée des créoles et parlers français d'outre-mer: Survivance et innovation, *Revue de linguistique romane* 147-148, 342-371.

Faltz, Leonard M. (1977), *Reflexivization: A study in universal syntax*, Los Angeles: University of California dissertation.

Heine, Bernd (1999), Polysemy involving reflexive and reciprocal markers in African languages, in Z. Frajzyngier, and T. Curl (eds.), *Reciprocals: Forms and functions*, Amsterdam / Philadelphia: John Benjamins, 1-29.

Heine, Bernd & Tania Kuteva (2003), On contact-induced grammaticalization, *Studies in Language* 27(3), 529-572.

Heine, Bernd & Tania Kuteva (2005), *Language contact and grammatical change*, Cambridge: Cambridge University Press.

Jourdain, Élodie (1956), *Du français aux parlers créoles*, Paris: C. Klincksieck.

König, Ekkehard (2007), Vers une nouvelle typologie des marques réfléchies, in A. Rousseau, D. Bottineau, and D. Roulland (eds.), *L'énoncé réfléchi*, Rennes: Presses Universitaires de Rennes, 107-130.

Lefebvre, Claire (1998), *Creole genesis and the acquisition of grammar: The case of Haitian Creole*, Cambridge: Cambridge University Press.

Michaelis, Susanne M., Philippe Maurer, Martin Haspelmath & Magnus Huber (2013), *Atlas of Pidgin and Creole Language Structures Online*, Leipzig: Max Planck Institute.

Ziegeler, Debra (2017), Historical replication in contact grammaticalization, in D. Van Olmen, H. Cuyckens, and L. Ghesquière (eds.), *Aspects of grammaticalization: (Inter)subjectification and directionality*, Berlin / Boston: Mouton de Gruyter, 311-352.

Doubling as a reflexivization strategy

Martin Everaert & Kārumūri V. Subbārāo
(Utrecht University & University of Hyderabad)

Keywords: Reflexivization, (complex) reflexive, doubling, Binding theory

A reflexive predicate can be described as involving a single participant (entity in the real world) that is related to an event in two different capacities, often described as thematic roles: in (1) John is 'the one who criticizes' and 'the one who is being criticized':

(1) John criticized himself.

Reflexivity can be expressed via an argument (reflexive) or on the predicate (reflexive verb) (Faltz 1977), or a combination of the two. Cross-linguistic descriptive generalizations of reflexivization strategies have been offered (Geniušiene 1987; König&Siemund 2000; Schladt 2000; Dimitriadis&Everaert 2004, a.o.). Such taxonomies summarize what was observed given the data set that was used/available, giving a perspective on possible limits to variation.

One of the strategies mentioned is 'doubling' exemplified in (2) from Manipuri:

(2) caoba-na masa-na masa-bu thagat-ce-i
Chaoba-nom REFL-NOM REFL-ACC praise-VREFL-PRES
'Chaoba praises himself'

This strategy is often only mentioned in the case of reciprocity strategies (Russian drug druga). Nedjalkov (2007: 200) even suggests that two-component reciprocal pronouns are a form of iconicity, reflecting the semantics of reciprocity and observes that in many languages such reciprocals are often a reduplication of the reflexive pronoun. We will illustrate that this strategy is also found for reflexivization across a number of languages/families: Indo-European (Bangla), Dravidian (Telugu, Kannada), Sino-Tibetan (Kokborok, Manipuri), Nakh-Daghestanian (Hinuq, Tsakhur), Austroasiatic (Juang), and many more. Iconicity, clearly, does not play a role in such cases. But note that König&Siemund (2000:62) observe that doubling in some languages, might only be used for other-directed predicates.

The question, of course, is whether we can come up with a theory that can help us understand why the morpho-syntactic encoding of reflexivity takes the form we observe. The Reflexivity approach as developed in Reuland (2011) tries to offer such a theory. The assumption is that reflexivization, in essence, triggers valency reduction at the level of logical form. This results in either using a place holder ('simple') reflexive, marking syntactic intransitivity, or using 'complex' reflexives, preserving

valency, both semantically and syntactically through a 'protection' strategy. In this paper I want to concentrate on a specific instantiation of this 'protection' strategy, the doubling strategy.

Reuland (2011) assumes that reflexive sentences as in (1), at the argument structure level represented as ($V(x_{\theta 1}, x\theta_2)$)), will translate as a 1-place semantic predicates ($\lambda x(Vx)$), unless we 'protect' one of the variables ($\lambda x(V(x,f(x)))$), saving the syntactic arity of the predicate. Protection would be the result of adding morpho-syntactic complexity: (V(x,[x Morph). 'Morph' can manifest itself through a.o. the use of a bodypart noun, adding an intensifier to a pronominal element, or 'doubling' (cf. 2).

We will discuss: (a) how this strategy precisely manifests itself across languages: obligatory/ optional (given the type of predicate?); with/without case-marking, case-copying and the subsequent possibility of 'swapping' (Subbarao 2012); (b) which elements are doubled: pronominal elements (Tsakhur?), anaphors (Telugu?), and whether this makes any difference on their distribution; (c) why the doubling strategy (discussed under (a-b)) would have the valency preserving effects Reuland attributes to complex reflexives.

References

Dimitriadis, Alexis & Martin Everaert (2004). Typological Perspectives on Anaphora. In P. Suihkonen & B. Comrie (Eds.), *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia*, 51-67. Iževsk: Udmurt State Univ.

Faltz, Leonard M. (1977) Reflexivization: A study in universal syntax. Doctoral Dissertation, University of California at Berkeley.

Geniušiene, Emma. (1987) *The typology of reflexives*. Berlin: Mouton de Gruyter.

Nedjalkov Vladimir (2007) Encoding of the reciprocal meaning, in Guentchéva, Z., Geniušiene, E. S., & Nedialkov, V.P. *Reciprocal Constructions, volume 1*, 147-2028Amsterdam: John Benjamins.

König, Ekkehard & Peter Siemund (2000) Intensifiers and Reflexives: A Typological Perspective, in: Frajzyngier, Zygmunt & Traci Curl (eds.), *Reflexives: Forms and functions*, 41–74. Amsterdam: Benjamins.

Reuland, Eric (2011) *Anaphora and Language Design*. Cambridge, Mass.: MIT Press

Schladt, Mathias (2000) The typology and grammaticalization of reflexives. In Frajzyngier, Zygmunt & Traci Curl (eds.), Reflexives: Forms and functions, 103–124. Amsterdam: Benjamins.

Subbarao, Kārumūri V. (2011) *South Asian Languages: A Syntactic Typology*. Cambridge: Cambridge University Press.

'Reflexives' coding point of view: The motivation for the emergence of a function

Zygmunt Frajzyngier
(Dept. of Linguistics, University of Colorado)

The main question to be considered in the present study is what is the function of (short) reflexive markers in Indo-European languages. Many studies, e.g. Dixon 2010, attribute to reflexives a valency reduction function. Some studies (see session proposal) attribute to reflexives an anti-passive function. Such studies stem from often tacit assumptions about the syntactic and semantic properties of the categories subject and object. Some studies also assume that coding coreferentiality of subject and object is synchronically and diachronically the primary function of (short) reflexive markers.

Terms used in the study: Goal orientation is an instruction to listener to consider what the subject does rather than what happens to the subject. In some Slavic languages, Dutch, there is a distinction between short and long reflexive markers. For such languages, this study is concerned only with the short reflexive marker. In Romance languages there is only one set of reflexive markers.

The hypothesis to be argued for here is that reflexive markers as defined above code the point of view of the subject (POS). Coding of the point of view of the subject instructs the listener to consider how the event affects the subject or how it concerns the subject (Stefanini 1962). Cross-linguistically , the point of view of the subject is a function that may be an inherent property of verbs or may be coded by a variety of formal means, e.g., by a dedicated verbal inflection. The evidence for the proposed hypothesis is provided by the following:

(1) If a verb inherently indicates the point of view of the subject, there is no morphological or syntactic coding of the point of view of the subject.

(2) If a verb inherently indicates goal orientation (GO), coding the point of view of the subject changes the point of view of the verb.

(3) Languages whose verbs do not inherently indicate goal orientation (GO) or do not have a morphological GO marker do not have a means of coding the point of view of the subject.

Implications for linguistic theory: The Point of View of the Subject hypothesis allows us to explain several inferences about relations between the predicate and arguments triggered by (short) reflexive markers, such as: detransitivizing function (adopted by many linguists), middle voice (Kemmer 1993), anti-passive (Janic 2016), coreferentiality of subject and object, reciprocal, and other syntactic and semantic functions (papers in Frajzyngier and Curl 1999 and in König and Gast 2008).

Changes in transitivity and argument structure, including the often-postulated passive and anti-passive functions, are merely by-products of changes in the coding of the point of view. The hypothesis also explains why some languages code point of view of the subject while other languages do not. In this way, the proposed hypothesis explains why some functions emerge in languages and not merely how they are coded.

Selected data

Language	Inherent POG	GO marker	POS marker
Slavic, Romance, some Germanic	yes	no	(short) reflexive
Moba (Gur)	no	no	no
Hdi (Chadic)	no	yes	ventive
Sino-Russian idiolects	no	no	no
English	not for many verbs	no	no

References

Dixon, R.M.W. 2010-2012. *Basic linguistic theory*. Vol. 1: *Methodology;* vol.2: *Grammatical topics*; vol. 3: *Further grammatical topics*. Oxford: Oxford University Press.

Frajzyngier, Zygmunt, and Traci Curl (eds.) 1999. *Reflexives: Forms and Functions*. Amsterdam/Philadelphia: Benjamins.

Janic, Katarzyna. 2016. On the reflexive-antipassive polysemy: typological convergence from unrelated languages. *Berkeley Linguistics Society* 36, 158-173.

Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam/Philadelphia: Benjamins.

König, Ekkehard, and Volker Gast. 2008. *Reciprocals and reflexives*: *Theoretical and typological explorations*. Berlin: Mouton de Gruyter.

Representing reflexivity

Bambang Kartono, Sally Wong, Eric Reuland & Martin Everaert
(Utrecht institute of Linguistics OTS)

Keywords: reflexivity, interpretation, local identity avoidance, valency, protection

Languages use a variety of means to express reflexivity, with reflexivity defined as (1):

(1) A predicate is reflexive iff two of its theta-roles are assigned to one argument.

This leads to questions like:

I. How are reflexive interpretations expressed?
II. What are cross-linguistic generalizations resulting from the work on reflexivity so far?
III. What is the relationship of reflexive markers with the domain of valency?
IV. What are other semantic effects associated with reflexive morphology?

Our discussion starts with the anaphoric systems of Indonesian and Mandarin, explaining them from a general perspective, offering answers to I-IV.

Indonesian has three reflexive strategies: i) affixal, V+*diri* 'body'; ii) complex *diri*+poss (diri+*nya*); iii) 'supercomplex' *dirinya sendiri.* A verb with *diri* has a reflexive interpretation. Unlike ii-iii), strategy i) is restricted: a) to a subset of agent-theme verbs; b) has no statue reading (Jackendoff, 1992); c) has only sloppy readings under VP ellipsis; d) object comparison is impossible - *John membasuh diri lebih sering daripada George* 'John washed himself better than George', does not mean that John washed himself better than John washed George.

Similarly, Mandarin has an affixal strategy *zi*-+V (*zi-sha* 'self-kill'), a complex anaphor *zi-ji* formed by *zi*- + a defective pronoun *–ji* (Reuland,Wong&Everaert 2018), and a 'supercomplex' anaphor *ta ziji*. Mandarin *zi*-verbs share properties a-d with *diri*-verbs.

Starting question*:*

(2) Why do languages use special means to represent reflexivity, instead of simply using co-argument binding (***John*** *admired* ***him***) as in (3), with two identical variables in the co-argument domain:

(3) *DP (λx ($V_{\theta 1, \theta 2}$ (x, x)))

Reuland (2011, 2017): (**3) violates 'local identity avoidance'** (e.g. Leben 1973, Farmer&Harnish 1987, Richards 2002, Abels 2003) →reflexivity must be licensed.
Moreover, if the bindee is a fully specified pronominal, (3) violates a condition on syntactic chains.

Answering I-IV*:* Reflexive interpretations are expressed by either:
A) differentiating the arguments, forming a complex structure as (4) ('protection' by *Morph*):

(4) DP (λx ($V_{\theta 1,\theta 2}$ (x, [Morph x])))

Complex markers in Indonesian *diri-nya (sendiri),* and Mandarin *(ta) zi-ji* (and many other languages) *preserve valency* → semantic arguments, not restricted (allow statue readings).

B) detransitivizing the predicate with bundled theta-roles (see Reinhart&Siloni 2005, Reinhart 2016, Dimitriadis&Everaert 2014 for diagnostics) as in (5):

(5) DP (λx ($V_{[\theta 1\text{-}\theta 2]}$ (x)))

Simplex markers license reduction (e.g. checking residual case)→ not semantic arguments and properties a-d follow (for clitics see Marelj&Reuland 2016).

Indonesian *diri*: falls under **B)** → not a semantic argument → subject to a-d.
Mandarin *zi-*: is a reflexivizing operator → not a semantic argument → subject to a-d.
Additional: *Zi-ji* allows both local and non-local binding, reflecting a timing difference involving the access to its constituents, see (6):

(6) Zhangsan xiang ni biaobai le ziji → Zhangsan xiang ni [(**zi**$_i$)[biaobai]] **zi**$_i$-ji
'Zhangsan unburdened himself to you.'

If it carries a suitable Person feature *(zi-)-ji* enters a non-local Agree-based dependency and long-distance binding ensues, subject to blocking (Giblin 2016). If it doesn't, *zi-* blindly operates on the verb and reflexivizes it, correctly bypassing any factors causing blocking.

References

Abels, Klaus (2003), Successive cyclicity, anti-locality, and adposition stranding. Storrs, CT: University of Connecticut PhD thesis.

Dimitriadis, Alexis, and Martin Everaert (2014), How many theta-roles in a reflexive verb? *Acta Linguistica Hungarica* 61.3, 247-269

Farmer, Ann & Robert Harnish (1987), Communicative reference with pronouns. In Verschueren, Jef & Marcella Bertuccelli-Papi (eds.), *The pragmatic perspective*, 547–565. Amsterdam: Benjamins

Giblin, Iain (2016), Agreement Restrictions in Mandarin Long-distance Binding. Cambridge, MA: MIT dissertation.

Jackendoff, Ray (1992), Mme. Tussaud meets the binding theory. *Natural Language and Linguistic Theory*, 10:1–32.

Leben, William (1973), Suprasegmental phonology. Cambridge, MA: MIT dissertation. Distributed by Indiana University Linguistics Club.

Marelj, Marijana & Eric Reuland (2016), Clitics and reflexives – reducing the lexicon-syntax parameter. In Everaert, Martin, Marijana Marelj, and Eric Reuland (eds.), *Concepts, syntax, and their interface: The Theta-system.* Cambridge, MA: MIT Press.

Reinhart, Tanya (2016), The Theta System. Syntactic Realization of Verbal Concepts. In Martin Everaert, Marijana Marelj, and Eric Reuland, eds. *Concepts, Syntax, and their Interface: The Theta System*, 1-111. Cambridge, MA: MIT Press

Reinhart, Tanya, and Tal Siloni (2005), The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36(3):389–436.

Reuland, Eric (2011), *Anaphora and language design*. Cambridge, MA: MIT Press.

Reuland, Eric (2017), Why is Reflexivity so Special? Understanding the World of Reflexives. *Studia Linguistica*, 12–59, DOI: 10.1111/stul.12070

Reuland, Eric, Sally Wong & Martin Everaert (2018), Mandarin *zi-ji* is complex. Paper presented at the 2018 SLE conference, Tallinn

Richards, Norvin (2002), A distinctness condition on linearization. *West Coast Conference on Formal Linguistics* (WCCFL) 20. 470–483.

Comparative concepts for reflexive constructions as a basis for testable universal claims

Martin Haspelmath
(MPI-SHH Jena & Leipzig University)

Keywords: anaphora, reflexive construction, universal, comparative concept, reflexive voice marker

This paper has a methodological focus, starting out from the idea that general claims about the grammatical properties of Human Language should be testable in the usual way, i.e. by studying a representative set of languages from around the world. This means that the relevant properties must be identified in the same way in all languages, on the basis of clearly defined comparative concepts. While a number of authors have proposed universals of reflexive constructions that are potentially testable (Faltz 1977: 107-133; Haspelmath 2008; Dixon 2012: Ch. 22), these authors have not emphasized clear definitions of comparative concepts (typically presupposing that traditional terms are reasonably clear), and they have not really engaged with the Reinhart-Chomsky tradition of theorizing about reflexive constructions (Reinhart 1983; Reuland 2011), even though there is a substantial literature in this tradition that contains interesting and cross-linguistically relevant ideas (e.g. Kiparsky 2002; Volkova 2014; Déchaine & Wiltschko 2017).

Here I propose cross-linguistically applicable definitions for a number of key terms in the domain of reflexive constructions: Reflexive pronoun, reflexive voice marker, reflexive argument marker, long-distance reflexive pronoun, obviative pronoun, and others. I also argue that the notions of "reflexive predicate" or "reflexivity" (e.g. Reuland 2011) should be replaced by a new term such as "autopathic predication", because it is confusing to use the term "reflexive" in two rather different senses (narrowly for coargument relations, and more broadly for forms specialized to express coreference within a sentence). I will not provide a definition for the term "anaphor", which is widely used in different traditions but with very different meanings (two different meanings in the Reinhart-Chomsky tradition, and a third meaning in the tradition of automatic anaphora resolution). I will also briefly discuss the idea that Canonical Typology provides a solution to the terminological and conceptual problems (Everaert 2013).

Finally, I will discuss a few of the universal claims made by the above-mentioned authors, addressing some of the problems that arise due to the vagueness of earlier terms and proposing solutions for resolving these problems. I will not in any detail discuss the alternative approach that assumes a rich set of innate categories of universal grammar, but will keep in mind that it is conceivable that a better understanding of cross-linguistic patterns might come from a theory that assumes a rich UG (so that there is no need to define comparative concepts).

References

Déchaine, Rose-Marie & Martina Wiltschko. 2017. A formal typology of reflexives. *Studia Linguistica* 71(1–2). 60–106. doi:10.1111/stul.12072.

Dixon, R. M. W. 2012. *Basic Linguistic Theory, volume 3: Further grammatical topics*. Oxford: Oxford University Press.

Everaert, Martin. 2013. The criteria for reflexivization. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 190–206. Oxford: Oxford University Press.

Faltz, Leonard. 1977. *Reflexivization: A study in universal syntax*. Berkeley: University of California at Berkeley PhD dissertation. https://escholarship.org/uc/item/3sv079tk.

Haspelmath, Martin. 2008. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40–63. doi:10.1349/PS1.1537-0852.A.331.

Kiparsky, Paul. 2002. Disjoint reference and the typology of pronouns. In Ingrid Kaufmann & Barbara Stiebels (eds.), *More than words: A festschrift for Dieter Wunderlich*. Berlin: Akademie.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Routledge.

Reuland, Eric J. 2011. *Anaphora and language design*. Cambridge, Mass: MIT Press.

Volkova, Anna. 2014. *Licensing reflexivity: Unity and variation among selected Uralic languages*. Utrecht: Utrecht University PhD dissertation. https://dspace.library.uu.nl/handle/1874/296571

WORKSHOP 2

Schedule: We 11.00 – Thu 12.25 (Room 4)

Computer-assisted approaches in historical and typological language comparison

Johann-Mattis List
(Max Planck Institute for the Science of Human History, Jena)

Abstract: The workshop invites papers that deal with *computer-assisted* (as opposed to pure computational or pure qualitative) approaches to historical and typological language comparison. Computer-assisted approaches are hereby understood as procedures involving different stages of qualitative *and* quantitative data analysis, ranging from the initial preparation of lexical or structural data, via automatic or manual annotation, up to qualitative or quantitative analysis, that yield a specific result, be it a linguistic reconstruction system linking proto-forms to aligned reflexes, a phylogeny that lists inferred word histories, or tools for exploratory data analysis. By focusing on *computer-assisted approaches*, we hope to foster a more intensive collaboration between classical and computational linguists. In addition to detailed descriptions of concrete tasks in historical and typological language comparison, we also encourage submissions dealing with data standards enhancing data sharing and reuse, as well as the presentation of purely qualitative approaches for which no computational solutions exist so far.

Keywords: computational historical linguistics, classical historical linguistics, comparative method, phylogenetic reconstruction, qualitative data analysis

Workshop description

By comparing the languages of the world historically, we can gain invaluable insights into human prehistory. By comparing them typologically, we can gain invaluable insights into the fundamentals of perception and cognition. The classical methods for historical and typological language comparison date back to the early 19th century and have been constantly refined and improved since then. Thanks to the comparative method for historical language comparison, linguists have made ground-breaking insights into language change in general and into the history of many specific language families in specific (Campbell and Poser 2008), and external evidence has often confirmed the validity of the findings (McMahon 2005). Thanks to large-scale approaches to typological comparison (Greenberg 1963; Dryer and Haspelmath 2013), we have gained many new insights into "universal" patterns recurring independently across the world's languages.

With increasing amounts of data, however, the methods to prepare, compare, and analyze data, which are largely based on manual labor, reach their practical limits. As a result, scholars are now increasingly trying to automatize different aspects of the classical comparative method in historical linguistics (Kondrak 2000; Prokić, Wieling, and Nerbonne 2009; List 2014), or to automatize the retrieval of typological information (Bender 2017; Malaviya et al. 2017). On the other hand, the last decade has seen a large number of attempts to analyze cross-linguistic data statistically, be it to uncover universal factors that shape linguistic diversity independently of language history (Everett et al. 2015; Blasi et al. 2016), to gain insights into the past of specific language families (Bouckaert et al. 2012; Chang et al. 2015), to understand the dynamics underlying lexical and grammatical evolution (Greenhill et al. 2017), or to arrive at a better understanding of areal factors in language history (**Cathcard et al. 2018**).

Purely computational applications, however, are not capable of replacing experts' experience and intuition, and given that most of the computational methods for data preparation still largely lag behind human judgments, it is not surprising that most of the computational analyses still rely on manually annotated data. In a situation where computers cannot replace experts and experts do not have enough time to analyze the increasing amounts of data, a new framework, neither completely computer-driven, nor ignorant of the help computers provide, becomes urgent. Such frameworks are well-established in biology and translation, where computational tools cannot provide the accuracy needed to arrive at convincing results, but do assist humans to digest large data sets.

This has led to a situation in which computational methods can only be carried out by a small number of experts who have a strong background in programming. Since computational experts do not necessarily always have a strong background or interest in linguistic topics, this has lead to a certain split in the field, with classical linguists being often dismissive and sceptical with respect to computer-based applications, and computational linguists being unsatisfied with the lack of interest in the multiple opportunities which quantitative and digital approaches have to offer.

That both classical and computational analyses could profit from each other has been increasingly demonstrated in *computer-assisted frameworks* in which classical linguists collaborate closely with computational linguists, with the data being analyzed both qualitatively and quantitatively (List 2016). In these frameworks, computational methods can be used in various ways to assist experts in qualitative analysis, be it (1) by pre-processing large datasets automatically before having experts manually correct the results (Hill and List 2017), (2) by visualizing large datasets in a convenient way that allows experts for a quick inspection (List et al. 2018; List 2017), (3) or by using automatic methods to check expert annotations for internal consistency (Kolipakam et al. 2018).

What is important for a successful application of computer-assisted methods are the detailed *workflows* that experts use to retrieve and analyze information both quantitatively and qualitatively. Since they usually require a complicated mixture of programming using different software packages, data annotation using different formats, and statistical analysis using different models, computer-assisted approaches are not (yet) easy to apply, especially for scholars with little experience in programming or data handling. What further exacerbates the more widespread sharing and reuse of computational and computer-assisted approaches that have been proposed in the past is that the information provided in the articles that discuss them is usually very sparse.

By bringing together scholars from the classical and the computational camps, we hope to foster a closer future collaboration that integrates both quantitative and qualitative approaches. Topics for papers include (but are not limited to):

- Computer-assisted approaches to study language contact in specific linguistic areas.
- Computer-assisted approaches to study language history in form of networks of phylogenies.
- Papers discussing the compilation of large annotated datasets in historical linguistics and language typology.
- Workflows for linguistic reconstruction (phonology, lexicon, syntax).
- Tools for exploratory data analysis in historical linguistics and language typology.
- Standards and best practices for data curation and reuse.
- Qualitative workflows and computer-assisted cases studies.
- Presentation of linguistic problems for which only qualitative workflows exist so far.

References

Bender, Emily. 2017. “Linguistic Typology in Natural Language Processing.” *Linguistic Typology* 20 (3): 645–60.

Blasi, Damián E., Søren Wichmann, Harald Hammarström, Peter Stadler, and Morten H. Christiansen. 2016. “Sound–meaning Association Biases Evidenced Across Thousands of Languages.” *Proceedings of the National Academy of Science of the United States of America* 113 (39):

10818–23.

Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Aalexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard, and Quentin D. Atkinson. 2012. "Mapping the Origins and Expansion of the Indo-European Language Family." *Science* 337 (6097): 957–60.

Campbell, Lyle, and William John Poser. 2008. *Language Classification: History and Method*. Cambridge: Cambridge University Press.

Cathcart, Chundra, Gerd Carling, Niklas Johansson, and Erich Round. 2018. "Areal Pressure in Grammatical Evolution." *Diachronica* 35 (1): 1–34.

Chang, Will, Chundra Cathcart, David Hall, and Andrew Garret. 2015. "Ancestry-Constrained Phylogenetic Analysis Ssupport the Indo-European Steppe Hypothesis." *Language* 91 (1): 194–244.

Dryer, Matthew S., and Martin Haspelmath, eds. 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/.

Everett, C., D. E. Blasi, and S. G. Roberts. 2015. "Climate, Vocal Folds, and Tonal Languages: Connecting the Physiological and Geographic Dots." *Proceedings of the National Academy of Sciences of the United States of America* 112 (5): 1322–7.

Greenberg, Joseph Harold. 1963. *The Languages of Africa*. Bloomington: Indiana University Press.

Greenhill, S. J., C. H. Wu, X. Hua, M. Dunn, S. C. Levinson, and R. D. Gray. 2017. "Evolutionary Dynamics of Language Systems." *PNAS* 114 (42): E8822–E8829.

Hill, Nathan W., and Johann-Mattis List. 2017. "Challenges of Annotation and Analysis in Computer-Assisted Language Comparison: A Case Study on Burmish Languages." *Yearbook of the Poznań Linguistic Meeting* 3 (1): 47–76.

Kolipakam, Vishnupriya, Fiona M. Jordan, Michael Dunn, Simon J. Greenhill, Remco Bouckaert, Russell D. Gray, and Annemarie Verkerk. 2018. "A Bayesian Phylogenetic Study of the Dravidian Language Family." *Royal Society Open Science* 5 (171504): 1–17.

Kondrak, Grzegorz. 2000. "A New Algorithm for the Alignment of Phonetic Sequences." In *Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference*, 288–95.

List, Johann-Mattis. 2014. *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.

———. 2016. "Computer-Assisted Language Comparison: Reconciling Computational and Classical Approaches in Historical Linguistics." Jena: Max Planck Institute for the Science of Human History.

———. 2017. "Using Network Models to Analyze Old Chinese Rhyme Data." *Bull. Chin. Ling.* 9 (2): 218–41.

List, Johann-Mattis, Simon J. Greenhill, Cormac Anderson, Thomas Mayer, Tiago Tresoldi, and Robert Forkel. 2018. "CLICS². An Improved Database of Cross-Linguistic Colexifications Assembling Lexical Data with Help of Cross-Linguistic Data Formats." *Linguistic Typology* 22 (2): 277–306.

Malaviya, Chaitanya, Graham Neubig, and Patrick Littell. 2017. "Learning Language Representations for Typology Prediction." *ArXiv E-Prints*, no. 1707.09569: 1–7. https://arxiv.org/abs/.

McMahon, April. 2005. "Special Issue: Quantitative Methods in Language Comparison - Introduction." *Transactions of the Philological Society* 103 (2): 113–19.

Prokić, Jelena, Martijn Wieling, and John Nerbonne. 2009. "Multiple Sequence Alignments in Linguistics." In *Proceedings of the EACL 2009 Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education*, 18–25.

ꝺ

Binominal fingerprints and Indo-European genetics

Steve Pepper
(University of Oslo)

Keywords: typology; language comparison; nominal modification; binominal lexemes; compounding

This paper examines the potential of *binominal lexemes* in computer-assisted approaches to historical and typological language comparison. It describes an onomasiologically structured lexical database and the use of statistical clustering to reveal patterns of genetic affiliation and language contact.

The comparative concept encompasses noun-noun compounds and their functional equivalents. A binominal lexeme is a "lexical unit that consists primarily of two thing-morphs (a thing-morph is a root or affix that profiles a thing, cf. Haspelmath's (2012) notion of 'thing-root') and whose function is to name a complex concept that involves an unstated (or under-specified) relation between two entities" (Pepper to appear). Indo-European examples denoting RAILWAY are given in (1).

	(a) German	(b) French	(c) Russian	(d) Slovak
(1)	*Eisen-bahn*	*chemin de fer*	*želez-naja doroga*	*želez-n-ica*
	iron-way	way of iron	iron-ADJZ way	iron-ADJZ-NMLZ

The binominals database is built around a set of 100 concepts, 84 of which are found in the Concepticon (List et al. 2018), and their denotations in 106 languages. The data were extracted from WOLD (Haspelmath & Tadmor 2009) and supplemented using dictionaries and questionnaires. Each binominal is annotated for construction (e.g. Ger. **Mod Head**, Fr. **Head PREP Mod**, etc.); order of constituents (here, **R** and **L**, respectively); semantic relation (here, COMPOSITION); and morphosyntactic strategy. The latter uses an eight-way typology: four types are exemplified in (**Fout! Verwijzingsbron niet gevonden.**): (a) compounding (**cmp**), (b) prepositional (**prp**), (c) adjectival (**adj**) and (d) derivational (**der**); the remaining four strategies are exemplified in (2) by items denoting NOSTRIL: (a) genitival (**gen**), (b) construct (**con**), (c) double-marking (**dbl**) and (d) classifier (**cls**).

	(a) Archi	(b) Turkish	(c) Takia	(d) Murui Huitoto
(2)	*muč-li-n kłan*	*burun deliğ-i*	*ŋdu-n awa-n*	*defo*
	nose-OBL-GEN hole	nose hole-3SG	nose-3SG mouth-3SG	nose.CL(cavity)

Statistics show that languages vary greatly in terms of the number and kinds of morphosyntactic strategies used: from one in Norwegian (**cmp**) to as many as five in Polish (**cmp**, **prp**, **gen**, **adj**, and **der**). Where there is competition, as in Polish, the relative frequency also varies. The data are exploited to create a binominal profile or "fingerprint" for each language, notated as [100_{cmp}] for Norwegian, [$4_{cmp}7_{prp}5_{gen}34_{adj}50_{der}$] for Polish, etc. These turn out to be highly indicative of genetic affiliation, at least in Indo-European, the family for which most data is currently available (Figure 1). Germanic, Romance and Slavic in particular show clear intragroup similarities and equally clear intergroup differences.

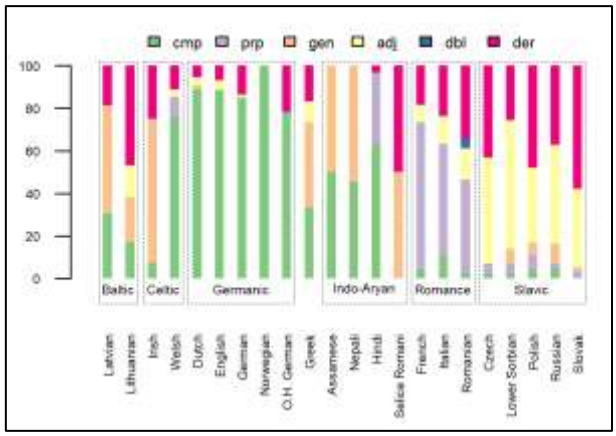


Figure 1

This is confirmed by statistical clustering performed using Principal Components Analysis (Figure 2). In the diagram, each of the six quadrants containing symbols provides a view into the same three-dimensional space of binominal types, as though looking through the different faces of a cube. Where languages from the same genus do not cluster together, as is the case with Celtic and Indo-Aryan languages, explanations are sought in grammaticalization and language contact, as will be shown in the talk. Thus the formal patterning of the binominal lexicon, arrived at by computer-assisted means, tells a story about the linguistic history of the language in question.

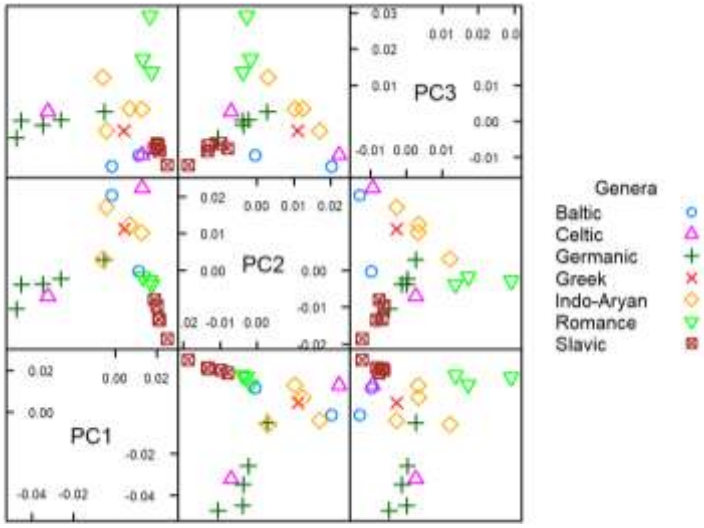


Figure 2

References

Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. *UCLA Working Papers in Linguistics, Theories of Everything* 17, Article 16. 109–130.

Haspelmath, Martin & Uri Tadmor (eds.). 2009. *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wold.clld.org.

List, Johann Mattis, Michael Cysouw, Simon Greenhill & Robert Forkel (eds.). 2018. *Concepticon*. Jena: Max Planck Institute for the Science of Human History. http://concepticon.clld.org.

Pepper, Steve. to appear. *The typology and semantics of binominal lexemes: Noun-noun compounds and their functional equivalents*. Oslo: University of Oslo PhD dissertation.

Lexical innovation and language grouping in Pano

Pilar Valenzuela, Roberto Zariquiey and Gabriela Tello

<pdf>

Automatic recognition of semantic proximity for cognate detection

John B. Lowe, Alex Francois, Martine Mazaudon & Charles Zhang
(UC Berkeley; CNRS Paris; CNRS Paris; & UC Berkley)

Since the time of the Neogrammarians, applying the comparative method in language reconstruction has been perceived mostly as studying the effect of sound change on the form of words over time. During the initial phase of constructing cognate sets and assessing sound laws, the semantic component of comparative studies has hardly been systematized: decisions based on meaning remain implicit and intuitive, as the linguist selects cognates on the basis of the semantic closeness of their glosses – e.g. grouping together items glossed 'plant', 'tree' and 'wood'.

Yet a methodological adjustment becomes necessary as we attempt to develop computational tools for doing language reconstruction – such as the *Reconstruction Engine* (RE) created by Lowe & Mazaudon (1994). The initial focus of RE was the implementation of regular sound rules, and the ability to automatize cognate recognition on the basis of explicit phonological regularities. Yet in order to process the results of automatic form recognition, it soon became obvious that a system of *semantic filters* was needed. Among several motivations for such a filter is the need to handle homophony, both in the modern languages and in the proposed reconstructions.

For example, among the several thousand candidate cognate sets output by the Reconstruction Engine for a corpus of over 6,900 lexical items in 8 Tamangic (Tibeto-Burman, Nepal) languages, *Table 1* brings together all the modern forms that could reflect a putative etymon ***gaŋ** (under proto-tone 'A'). The human mind can group together forms on the basis of their glosses: 'hill' goes with 'mountain', 'mountain range', and possibly 'snow' (via 'glacier'). A separate cognate set would group 'leg' with 'long bone' ('os long' in the French source); as for 'rainshield', it would become isolated. So, once meaning is taken into account, the forms contained in *Table 1* must be split into three separate cognate sets.

***Table 1* – *^{A} gaŋ**

GHA	³kũ	*rainshield*
RIS	³kaŋ	*mountain*
SAH	³kaŋ	*mountain range*
MAN	³kaŋ	*snow*
MAR	^{X}kaŋ	*os long*
TAG	^{X}kaŋ	*leg*
TUK	³kaŋ	*hill*

In the initial stage of RE, the work of filtering and cleanup was done manually by the linguist; yet this time-consuming effort was ad hoc, and hardly reusable on other datasets. The aim of our talk is to explore the possibility of automatically generating and applying the semantic filters we need, based on extant meaning-based resources, including [1] *WordNet* (Miller 1995); [2] *CLiCS*, a database of colexification (List *et al.* 2018); and [3] *DatSemShift*, a catalogue of semantic shifts (Zalizniak 2018).

The first two resources are synchronic in nature, but our hypothesis is that synchronic semantic associations such as colexification (François 2008) can be used as a proxy to diachronic semantic shifts. For example, the detection of languages that colexify 'hill' and 'mountain' will help pick up automatically some items of *Table 1*, and assist in putting together a valid cognate set.

The experiment will be run on several datasets – Tamangic, Romance, Oceanic from Vanuatu – and tested with different online resources. We will measure the success of our automatic experiments by

comparing them with the linguist's manual work. Each option will be evaluated quantitatively, by computing measures e.g. of Precision and Recall, compared to a "gold standard". Our preliminary tests suggest the need to control for the granularity of our semantic tools at different stages of the research.

References

François, Alexandre. 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. *From Polysemy to Semantic change: Towards a Typology of Lexical Semantic Associations*, ed. by M. Vanhove, 163-215. Amsterdam: Benjamins.

List, Johann-Mattis, Simon J. Greenhill, Cormac Anderson, Thomas Mayer, Tiago Tresoldi & Robert Forkel. 2018. CLICS²: An Improved Database of Cross-Linguistic Colexifications. Assembling Lexical Data with the Help of Cross-Linguistic Data Formats. *Linguistic Typology* 22.277–306.

Lowe, John B. & Martine Mazaudon. 1994. The Reconstruction Engine: A Computer Implementation of the Comparative Method. *Association for Computational Linguistics* 20.381-417.

Miller, George A. 1995. WordNet: a lexical database for English. *Communications of the ACM* 38.11: 39-41.

Zalizniak, Anna. 2018. The Catalogue of Semantic Shifts: 20 Years Later. *Russian Journal of Linguistics* 22: 770-87.

Computer-assisted approaches to semantic maps: A qualitative approach to large-scale lexical datasets

Thanasis Georgakopoulos & Stéphane Polis

A semantic map is a way to visually represent the relationships between meanings based on patterns of co-expression across languages (Georgakopoulos & Polis 2018). In this talk, we will focus on the so-called 'classical maps' (Haspelmath 2003; van der Auwera, 2013), which typically take the form of a graph—with nodes standing for meanings and edges between nodes standing for relationships between meanings. Up until recently, the classical maps were plotted manually and based on relatively small datasets. The very possibility of plotting such maps automatically has been questioned, because the model has been considered "not mathematically well-defined or computationally tractable, making it impossible to use with large and highly variable crosslinguistic datasets" (Croft & Pool 2008: 1). However, Regier et al. (2013) showed that an efficient algorithm exists in order to infer semantic maps "that approximates the optimal solution nearly as well as is theoretically possible."

The first goal of this talk is to show that the algorithm of Regier et al. (2013) produces high quality maps, but that these need to be revised and emended based on qualitative semantic analyses (a) when the connection suggested between two meanings is possible from a mathematical point of view, but at the same time an alternative connection is equally possible and semantically more satisfying, and (b) when historical data show that there is an indirect relationship between two meanings that appear directly related in synchronic datasets. In order to do so, we resort to CLICS^2 (List et al. 2018; https://clics.clld.org), which is an online database of synchronic lexical associations that provides information about 2638 distinct colexification patterns in 1220 language varieties. As case studies, we explore the structured network of two semantic fields, the one of PERCEPTION/COGNITION and the one of EMOTIONS/VALUES, taking as a point of departure concepts that belong to these ontological categories in the Concepticon (https://concepticon.clld.org). In a second step, we expand the algorithm in such as way that it takes into account qualitative historical data (leading to diachronic semantic

maps, i.e., oriented or mixed graphs). Information regarding the evolutionary paths of the lexemes is based on the database of semantic shifts in the languages of the world (DatSemShift; http://semshifts.iling-ran.ru/) as well as on data from Ancient Greek (8th – 1st c. BC) and Ancient Egyptian (26th c. BC – 10th c. AD), two languages with significant diachronic material. We conclude by arguing that classical semantic maps can be combined with the formal concept lattices introduced by Ryzhova & Obiedkov (2017). Formal concept analysis produces hierarchical graphs, which visualize in a principled way the mapping of polysemic items onto meanings. This approach therefore allows one to interpret the universal semantic networks of the classical semantic maps (a kind of black-box when considered in isolation) from both a genetic and an areal viewpoint, since the clustering of polysemic items according to parameters such as lineage and geography can be straightforwardly observed.

References

van der Auwera, J. 2013. Semantic maps, for synchronic and diachronic typology. In A.G. Ramat, C. Mauri, & P. Molinelli (eds.), *Synchrony and diachrony. A dynamic interface*, 153–176. Amsterdam: Benjamins.

Croft, W., & Poole, K.T. 2008. Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics* 34(1). 1–37.

Georgakopoulos, Th. & St. Polis. 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12(2). 1–33.

Haspelmath, M. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (ed.), *The new psychology of language*, 211–242. Mahwah, NJ: Lawrence Erlbaum.

List, J.-M., S. Greenhill, C. Anderson, Th. Mayer, T. Tresoldi & R. Forkel. 2018. CLICS2: An improved database of cross-linguistic colexifications assembling lexical data with the help of cross-linguistic data formats. *Linguistic Typology* 22(2). 277–306.

Regier, T., N. Khetarpal & A. Majid. 2013. Inferring semantic maps. *Linguistic Typology* 17. 89–105.

Ryzhova, D. & S. Obiedkov. 2017. Formal Concept Lattices as Semantic Maps. In Ek.L. Chernyak (ed.), *Computational linguistics and language science*, 78–87. Aachen CEUR Workshop Proceedings.

Network phylogenies of Abui

Gereon A. Kaiping
(Universiteit Leiden)

Timor-Alor-Pantar languages, Bayesian inference, Phylogenetic network, Loanword detection, Abui

The Alor-Pantar languages are one branch of a small family of non-Austronesian languages spoken on the eponymous islands in Eastern Indonesia. They are supposed to have a long, dense history of internal contact (Klamer 2014). This impedes historical reconstruction, but without established reconstructions and sound correspondences as a base line, a focused study of the borrowing history is nearly impossible. The LexiRumah lexical database (Kaiping & Klamer 2018), together with recent results concerning automatic cognate judgement methods for phylogenetics (Rama et al. 2018), provides a broad, synchronic lexical dataset including items beyond basic vocabulary, which can be used for language comparison with a focus on lexical borrowing. This allows us to study contact based on the signal of lexical change, instead of the (ir-)regularity of sound changes in individual lexical

items. Using a pseudo-Dollo model (Bouckaert & Robbeets 2017) in Bayesian phylogenetic network inference methods (Zhang et al. 2018), we study the contact history of Abui [abui1241], an Alor-Pantar language spoken in central Alor, and some of its neighbors. We find a strong signal of proto-Klon [kelo1247] as donor language (γ=50%) for borrowings into Kafoa [kafo1240], which is otherwise very closely related to Abui, possibly even part of the Abui dialect chain. A dated phylogenetic network, created using this methodology, may paint a clearer picture of the language history in the region. But more importantly, this approach gives the inheritance and borrowing history of every lexical form, automatically grouping them into different strata. Grouping the forms by their shared vs. non-shared history will give a clearer picture of systematic sound correspondences and therefore help bottom-up reconstruction.

References

Bouckaert, Remco & Robbeets, Martine. 2017. Pseudo Dollo models for the evolution of binary characters along a tree. *bioRxiv*. 207571. doi:10.1101/207571.

Kaiping, Gereon A. & Klamer, Marian. 2018. LexiRumah: An online lexical database of the Lesser Sunda Islands. *PLOS ONE* 13(10). e0205250. doi:10.1371/journal.pone.0205250.

Klamer, Marian. 2014. *The Alor-Pantar languages: history and typology*. Berlin: Language Science Press.

Rama, Taraka & List, Johann-Mattis & Wahle, Johannes & Jäger, Gerhard. 2018. Are Automatic Methods for Cognate Detection Good Enough for Phylogenetic Reconstruction in Historical Linguistics? *arXiv:1804.05416 [cs]*. http://arxiv.org/abs/1804.05416 (14 July, 2018)

Zhang, Chi & Ogilvie, Huw A. & Drummond, Alexei J. & Stadler, Tanja. 2018. Bayesian Inference of Species Networks from Multilocus Sequence Data. *Molecular Biology and Evolution* 35(2). 504–517. doi:10.1093/molbev/msx307.

Computer-assisted language comparison with reflex

Guillaume Segerer

The RefLex database (www.reflex.cnrs.fr/database) is devoted to lexical data from African languages. So far (as of January, 2019), it contains 1,168,209 lexical entries from 1,312 sources covering 799 languages. Besides its importance as a (freely available) online Reference Lexicon (nearly all entries are linked with an image of the original page where it appears), RefLex proposes many tools designed to help the linguist perform various taslks pertaining to language comparison, language typology or historical linguistics. The way by which the tools may help are best summed up as follows: Counting; Organising; Retrieving.

After a quick overview, my talk will present some of these tools:

1. **Counting: the Statistical tool**. It allows to count all kind of combinations in a given field, and to display the result as a contingency table where the values that are well above or well under the expected ones are highlighted. The statistical tool is also able to count combinations of values in two different fields (as, for instance, Part of Speech and Tone Pattern). As a side effect, this tool is also very useful as an error finder, be it errors in the original source or in the coding of the data. Examples will show how, for instance, specific counts can be helpful in suggesting regular sound changes.

2. **Organising: the Reconstruction tool**. It consists of a set of user-friendly interfaces designed to select sources, then to select words, then to align words phonetically and finally to manage the

correspondence sets thus created. There is nothing automatic here: the choice of cognates as well as the details of phonetic alignments are left to the linguist alone, but the interface makes it very easy to handle. In addition, it is also possible for several registered users to work collaboratively on the same dataset.

3. **Retrieving: the Loanword tool**. Currently under development, this very specific tool will take advantage of the fact that whenever the information is available in the original source, the borrowing status of words is hard-coded in the database. In addition, it is possible for any user to add borrowing information to any record. So far, more than 20,000 entries are identified as loanwords. The Loanword tool will make it possible to know what are the prominent donor languages, which notions are borrowed most and where, and other similar facts. It will of course make use of the mapping features already present in RefLex.

These three features of RefLex illustrate how an online lexical database can contribute to a better knowledge of language history, by going far beyond the mere display of lexical material.

Inferring the Stability of Grammatical Structures in Major Language Families

Cara Evans, Hannah Haynie, Hedvig Skirgård and Simon Greenhill

The matter of how grammatical features change over time has only relatively recently come into focus in historical linguistics. Yet understanding grammatical change is at the heart of many important questions regarding language histories and the global distribution of linguistic diversity, for example: Do grammatical features reliably track historical patterns of inheritance? Which grammatical features are most stable, which change most rapidly, and why? And are the relative rates and patterns of grammatical change and stability similar across different language families? Previous suggestions and findings indicate that specific types of grammatical information (e.g. morphosyntactic alignment, head marking) are highly stable over time, either through resistance to borrowing or repeated areal convergence (Nichols 1992; Dunn et al. 2008). Yet other work has argued that there is more diffusion of grammatical structures than previously thought (horizontal transfer), and that repeated co-evolution and a restricted design space lead to higher chances of spurious similarity in grammatical data (Thomason and Kaufman 1988; Ross 1996; Greenhill et al. 2017).

In this paper we use a new and well-sampled global database, *Grambank*, which contains information about 195 grammatical features across ~1400 languages. These features describe similar morphosyntactic phenomena to the World Atlas of Language Structures (Dryer and Haspelmath 2013) and have been consistently coded across families and regions. Using computational phylogenetic methods, we reconstruct probable grammatical histories, and investigate grammatical stability for five of the world's major language families (Austronesian, Bantu, Indo-European, Pama-Nyungan, and Uto-Aztecan). We use Stochastic Character Mapping (Bollback 2006) to infer trait stability, as well as feature gain and loss rates over time.

Our results shed light on the stability of individual features within language families and correlations in the relative stabilities of features across families. We also evaluate the robustness of the grammatical systems inferred for proto-languages, which may be important for understanding our confidence in the reconstruction of grammatical features more generally.

References:

Bollback, Jonathan P. 2006. SIMMAP: Stochastic Character Mapping of Discrete Traits on Phylogenies. *BMC Bioinformatics* 7: 88. https: //doi.org/10.1186/1471-2105-7-88.

Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink, and Angela Terrill. 2008. Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language* 84 (4): 710–59.

Farris, J S. 1989. The Retention Index and the Rescaled Consistency Index. *Cladistics* 5: 417–19.

Greenhill, Simon J., Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen C. Levinson, and Russell D. Gray. 2017. Evolutionary Dynamics of Language Systems. *Proceedings of the National Academy of Sciences*, 201700388. https: //doi.org/10.1073/pnas.1700388114.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.

Ross, Malcolm D. 1996. Contact-Induced Change and the Comparative Method: Cases from Papua New Guinea. In *The Comparative Method Reviewed*, edited by Mark Durie and Malcolm D. Ross, 180–218. 24. Oxford: Oxford University Press.

Thomason, S G, and T Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. California: University of California Press.

Towards refined phylogenies of the Tukanoan languages: Comparing phonological and lexical data

Thiago Costa Chacon, Tiago Tresoldi & Johann-Mattis List
(University of Brasília, MPI-SHH; MPI-SHH; & MPI-SHH)

Keywords: Phylogenetics, Computer Assisted Language Comparison, Tukanoan family, Sound Change, Cognate

The Tukanoan family has been classified based on the application of the comparative method and sound innovations by Chacon (2014). Building on this earlier work, Chacon and List (2015) developed a model of *phonetic transition networks*, which explicitly defined – for all correspondence sets analyzed by Chacon (2014) – phonetic diachronic pathways from proto-sounds, to intermediate proto-sounds and attested states, and applied a parsimony framework in order to infer different phylogenetic trees of the Tukanoan family that coincided well with phylogenies that were proposed independently before, allowing for a rather transparent inspection of the proposed sound change events plotted on phylogenetic trees. While this approach reflects the traditional way by which linguists subgroup languages rather closely, it is not clear how well its result coincide with the now classical reconstruction of language phylogenies based on cognate sets distributed across lexical data.

In this talk, we fill this gap by discussing different approaches to infer phylogenetic trees from a dataset of 19 Tukanoan languages across 375 concepts based on Huber and Reed (1992). We use state-of-the-art methods for computer-assisted language comparison to infer and annotate cognate sets (List 2017) along with sound correspondence patterns (List 2019). The high degree of computer-assisted annotation allows us to derive both phonological characters (including sound transition networks in the sense of Chacon and List 2015) and lexical characters (cognate sets inside the same meaning slot and across different meanings) in different flavors. In the talk, we illustrate how we use these data to reconstruct different phylogenies of the Tukanoan language family, analyzing their specific advantages and disadvantages, and discussing their differences.

References

Chacon, Thiago. 2014. 'A Revised Proposal of Proto-Tukanoan Consonants and Tukanoan Family Classification'. *International Journal of American Linguistics* 80 (3): 275–322. https://doi.org/10.1086/676393.

Chacon, T. and J.-M. List (2015): Improved computational models of sound change shed light on the history of the Tukanoan languages. *Journal of Language Relationship* 13.3. 177-204.

Huber, Randall and Robert B. Reed. 1992. Vocabulario comparativo: palabras selectas de lenguas indigenas de Colombia. Bogotá: ILV.

List, J.-M. (2019): Automatic inference of sound correspondence patterns across multiple languages. *Computational Linguistics* 45.1. 1-24.

List, J.-M. (2017): A web-based interactive tool for creating, inspecting, editing, and publishing etymological datasets. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. System Demonstrations. 9-12.

The lost city of "Mande": Linguistic and genetics perspectives

Abbie Hantgan-Sonko & Hiba Babiker
(CNRS-LLACAN, France & Department of Linguistic and Cultural Evolution, Max-Planck Institute for the Science of Human History, Jena)

Keywords: Mande expansion, linguistic-genetics relationship, language contact, Dogon & Bangande

Beyond Bantu, or perhaps in parallel, that which has been referred to as the "Mande Expansion" (Brooks 1993) is thought to have taken place before 4000 BC, likely from a position in modern-day Mauritania (Vydrin 2009). At the same time, the term Mande is used in the linguistic and genetics literature (Cavalli-Sforza et al. 1994; Güldemann 2018) in reference to a specific group of related languages and peoples living across lower latitudes from Guinea to Mali. Specifically within Mali, but outside of the boundaries of the Mande ethno-linguistic family, such as the Songhay, Dogon, and linguistic-genetic isolate group Bangande, also all recount a common ancestry from a somewhat mythical place also referred to as "Mande" (Bouju 1995), described as being located in the extreme southwest of Mali (Kutsenkov 2018). Arazi (2005) reveals archeo-ethnic results from the famous city of Dia that may indicate an ancient cultural continuum from whence Dogon could have encountered other Mande groups following the fall of the Ghana Empire, around the 1st millennium BC, before moving on to occupy the Bandiagara Escarpment some thousand years later (Mayor et al. 2005). Thus, the question presents itself as to which of these indications of "Mande" included the dispersed peoples, where it was located, and how many waves succeeded and reached the current locations.

We present evidence based on Bayesian inferences of linguistic borrowings combined with genome-wide genetic results and Y-chromosome haplogroups distribution for the positioning of Mande as well as propose a time depth for the source from which each of these now distantly related groups emerged. In addition to primary postulations as to the place and time of language contact among pre-Dogon and Mande populations, we provide a detailed overview of our interdisciplinary data collection approaches, as well as our computer-assisted workflow. Although separating linguistic borrowings from true cognates is a notoriously difficult task, we employ the use of computer-assisted technologies to separate contact from cognates by relying on regular sound correspondences in the latter. Further, we source archaeological evidence in order to estimate dates of introduction for domestic animals and crops, thus providing us with an integrated picture of not only from which language but also when and potentially where a borrowing was introduced, effectively drawing a migration map to the source of

pre-history. This work sets a new perspective to the role of multidisciplinary studies in resolving the linguistic paradox in western African.

Acknowledgements:

This project has received support and funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 758232), the Department of Linguistic and Cultural Evolution at the Max Planck Institute for the Science of Human History, and by the Dogon and Bangime Linguistics Project under NSF Grant award #BCS-1263150.

References

Arazi, N. (2005). "Tracing History in Dia, in the Inland Niger Delta of Mali - Archaeology, Oral Traditions and Written Sources". University College London Institute of Archaeology.

Bouju, J. (1995). "Qu'est-ce que "l'ethnie" Dogon?" In: *Identites et appartenances dans les societes saheliennes*. Ed. by C. Fay. Vol. 31. Cahiers des Sciences Humaines 2, pp. 329–363.

Brooks, G. E. (1993). *Landlords and Strangers: Ecology, Society, and Trade in Western Africa, 1000–1630. (African States and Societies in History.)* Boulder, Colorado: Westview.

Cavalli-Sforza, L., P. Menozzi, and A. Piazza (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Güldemann, T. (2018). "Historical linguistics and genealogical language classification in Africa". In: *The languages and linguistics of Africa*. Ed. by T. Güldemann. Berlin: Mouton de Gruyter.

Kutsenkov, P. A. (2018). "The oral history of Dogon villages and migrations' to Dogon Plateau". In:

Vestnik of Saint Petersburg University. Asian and African Studies 10 (3), pp. 330–339.

Mayor, A., E. Huysecom, A. Gallay, M. Rasse, and A. Ballouche (2005). "Population dynamics and Paleoclimate over the past 3000 years in the Dogon Country, Mali". In: *Journal of Anthropological Archeology* 24, pp. 25–61.

Vydrin, V. (2009). "On the Problem of the Proto-Mande Homeland". In: *Journal of Language Relationships*, pp. 107–142.

Interactive Etymological Inference via Statistical Relational Learning

Johannes Dellert
(University of Tübingen)

Keywords: computational historical linguistics, interactive theory development, etymological inference, statistical relational learning, probabilistic soft logic

This talk presents our ongoing work on a new interactive system for computer-assisted etymological inference. Unlike existing well-developed toolchains for computational historical linguistics such as LingPy (List ea., 2018) in which every layer of analysis (e.g. phylogenetic tree inference) builds on an analysis of the previous layer (e.g. cognate sets) that is assumed to be complete, our system is designed to work with partial theories which span various layers of description such as phylogeny and language contacts, sound laws and analogical change, (partial) cognates and loanwords, as well as semantic shifts. A graphical interface makes it possible for the user to interact with such partial theories in an iterative refinement workflow. For instance, a linguist might want to annotate an unanalysed input database with some tree constraints, a couple of obvious loanwords, and a few initial suspicions about sound laws. The system will then generate a complete theory that does not violate

any of these specifications, but otherwise follows the comparative method as encoded by the reasoning engine. After each refinement cycle, the current state of the theory can be inspected and additional constraints can be added, e.g. to reject cognacy judgments that are implausible according to the linguist's background knowledge, causing the model to adapt accordingly in the next cycle.

To achieve this flexibility, our reasoning engine combines existing and new efficient algorithms for subtasks such as cognate detection or ancestral state reconstruction by means of Statistical Relational Learning (SRL). Probabilistic Soft Logic (PSL) as described by Bach ea. (2017), our SRL framework of choice, allows to combine inviolable constraints, which we use for enforcing the principles of the comparative method, with weighted relational rules, which we use to model the less precise heuristics commonly employed in the field, such as intuitions about which sound changes or semantic shifts are plausible. Inference over such combined models translates into efficiently solvable optimization problems, the solutions to which assign belief values to meaningful elementary judgments about the observed data. The resulting structured explanations mirror the reasoning patterns of classical historical linguistics much more closely than current methods e.g. for automated cognate detection, where it is generally impossible to explain in classical terms why a certain elementary judgment (like a cognate pair) was inferred.

```
SoundCorrespondence(Lang1, Sound1, Lang2, Sound2) &
Alignment(Lang1_Word1, Pos1, Lang2_Word2, Pos2) &
SoundAtPosition(Lang1_Word1, Pos1, Sound1) &
SoundAtPosition(Lang2_Word2, Pos2, Sound3) &
Sound2 != Sound3
->
!Cognate(Lang1_Word1, Lang2_Word2)
```

Example: PSL formalization of the heuristic that non-adherence of aligned segments to a sound correspondence makes two words less likely to be cognates.

In my talk, I will describe the overall system architecture, showcase how we model etymological reasoning in PSL (see example), and present preliminary versions of our modules for the interactive inference of sound laws, loanwords, and recurring morphological material, with the goal of receiving from experts who work on different language families feedback about the current state, and extending our list of desirable additional features for future development.

Acknowledgements

This research has been supported by the intramural Program for the Promotion of Junior Researchers of the University of Tübingen, complemented by a RiSC grant from the Baden-Württemberg Ministry of Education. The author was funded by the DFG Center for Advanced Studies "Words, Bones, Genes, Tools: Tracking Linguistic, Cultural and Biological Trajectories of the Human Past".

References

Bach, Stephen; Broecheler, Matthias; Huang, Bert; Getoor, Lise (2017), Hinge-Loss Markov Random Fields and Probabilistic Soft Logic. Journal of Machine Learning Research (JMLR) 18: 1–67.

List, Johann-Mattis; Greenhill, Simon; Tresoldi, Tiago; and Forkel, Robert (2018), *LingPy. A Python library for historical linguistics*. Version 2.6.4. URL: http://lingpy.org. Jena: Max Planck Institute for the Science of Human History.

Recurrent neural net approaches to sound correspondence and change

Chundra Cathcart and Damián Blasi

<pdf>

A historical linguistics machine

Søren Wichmann
(Leiden University & Kazan Federal University & Beijing Language University)

Keywords: computational historical linguistics, ASJP, phylogenetics, linguistic distance, R programming

During the past decade the people behind the Automated Similarity Judgment Program or ASJP have continuously updated their database, with the latest version (Wichmann et al. 2018) now covering 7,655 doculects; and papers have published developing methods serving three of the major objectives of historical linguistics, namely (1) language classification (Wichmann et al. 2010a), (2) dating of language group divergence (Holman et al. 2011), and (3) the inferencing of hypothetical linguistic homelands (Wichmann et al. 2010b).

While the ASJP database is easily accessed, software that implements methods (1)-(3) above is available only to varying degrees. For computing the standardly used modified Levenshtein distance called LDND (Wichmann et al. 2010a) there is fast software by Holman (2011a), but this requires the correct preparation of an input file, which is tricky for many users, and to get a phylogeny, a particular, separate piece of software is required (MEGA, downloadable at https://www.megasoftware.net/, cf. Kumar et al. 2016).

As regards measurements of the time depth of language groups, Holman et al. (2011) provide tables for most of the world's language families and major subgroups, and there is also software available for calculating new dates as the database expands (Holman 2011b). However, preparing a input file for this software is not just tricky, but a real challenge that minimally requires using macros in a text editor and presupposes a perfect understanding of the structure of the ASJP database.

Finally, for inferring homelands, using the method of Wichmann et al. (2010b), software is not available, and the expanding ASJP database would make it possible to add homelands for many language families to the ones listed in Table 1 of the paper, not to speak of subgroups.

Thus, a member of community of historical linguists or someone else, such as an economist carrying out a study using linguistic distances (e.g., Adserà and Pytliková 2015), cannot harvest the full benefit of the data and methods developed in the ASJP project without having to go through various hurdles. Moreover, someone interested in applying these methods to their own dataset, such as, for instance, a lexical survey of some region, would be severely challenged.

In this presentation I introduce an interactive, platform-independent program that implements the production of Neighbor-Joining trees, calculates time depths, identifies homelands, and more. The program allows users to not only work with existing ASJP data but also their own wordlists in a standard transcription such as IPA or Americanist or in even a customized transcription. Thus, from raw data the user can directly get a tree, a date or a homeland in a matter of minutes (although it could be hours if the dataset is large). While written in R (R Core Team 2017), the user is not required to have any familiarity with this language. It only requires the user to have downloaded R, the rest of the

procedure is self-explanatory. Members of the audience are encouraged to have downloaded R, since copies of the program can be distributed before the talk.

References

Adserà, Alícia and Mariola Pytliková (2015), The role of language in shaping international migration, *Economic journal* 125, 586: F49-F81.

Holman, Eric W. (2011a), Programs for calculating ASJP distance matrices (version 2.1). Available at https://asjp.clld.org/software.

Holman, Eric W. (2011d), Program for calculating ASJP dates (version 1.1). Available at https://asjp.clld.org/software.

Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, and Dmitry Egorov (2011), Automated dating of the world's language families based on lexical similarity, *Current Anthropology* 52(6), 841-875.

Kumar, Sudhir, Glen Stecher, and Koichiro Tamura (2016), MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets, *Molecular Biology and Evolution* 33(7), 1870-1874.

R Core Team (2017), R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Wichmann, Søren, Eric W. Holman, Dik Bakker, and Cecil H. Brown (2010a), Evaluating linguistic distance measures, *Physica A* 389, 3632-3639.

Wichmann, Søren, Eric W. Holman, and Cecil H. Brown (eds) (2018), The ASJP Database (version 18). http://asjp.clld.org/.

Wichmann, Søren, André Müller, and Viveka Velupillai (2010b), Homelands of the world's language families: A quantitative approach, *Diachronica* 27(2), 247-276.

A cross-linguistic computational approach on chance resemblances

Tiago Tresoldi
(Max-Planck Institute for the Science of Human History)

Chance resemblances can be an anathema in linguistics, given the difficulties in identifying them, and, occasionally, their usage as evidence against the comparative method. Conversely, they also tend to be among the objections towards computational approaches, as only expert knowledge and supplementary evidences would be able to confidently distinguish among vertical transmission, horizontal transmission, and chance resemblance. Limitations of this kind have been advanced since the beginnings of the comparative method, such in the frequently cited false correspondence between Latin *deus* and Ancient Greek *θεός* (both meaning "god"). The reduction, or rather expansion, *ad absurdum* of such difficulty is demonstrated both by the long tradition of folk etymologies, motivated by the assumption that surface similarities are too strong to be due to chance, and by the recurrent claims of amateur linguists on impossible relationships such as, for example, between Ainu and Etruscan.

Among the sources for such difficulties is the fact that we have no clear definition on chance similarity, in general loosely defined as "words that sound similar", particularly when there is a limited semantic leeway and preferably when judged as such with the support of the phonotactics of

the languages involved. The uncertainties in terms of definition translate into limited sets of concrete examples, leading to the absence of baselines.

In this talk, we investigate the question of how to create a baseline for expected probability of chance resemblance according to different typological parameters. As such, we will present the results of a cross-linguistic and computational inquiry on chance resemblances, following three different experiments. In the first, developing on Rosenfelder (2002), is a purely mathematical modeling that calculates the probability of random correspondences on a set of linguistic models of very simplified phonological and semantic assumptions. The second applies state-of-the-art algorithms for automatic cognate detection (List et al., 2018b) on languages randomly generated from phonological and semantic parameters collected from real languages of different typologies, in a massive comparison that allows to highlight which factors contribute the more to the perception of
similarity; in fact, we will also explore the possibility of later re-using the dataset to collect resemblance judgments according to experts. The third and most important experiment uses actual linguistic data from a cross-linguistic database, Lexibank (forthcoming), by applying the same methods to languages pairs of varying phylogenetic relationships (Hammarström et al., 2018), which, combined with semantic information linked to Concepticon (List et al., 2018a) and CLICS (List et al., 2018c).

Our results support the hypothesis that chance resemblances, even across unrelated languages, are common and indeed expected even with minor leeway, and that chance resemblances need to judged on a per-language pair and not per-potential cognate pair baasis. This should allow to orient future experiment setup, besides offering a preliminary baseline on the expectancy of chance resemblance according to given sets of parameters, including language proximity in terms of lineage, possibly being incorporated into future work on automatic borrowing detection.

References

Guy, J. B. M. (1994). An algorithm for identifying cognates in bilingual wordlists and its applicability to machine translation. *Journal of Quantitative Linguistics* 1.1. 35-42.

Hammarström, H., R. Forkel, M. Haspelmath (2018). Glottolog 3.3. Jena: Max Planck Institute for the Science of Human History.

List, J.-M., M. Cysouw, S. Greenhill, R. Forkel (eds., 2018a). Concepticon. Jena: Max Planck Institute for the Science of Human History.

List, J.-M., M. Walworth, S. Greenhill, T. Tresoldi, and R. Forkel (2018b): Sequence comparison in computational historical linguistics. *Journal of Language Evolution* 3.2. 130–144.

List, J.-M., S. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2018c): CLICS^2. An improved database of cross-linguistic colexifications assembling lexical data with help of cross-linguistic data formats. *Linguistic Typology* 22.2. 277-306.

Nichols, J. (1996) The comparative method as heuristic. in *The comparative method reviewed* (ed. M Durie). New York: Oxford University Press.

Ringe, D. A. (1992) On calculating the factor of chance in language comparison. T*ransactions of the American Philosophical Society* 82.1 1-110.

Rosenfelder, Mark (2002). How likely are chance resemblances between languages? http://www.zompist.com/chance.htm

WORKSHOP 3

Schedule: Fri 9.00 – Sat 12.55 (Room 3)

Discourse particles and epistemic stance

Diana Forker & Angeliki Alvanoudi
(University of Jena; Aristotle University of Thessaloniki & James Cook University)

Keywords: discourse particles, epistemic stance, degree of certainty, epistemic modality, typology

Discourse particles are an under-examined area in typological and descriptive studies of minority languages, but even of many majority languages, as they are often treated as extra-grammatical items. Discourse particles/markers are short and phonologically reduced items, which routinely occur in oral speech, have little or no propositional meaning and display textual and interpersonal pragmatic functions, such as connecting current with prior talk, claiming the hearer's attention, organizing discourse (e.g. indicating new topic, initiating or closing discourse), and indexing the speaker's stance, attitudes and evaluation towards the addressee and his/her contribution (e.g. see Blakemore 2003, Brinton 1996, Norrick 2009, Schiffrin 1987, Zimmermann 2011). This panel targets the latter function of discourse particles, namely their function as markers of epistemics/epistemic stance. Following Ochs (1996), we understand epistemic stance as a central meaning component of social acts and social identities that refers to knowledge or belief vis-à-vis some focus of concern including degrees of certainty of knowledge, degrees of commitment to truth of propositions, and sources of knowledge among other epistemic qualities. The study of epistemic markers has largely focused on grammaticalized evidentiality as a category of verbs (e.g. Aikhenvald 2004, Chafe and Nichols 1986, Johanson and Utas 2000) and lexical or modal markers (e.g. Biber and Finegan 1989; Ifantidou 2001, Kamio1998). Discourse particles that express epistemic stance are well-known from Germanic and Slavonic languages (Abraham 1991, Jucker and Ziv 1998), e.g. German *doch*, Russian *že*, and a few other languages (see the analysis of Hebrew *harey* by Ariel 1998, the studies in Heritage and Sorjonen 2018). For example, in Japanese *a*-prefaced turns convey that the action or content of the prior turn was unexpected (Hayashi and Hayano 2018).

We know little about the epistemic functions of discourse particles in minority languages and largely lack a coherent cross-linguistic perspective due to the lack of typological studies (but see Auer and Maschler 2016 for a comparative study of the uses of the members of the NU/NÅ discourse marker family across European and other languages). Linguists in the field often find it difficult to tackle the pragmatic diversity of discourse particles and tend to throw these small items in the grammar's dustbin. Yet, these 'trivial' items play an important role in communicating cultural meanings in speech communities. This panel aims to navigate the complex domain of discourse particles and epistemics by identifying patterns of commonality and diversity in particles expressing epistemic stance across languages. We want to bring together scholars working across various subfields of linguistics, including typology, pragmatics, interactional linguistics, and using different methods and theories.

Panel contributions address the following topics:

(i) Descriptive accounts of the (functional, discourse-pragmatic, morphosyntactic) properties of discourse particles in lesser-known languages.
(ii) Typological studies of the functions of discourse particles expressing epistemic stance.
(iii) Diachronic studies on the origin and evolution of discourse particles.
(iv) Corpus-based studies on the use of discourse particles in natural speech.

(v) Experimental studies on the functions of discourse particles.
(vi) Areal and language contact perspectives on the functions of discourse particles.

References

Abraham, Werner (ed.) (1991), *Discourse particles: Descriptive and theoretical investigations on the logical, syntactic, and pragmatic properties of discourse particles in German*, Amsterdam: John Benjamins.

Aikhenvald, Alexandra Y. (2004), *Evidentiality*, Oxford: Oxford University Press.

Ariel, Mira (1998), Discourse markers and form-function correlations, in A. H. Jucker, and Y. Ziv (eds), (1998), *Discourse Markers: Descriptions and Theory*, Amsterdam: John Benjamins, 223-259.

Auer, Peter, and Yael Maschler (2016), The family of NU and NÅ across the languages of Europe and beyond: Structure, function, and history, in P. Auer, and Y. Maschler (eds), (2016), *NU / NÅ: A Family of Discourse Markers across Europe and Beyond*, Berlin: de Gruyter, 1-47.

Biber, Douglas, and Edward Finegan (1989), Styles of stance in English: Lexical and grammatical marking of evidentiality and affect, *Text* 9, 93-124.

Blakemore, Diane (2003), *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*, Cambridge: Cambridge University Press.

Brinton, Laurel J. (1996), *Pragmatic markers in English: Grammaticalization and discourse functions*, Berlin: Mouton de Gruyter.

Chafe, Wallace L., and Johanna Nichols (eds) (1986), *Evidentiality: The linguistic coding of epistemology*, Norwood, New Jersey: Ablex Publishing Corporation.

Du Bois, John W. (2007), The stance triangle, in R. Englebretson (ed.), (2007), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, Amsterdam: John Benjamins, 139-182.

Hayashi, Makoto, and Kaoru Hayano (2018), A-prefaced responses to inquiry in Japanese, in J. Heritage, and M.-L. Sorjonen (eds), (2018), *Between Turn and Sequence: Turn-Initial Particles across Languages*, Amsterdam: John Benjamins, 191-224.

Heritage, John, and Marja-Leena Sorjonen (eds) (2018), *Between turn and sequence: Turn-Initial particles across languages*, Amsterdam: John Benjamins.

Ifantidou, Elly (2001), *Evidentials and relevance*, Amsterdam: John Benjamins.

Johanson, Lars, and Bo Utas (eds) (2000), *Evidentials: Turkic, Iranian and neighbouring Languages*, Berlin: Mouton de Gruyter.

Jucker, Andreas H., and Yael Ziv (eds) (1998), *Discourse markers: Descriptions and theory*, Amsterdam: John Benjamins.

Kamio, Akio (1998), Evidentiality and some discourse characteristics in Japanese, in A. Kamio (ed.), (1998), *Directions in Functional Linguistics*, Amsterdam: John Benjamins, 145-171.

Norrick, Neal R. (2009), Interjections as pragmatic markers. *Journal of Pragmatics* 41, 866-891.

Ochs, Elinor (1992), Linguistic resources for socializing humanity, in J. J. Gumperz, and S. C. Levinson (eds), (1992), *Rethinking Linguistic Relativity*, Cambridge: Cambridge University Press, 407-437.

Schiffrin, Deborah (1987), *Discourse markers*, Cambridge: Cambridge University Press.

Zimmermann, Malte (2011), Discourse particles, in K. v. Heusinger, C. Maienborn, and P. Portner (eds), (2011), *Semantics: An International Handbook of Natural Language Meaning*. vol. 2, Berlin: Mouton, 2012-2038.

Finnish *hän* and *se*: from pronouns to enclitic discourse markers

Seppo Kittilä & Chingduang Yurayong
University of Helsinki

In Finnish, the pronouns *hän* 's/he' and *se* 'it' have developed into discourse enclitics. Originally, *se* refers to inanimate entities, while *hän* codes animate entities. However, such animacy distinction has been collapsed in modern spoken Finnish, and neutralised when used as enclitic discourse markers. Moreover, *hän* is a speaker-oriented logophoric pronoun, while *se* is an addressee-oriented general pronoun (Saukkonen 1967, Laitinen 2002, 2005; Priiki 2017). Our study focuses on the non-referential uses of these enclitics and we will propose an explanation for their differences. As finding two identical examples that are distinguished only by the presence of these two enclitics would be virtually impossible based on natural spoken corpus data alone, we also use elicited examples, which make a comparison of functional differences between the two markers possible.

First, the examination touches upon their discourse-deictic uses, especially in response to an uttered statement. Functionally, the addressee-oriented *=se* continues the reference flow from the preceding statement, while the speaker-oriented *=hAn* raises the speaker's stance, often as denial, disagreement or argument against the preceding statement (cf. Duvallon & Peltola 2012, 2013; Duvallon 2014). Animacy may play a role here in that we feel more empathy for animate referents than inanimate ones (see, e.g., Næss 2004). We may thus say that *=hAn* functions as the speaker's stance-taking device when talking about more important topics, especially when the speaker seeks a confirmation for a fact.

Other discourse-pragmatic features are also relevant for determining the use of the enclitics: 1) the speaker's expectations about the hearer's information, 2) mirativity, and 3) the nature of the information source. First, *=hAn* is used for talking about shared knowledge between interlocutors, and also for avoiding epistemic authority (the speaker is often looking for the hearer's confirmation with *=hAn*). In contrast, by using *=se* we assume epistemic authority when making a claim based on a direct observation about a given state-of-affairs. This renders the status of shared knowledge among interlocutors less relevant. Second, *=hAn* is often related to mirativity, while this is less frequent (albeit possible) with *=se*. With *=se*, the speaker usually expects a certain state-of-affairs to occur, but what happens may exceed his/her expectations. Third, *=se* is usually associated with concrete observable evidence, while *=hAn* can be used regardless of the evidence of the speaker. Following from this, *=hAn* appears whenever either of the enclitics is used with facts.

We will argue that the original semantics of the two particles (as pronouns) can, at least partly, account for their use as discourse enclitics. For instance, the addressee-related vs. speaker-related nature of the pronouns may explain some of the differences, and differences in animacy may be related to their uses as enclitics. For example, *=hAn* is used for facts, and we may say that facts are closer to us than events that have happened previously in that facts have become a part of our knowledge of the world. The above-mentioned empathy which is stronger with *=hAn* is probably also relevant to the use of the enclitics.

References

Duvallon, Outi & Peltola, Rea. (2012). Prises de distance en finnois: le mode jussif et le marqueur discursive *-hAn*. Linx. *Revue des linguistes de l'université Paris X Nanterre* (66-67). 101–116.

Duvallon, Outi, & Peltola, Rea. (2013). Le mode jussif et le marqueur discursive *-hAn* en finnois: deux procédés de distanciation. In *Colloque International de Linguistique GReG PLS III Sens attendu/inattendu dans les langues*.

Duvallon, Outi. (2014). Le marqueur discursif finnois *-hAn*: double repérage énonciatif. *Faits de langues* 43(1), 87–114.

Laitinen, Lea. (2002). From logophoric pronoun to discourse particle. In I. Wischer & G. Diewald (Eds.), *New reflections on grammaticalization*. John Benjamins Publishing Company. 327–344.

Laitinen, Lea. (2005). *Hän*, the third speech act pronoun in Finnish. Minimal reference in Finnic: the use and interpretation of pronouns and zero in Finnish and Estonian discourse. In R. Laury (Ed.), *Minimal Reference: The use of pronouns in Finnish and Estonian discourse*. Finnish Literature Society. 75–106.

Næss, Åshild (2004). What markedness marks: the markedness problem with direct objects. *Lingua* 114. 1186–1212.

Priiki, Katri. (2017). *Hän, se, tää vai toi?: Vuorovaikutussosiolingvistinen tutkimus henkilöviittauksista Kaakkois-Satakunnan nykypuhekielessä*. Doctoral dissertation, University of Turku.

Saukkonen, Pauli. (1967). Persoonapronominien *hän*: *se*, *he*: *ne* distinktiivi oppositio. *Virittäjä* 71(3). 286–292.

Evidential and epistemic semantics of modal particles in Northern Selkup

Svetlana Orlova, Svetlana Toldova & Maria Brykina
(Lomonosov Moscow State University; HSE University, Moscow & University of Hamburg)

Keywords: Selkup, modal particles, evidentiality, epistemic particles, mirativity

This work is devoted to the evidential and epistemic semantics of particles in Selkup, a minority Samoyedic language. It uses the data of the INEL Selkup corpus (Brykina et al. 2018) based on the archive of A.I. Kuzmina (collected in 1961–1977).

The category of evidentiality has been intensively investigated in recent years in a variety of languages, including Uralic (Künnap 2002, Буркова 2004, 2017; Храковский 2007: 415-468; Ильина 2002, 2006; Урманчиева 2014, 2015, among others). Besides grammatical evidentiality, the lexical marking is also of a great interest (Wagner-Nagy, Szeverényi 2017).

In this research we analyze the lexical evidential markers *montɨ, mompa, mɨta* in Northern Selkup. A short mention of them can be found in the reference grammar of Selkup (Кузнецова и др. 1980: 309). Our corpus data show that they refer to different kinds of sources of reported information such as visual (*montɨ*, *mɨta*), inferential (*montɨ*), quotative (*mɨta*, *mompa*) or hearsay (*mompa*). As in many other languages (Aikhenvald 2004, Hopper and Traugott 2003: 13–14) these particles represent the typical paths of development from epistemic and utterance verbs. Our research shows that in Northern Selkup it is often not possible to separate evidential and epistemic meaning components in the semantics of particles. The particles in question deal with the epistemic modality shifting the responsibility for the veracity of the utterance to the protagonist or to the author of the reported speech, while it can also bring some additional effects such as mirativity. Moreover, there are some more complicated ways of expressing evidentiality with epistemic "flavour" in Northern Selkup represented by different combinations of particles such as *mompa mɨta*, *montɨ mɨta* etc.

The Selkup data provides new evidence for debates on whether one should keep the evidentiality apart from the epistemic stance (cf. Givón 1982, Chafe and Nichols 1986). We claim that both components belong to the semantics of the described particles but the question is in their proportion in particular contexts and which of them can be in the main focus.

Example (*mɨta:* quotative function and the subjectivity of information):
Təptɨl' qarɨ-t nɨŋ εsa tɨntena ima-ntɨ-nɨk:

next morning-ADV.LOC such-ADVZ say.(3SG.S) that wife-OBL.3SG-ALL
Mat ***mɨta*** *n'änna* *qum-iː-qäk* *qoːnɨ-l'ä-k.* *Qaj*
I.NOM as.if upstream human.being-PL-1SG.ILL go.to-OPT-1SG.S either.or
mɨta *mɔːt-ɨ-n-pɔː-tɨt* *qaj* *qaːtɔː-tɨt.*
as.if tent-EP-VBLZ-PST.NAR-3PL either.or where.to.go-3PL
'Next morning he says to his wife: "I will go upstream to my people. Either they put a tent or they went away."' (NEP_1965_OrphanBoyAndPanOldMan1_flk.073-074)

Abbreviations

ADV – adverbial case, ADVZ – adverbializer, ALL – allative case, EP – epenthetic element, ILL – illative, LOC – locative, NOM – nominative, OBL – oblique case, OPT – optative mood, PL – plural, PST.NAR – past narrative tense, S – subjective conjugation, SG – singular, VBLZ – verbalizer.

Acknowledgements

This publication has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the German Academies of Sciences and Humanities.

References

Aikhenvald, A. Y. (2004), Evidentiality (Oxford Linguistics). Oxford: Oxford University Press.

Brykina, Maria; Orlova, Svetlana; Wagner-Nagy, Beáta (2018), INEL Selkup Corpus. Version 0.1. Publication date 2018-12-31. Archived in Hamburger Zentrum für Sprachkorpora. http://hdl.handle.net/11022/0000-0007-CAE5-3. In: Wagner-Nagy, Beáta; Arkhipov, Alexandre; Ferger, Anne; Jettka, Daniel; Lehmberg, Timm (eds.), (2018), *The INEL corpora of indigenous Northern Eurasian languages*.

Chafe, W. L., and Nichols, J. (eds.) (1986), *Evidentiality: The linguistic coding of epistemology*, Norwood, New Jersey: Ablex Publishing Corporation.

Givón, T. (1982), Evidentiality and epistemic space, *Studies in Language* 6, 23–49.

Hopper, P. J. and Traugott, E. C. (2003), *Grammaticalization*, Cambridge: Cambridge University Press.

Künnap, A. (2002), *Main Language Shifts in the Uralic Language Group*, LINCOM Studies in Asian Linguistics 45: Lincom Europa.

Wagner-Nagy, B. and Szeverényi, S. (2017), Evidentiality in Nganasan: A Corpus Based Study, In *Комплексное изучение языков и культур аборигенных народов Сибири. Сборник тезисов Межд. науч. конф. XXVIII-е Дульзоновские чтения*, Томск: ТГПУ, 10-11.

Буркова, С.И. (2004), Эвиденциальность и эпистемическая модальность в ненецком языке, in В.А. Плунгян, Ю.А. Ландер, А.Ю. Урманчиева (ред.) (2004), *Исследования по теории грамматики 3: Ирреалис и ирреальность*, 353-374.

Буркова, С.И. (2017), Эвиденциальность в самодийских языках, in *Комплексное изучение языков и культур аборигенных народов Сибири. Сборник тезисов Межд. науч. конф. XXVIII-е Дульзоновские чтения* (2017), Томск: ТГПУ, 5-6.

Ильина, Л.А. (2002), *Эволюция глагольной категории эвиденциальности (системно-диахроническое моделирование на материале селькупского языка)*. Дис. ... канд. филол. наук, Новосибирск.

Ильина, Л.А. (2006), Эволюция эвиденциальных высказываний в самодийских языках, in Н.Н. Широбокова, А.А. Мальцева (ред.) (2006), *Языки коренных народов Сибири 18. Аналитические структуры в простом и сложном предложении*, Новосибирск.

Кузнецова А. И., Хелимский Е. А., Грушкина Е.В. (1980), *Очерки по селькупскому языку. Тазовский диалект*. Т. 1, М.: Изд-во Московского ун-та.

Урманчиева А.Ю. (2014), Эвиденциальные показатели селькупского языка: соотношение семантики и прагматики в описании глагольных граммем, *Вопросы языкознания 4*, 66–86.

Урманчиева А.Ю. (2015), Как грамматическая система управляет семантической эволюцией показателей, *Вопросы языкознания 6*, 52–75.

Храковский В. С. (ред.) (2007), *Эвиденциальность в языках Европы и Азии. Сборник статей памяти Наталии Андреевны Козинцевой*, СПб.: Наука.

Sentence final discourse particles in Amdo-Tibetan

Camille Simon

<pdf>

The Akan discourse particle *na*

Dorothee Beermann & Abigail Agyewaa
(NTNU, Trondheim; Polytext, Oslo & NTNU, Trondheim; Polytext, Oslo)

Keywords: modular discourse analysis, focus, questions, Akan, Minsky frames

The Akan particle *na* has been associated with two tones representing different grammatical functions. *Nà* with a low tone has been analysed as focus marker while *na* as a conjunction has been associated with additional effects on the logical or temporal interpretation of the clause. (Boadi 1974, 2008; Essegbey and Hatav 2018; Amfo 2007). Building on Van Dommelen and Beermann (2019) we will argue that low tone *na* is an interpretation operator which marks focus and questions. Na as a conjunction is associated with different tones and fulfils several functions within a modular discourse structure. In its most basic function it is a **discourse connective** in the sense of Halliday and Hasan (1976). In addition it is a **frame marker** in the sense of Minsky (1975).

Using a 98 000 word corpus of Akan composed of data from traditional narratives, movie audiostreams and code-switching dialogues our presentation will focus on the framing use of *na*. **Frames in the sense of Minsky** provide **sitational moulds.** They correspond to larger units of text, and can often only be detected when considering units larger than the sentence. In a pilot annotation task we have identified purviews, and perceptual and psychological frames. Frames allow for a subjectivated processing of information on top of the processing of text based on basic coherence. Frames in that sense might be a concept similar to commonsense entailments (Lascarides and Asher,1993). On the basis of our corpus, we further noticed that *na* serves to express confidence in the **contexts of multilogs like public discussions.** Evidence for the latter claim comes from examples taken from a Ghanaian radio dialogue where the radio hosts and their guests code-switch extensively between English and Akan. Consider (1)

(1) Person 1: **Na** [$_{p1}$ɔmo yɛ team a they have to prepare]. Team biara pɛ sɛ ɔmo prepare.
[$_{p1}$3.PL COPteam Pnrel they have to prepare]
Calander is meant for the team a ɛsɛ sɛ ɔmo prepare. **Na** [$_{p2}$ yɛ-te hɔ dadaada **na** afei]
[$_{p2}$ 1PL-sit there longtime and now]
'[$_{p1}$They are a team, they have to prepare.] Every team wants to prepare. A calender is meant for the team for them to prepare. [$_{p2}$We have been sitting for long and it's now.]'

In (1) the first two occurrences of *na* together mark a cause-effect frame. Set initial to their clauses *na* also functions to express **speaker certainty** and claims the hearer's attentions. This latter function seems to be speech act specific. Different from Framing the last occurrence of *na* in (1) is an instance of its function as discourse connective, corresponding in meaning to English 'and'. In support of our claims we will present in addition to the standard corpus of Akan a small Akan subcorpus labeled for discourse senses.

References

Amfo, Nana Aba A. (2007). Clausal conjunction in Akan. *Lingua* 117: 666–84

Beermann, Dorothee and Lars Hellan. (2018). Creation of Digital Resources for Two Related Kwa Languages: Akan and Ga. Presentation, 3rd Symposium of West African Languages University of Warsaw, 26-29 September, 2018.

Bearth, T. (1999). The contribution of African Linguistics towards a general theory of focus. Update and critical review. JALL 20: 121-156.

_____ (2002). Fokus im Akan – Outline. Handout of a presentation at the Linguistic Colloquium of the Seminar of African Studies at the Humboldt University.

Boadi, L.A. (1974). Focus-Marking in Akan. Linguistics 140: 5-57.

Boadi, L.A. (2008). Tense, Aspect and Mood in Akan. In Ameka, F. and Kropp Dakubu, M.E. (eds.) Aspect and modality in Kwa languages. Amsterdam/Philadelphia: John Benjamins.

Essegbey, James and Hatav, Galia. (2018). Towards a unified analysis of *na* in Akan. Ms. University of Florida.

Halliday, M.A.K. and Ruqaiya Hasan (1976) Cohesion in English. Routledge, London.

Lascarides, A. and N. Asher (1993) Temporal Interpretation, Discourse Relations and Commonsense Entailment, Linguistics and Philosophy, 16(5), pp437-493, Kluwer Academic Publishers, Dordrecht, Holland.

Minsky, M. (1974). A Framework for Representing Knowledge. Reprinted in The Psychology of Computer Vision, P. Winston (Ed.), McGraw-Hill, 1975.

Schwarz, A. and Fiedler, I. (2007). Narrative strategies in Gur and Kwa. In E.O. Aboh, K. Hartmann and M. Zimermann (eds.) Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic. Berlin: Mouton de Gruyter. 139-160

On the rise of epistemic meanings from deictic expressions: The case of Hebrew *zehu* 'that's it'

Anna Inbar & Leon Shor
(Tel Aviv University)

Keywords: demonstratives, deixis, epistemic independence, spoken Israeli Hebrew

In this talk, we will focus on the particular uses of the deictic expression *zehu* 'that's it' in spoken Israeli Hebrew, in which it conveys a stance towards what the previous speaker has said.

Morphologically, *zehu* consists of the masculine singular demonstrative pronoun *ze* 'this' and the enclitic pronoun *hu* 'he'. When *zehu* is used for identification of people and objects or as a discourse-deictic demonstrative that introduces new referents into discourse, it has a gender-number inflection, and it is integrated within the structure of the clause. However, *zehu* may have various pragmatic-discursive uses, in which it usually constitutes a syntactically and prosodically independent utterance and does not inflect for number and gender (Bardenstein, Shor, & Inbar, in press). For example, *zehu* can be used to convey *epistemic independence* (Heritage, 2002), when uttered in response to information provided by the prior speaker, to markedly indicate access to a state of affairs regardless of the fact that the previous speaker was the first one to verbalize it. Consider the following example:

1 sp1 *ve ha=otobus | kmo ma nira /*
and DEF=bus | like what it.looks /
'And the bus, what does it look like?'

2 sp2 (1.1) *ha=otobus | kmo ma ʃe haja be=ʃnot.ha.arbaim || kaze ||*
(1.1) DEF=bus | like what that was in=the.forties || like.this ||
'(1.1) The bus, similar to what was in the forties. Like that.'

3 sp1 ***zehu*** || (1.4) *im kaze af ||*
this.he || (1.4) with like.this nose ||
'**That's it**. (1.4) With a nose like that.'

4 sp2 *ken ||*
yes ||
'Yes.'

(OCh_sp2_068-071, sp1_194-198)

In line (1), sp1 asks what the buses in Mongolia look like. After the short description provided by sp2 (line 2), sp1 produces a standalone *zehu*, which indicates his epistemic independence with regard to the information he requested; that is, it appears that sp1 had a general idea as to what buses in Mongolia look like, and sp2's response confirmed that idea. Then, he verifies his epistemic independence by providing a description of the buses by himself: 'With a nose like that' (line 4).

The aim of this study is to analyze the functional polysemy of the deictic expression *zehu* 'that's it' in Israeli Hebrew, as well as to suggest possible paths of development from a discourse-deictic demonstrative to a discourse marker with epistemic meanings. Furthermore, we will show that the phenomenon of epistemic modality rising from deixis is widespread in Hebrew as well as in other languages (e.g. Frajzyngier, 1991; Frajzyngier and Jasperson, 1991).

References

Bardenstein, Ruthi, Shor, Leon & Inbar, Anna (in press), *Zehu* 'that's it' as a discourse marker in spoken Israeli Hebrew. *Hebrew Linguistics*. Bar-Ilan University Press. [In Hebrew]

Frajzyngier, Zygmunt (1991), The *de dicto* domain in language, in E. C. Traugott, and B. Heine (eds), (1991), *Approaches to Grammaticalization*, Vol. 1 [*Typological Studies in Language 19*], Amsterdam: John Benjamins, 219-251.

Frajzyngier, Zygmunt and Robert Jasperson (1991), *That*-clauses and other complements, *Lingua* 83, 133-153.

Heritage, John (2002), Oh-prefaced responses to assessments: a method of modifying agreement/disagreement, in C. Ford, B. Fox, and S. Thompson (eds), (2002), *The Language of Turn and Sequence*, New York: Oxford University Press, 196-224.

ꕥ

Presentation and negotiation: Discourse particle sequencing in Ainu

Katsunobu Izutsu & Mitsuko Narita Izutsu
(Hokkaido University of Education & Fuji Women's University)

Keywords: Ainu, discourse particle, particle sequence, stance marking, speech event conception

This study addresses sequences of utterance-final discourse particles in Ainu, an indigenous language of Japan. Using *Ainugo Onsei Shiryo* (*A Spoken Corpus of the Saru Dialect of Ainu*), we demonstrate that the particle sequences largely exhibit an ordering summarized as PRESENTATION FOLLOWED BY NEGOTIATION: a particle concerning the "presentation" of a speech act is followed by another particle representing the "negotiation" with the addressee(s) for the act. We also analyze these "presenting" and "negotiating" functions with respect to the speech event conception that underlies those functions.

Ainu allows for particle sequences like *yan hani*, *yan un*, and *na hani*, as illustrated in (1). The particle *yan* can follow imperative clauses with an honorific or plural addressee(s), as in (1a-b). It serves to encode the speaker's unilateral stance of presenting the directive force to the addressee(s) but, at the same time, it adds a more polite or reverent nuance to the overall utterance. On the other hand, the particle *na* is used in declarative clauses when the utterances are likely to imply the speaker's directive attitude toward the addressee(s). For instance, (1c) is intended to get the addressee to be prepared for being picked up by the speaker.

(1) a. *suy* *sinewe* ***yan*** ***hani***. (Tamura 1977: 42)
again visit DP DP
'(Please) come again.'

b. *hokure* *ahunke* ***yan*** ***un***. (Tamura 1977: 41)
quickly bring.in DP DP
'(Please) bring (him) in quickly.'

c. *eci-ekanok* *kusu* *ne* ***na*** ***hani!*** (Tamura 1977: 135)
you.I-go.pick.up intending be DP DP
'I will come and pick you up.'

Sometimes following *yan* or *na*, the particles *hani* and *un* index the speaker's intention to negotiate the illocutionary force indicated by the clauses with *yan* or *na*, as shown in (1a-c): *hani* is added to confirm the addressee's reaction to the illocutionary force, while *un* is attached to convey the speaker's responsive stance to the addressee. Therefore, the utterances more explicitly signal some negotiatory meanings, which can be rendered in English as: '(Please) come again, will you?' in (1a), '(Please) bring him in quickly, then' (or more specifically, '(If you say there is a young man waiting outside,) please bring him in quickly.') in (1b), and 'I will come and pick you up, so be ready by that time, is it alright?' in (1c), respectively.

These particle sequences reflect the speaker's stance marking procedure, where the speaker's unilateral force exertion is coded before the negotiation with the addressee(s) for the force. We argue that this procedure manifests a cross-linguistically observed pattern of final discourse particle sequencing: a more subjective particle is produced before a more intersubjective one (Saji 2001,

Shinzato 2007, inter alia). This is attested in sentence-final sequences of Japanese final particles (e.g., *-yo-ne*) or English discourse markers (e.g., *please, will you?*). We finally propose this stance-marking procedure as a cross-linguistically applicable facet of the speech event conception, which has been described more schematically in earlier studies such as Langacker (1991: 283-284) and Ide (2005: 51-52) among others.

References

Ide, Sachiko (2005), How and why honorifics can signify dignity and elegance: The indexicality and reflexivity of linguistic rituals, in Lakoff, Robin T. and Sachiko Ide (eds.), *Broadening the Horizon of Linguistic Politeness*, Amsterdam: John Benjamins, 45-64.

Langacker, Ronald W. (1991), *Foundations of Cognitive Grammar, vol. II: Descriptive Applications*, Stanford: Stanford University Press.

Saji, Keizo (2001), Nihongo kyooiku to bunpoo gakusetsu [Japanese language teaching and theories of grammar], *Nihongogaku* [*Japanese Linguistics*] 20, 6-10.

Shinzato, Rumiko (2007), (Inter)subjectification, Japanese syntax and syntactic scope increase, *Journal of Historical Pragmatics* 8(2), 171-206.

Tamura, Suzuko (1977), Ainugo Saru hoogen no meirei hyoogen [Imperative expressions in the Saru dialect of Ainu], *Asian and African Linguistics* 5, 22-141.

Waseda Daigaku Gogaku Kyoiku Kenkyujo (ed.) (1984-1989), *Ainugo Onsei Shiryo* [Ainu language phonetic material] vols.1-6, Waseda University.

Discourse Particles and Information Structuring: Epistemic stance in oral languages

M.M.Jocelyne Fernandez-Vest
(CNRS & Université Sorbonne Nouvelle – Paris 3)

Keywords: discourse particles, orality, epistemic stance, implicitness, typology

Enunciation theories that favor language subjectivity have documented from diverse perspectives the contribution of Discourse Particles (DIPs) to Information Structuring (IS) – Benveniste 1966, Kerbrat-Orecchioni 1984, Fernandez-Vest 1994, Montaut 2002, Belyaev & Forker 2016. This paper intends to focus on the IS role of epistemic DIPs in the oral registers of some typologically different languages, Western Uralic and French.

In orally transmitted Northern Sami (Samic) for instance, the negotiation of a distance was operated by numerous polysemic DIPs,

(1) – *Goal mo bat dal dat lea? Galhan dat lea vissa beannot miilla vai.*
'How much could it in fact actually be? Yes indeed it is surely one and a half mile or what.
(Ohcejohka corpus, Fernandez-Vest 1987)

But after the recent standardization, DIPs have been reduced to a few core functions, e.g. epistemic stances in electronic conversations

(2) – *Doppe han eai máhte šat 'č' jiena dadjat.*
'There, as we know, they are not able any more to pronounce the *'č'* sound'
(Facebook corpus, Fernandez-Vest 2018).

The analysis of another type of internal contrastivity (oral-written) in Finnish (Finnic), shows how one can change a recorded conversations's experiential subjectivity into a neutral impersonal style through obliterating in its edited version two epistemic DIPs, *sitä*, denoting a speaker's discrete participation, and *-hAn*, appealing to shared knowledge:

(3) – *Nuorena sitähän* [DIP+DIP] *elää* *niinku* [DIP]
as-young [ESS] that's true [DIP+DIP] one lives [PRS.3SG] like [DIP]
semmosessa / *rakastumisen* *ilmapiirissä.*
in-sort of [INESS.SG] / of-falling in love [GEN.SG] in-atmosphere [INESS.SG]
'When one is young it's sure one lives like in a sort of / in an atmosphere of passion.'
> – *Nuori ihminen elää rakastumisen ilmapiirissä.*
'A young person lives in an atmosphere of passion.'
(Fernandez-Vest 2015:213–216)

Supplementing studies dedicated to turn-initial particles (Auer & Maschler 2016, Heritage & Sorjonen (eds.) 2018)), my observations concentrate on the right periphery, a syntactic position semantically essential for oral registers, e.g. in a Southwestern French corpus

(4) – F1 *C'est quand même* [DIP] *dangereux* / *les moules*.
– F1 'It is all the same [DIP] dangerous / the mussles [Final Detachment].'
– F2 *Mais c'est si frais chez nous quand même* [DIP].
– F2 'But it is so fresh in our place all the same [DIP].
(Fernandez-Vest 2015:130–136)

With the same contesting DIP, but moved into final position, F2 turns round F1's argument: maybe dangerous, but fresh here. Their common DIP expresses an impersonal implicit judgement. Epistemic stances, that generally build an evaluative frame through overt communicative means (Du Bois 2007), are reputed to use "strategies of mystification such as implicitness" in political discourses (Marín-Arrese 2011). I will argue, relying upon several corpus-based comparisons, that implicitness is equally conveyed by epistemic DIPs in impromptu conversations that entail a familiarity of communication.

My conclusion emphasizes the necessity to give a priority to register typology for a functional inventory of epistemic DIPs.

References

Auer, Peter, and Yael Maschler (2016), The family of NU and NÅ across the languages of Europe and beyond: Structure, function, and history, in P. Auer, and Y. Maschler (eds.), *NU / NÅ: A Family of Discourse Markers across Europe and Beyond*, Berlin: de Gruyter, 1–47.

Belyaev, Oleg, and Diana Forker (2015), Word order and focus particles in Nakh-Daghestanian languages, in M.M.J. Fernandez-Vest & R. D. Van Valin, Jr. (eds.), *Information Structure and Spoken Language from a Cross-Linguistic Perspective*, Berlin / Boston: De Gruyter Mouton (TILSM 283), 239–261.

Benveniste, Émile (1966 [1958]), De la subjectivité dans le langage, *Problèmes de Linguistique Générale*, vol. 1, Paris: Gallimard, 258–266.

Du Bois, John W. (2007), The stance triangle, in R. Englebretson (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, Amsterdam: John Benjamins, 139–182.

Fernandez-Vest, M.M.Jocelyne (1987 [1984]), *La Finlande trilingue. 1. Le discours des Sames. Oralité, contrastes, énonciation*, Paris: Didier Erudition.

Fernandez-Vest, M.M.Jocelyne (1994), *Les Particules Énonciatives dans la construction du discours*, Paris: PUF (Linguistique nouvelle).

Fernandez-Vest, M.M.Jocelyne (2015), *Detachments for Cohesion – Toward an Information Grammar of Oral Languages*, Berlin / Boston: De Gruyter Mouton (EALT 56).

Fernandez-Vest, M.M.Jocelyne (2018), Typological evolution of Northern Sami from oral tradition to electronic conversations: Analyticity vs. syntheticity revisited, in L. Lindström, H. Metslang and A. Kalnača (convenors), *Circum-Baltic languages: varieties, comparisons, and change,* Workshop at *SLE 2018,* Tallinn, 28/08-01/09/2018.

Heritage, John, and Marja-Leena Sorjonen (eds.), (2018), *Between turn and sequence: Turn-Initial particles across languages*, Amsterdam: John Benjamins.

Kerbrat-Orecchioni, Catherine (1984), *L'énonciation. De la subjectivité dans le langage* Paris: Armand Colin (Collection Linguistique).

Marín Arrese, Juhana I. (2011), Effective vs. Epistemic Stance and Subjectivity in Political Discourse – Legitimising Strategies and Mystification of Responsibility, in Chris Hart (ed.), *Critical Discourse Studies in Context and Cognition*, Amsterdam: John Benjamins, 193–223.

Montaut, Annie (2002), La particule énonciative *to* en hindi/ourdou, *Cahiers de linguistique de l'INALCO* (4), 111–134.

Epistemic authority in reported discourse: Expression of epistemic stance by the quotative marker *cāà* in Wan

Tatiana Nikitina
(CNRS, LLACAN)

Keywords: epistemic stance, reported discourse, quotatives, egophoricity, Mande

This study explores the use of an unusual particle attested in constructions with reported speech in Wan (Mande). The particle *cāà*/*cāáà* has no propositional meaning, and expresses epistemic stance in the sense of Ochs (1996). With declaratives and commands, it is used to report statements that are likely to seem surprising or unexpected to the person to whom they are addressed. In (1), the character gives his wife instructions that she is likely to find surprising (given that sauce is considered an obligatory part of a decent meal):

(1)	è	gé	é	nɔ̀	lɛ̀ŋ̀	cāáà		ké	lā	pɔ̄		é
	drō	gɛ̄	ó								
	3SG	said	REFL	wife	to	QUOT	if	2SG	thing		DEF	
	cooked	PRT	PRT									
	lā	à	ɓlé		wō		wá	ō				
	2SG	3SG	sauce	make	not	PRT						

'He said to his wife: When you cook the food, do not make sauce!'
(contrary to your expectation)

The particle has an unusual distribution: it does not appear outside reported speech. This distribution suggests that the marker is used to express the epistemic stance of a reported – but not the current – speaker. This explains why it is not taken to be part of the reported discourse (it is not part of what is perceived as "original discourse" in 1). The marker is in this sense an instance of a previously unattested category – "embedded", or "second-order" epistemic stance.

The marker's interpretation depends on the utterance's mood. With declaratives and commands, it signals unexpectedness on behalf of the reported addressee; yet with questions, it is used to express surprise on behalf of the reported speaker. In (2), the character is baffled at finding his wife dancing like crazy in the bush:

(2) ɓílā lē ɓálè jēī-jēī lā zō lā gbè nɔ̀ klā cāà

2SG woman big old 2SG came 2SG son wife after QUOT
ɓé è wō á yɛ̱?
then 3SG became FOC how
'You old woman, you went to look for your son's wife, and how come this is happening?'

This difference in interpretation finds parallels in *egophoric* systems of encoding personal knowledge, which are also highly sensitive to mood (Floyd et al. 2018). It reflects a difference in the attribution of epistemic authority in declaratives (where the authority rests with the speaker) and questions (where it shifts to the addressee). The parallel with egophoricity helps explain the shifting interpretation of *cāà* as a marker of epistemic conflict between the interaction participants: in declaratives and commands, it signals unexpectedness on behalf of the (less knowledgeable) addressee, but in questions, it xpresses the reported speaker's surprise with something about which s/he requests more information from the more authoritative addressee.

I describe the use of the marker based on a collection of narratives, discuss its implications for the study of epistemic stance, and suggest that it should be treated in the context of other devices used by West African storytellers to manipulate the roles of the current and the reported speaker, most importantly logophoricity (Nikitina 2012a,b).

References

Floyd, Simeon, Elisabeth Norcliffe & Lila San Roque, eds. (2018), *Egophoricity*, Amsterdam: John Benjamins.

Nikitina, Tatiana (2012a), Personal deixis and reported discourse: Towards a typology of person alignment, *Linguistic Typology* 16(2), 233-263.

Nikitina, Tatiana (2012b), Logophoric discourse and first person reporting in Wan (West Africa), *Anthropological Linguistics* 54(3), 280-301.

Ochs, Elinor (1996), Linguistic resources for socializing humanity, in J. J. Gumperz & S. C. Levinson (eds.) *Rethinking Linguistic Relativity*, Cambridge: Cambridge University Press, 407-437.

Emotional Event-Linking Pragmatic Markers of Negativity in Spoken English and Polish: A Corpus-Based Approach

Paul A. Wilson and Barbara Lewandowska-Tomaszczyk

Research objective. The paper aims to look into one category of Event-Linking pragmatic markers, more specifically into the concepts including negativity in emotional contexts. The study is based on empirical data derived from spoken materials in the referential corpora of Polish (nkjp.pl), as well as the spoken component of the British National Corpus, accompanied by relevant recordings and pitch profiles. To analyse the contrasts more broadly translational (parallel) English-to-Polish and Polish-to-English corpora (Cartoni et al., 2013) are consulted, available at http://pelcra.clarin-pl.eu/. The focal research questions refer to the identification of the negativity elements in the linking devices,

particularly those which convey higher emotional arousal (AUTHOR 1 & AUTHOR 2, 2011 on emotion dimensions). The main objective is to identify functions of these markers and propose their categorization in terms of the epistemic stance (Ochs 1996) and the speaker's evaluative judgment expressed.

Negativity & emotions. Negativity is understood here as the presence of some elements of truth-conditional negation either in terms of direct negative markers (Eng. no, not, Pol. nie, affixal markers) or as indirectly conveyed negative meanings (presupposition or implicative, evaluative). Elements which possess negative properties are reported to be more salient cognitively as well as more powerful rhetorically than less marked corresponding positive forms (Tice, Bratslavsky, and Baumeister 2001), and intense negative emotional states affect self-control by significantly releasing it. Therefore the tracing of the elements of negativity in discourse markers can contribute to the emotion and negation research as well as uncover new vistas in the analysis of pragmatic markers, also contrastively.

Research methodology.The research methods used are both quantitative (frequencies of use), as well as qualitative, i.e., involving the cognitive frame-based linguistic and discourse perspective. The study focuses on the analysis of Polish corpus data of one of the negativity clusters for spoken discourse relations such as not that (e.g., it's not that he didn't take it seriously), usually followed by the adversative but clause, otherwise (You have to see the light you know, otherwise you won't see it!), oh no, not at all, and by no means as contrasted to Pol. co(ś) ty/no co ty, present exclusively in conversational materials, roughly equivalent to 'but', and (no) nie ('oh no', 'but not'), displaying a whole range of polysemously linked senses, in terms of knowledge frames (Fillmore, 1982), embracing particular Emotion Events (AUTHOR 1, 2011 on the concept of Event; AUTHOR 1, 2015 on negative discourse markers). The stress-pitch patterns are also investigated for these uses (AUTHOR 1 2015) to identify their prosodic structure and determine properties characteristic for such emotional utterances. The parallel corpus data and their functional interpretation are analysed to contrast inter-language equivalence clusters (compare Cartoni et al., 2013).

Expected results. Categories of the relevant constructions and uses are identified in the Polish and English spoken corpora (http://spokes.clarin-pl.eu/) and their criterial properties and cross-linguistic typology proposed. They serve to provide a more complete description of the functions of emotional pragmatic markers expressing negative epistemic stance for this type of Event-Linking markers.

References

Cartoni, B., Zufferey, D. & Th. Meyer. (2013). Annotating the Meaning of Discourse Connectives by Looking at their Translation: The Translation Spotting Technique. In Discourse and Dialogue. Beyond semantics: the challenges of annotating pragmatic and discourse phenomena. Eds. S. Dipper, H. Zinsmeister, & B.Webber Vol. 4, No 2, pp.65-86.

Fillmore, Ch. (1982). Towards a descriptive framework for spatial deixis. In R. J. Jarvella and W. Klein (eds.). Speech, Place and Action. London: John Wiley, pp.31-59.

AUTHOR 1 (2015).

AUTHOR 1 & AUTHOR 2 (2011).

Ochs, E. (1996). Linguistic resources for socializing humanity. In JJ. Gumperz and S.C. Levinson: Rethinking Linguistic Relativity. CUP: Cambridge. 407-437.

Tice, D. M., E. Bratslavsky & Roy F.

Baumeister. (2001). Emotional distress regulation takes precedence over impulse control. Personality and Social Psychology 80(1). 53–67.

Epistemic markers in colloquial Komi

Gerson Klumpp, Denys Teptiuk & Eda-Riin Leego
(University of Tartu)

Keywords: epistemic stance, epistemic particles, language contact Komi–Russian, Family Problems Picture Task

In this study, we investigate the use of discourse markers expressing epistemic stance in colloquial Komi (Permic branch of Uralic). For this, we employ experimental data recorded with 36 Komi speakers solving the Family Problems Picture Task (Barth & Evans 2017). Our material contains approximately 6 hours of spoken data recorded in Syktyvkar (Komi Republic, Russia) in April 2018. The speakers represent several dialect areas of the two literary standards of Komi (cf. Hausenberg 1998), and all are bilingual Komi-Russian.

The Family Problems Picture Task was designed the way that it encourages speakers to interact in a naturalistic way by solving a narrative problem (Barth & Evans 2017: 3). To find a compromise in the description of the set of pictures resulting in a narrative, speakers turn to the use of epistemic particles marking different degrees of "the epistemic states of the speaker, or her interlocutors, or both, with respect to the descriptive, or prepositional, content of an utterance" (Zimmermann 2011: 2012):

(1) *tani* *n'in* ***naverno*** *sijö* ***gaškö*** *(.)* *osoznovajtö* *tajö*
here already maybe 3SG maybe realize.PRS.3SG DEM
vot *(1.0)* *mužyćöjys*.
PTCL man.3SG
'Here already **maybe** he **maybe** realizes it, this, well, … man' (fulab2, 07:08–07:14).

(2) *tajö* *arbuz* ***žö,*** *abu?*
DEM watermelon still isn't
'**But** this is a watermelon, isn't it?' (rejecting interlocutor's idea that it's a pumpkin) (MI, 10:54–10:55).

(3) *ta* ***vid'ima*** *myjkös* *koröma*
DEM apparently something.ACC request.PST.EVID.3SG
'This one **apparently** requested something' (fulab2, 21:11–21:13).

Due to unidirectional influence of Russian on Komi and extensive use of the former in everyday life by Komi speakers (cf. Leinonen 2006, 2009; Nekrasova 2013), it goes without any surprise that besides indigenous means (*gaškö* 'maybe' in (1)), or older borrowings (*žö* 'still' in (2)) expressing epistemic stance, in colloquial Komi one can also observe markers replicated recently from Russian, e.g. *vid'ima* 'apparently' in (3) (not listed in the Komi-Russian dictionary Beznosikova et al. 2000). Thus, the following research questions arise: (i) What kind of indigenous and Russian epistemic particles can be observed in colloquial Komi (e.g. in relation to the stance triangle, Du Bois 2007)? (ii) What triggers the employment of Russian epistemic particles? First results show that among the replicated markers we find those that are functionally interchangeable with their indigenous counterparts (*gaškö* vs. *naverno* in (1)). Thus, our initial hypothesis suggests that the appearance of some epistemic markers is motivated merely by frequent code-switching in spoken Komi. However, several Russian markers are likely to appear in specific pragmatic settings, e.g. visual inferential meaning (*vid'ima* 'apparently'), cf. (3), or also in reported speech constructions (*t'ipa* 'like', *byt't'ökö* 'like, as if'). Thus, their use might be triggered by additional functions of such markers, not observed among the indigenous counterparts. To answer the research questions, we conduct a qualitative descriptive study of the functional, discourse-pragmatic, and morphosyntactic properties of Komi

discourse particles, the results of which provide insights into epistemic markers in contemporary spoken Komi under a language contact perspective.

References

Barth, Danielle & Nicholas Evans. 2017. "SCOPIC: Design and Overview". In: Barth, Danielle & Nicholas Evans, *Social Cognition Parallax Interview Corpus (SCOPIC).* Language Documentation & Conservation Special Publication No. 12, 1–21. Honolulu: University of Hawai'i Press.

Beznosikova, Ljucija M. & Ajbabina, Evgenija A. & Kosnyreva, Raisa I. 2000. *Komi-russkij slovar' – Komi-roč kyvčukör.* Syktyvkar: Komi knižnoe izdat.

Du Bois, John W. 2007. "The Stance Triangle". In: Englebretson, Robert (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, 139–182. Amsterdam: Benjamins.

Hausenberg, Anu-Reet. 1998. Komi. In: Abondolo, Daniel (ed.), *The Uralic Languages*, 305–327. London & New York: Routledge.

Leinonen, Marja. 2006. "The Russification of Komi". *Slavica Helsingiensia* 27, 234–245. Available at: http://www.helsinki.fi/venaja/nwrussia/eng/Conference/pdf/Leinonen.pdf

Leinonen, Marja. 2009. "Russian influence on the Ižma Komi dialect". *International Journal of Bilingualism* 13(3), 309–329.

Nekrasova, Galina A. 2013. "Dublirovanija v televizionnoj reči komi bilingva [Duplications in a television speech of bilingual Komi]". *Jazyk i kul'tura* 3(23), 73–81.

Zimmermann, Malte. 2011. Discourse particles. In: Heusinger, von Klaus, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning.* Vol. 2, 2012–2038. Berlin: Mouton.

Marking epistemic stance with a borrowed marker: *ki* in Pharasiot Greek

Metin Bagriacik
(Ghent University)

Keywords: discourse particle, epistemic vigilance, language contact, Pharasiot Greek, Turkish

Discourse particles (DPs) are listed among the linguistic elements that are highly vulnerable to borrowing in contact/bilingual situations (Brody 1987, Sankoff *et al.* 1997, Matras 2009). This cross-linguistic observation is built on elements functioning as DPs in the donor language; however, in principle, elements belonging to other linguistic categories in the donor language could also be accommodated as DPs in the recipient language. This study provides an exemplar case, showing that a grammatical element in the donor language has been accommodated into the syntax of the recipient language as a well-defined DP.

Pharasiot Greek (PhG), an Asia Minor Greek dialect which was in intensive contact with Turkish for centuries, possesses a certain marker, *ki*, borrowed from Turkish (Dawkins 1916, Anastasiadis 1976). *Ki* in Turkish is generally claimed to be a complementizer introducing right-branching complement clauses (Kornfilt 1997, Erkman-Akerson and Özil 1998) or a coordinator connecting *ki*-clause to the matrix one (Griffiths and Güneş 2014). In PhG, *ki* appears as an optional marker after epistemic adverbs (1), a coordinator combining two full-fledged sentences (2), a subordinator after matrix predicates introducing a complement clause (3), and an emphatic marker in clause-final position (4).

(1) Epistemic adverb + KI + declarative main clause
(2) Clause 1 + KI + declarative main clause
(3) Declarative main clause + KI + declarative complement clause
(4) Declarative main clause + KI

Based on speaker judgments, I show that in (1), *ki* blocks objective modal reading of the adverb, which is otherwise available. In (2), the first clause functions as an adverbial, by which the speaker presents justification for the truth of the proposition she expresses with the second clause. In (3) and (4), the speaker attempts to raise her credibility with respect to the (main) proposition in the clause she utters. I claim that the functions of *ki* in (1–4) can be subsumed under a unique interpretive function: *ki* is a speaker-oriented discourse marker geared to influencing the hearer's epistemic vigilance (Sperber *et al*. 2010). It is employed by the speaker to signal to the hearer her authority and competence with respect to the content of her assertion and/or to show the hearer her benevolence and trustworthiness with respect to the content of her assertion. Adopting Speas and Tenny's (2003) hypothesis that the pragmatic roles of "speaker" and "hearer" are encoded in a discourse domain in the syntax, I identify this discourse domain as Speech Act Phrase (SAP). I propose that *ki* is the overt exponent of SA^0, the functional head endowed with a [+sentience] feature identifying the "speaker" as the unique "sentient mind". SAP dominates ForceP, whose head position is not overtly lexicalized in a declarative main clause. The apparent surface differences between (1)–(4) then reduce to whether the [+sentience] feature is checked by an internally- or externally-merging category.

References

Anastasiadis, Vasilios (1976), I sintaksi sto Farasiotiko idioma tis Kappadokias [The syntax of the dialect of Pharasa in Cappadocia], PhD dissertation, University of Ioannina.

Brody, Jill (1987), Particles Borrowed from Spanish as Discourse Markers in Mayan Languages, *Anthropological Linguistics* 29(4), 507–532.

Dawkins, Richard (1916), *Modern Greek in Asia Minor: A study of the dialects of Síli, Cappadocia and Pharasa, with grammar, texts, translation and glossary*, Cambridge, UK: CUP.

Erkman-Akerson, Fatma and Şeyda Özil (1998), *Türkçe'de niteleme: Sıfat işlevli yan tümceler* [Modification in Turkish: Adjectival complement clauses]. Istanbul: Simurg Yayınları.

Griffiths, James and Güliz Güneş (2015), Ki issues in Turkish: Parenthetical coordination and adjunction, in M. Kluck, D. Ott, and M. de Vries, (eds), (2015), *Parenthesis and Ellipsis: Cross-Linguistic and Theoretical Perspectives*, Berlin: Mouton de Gruyter, 173–218.

Kornfilt, Jaklin (1997), *Turkish*, London / NY: Routledge.

Matras, Yaron (2009), *Language contact*, Cambridge, UK: CUP.

Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, Hélène Blondaeu, Marie-Odile Fonollosa and Lucie Gagnon (1997), Variation in the Use of Discourse Markers in a Language Contact Situation, *Language Variation and Change* 9(2), 191–217.

Speas, Margaret and Carol Tenny (2003), Configurational properties of point of view roles, in A. M. Di Sciullo (ed), (2003), *Asymmetry in Grammar*, Amsterdam / Philadelphia: John Benjamins, 315–343.

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi and Deirdre Wilson (2010), Epistemic Vigilance, *Mind and Language* 25(4), 359–393.

The markers of uncontroversial information in Europe

Vladimir Panov

This contribution presents a cross-linguistic study of elements whose functions are comparable to the ones of the German modal particle *doch* or the Russian *že* (in one of its functions). It also explores the areal distribution of such elements in Europe.

Doch and other modal particles of German as well as the Russian *že* and *ved'* have been extensively studied. There are innumerable language-particular studies of these within different frameworks but also some contrastive studies, e.g., German *vs* Russian (Orlova 2012), German *vs* Dutch (Foolen 2006), German *vs* Czech (Rinas 2006). These studies provide fine-grained pictures of the functioning of particles and highlight the differences between the compared elements of different languages.

By contrast, in this typological and areal-typological study, I focus on cross-linguistic similarities rather than differences. The study pursues two goals.

The first goal is to define cross-linguistically the discourse function that is, as I demonstrate, overtly expressed by specific markers in a number of unrelated and geographically remote languages, and to provide its semantic map. According to Forker (forthcoming), the element *=q'al* of Sanzhi Dargwa signals that the proposition to which it refers is *uncontroversial*. I find this characterization also cross-linguistically applicable, although this function typically goes hand in hand with other closely related ones such as the marking of *given information*, the information *shared by the speaker and the addressee*, or the indication of *causal relations*, and *contrast*. I propose the label *enimitive* for this kind of discourse functions: it appeals to one of the core uses of the Latin *enim* (Kroon 1995: 185). Languages lacking specific enimitive markers are normally able to express this group of functions periphrastically or by means of intonation.

Although enimitive markers are found worldwide, I argue that there are some areal tendencies in their geographical distribution. The second goal of this talk is to demonstrate that the presence of highly specialized enimitive markers is characteristic of a macroarea that covers a small part of Western Europe, almost the whole of Central, Eastern and Northern Europe, including the Finno-Ugric and Turkic languages of the European part of the Russian Federation. The marking of enimitive seems to form an uninterrupted isogloss.

The origin of the "European" enimitive isogloss is rather obscure. The sources of enimitive markers vary significantly, material or pattern borrowing is well attested in some cases. The areal pattern itself, however, is not uniform, and various sub-areas may be identified on the basis of the number of enimitive markers, their position in the sentence, polysemy model and frequency in speech.

In this research I do not adhere to any particular descriptive framework and try to use intuitively clear formulations which allow for cross-linguistic comparison using the semantic map method. However, I make extensive use of existing language-particular descriptions done within different frameworks. Other sources of data include parallel corpora and, in some cases, examples provided by native speakers that were asked to describe certain situations.

References

Foolen, Ad. 2006. Polysemy patterns in contact: the case of Dutch *toch* and German *doch*. In Aijmer, Karin & Anne-Marie Simon-Vandenbergen (eds.) Pragmatic markers in contrast: 59-72.

Forker, Diana (forthcoming). More than just a modal particle: The enclitic *=q'al* in Sanzhi Dargwa. To be published in *Functions of Language*.

Kroon, Caroline. 1995. Discourse particles in Latin: a study of *nam*, *enim*, *autem*, *vero* and *at*. Amsterdam: J. C. Gieben.

Orlova, Svetlana N. 2012. Sopostavitel'nyj analiz semantiki nemeckix častic denn, doch i ja i ix

russkix perevodnyx ekvivalentov [A comparative analysis of the German particles denn, doch, and ja and their Russian translation equivalents]. A PhD thesis. Moskva: Moscow State University.

Rinas, Karsten. 2006. Die Abtönungspartikeln „doch“ und „ja“. Semantik, Idiomatisierung, Kombinationen, tschechische Äquivalente (Europäische Hochschulschriften XXI, 302). Frankfurt am Main u. a.: Peter Lang.

Discourse from motion: Cooptation or grammaticalization?

Aicha Belkadi
(SOAS University of London)

Keywords: discourse markers, consecutive markers, motion, grammaticalization, pragmaticization, cooptation

It is cross-linguistically common for discourse markers (DM) to develop from lexical sources. The origin of most English DM, for instance, can be traced back to nouns (e.g. *in fact*), adverbs (*frankly*, *well*) or more complex units (e.g. *You know what I mean*) (Traugott 1995). Despite a growing interest in their diachrony, it is not yet established how exactly DM emerge from their sources and how this process of change operates. Three contrastive analyses have so far been proposed. According to the first, the process is grammaticalization since it involves desemanticization, phonological reduction and decategorialisation, features which are typical of grammaticalization (Traugott 1995, 2003, Giacalone Ramat & Mauri 2009, Diewald 2011). The second view holds that the process is pragmaticization as it involves changes, such as increase in syntactic freedom and semantic-pragmatic scope, which are incompatible with grammaticalization and it does not give rise to grammatical units (Norde 2009). According to the third proposal, DM arise by cooptation, a non-gradual process whereby lexical items are simply borrowed from sentence grammar and directly mapped onto discourse grammar (Heine 2013).

My talk aims to contribute to this research by evaluating the cooptation vs. grammaticalization/ pragmaticalization proposals, focussing on the development of connectives/consecutives from motion verbs (Heine 2000, Heine & Kuteva 2002, Ebert 2003, Bourdin 2008). My primary data will be of verb *uɣal* ‘go.back’ in Taqbaylit Berber (Afro-Asiatic: Algeria). The verb lexicalises return-shaped motion (1) and has grammaticalised to encode inchoative semantics (2).

(1) **uɣal-(ə)ɣ=d** alami d amkan nni i qqim-(ə)ɣ
go.back.PFV-1SG=VEN until COP place DEM REL sit.PFV-1SG
*‘I **went back** to that place where I had been staying’.*

(2) y-pwḍ=(ə)d meskin. **Y-uɣal** d awəraɣ
3SGM-arrive.PFV=VEN poor.one 3SGM-become.PFV COP yellow
‘*He arrived, the poor man. He **became** yellow*.

Inside narratives, it is often also used as a DM, linking sets of related events to prior discourse (3). These events are often new or unexpected (4) but can be the consequence/culmination of previous events or simply a return to the topic of conversation after an aside.

(3) tə-gzəm=as azərmuc.
3SGF-cut.PFV=3SG.DAT umbilical cord
T-uɣal tə-ttəl aqcic

3SGF-go.back.PFV 3SGF-bandage.PFV boy
*'She cut his umbilical cord. **And then** she bandaged the boy.'*

(4) **Y-uɣal** yə-fka=yas=t
3SGM-go.back.PFV 3SGM-give.PFV=3SGM.DAT=3SGM.ACC
'**Finally** (**and contrary to what he had said**), he gave it to him.'
(Chaker, 1997: 115)

My talk will show that, although it does not display characteristic signs of verbal grammaticalization – such as reduction, fusion, loss of agreement and ability to inflect for TAM – there are signs that the discourse construction involving discourse *uɣal* is developing gradually and does not seem to be simply borrowed and mapped somewhere else in the clause with another function, as proposed by cooptation.

References:

Bourdin, Philippe. 2008. On the grammaticalization of 'come' and 'go' into markers of textual connectivity. *Typological Studies in Language*, Vol. 76. pp. 37-59. John Benjamins Publishing Company.

Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2). 365–390.

Giacalone Ramat, Anna & Caterina Mauri. 2010. Dalla continuità temporale al contrasto: La grammaticalizzazione di *tuttavia* come connettivo avversativo. In *Atti del X Congresso della Società Internazionale di Linguistica e Filologia (SILFI), Basilea 30.6–2.7 2008;* 449–470. Florence: Franco Cesati.

Heine, Bernd. 2000. On the rise of new-event markers in Kxoe. In Vossen, Rainer, Angelika Mietzner and Antje Meiβner (eds) *Mehr als nur Worte: Afrikanistische Beiträge zum 65. Geburtstag von Franz Rottland*. Cologne: Rüdiger Köppe Verlag.

Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else?. *Linguistics,* Vol. 51. 10.1515/ling-2013-0048.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Norde, Muriel. 2009. *Degrammaticalization.* Oxford: Oxford University Press.

Traugott, Elizabeth C. 1995. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the International Conference of Historical Linguistics XII, Manchester, August.

Traugott, Elizabeth C. 2003. Constructions in grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 624–647. Oxford: Blackwell.

WORKSHOP 4

Schedule: We 11.00 – We 17.55 (Room 12)

Elevation as a deictic category. Typological and diachronic perspectives

Antoinette Schapper & Gilles Authier

Keywords: elevation, spatial deixis, linguistic typology, diachrony, language diversity

Elevation in a spatial deictic system is where a referent's location or trajectory is identified as being at a certain elevation relative to the deictic centre. Although the topic of spatial language has subject to extensive cross-linguistic study in the last two decades (e.g., Talmy 1983, 1985, Wilkins and Hill 1995, Slobin 1996, Diessel 1999, Levinson 2003, Levinson & Wilkins 2006, Dixon 2013, Levinson et al. 2018), there has little comparative treatment or typologisation of elevation components in deictic systems. The lack of systematic treatment is most patent in the many terms that have been used to describe elevation components in spatial deictic systems in the typological and descriptive literature, including: "spatial coordinate systems" (Burenhult 2008), "topographical deixis" (Post 2011), "elevational deixis" (Schapper 2014), "environmental space deixis" (Bickel 2001), "altitude" (Gerner 2009), "altitudinal case markers" (Ebert 2003), "height" (Dixon 2003), "verticality" (Imai 2003) and "vertical case" (Noonan 2006).

While a relatively rare feature world-wide (Diessel 2013), elevation is a common component of systems of spatial reference in several mountainous areas of the world. It is pervasive in the languages of the Caucasus (e.g., East Caucasian languages, Alekseev 1997, Schulze 2003), of the Himalayas (Cheung 2007, Bickel & Gaenszle 1999, Post 2011, Heegård & Liljegren 2018) and of the Melanesian area (Heeschen 1982, Senft 1997, 2004, Aikhenvald 2015). It is less common but recurrent in many other parts of the world, including pockets of the Americas (e.g., Inuit languages, Fortescue 1988, Uto-Aztecan languages such as Guarjío, Miller 1996), of northern Southeast Asia (e.g., Gerner 2009), of Ethiopia (e.g., Mous 2012).

This workshop will bring together descriptive and historical comparative linguists working on languages and language families in both well-known "elevational" areas such as the Caucasus and New Guinea, but also on elevation-coding languages in regions where it is not the norm. The aim of the workshop is to explore the various ways in which spatial deictic systems incorporate elevation as a category through in-depth case studies of elevational coding. The ultimate goal is create a cross-linguistic picture of the diversity of elevation marking systems in terms of their grammatical and semantic properties. The following synchronic questions are among those that will be considered by contributors:

1. In what parts of the grammar is elevation marked? Which of the following are cross-linguistically common, which unusual? spatial information is frequently marked by the following:
    - Demonstratives
    - Pronouns
    - Case markers
    - Adpositions
    - Locational adverbs
    - Verbal affixes
    - Other…

2. What elevation distinctions are coded cross-linguistically? In some languages we find a binary elevation distinction of HIGH/LOW, and in others a tripartite elevation distinction HIGH/LOW/LEVEL. Finer distinctions are also found in some languages, with, for example, steepness of the slope playing a role in some elevation coding distinctions.
3. In some languages elevation is tied to very specific aspects of geography such as mountains (UPHILL/DOWNHILL), but in others it refers to any higher or lower location ('global' elevation in Burenhult's (2008) term). Are there systematic differences between elevation coding systems that are tied to specific geographic features and those with global reference?
4. What can be said about the relationship of elevation coding to other parts of spatial deictic coding? For example, what is the relationship between elevation and the PROXIMAL/DISTAL distinctions found in many spatial deictic systems? In some languages, elevation is only coded in the distal, whereas in others elevation is marked separately and can be used irrespective of distance.
5. How are elevational deictic systems used in real-world place- and path-finding? Are they subject to conventionalization at different scales of deictic reference? That kinds of social and cultural patterns might underpin use of elevational markers in ways that do not strictly conform to real-world elevation, e.g., in the house (cf. Yupno, Cooperrider et al. 2016), in relationship to important places such as religious or economic centres (cf. Bunaq, Schapper 2010)?
6. Are elevationally marked terms readily extended to non-spatial uses? If yes, what kinds of extensions beyond the realm of space are synchronically attested? To time (e.g., in Nungon, Sarvasy 2017), to narrative/discourse (e.g., in Siroi, van Kleef & van Kleef 2012), to social status (e.g., in Kryz, Authier 2009), to epistemic categories (e.g., in Blagar, Steinhauer 1991)?

The fact that elevation is frequent in certain parts of the world, but largely absent in others suggests that areal (including geographical/geo-physical) as well as genealogical factors play a role in explaining its distribution worldwide. As a result the workshop will also address comparative and diachronic questions on elevation marking systems, such as the following:

1. Are elevational deictic systems stable in families and/or areas? In what circumstances are elevational systems lost, maintained or innovated?
2. Where elevational deictic systems are innovated, what are the diachronic sources of elevational terms?
3. Where elevational systems are extended, are there patterns in the pathways for extension? Compounding of morphemes with one another to create complex reference is one attested pathway (e.g., in East Alor languages, Schapper 2014).
4. To what extent does Palmer et al's (2017) *Topographic Correspondence Hypothesis* (i.e., that languages spoken in similar topographic environments tend to have similar systems of absolute spatial reference) hold in relation to elevational deictic systems?

We aim to develop papers out of this panel into a collected volume to be submitted to Language Science Press.

References

Aikhenvald, Alexandra. 2015. Distance, Direction, and Relevance: How to Choose and Use a Demonstrative in Manambu. *Anthropological Linguistics*, 57.1: 1-45.

Alekseev, Mikhail. 1997. Reconstruction of the Proto-East-Caucasian locative morphemes. In Helma van den Berg (ed.), *Studies in Caucasian Linguistics*.

Authier, Gilles. 2009, *Grammaire kryz, langue caucasique d'Azerbaïdjan, dialecte d'Alik*. Peeters: Leuwen.

Bickel, Balthasar. 2001. Deictic transposition and referential practice in Belhare. *Journal of Linguistic Anthropology* 10(2). 224–247.

Bickel, Balthasar & Martin Gaenszle (eds.). 1999. Himalayan space: Cultural horizons and practices. Völkerkundemuseum Zürich.

Burenhult, Niclas. 2008. Spatial coordinate systems in demonstrative meaning. *Linguistic Typology* 12. 99–142.

Cheung, Candice Chi-Hang. 2007. On the noun phrase structure of Jingpo. *USC Working Papers in Linguistics* 3. 32–56.

Cooperrider, Kensy, James Slotta & Rafael Núñez. Uphill and Downhill in a Flat World: The Conceptual Topography of the Yupno House.

Diessel, Holger. 1999. *Demonstratives: form, function, and grammaticalization*. (Typological Studies in Language). Amsterdam: John Benjamins.

Diessel, Holger. 2013. Distance Contrasts in Demonstratives. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: http://wals.info/chapter/41

Dixon, R. M. W. 2003. Demonstratives: a cross-linguistic typology. *Studies in Language* 27(1). 61–112.

Ebert, Karen H. 2003. Kiranti languages: an overview. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages* (Routledge Language Family Series), 505–517. London: Routledge.

Fortescue, Michael. 1988. Eskimo orientation systems. Meddelelser om Grønland, *Man and Society* 11.3-30.

Gerner, Matthias. 2009. Deictic Features of Demonstratives: A Typological Survey with Special Reference to the Miao Group. *The Canadian Journal of Linguistics* 54(1): 43-90.

Heegård, Jan & Henrik Liljegren. 2018. Geomorphic coding in Palula and Kalasha, Acta Linguistica Hafniensia, 50.2: 129-160.

Heeschen, Volker. 1982. Some systems of spatial deixis in Papuan languages. In Jürgen Weissenborn & Wolfgang Klein (eds.), *Here and there. Cross-linguistic studies in deixis and demonstration.* 81–109. Amsterdam: John Benjamins.

Imai, Shingo. 2003. Spatial Deixis. PhD thesis, State University of New York at Buffalo.

Levinson, Stephen C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Levinson, Stephen C. & David Wilkins (eds.). 2006. *Grammars of space: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Levinson, Stephen, Sarah Cutfield, Michael Dunn, Nicholas Enfield, Sérgio Meira, & David Wilkins, (eds.). (2018). *Demonstratives in Cross-Linguistic Perspective* (Language Culture and Cognition). Cambridge: Cambridge University Press.

van Kleef, Sjaak & Jacqueline van Kleef. 2012. The use of a conceptual metaphor in the Siroi language of Papua New Guinea. Narrative is climbing a mountain. In Anna Idström & Elisabeth Piirainen (eds), Endangered Metaphors. [Cognitive Linguistic Studies in Cultural Contexts 2], 161–184. Amsterdam: John Benjamins.

Mous, Maarten. 2012. Cushitic. In Frajzyngier Zygmunt & Shay Erin (eds.), *The Afroasiatic languages*, 342-422. Cambridge: Cambridge University Press.

Palmer, Bill, Jonathon Lum, Jonathan Schlossberg & Alice Gaby. 2017. How does the environment shape spatial language? Evidence for sociotopography. *Linguistic Typology* 21(3). 457–492.

Post, Mark. 2011. Topographical deixis and the Tani languages of North-East India. In G. Hyslop, S. Morey & M. Post (eds.), *North East Indian linguistics*, 137–154. New Delhi: Cambridge University Press India.

Sarvasy, Hannah. 2017. *A Grammar of Nungon. A Papuan Language of Northeast New Guinea*. Leiden: Brill.

Schapper, Antoinette. 2010. Bunaq, a Papuan language of central Timor. PhD thesis, The Australian National University.

Schapper, Antoinette. 2014. Elevational systems in Alor-Pantar languages. In Marian Klamer (ed.), *The Alor-Pantar languages: History and typology*, 247–285. Berlin: Language Science Press.

Schulze, Wolfgang. 2003. The diachrony of demonstrative pronouns in East Caucasian. In Dee Ann Holisky & Kevin Tuite (eds.), *Current trends in Caucasian, East European and Inner Asian linguistics: Papers in honor of Howard I. Aronson*, 291–348. Amsterdam: Benjamins.

Senft, Gunter. 1997. *Referring to space: studies in Austronesian and Papuan languages* (Oxford Studies in Anthropological Linguistics 11). Oxford: Clarendon Press.

Senft, Gunter. 2004. *Deixis and demonstratives in Oceanic languages*. Canberra: Pacific Linguistics.

Slobin, Daniel I. 1996. Two ways to travel. In Matt Shibatani and S. A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 195-220. Oxford: Clarendon Press.

Steinhauer, Hein. 1991. Demonstratives in the Blagar language of Dolap (Pura, Alor, Indonesia). In Tom Dutton (ed.), *Papers in Papuan linguistics*, 177–221. Canberra: Pacific Linguistics.

Talmy, Leonard. 1983. "How language structures space". In Herbert Pick (ed.), *Spatial orientation: Theory, research, and application*, 225-282. New York: Plenum Press.

Talmy, Leonard. 1985. "Lexicalization patterns: Semantic structures in lexical forms". In Timothy Shopen (ed.*), Language typology and syntactic description*, 57-149. Cambridge: Cambridge University Press.

Wilkins, David & Deborah Hill. 1995. When GO means COME: Questioning the basicness of basic motion verbs. *Cognitive Linguistics* 6: 209-259.

Ups and downs: the history of elevational roots in East Caucasian

Gilles Authier
(Ecole Pratique des Hautes Etudes, Paris)

Keywords: East Caucasian, spatial cases, deixis, elevation, pronouns, preverbs

The East Caucasian family is deservedly famous for its elaborate systems of spatial cases on nouns, and among the most stable morphemes are reflexes of **-ƛ:'* and **-ƛ'* (tense and lax ejective lateral stops, still preserved in this phonetic shape in the Andic branch of the family), usually glossed 'SUB' = 'under' and 'SUPER' 'on (a horizontal surface, higher up)' respectively, which can be described as "elevational" markers.

Like most other East Caucasian case markers, these two markers commonly have reflexes in various branches of the family: as preverbs, as defective position verbs (/ copulas), and as adverbs, which points to an origin as free-standing elements. The most basic syntactic status of the former root remains unclear, but as for the latter, its origin is demonstrably lexical, as shown by related lexical items like 'head' or 'roof'. And whereas Avar has grammaticalized only SUB as a case ending, leaving to SUPER its status of free-standing adverb, neither of the two morphemes has ever become bound in the Nakh branch, which presumably split off earlier from the rest of the family. In other (Daghestanian) branches, however, outcomes of both roots are found more or less grammaticalized, and a chronology can be attempted, with consequences for internal classification.

After giving a brief overview of the different systems in which these two roots take part and their evolution across branches, we will also endeavour to describe in some detail the semantic map of the 'SUB' and 'SUPER' morphemes, comparing them with alternative or (synchronically) nearly synonymous roots also found meaning 'on, up' (as opposed to 'on a *vertical* surface' or 'down, below'. Inside each branch, different developments have taken place, and both markers have undergone various phonetic processes in onset and coda position, as well as rounds of paradigmatic replacement. Sometimes other markers have been added to the original elevational system (such as the 'up' and 'down' specification added to basic locations in the Dargwa branch, which are not related to the pair described above). In other instances, the same original morpheme has come to take on synchronically unrelatable suppletive shapes in prefixal and suffixal position.

Finally, we will contrast the intricate adventures of these two roots with the much simpler history of the root **d-* (mainly adverbial/pronominal), meaning either 'level' or 'on level ground', which appears to be more stable in diachrony. This 'unmarked level' (neither above or below) belongs not only in the elevational system, but usually also denotes 'distality', being submitted to the pressure of various 'proximal' roots. Ultimately, if related to the default ergative marker **-di* and to the most common singular substantivizing suffix **-d*, this 'level' and 'distal' deictic marker might be taken to represent a former marker of topicality.

ø

Elevation as a deictic category in Kartvelian

Merab Chukhua and Alexander Rostovtsev-Popiel

Keywords: elevation, localization, spatial deixis, preverbs, Kartvelian

The expression of spatial semantics in Kartvelian rests upon significantly diversified grammatical means in the adverbial, pronominal, and nominal domains, jointly conveying deictic, orientational, and localizational meanings. The semantic spheres of *high* and *low* particularly distinguish between absolute and relative oppositions, cf. Georgian *dabla* 'down(stairs)' and *maγla* 'upstairs/on top' vs. *kvemot* 'down(wards)/lower' and *zemot* 'up(wards)/higher,' respectively. It is remarkable that these expressions not only possess established etymologies in Common Kartvelian, but also have cognates throughout the family with next to identical contents in the systems of spatial terms (Klimov 1998; Rostovtsev-Popiel 2012).

Furthermore, such spatial categories as deixis, orientation, and localization have become grammaticalized in the elaborate systems of preverbs, the latter both sharing common properties across the family and representing remarkable diversity in terms of their internal organization (Harris 1991; Kutscher 2003, 2011; Lacroix 2009; Reseck 2014; Rostovtsev-Popiel 2012, 2016; Kobalava et al. 2017).

Considering the lower number of preverbs, Georgian and Svan keep relative parity between the nominal and verbal domains in the expression of spatial relations, whereby Megrelian and, especially, Laz data demonstrate significant imbalance in favor of the verbal domain. Being in possession of by far richer and "spatially-precise" systems of preverbs, these languages mark deixis and orientation/localization not only on verbs of motion, as is the case in Georgian and Svan, but also on verbs of position:

(1) *baγana cxvrins amaǯanu*. (Megrelian)

baγana-Ø	cxvin-s	**ama**-Ø-ǯan-u[n]-Ø
child-NOM	attic-DAT	**PRV**-S_3-lie-SM-S_3SG

'The child lies (**upstairs**) in the attic'.

(2) *bere otva-s gelažans* (Laz)
bere-Ø otva-s **gela**-Ø-ǯan-s
child-NOM roof-DAT **PRV**-S_3-lie-S_3SG
'X lies (**atop**) on the roof'

Furthermore, the expression of the semantics of elevation (as well as downward motion) includes such discrete nuances as travelling the path/carrying something up vertically (e.g. on the ladder) and gradually (e.g. up the stairs):

(3) *k'očik ǯaša ešelu*. (Megrelian)
k'oč-i-kǯa-ša **eš**[**a**]-Ø-i-[w]l-u
man-B-ERG tree-ALL **PRV**-S_3-VER_S-go-S_3SG.PST
'The man climbed **up** the tree'.

(4) *k'očik golaša ek'ilu*. (Megrelian)
k'oč-i-kgola-ša **ek'**[**o**]-Ø-i-[w]l-u
man-B-ERG mountain-ALL **PRV**-S_3-VER_S-go-S_3SG.PST
'The man went **up on** the mountain'.

This occurs both in purely elevation-oriented contexts and in combinations with numerous specific values, such as *inwards*, *outwards*, *over an obstacle*, whereby the spatial deictic contrast comes usually into play as well, cf. ***guma**-re(n)* 'X is/stands **upstairs <u>here</u>**' vs. ***gima**-re(n)* 'X is/stands **upstairs <u>there</u>**' (Rostovtsev-Popiel 2012).

Basing upon available corpora and recently collected fieldwork data, this study aims to provide a comparative account of the expression of the semantic domain of elevation in Kartvelian with a special emphasis on its deictic semantic component.

Abbreviations

3 – 3rd person; ALL – allative; B – base; DAT – dative; ERG – ergative; NOM – nominative; PRV – preverb; PST – past; S – subject person marker; SG – singular; SM – series marker; VER_S – subjective versionizer.

References

Harris, Alice (ed.) (1991), *The Indigenous Languages of the Caucasus, vol. 1. The Kartvelian Languages*, Delmar/New York: Caravan Books.

Klimov, Georgij (1998), *Etymological Dictionary of the Kartvelian Languages*, Berlin/New York: Mouton de Gruyter.

Kobalava, Izabela, Topadze Gäumann, Manana, and Rusudan Gersamia (2017). Rendering Spatial Relations by Means of Preverbs in Megrelian and Laz. In: Authier, Gilles, Gérardin, Hélène, Magomedov, Magomed, and Timur Maisak (eds.), *Historical Linguistics of the Caucasus. Book of Abstracts*. Makhachkala: 90–95.

Kutscher, Silvia (2003), Raumkonzeptionen im lasischen Verb: Das System der deiktischen und topologischen Präverbien. In: Boeder, Winfried (ed.), *Kaukasische Sprachprobleme*, *Caucasica Oldenburgensia 1*. Bibliotheks- und Informationsystem der Universität Oldenburg: 223–245.

Kutscher, Silvia (2011), On the Expression of Spatial Relations in Ardeşen-Laz, *Linguistic Discovery* 9, 2: 49–77.

Lacroix, René (2009), *Description du dialecte laze d'Arhavi (caucasique du sud, Turquie). Grammaire et textes*, Lyon PhD Dissertation, Ms.

Reseck, Tamar (2014), *Präverbien im Megrelischen (Diversitas Linguarum 37)*, Bochum: Brockmeyer.

Rostovtsev-Popiel, Alexander (2012), *Grammaticalized Affirmativity in Kartvelian*, Frankfurt PhD Dissertation, Ms.

Rostovtsev-Popiel, Alexander (2016), Kartvelian Preverbs in Cross-Linguistic Perspective, VII International Symposium on Kartvelian Studies. In: *Georgia in the Context of European Civilization*, Tbilisi: Tbilisi University Press: 94–97.

ꕥ

On the development of elevation deictics in Hmong-Mien

Martha Ratliff
(Wayne State University, Detroit)

Keywords: deixis, linguistic geography, grammaticalization, tone, Hmong-Mien

The Hmong-Mien (Miao-Yao) languages are spoken in the mountainous southern provinces of China and the mountainous north of Vietnam, Laos and Thailand. For those languages that have been well-studied, I show that some (Hmong, A-Hmao, A-Hmø, Iu Mien) have elevation deictics, and some (Xong, Hmu) do not. Two HM elevation deictics are derived from nouns meaning "mountain" and "plain", while another meaning "opposite, at the same elevation" (whether the opposite mountain or the opposite side of the river) seems to be a basic, nonderived deictic.

Elevation deictics have developed into demonstratives in two West Hmongic languages: White Hmong of Southeast Asia (Downer 1967, and Ratliff 2010) and A-Hmao of Weining County, Guizhou Province, China (Wang & Wang 1982). The development is from landscape noun to adposition via a class-marking tone change, followed by the optional absorption of a following demonstrative. This last change created new elevation-based demonstratives as part of a larger set. For example, in White Hmong, an archaic noun /**pe**$^{\mathbf{22}}$/ 'mountain' first became one of a set of tonally-defined adpositions:

(1) /**pe**$^{\mathbf{21ʔ}}$ ʈoŋ^{55}/ 'higher-on-slope mountain' ('up on the mountain')
(2) /**pe**$^{\mathbf{21ʔ}}$ no^{33}/ 'higher-on-slope that' ('that up-there')

The derived adposition absorbs a following demonstrative in fast speech, creating a new elevation-based demonstrative with a new minor tone:

(3) /**pe**$^{\mathbf{21ʔ}}$ no^{33}/ 'that up-there' > /**pe**$^{\mathbf{213}}$/ 'that-up-there'

A similar development took place in A-Hmao, but instead of creating a new minor tone (21ʔ + 33 = 213), the combination of adposition tone and demonstrative tone yielded a tone that already existed in the inventory (31 + 55 = 35).

Two facts suggest that HM elevation deictics have had a relatively shallow history. One is that landscape nouns recruited for grammaticalization in this way are not cognate across the family, and the other is that some languages have not developed elevation deictics at all. They appear to be accommodation to life in the mountains undertaken independently by small speech communities.

References

Downer, Gordon B (1967), Tone-change and tone-shift in White Miao, *Bulletin of the School of Oriental and African Studies* 30, 589-99.

Ratliff, Martha (2010), *Meaningful Tone: A study of tonal morphology in compounds, form classes, and expressive phrases in White Hmong*, DeKalb: Northern Illinois University Press.

Wang Fushi & Wang Deguang (1982), Guizhou Weining Miaoyu de fangweici [The localizers of Miao in Weining, Guizhou Province], *Minzu Yuwen* 4, 20-34.

The expression of ELEVATION in demonstratives of the Omotic-Lowland East Cushitic contact zone (Ethiopia)

Yvonne Treis
(CNRS-LLACAN)

Keywords: Omotic, Cushitic, demonstratives, elevation, language contact

In the southwestern highlands of the Ethiopian plateau we find an area in which languages are spoken that encode ELEVATION in their demonstrative systems. Interestingly, these languages are only distantly, if at all, related. The paper starts with an analysis of the demonstrative system of Baskeet (North Omotic). Baskeet differentiates formally between demonstrative determiners/pronouns and adverbs. Both morphosyntactic types encode six deictic degrees, which are exemplified here with the masculine singular pronouns (in *-áá*): DEM1 *ɦá[á]* (proximal), DEM2 *yíí* (proximal, invisible), DEM3 *gedáa* (medial), DEM4 *sekáa* (distal), DEM5 *lokáa* (high), DEM6 *yetáa* (low). Determining/pronominal demonstratives inflect for gender (masculine/feminine), number (singular/plural) and case (primary: nominative, accusative, predicative, secondary: ablative, allative, locative, instrumental). Two uses of the adverbial demonstratives need to be distinguished: (i) reference to locations and directions in the environs of the speech situation and (ii) reference to locations and directions in the wider geographical context. Places in and outside the Baskeet area are located as 'up' (DEM5), 'down' (DEM6) and 'over there' (DEM4), if they are to the East, to the West or to the North/South, respectively. Places that are 'up' or 'down' are not necessarily at a higher or lower elevation, but found in the direction of the rising and the setting sun.

The description of Baskeet is based on a corpus of texts (natural conversations, narratives) collected in the field and supplemented through elicitation.

In the second part of the paper, the Baskeet demonstrative system is compared to that of other assumed ELEVATION-marking languages in the area: Dime and Hamar (South Omotic), Dawuro, Gamo, Yemsa, Maale and Bench (North Omotic), Gawwada, Gidole and Konso (Lowland East Cushitic). I review a hypothesis brought forward in the literature (e.g. in Tosco 2012, 2013, Mulugeta 2008), which states that ELEVATION-marking is a common feature of the Omotic-Lowland East Cushitic contact zone. A review of all available data shows that the demonstrative systems in the area are very different and not readily comparable. Only few languages encode ELEVATION as consistently in their system of demonstrative determiners/pronouns as Baskeet, but it seems that the expression of elevation is predominately expressed in demonstrative adverbs.

Acknowledgements
Research on Baskeet has been sponsored by the Endangered Language Documentation Programme (Individual Postdoctoral Fellowship 2011–2013: "Documentation of Baskeet song, verbal art and ceremonial language").

References

Mulugeta Seyoum (2008), *A grammar of Dime*, Utrecht: LOT.

Tosco, Mauro (2012), The grammar of space of Gawwada, in: M. Brenzinger & A.-M. Fehn (eds), *Proceedings of the 6th World Congress of African Linguistics, Cologne, 17-21 August 2009*, Cologne: Köppe, 523-532.

Tosco, Mauro (2013), Le relief dans la tête : un système à repérage absolu de l'Ethiopie méridionale, *Faits de Langues* 42 (*Sémantique des relations spatiales*), 153-166.

ꕥ

Elevational deixis in Pacaraos Quechua

Willem Adelaar

Keywords: Quechua, deixis, demonstratives, verbal morphology, elevation marking

Quechuan languages have a straightforward nominal subclass of demonstratives featuring either two or three deictic categories: PROXIMATE /NON-PROXIMATE or PROXIMATE / INTERMEDIATE / DISTAL. Pacaraos Quechua, a variety spoken in a local Andean community in Central Peru until the 1980s, is particular in that it features a more complex system of six demonstratives consisting of a PROXIMATE, two INTERMEDIATES (possibly anaphoric and non-anaphoric), and three DISTALS. The elevational distinctions, HIGH, LOW and LEVEL, are limited to the DISTAL dimension. Whereas the LOW DISTAL demonstrative is linked to the DISTAL category in several other Quechuan languages (Ancash Quechua, Cuzco Quechua, etc.) the HIGH and LEVEL DISTALS of Pacaraos Quechua are expressed in formally unique ways, and their etymological background remains unexplained.

From a semantic perspective both the deictic categories and their elevational extensions are furthermore partly matched by derivational elements proper to the verbal base. The HIGH / LOW distinction can be discerned in a set of DIRECTIONAL verbal markers (UPWARDS / DOWNWARDS), whereas the PROXIMATE / NON-PROXIMATE distinction is reflected in a verbal affix that indicates motion or orientation towards the speaker. From a formal point of view, however, there is no relation at all between these verbal extensions and the nominal demonstratives, and no indications of grammatical agreement have been attested so far.

ꕥ

Elevational and riverine orientation system in Alaska, and their consequence for understanding landscape

Gary Holton

<pdf>

ꕥ

The origin and development of spatial coordinate systems in South Halmahera-West New Guinea

Laura Arnold

Keywords: Spatial deixis, spatial coordinate systems, Austronesian, language contact, grammaticalization

South Halmahera-West New Guinea (SHWNG), a little-known subbranch of Austronesian and the putative sister of Oceanic, comprises approximately 40 languages, spoken in and to the west of New Guinea (Kamholz 2014). At least ten SHWNG languages have spatial coordinate systems of spatial deixis, in which deictic forms encode angular search domains (e.g. UP, DOWN, LAND, SEA; Burenhult 2008). Holton (2017) is an in-depth look at spatial coordinate systems in the SHWNG languages spoken in the Halmahera archipelago. This presentation will expand this work, by proposing a first typology of spatial coordinate systems across the SHWNG subbranch, and discussing the historical development of these systems.

A typology of spatial coordinate systems in SHWNG is given below. This typology is based on the number and type of categories distinguished, and the relationship of the spatial coordinate system to other categories of spatial deixis.

1. UP, DOWN (Waropen, Wooi). Type 1 systems distinguish elevational categories, which are paradigmatic with other, non-elevational categories of spatial deixis. Thus, in Wooi, the non-elevational deictics (*ne* 'PROX', *vaw* 'NEUT', *pa* 'DIST') are paradigmatic with the elevational *pe* 'UP' (Sawaki 2017).
2. UP, DOWN, LAND, SEA, ACROSS (Taba, Buli, Gane, Patani). In Type 2 systems, spatial coordinate categories operate in a separate grammatical subsystem from other categories of spatial deixis. For example, in Taba, spatial coordinate terms (e.g. *ya* 'UP', *po* 'DOWN') occur with affixes marking motion and position (e.g. *ak-* 'ALL', *-ma* 'VEN'); non-spatial coordinate demonstratives, however, do not enter into these paradigms (Bowden 2001).
3. UP, DOWN, LAND, SEA, IN, OUT, etc (Biak, Maʿya, Ambel, Sawai). Languages with Type 3 systems have at least six spatial coordinate categories, which operate in a system syntagmatic to, but dependent upon, other categories of spatial deixis. In Ambel, for example, spatial coordinate prefixes (e.g. *i-* 'UP', *pu-* 'DOWN') only occur when they attach to any of the four non-spatial coordinate demonstrative roots (e.g. *ne* 'PROX', *pa* 'MID'; Arnold 2018).

These three types are not cognate, but have emerged independently, subsequent to the break-up of proto-SHWNG – shown, for example, by the diversity in the categories distinguished, the status in the grammar, and the non-cognacy of forms. Comparison with local Papuan languages suggests that the systems have developed in SHWNG as the result of contact: with the Papuan languages spoken on the east of the Bird's Head Peninsula in the case of Type 1; with languages from the North Halmahera family in the case of Type 2 (cf. Holton 2017); and, for Type 3, with Tidore, which was historically influential throughout the region.

In the case of Type 2 and 3 systems, there is a transparent relationship between grammatical and lexical spatial terms, suggesting grammaticalization occurred comparatively recently: for example, in Ambel, the prefixes *i-* 'UP' and *pu-* 'DOWN' have developed from the directional nouns *il* 'upwards' and *pul* 'downwards'. This presentation will therefore additionally have implications for theories of contact-induced grammaticalization, for example the type of lexical material that speakers draw on in such contexts (see Heine & Kuteva 2003).

Abbreviations

ALL	allative	PROX	proximal demonstrative
DIST	distal demonstrative	UP	upwards elevational
DOWN	downwards elevational	VEN	venitive
NEUT	neutral-distance demonstrative		

References

Arnold, Laura (2018), *A grammar of Ambel: An Austronesian language of Raja Ampat, west New Guinea*. PhD thesis, University of Edinburgh.

Bowden, John (2001), *Taba: Description of a South Halmahera language*. Canberra: Pacific Linguistics.

Burenhult, Niclas (2008), Spatial coordinate systems in demonstrative meaning, *Linguistic Typology* 12: 99-142.

Heine, Bernd & Tania Kuteva (2003), On contact-induced grammaticalization, *Studies in Language* 27: 529-572.

Holton, Gary (2017), A unified system of spatial orientation in the Austronesian and non-Austronesian languages of Halmahera, *NUSA* 62: 157-89.

Kamholz, David (2014), *Austronesians in Papua: Diversification and change in South Halmahera-West New Guinea*. PhD thesis, University of California, Berkeley.

Sawaki, Yusuf (2017), *A grammar of Wooi: An Austronesian language of Yapen Island, Western New Guinea*. PhD thesis, Australia National University.

Deictic prefixes in Chechen

Zarina Molochieva

The paper deals with the deictic prefixes in Chechen (Nakh, Nakh-Daghestanian). The deictic prefixes have a complex semantics and has been not yet studied in this language. There are four deictic prefixes in Chechen, which are responsible for horizontal and vertical deixis. The primary meaning of which is as follows:

ħal- up/upwards
voħ- down/downwards
dʕa- away/there
ħa- towards/here

The preverbal prefixes are used to express the spatial deixis. They show the location and the direction of the speaker relatively to the addressee. However, they can be used with almost all kind of verbs and very frequently used in the everyday conversations, narration, and written texts. The prefixes, if they are used with the motion verbs express the goal of the motion, which may be also based on aspects of geography, for instance, *ħal-* 'up/upwards' is used to express the motion towards the goal in the mountains (hills, etc.), or *voħ-* 'down/downwards' is used for the goal downhill. The verbs indicating motion, for instance, such as open *della* may take all four deictic prefixes. In such cases the motion is based to the direction of the person opening the window/door, for instance, if the windows open in direction to the person opening the window the prefix *ħa-* is used, otherwise the other prefixes are used depending on motion of the window and location of the person.

However, with the verbs of transfer such as give *dala*, send *daxita*, bring *daa*, etc. the semantics are more complex than with the verbs of motion. In such cases the horizontal deixis is person-based, whereas the vertical deixis can be based on the geography.

With other verbs these prefixes are also not tied to the specific aspects of the geography. For instance, they can have metaphorical meaning. *ħal-* 'up/upwards' is used for the place of honor for the guest in a house, which is usually on the opposite side than the entrance door of the house/room.

In addition, in certain cases, deictic prefix *ħal-* 'up/upwards' can enforce a telic interpretation of an otherwise atelic verbs, for instance, *eating a bit of the food* vs. *eating all the food*, however, not all

verbs allow such interpretation. For some verbs the other prefixes must be used, for example, for the verb *dila* 'wash' must be used the prefix *voħ-* 'down/downwards', otherwise it is ungrammatical (1).

(1) kyigaʃ **ħal*/ voħ-d-ila.
hands up-down-D-wash.IMP
'Wash your hands.'

In this paper I discuss the properties of geographically and non-geographically based semantics of the deictic prefixes based on the corpus. The verbs will be also classified into the groups: a) verbs using all four deictic prefixes; b) verbs using either only horizontal deixis or vertical deixis; and c) verbs using The semantics of other local prefixes: *dyħal-* 'against', *juxie-* 'close to, near', *irx-* 'upwards', *t'ieħa-* 'behind' will be also shortly discussed in the paper.

WORKSHOP 5

Schedule: We 11.00-17.55 (Room 10)

‘External’ agreement with unexpected targets

Marina Chumakina, Oliver Bond, & Steven Kaye
(Surrey Morphology Group, University of Surrey)

Key words: agreement, clause structure, locality, syntactic relations, typology

Across the world’s languages, agreement is generally limited to relations between a verb and its arguments (clausal agreement) or a noun and its dependents (nominal agreement), and it usually occurs between elements belonging to particular parts of speech within the boundaries of established syntactic constituents.

We focus on a radically different type of agreement, where the relation between the controller and the target is typologically and theoretically unexpected. This is typically seen when the controller and target of agreement are in a ‘non-local’ syntactic relation.

Examples have been registered in abundance in one linguistic family, Nakh-Daghestanian, and have been sporadically reported for other languages of the world (Antrim 1991, Fábregas and Pérez-Jiménez 2008, Ledgeway 2011 among others). Consider example (1), from the Nakh-Daghestanian language Avar, where the postposition *žaniw* ‘inside’ has the neuter noun *tusnaq’* ‘prison’ as its complement, yet the agreement on the postposition is controlled by one of the verb’s arguments, the object Rasul:

(1) tusnaq-al-da žani-**w** t’amuna niže-c:a Rasul
prison(N)-SG.OBL-SUP in-M.SG put.PST 1PLEXCL-ERG Rasul(M)[SG.ABS]
‘We put Rasul in prison.’

The agreement represented in (1) is striking: first of all, it is an unusual part of speech that shows agreement. Secondly, the agreement happens with an unexpected controller: not the complement of the postposition, but the object of the predicate. The target and controller in (1) do not form a local domain either in terms of strict locality (sisterhood) or high locality (such as verb-argument relations), as defined in Alexiadou et al. (2013: 3-4).

Languages of the Andic group present an even more striking example: clause-level agreement between nouns, as illustrated by Andi. In (2) and (3), the noun in the affective case (‘mother’), which encodes the experiencer argument of the verb *haɢo* ‘see’, has a morphological slot for agreement and agrees in gender and number with the absolutive argument of the clause:

(2) ilu-**b**-o q’inkom haɢo
mother(II)-III.SG-AFF bull(III)[SG.ABS] see.AOR
‘Mother saw a bull.’

(3) ilu-**r**-o c’ul haɢo
mother(II)-V.SG-AFF stick(V)[SG.ABS] see.AOR
‘Mother saw a stick.’

These examples demonstrate that, first, agreement affects a much wider range of grammatical elements than previously thought and, second, agreement can ‘sidestep’ syntactic relations, contrary to what has been generally assumed. To capture the latter property of this phenomenon, we call it

'external agreement'. This is, however, just a shorthand for a more accurate description: 'agreement of non-verbal targets outside their minimal syntactic phrase, yet within the clause'.

The aim of this workshop is to expand our understanding of how agreement works by investigating agreement phenomena such as arguments agreeing with other clause-level arguments, adpositions agreeing outside the adpositional phrase, and agreeing adverbs and discourse particles, as represented in the world's languages. Our overarching research question is: **What does external agreement tell us about the structure of human language?** Answering this question involves considering the following:

- What types of controllers and targets are involved in external agreement?
- What are the structural constraints on non-local agreement?
- What are the morphosyntactic properties of external agreement?
- How does external agreement develop?

References

Alexiadou, Artemis, Tibor Kiss and Gereon Müller (eds) (2013). *Local Modelling of Non-Local Dependencies in Syntax.* Series: Linguistische Arbeiten 547. Berlin: De Gruyter.

Antrim, Nancy Mae (1991). 'Italian adverbial agreement', in Michael L. Mazzola (ed.), *Issues and Theory in Romance Linguistics: Selected Papers from the Linguistic Symposium on Romance Languages XXIII, April 1–4, 1993.* Washington D.C.: Georgetown University Press, 129-40.

Baker, Mark (2014). 'On dependent ergative case (in Shipibo) and its derivation by phase', *Linguistic Inquiry* 4/3: 341-79.

Bond, Oliver, Greville G. Corbett, Marina Chumakina and Dunstan Brown (eds) (2016). *Archi: Complexities of agreement in cross-theoretical perspective*, Oxford: Oxford University Press.

Carstens, Vicki, and Michael Diercks (2013). 'Agreeing how? Implications for theories of agreement and locality', *Linguistic Inquiry* 44/2: 179-237.

Corver, Norbert (2007). 'Dutch *'s*-prolepsis as a copying phenomenon', in Norbert Corver and Jairo Nunes (eds), *Copy Theory of Movement*. Amsterdam: John Benjamins, 175-215.

Csirmaz, Aniko (2008). 'The accusative case and aspect', in Katalin É. Kiss (ed.), *Event Structure and the Left Periphery: Studies on Hungarian.* Dordrecht: Springer, 159-200.

Fábregas, Antonio, and Isabel Pérez-Jiménez (2008). 'Gender agreement on adverbs in Spanish', *Journal of Portuguese Linguistics* 7: 25-45.

Haug, Dag T.T. and Tatiana Nikitina (2016). 'Feature sharing in agreement', *Natural Language and Linguistic Theory* 34: 865-910.

Ledgeway, Adam (2011). 'Adverb agreement and split intransitivity: Evidence from southern Italy', *Archivio glottologico italiano* 96: 31-66.

Souag, Lameen (2015). 'How to make a comitative preposition agree it-with its external argument: Songhay and the typology of conjunction and agreement', in Jürg Fleischer, Elisabeth Rieken and Paul Widmer (eds), *Agreement from a diachronic perspective*. Berlin: De Gruyter, 75-100.

Nominals as agreement targets in Andi (Nakh-Daghestanian): 'Agreement between arguments' revisited

Steven Kaye
(University of Surrey)

Keywords: agreement; argument structure; subjecthood; morphosyntax; Nakh-Daghestanian

A striking type of 'external' agreement is represented in several languages of the Nakh-Daghestanian family, whereby nominals in certain case forms are required to display an agreement marker sensitive to the features of another nominal. Of particular syntactic interest is the situation found in some Andi varieties, where verbs of perception and cognition such as *see*, *know*, *find* do not behave as normal (ergatively-aligned) transitives but require so-called 'affective' case on the experiencer and absolutive case on the stimulus. Nominals in the affective case must agree in gender and number with the absolutive argument (note that whether the verb itself also shows agreement is a purely lexical matter):

(1) ilu-r-o **c'ul** haɢo
mother(II).SG.OBL-V.SG-AFF stick(V)[SG.ABS] see.AOR
'Mother saw a **stick**.' (Salimov 2010: 105)

(2) hege-lu-b-o b-ison-d:u se-b k:ub-ʟi **torba**
DEM-I.PL.OBL-IV.SG-AFFIV.SG-find-PRF one-IV.SG iron(IV).SG.OBL-GEN
pipe(IV)[SG.ABS]
'They found an iron **pipe**.' (*The pitiless father*, Alisultanova 2010)

Agreement **controllers** are in bold and relevant targets are underlined; genders are glossed by Roman numerals.

The agreement phenomenon seen in (1-2) has been described for different Andi dialects by Sulejmanov (1957) and Salimov (2010) and presented alongside other Nakh-Daghestanian external agreement phenomena by Chumakina & Kaye (2018). However, within all major syntactic frameworks the existence of agreement between the arguments of a predicate is unexpected or predicted not to be possible, and accordingly attempts have been made to demonstrate that apparent instances of this behaviour do not genuinely exemplify it (cf. Polinsky et al. 2017).

Building on these observations, I examine the behavioural properties of agreeing nominals in Andi to elucidate the syntactic status of the items expressing the experiencer and the stimulus in examples such as (1-2).

In other Nakh-Daghestanian languages possessing a set of syntactically divergent experiencer verbs, both of the nominals involved are treated as arguments in the full sense. A particularly careful analysis of this area has been carried out for Tsez, a language belonging to the same (Avar-Andic-Tsezic) subgroup of the family as Andi itself. On syntactic and morphological grounds, Polinsky (2015) divides Tsez experiencer verbs into two categories. Each type assigns a different grammatical function to the experiencer, but in either case she considers both nominals to be arguments of the predicate.

On the basis of similar structural tests, I claim that the situation in Andi is ultimately no different: that is, there are no substantial grounds for denying argumenthood to at least some Andi nominals expressing experiencer roles. If the possibility of agreement between arguments is treated as an empirical question rather than being ruled out as a matter of principle, sentences such as (1-2) on the face of it provide evidence that it can indeed be found. The Andi facts thus pose a non-trivial problem for syntactic theories which disallow this configuration.

References

Alisultanova, M. A. (2010). *Dunjalla baχunnirs:i χalɢilʟol ɢwanab mic:'illas:ol muχal [Tales of the world's peoples in the Andi language]*. Makhachkala: Nurul' Irshad.

Chumakina, M. & S. Kaye (2018). 'Multiple agreement targets in Nakh-Daghestanian languages.' Poster presented at *MAD, Multiple Agreement across Domains*, Berlin, 8-9 November 2018.

Polinsky, M. (2015). *Tsez syntax: a description* (draft, February 2015). Manuscript; available online at https://scholar.harvard.edu/mpolinsky/publications/tsez-syntax-description-manuscript

Polinsky, M., N. Radkevich & M. Chumakina (2017). 'Agreement between arguments? Not really.' In: Roberta D'Alessandro, Irene Franco & Ángel J. Gallego (eds.), *The verbal domain* [Oxford Studies in Theoretical Linguistics 64]. Oxford: Oxford University Press, 49-84.

Salimov, Kh. (2010). *Gagatlinskij govor andijskogo jazyka [The Gagatl variety of the Andi language]*. Makhachkala: DNC RAN.

Sulejmanov, Ja. G. (1957). *Grammatičeskij očerk andijskogo jazyka (po dannym govora g. Rikvani) [Grammatical sketch of the Andi language (on the basis of data from the Rikwani variety)]*. PhD thesis, Institute of Linguistics of the USSR Academy of Sciences.

Clausal agreement in Zilo Andi: Number

Aigul Zakirova
(Independent researcher, Moscow)

Keywords: East Caucasian, syntax, number, agreement, adverb

This talk aims to give an account of clausal number agreement in Zilo Andi (NE Caucasian, Andic) in terms of its possible targets, which besides verbs include such unexpected targets as adverbs and postpositions.

The Zilo dialect of Andi exhibits gender and number agreement both in DPs and in the clausal domain. In DPs modifiers agree with the head noun. Gender markers are presented in Table 1 (there are five genders in total)

Table 1. Gender exponents in Zilo Andi.

	M	F	AN	INAN1	INAN2
SG	w	j	b	b	r
PL	w	j	j	b	r

Number agreement is realized by a suffix *-l* attaching to the plural stem. Whereas the presence of a gender slot is lexically determined, the suffix *-l* allows for more variation: it is obligatory with some targets whereas for others it is optional or impossible:

-l is obligatory on imperative forms of intransitives and on some adverbs, as (1) shows:

(1) *hege-w-ul* *ʃu-w-ul /* **ʃu-w* *helli-r*
DEM-M-PL good-M-PL good-M run-PROG
'They run well'.

-l is optional on some other low adverbs (2) and postpositions (3); and on some verb forms not discussed here.

(2) *hege-w-ul* *berq'o-l/* *berq'a* *w-atʃ':-i-r*
DEM-M-PL early.OBL-PL early M-rise_up-PST-PROG
'They wake up early'.

(3) *iʃːil* *reło-tʃʼu* *tłeru* / *tłeru-l* *w-aχ-u-mallo*
we.EXCL sea-AD2 close close-PL M-live-PST-PROG
'We live close to the sea'.

In her analysis of Archi, Polinsky (2016) proposes that only vP adverbs display agreement. Likewise, in Zilo Andi high adverbs never attach the suffix *-l*. However, high adverbs are scarce and there may be morphological reasons for their not taking the plural suffix. Thus, for example, *taleʕeldi* 'fortunately' is still analyzable as an ergative case form.

Adverbs: possible analysis

Polinsky (2016) proposes four possible analyses of agreeing adverbs in Archi:

a. Agreeing adverbs represent dedicated functional heads in vP.
b. Agreeing adverbs are adjuncts with unvalued phi-features adjoined to vP.
c. Agreeing adverbs are adjuncts that undergo head movement to v.
d. Agreeing adverbs are modifiers of an absolutive DP.

The analysis in (c) implies head movement of adverbs to v. Rudnev (2018) shows that no such movement happens to Avar postpositions, the evidence being the absence of head movement constraint (Travis 1984). Head movement constraint is also absent in Zilo Andi, from which I conclude that adverbs and postpositions do not undergo head movement to v.

(4) *iwa-l* *hege-w-ul* *wo-in-ne?*
where_to.H-PL DEM-M-PL go-NPST-HAB
'Where do they go?'

The analysis in (d) implies that the adverb will modify the absolutive argument. Whereas this might work with verbs of creation such as 'build' or 'cook', the interpretation of adverbs with verbs like 'know' is obviously different:

(5) *di-j-o* *hege-j* *ʃu-j* *ts'inn-e,* *amma* *hege-j* *ʃu-j-sːu*
I.OBL-F-AFF DEM-F good-F know-HAB but DEM-F good-F-NEG
'I know her well, she is not good'.

In my talk I will elaborate on the analysis of plural agreement of adverbs in Zilo Andi.

This research was conducted as part of the project "The Andi morphosyntax in a typological perspective" supported by the Russian Foundation for Basic Research (RFBR), grant No. 18-012-00852a.

All data is drawn from my fieldwork (2017-2018) in the village of Zilo, Botlikh district, Daghestan, Russia

References

Polinsky, Maria (2016), Agreement in Archi from a minimalist perspective, in O. Bond et. al. (eds), (2016), *Archi: Complexities of Agreement in a Cross-Theoretical Perspective*, Oxford University Press, 184-232.

Rudnev, Pavel (2018), Agreeing adpositions in Avar and the direction of valuation. NRU HSE. Manuscript.

Travis, Lisa DeMena (1984), Parameters and effects of word order variation. Massachusetts Institute of Technology dissertation.

Ordinal numerals and animacy in Botlikh

Samira Verhees
(Linguistic Convergence Laboratory NRU HSE Moscow)

Keywords: agreement, animacy, ordinals, East Caucasian languages, Botlikh

East Caucasian languages are known for gender (also ‘noun class’) marking on unusual targets, which include postpositions, adverbs and inflectional endings (alongside less exotic targets like verbs and adjectives). Botlikh (a language of the Avar-Andic branch of East Caucasian) features a special set of animacy markers alongside the common East Caucasian noun class system. These animacy markers constitute an independent system and show up as part of negative copulas, interrogative clitics, ordinal numeral suffixes, attributivizing clitics and participle suffixes. Animacy also shows up as a parameter in the noun class system. There are three classes in singular: masculine human, feminine human, and a third residual class for everything else, as is rather common in East Caucasian. In plural, most languages distinguish human and non-human, whereas cognate markers in Botlikh distinguish animate vs. inanimate.

The existence of dedicated animacy markers is mentioned in available descriptions of Botlikh (see, e.g., Gudava (1967)), but little is known about their role in agreement and the extent of their use. This paper focuses on one type of target, namely ordinal numerals. Ordinal numerals are formed by attaching a suffix *-ɬa-CM // -ɬo-CM* for animate referents, or *-χo-CM* for inanimates. Note that while the choice of the suffix expresses agreement with animacy, it also contains a slot for noun class agreement. Any ordinal is therefore marked for two types of agreement. Agreement is typically controlled by the nominal head which is modified by the ordinal. This concerns the choice of the ordinal suffix itself, as well as the noun class marker.

(1) hac’a.j-**χo-b** ziw
ten-INAN.ORD-N day
‘the tenth day’

(2) buʁu.j-**ɬo-w** waša
four-AN.ORD-M son
‘the fourth son’

In a clause, agreement can be controlled by the head of the noun it modifies (3) or the absolutive argument of the main verb, similar to postpositions in East Caucasian (4):

(3) ješi j-aʔ-a ida k’e.ji-**χo-b** kalasa-ɬi
girl(F)[SG.ABS] F-go-CVB COP two-INAN.ORD-N grade(N)-IN
‘[Our] daughter is in the second grade.’

(4) ištu.j-**ɬa-b** kalasa-ɬi hiƛ’-i w-eχ̣-u
five-AN.ORD-N grade(n)-IN down-SUP M-stay-AOR
‘He stayed in the fifth grade.’ (i.e. he had to repeat the fifth grade for a second year)

In example (3), both the ordinal suffix and the class marker agree with the nominal head ‘grade’. In example (4), the noun class marker *-b* agrees with the nominal head ‘grade’, while the ordinal suffix itself agrees with the covert subject / absolutive argument of ‘stay’, a male human as is clear from the verb agreement. Elicitation of constructions with ordinals with different native speakers showed that

the choice of controller for the ordinal suffix does not depend on clause type or other morphosyntactic parameters, although variation is possible only in adverbial phrases (e.g. 'in the fifth grade').

The examples presented above are from the Botlikh dictionary (Saidova & Abusov 2012).

The paper was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

References

Gudava, Togo E. 1967. 'Botlixskij jazyk [Botlikh]'. Jevgenij A. Bokarev, Ketevan Lomtatidze (eds.). *Jazyki narodov SSSR. Vol. 4: Iberijsko-kavkazskie jazyki*, 293−306. Moscow: Nauka.

Saidova, Patimat A. & Magomed G. Abusov. 2012. *Botlixsko-russkij slovar'* [Botlikh-Russian dictionary]. Makhachkala: IJaLI.

'External' agreement in Ripano. A corpus-based account

Tania Paciaroni

<pdf>

Agreeing adverbs in Forest Enets

Marina Chumakina and Olesya Khanina

In this paper we discuss the data from Forest Enets (Uralic), which show that grammatical agreement is possible between an argument and an adverb. Agreeing adverbs in Enets form a closed sub-class. It is possible to trace some of them back to the frozen case forms of nouns and adjectives, but they have lost all grammatical properties of the nominals, and the etymology is not always transparent.

All agreeing adverbs are subject-oriented and agree with a subject argument. In (1) the adverb *barimagu-nʲiʔ* 'barely' agrees in person/number with the covert subject 'we', while in (2) the agreement is with the subject of the passive:

(1)	barimagu-nʲiʔ	dʲipra-jʔ	ɔdu	mi-n
	barely-OBL.SG.1DU	untie(PFV)-1DU.SG	boat	into-LOC
	'We untangled it in the boat with difficulty.'			

(2)	**minxu-da**	stada-d	ɛzta-ra-e-zʔ
	right_away-OBL.SG.3SG	herd-DAT.SG	send(PFV)-PASS-M-3SG.M
	'He was sent to the herd at once.'		

The agreement behaviour of the adverbs do not depend on the agreement of the verb: in the non-finite clauses the adverbs retain the possibility of agreement independently of whether the verbal form agrees or not. Thus, in (3) the imperfective converb *kuraxad* 'pursuing' does not agree whereas conditional form *kajibuʔduʔ* 'if they remain' in (4) does; whereas the adverb *tɔminidʲiʔ* 'just' agrees with the clause subject in both examples:

(3) tʃike ʃize kasa ɛdʲuku-r tɔminidʲiʔ tɔ
this two man child-NOM.SG.2SG just-OBL.SG.3DU that
ɛɛdʲiʔ-nɔlʲkusʃ kuraxad
mother-OBL.SG.3DU pursue(IPFV)-CVB
'While these two boys were running after their mother….'

(4) tɔmini-duʔ kaji-buʔ-duʔ kɔnʲeʃnɔ sɔjza
just-OBL.SG.3PLremain(PFV)-CVB.COND-OBL.SG.3PL of_course good
'If they just stay as they are, that's good, of course.'

This behaviour contrasts Forest Enets against Tundra Nenets, a related language, where adverbs can only host agreement in the presence of agreement on the verb. In the situation where the verbs in Nenets do not agree, the adverbs cannot agree either, which suggests that adverbial agreement functions as some kind of concord. Thus, agreeing adverbs are impossible in dependent clauses headed by unchangeable converbs.

Some Enets adverbs allow variation in agreement; compare (5) and (6) where the same adverb *tɔnane* 'once' is used in the agreeing (5) or non-agreeing (6) form:

(5) tɔnane-da igarka-xon dʲiri-j entʃeu ɛ-bi-Ø
then-OBL.SG.3SG Igarka-LOC.SG live(IPFV)-PTCP.ANT person be(IPFV)-PRF[3SG.S]
'Formerly this person lived in Igarka'.

(6) ʃeeaʔ man-dkoda tɔnane
who-NOM.SG.1PL say(PFV)-HYPOT.3SG.S once
'Who would say (it) to us then?'

Since the adverbs in Forest Enets are not associated with a particular position, the configurational account of these data is unlikely.

In this paper, we will discuss factors influencing the choice of agreement strategy and present a contrastive analysis of agreeing adverbs behaviour in Nenets and Enets, with some comparison from another Samoyedic language, Nganasan. We will also be discussing the possible historic origins of adverbial agreement in these languages.

Agreement of essive adverbials in Tanti Dargwa

Nina Sumbatova

<pdf>

What Else Can Rally Case-Shift in Kartvelian? A Case Study on Megrelian Adverbs

Alexander Rostovtsev-Popiel

Keywords: inflecting adverbs, case marking alignment, external agreement, affective verbs, Megrelian

Kartvelian languages are known for the TAME Series-based case-shift in transitives and unergatives (Classes I and III, respectively), whereby the case-marking alignment of core arguments depends on the TAME that the verb predicate of the clause belongs to.

In Megrelian, unlike Georgian, case-shift also involves Class II verbs and, partially, Class IV inverted statives. The overall picture can be summarized in the following table.

Table 1. Case marking alignment throughout the TAME groups, aka Series

	PRS-FUT Series	AOR Series	PERF Series	EVID Series
Class I	NOM–DAT(–DAT)	ERG–NOM(–DAT)	DAT–NOM(–ALL)	NOM–DAT(–DAT)
Class II	NOM(–DAT)	ERG(–DAT)	DAT (–DAT/ALL)	NOM(–DAT)
Class III	NOM(–DAT)	ERG(–NOM)	DAT(–NOM)	NOM(–DAT)
Class IV	DAT(–NOM)	**DAT(–ERG)**	**--**	DAT(–NOM)

In focus of the present talk are adverbs of degree and measure that surprisingly appear to be sensitive to case-shift in the AOR Series of affective verbs and react to it similarly to the 2nd argument (which has to be coded by the ERG), thus representing a specific instance of the so-called "external agreement".

(1) *osurc xolo keʔoropə te ndemk ʒalamk.*
osur-s xolo k[o]-Ø-e-ʔor-op-u te
woman-DAT ADD AFF-IO_3-VER_R-love-AUG-S_3SG.PST DEM.PROX
ndem-k ʒalam-**k**
daeva-ERG very.much-**ERG**
'The woman, in her turn, fell **greatly** in love with this daeva'. (Xubua 1937: 96.23)

(2) *mara gamocdilebak dak'ordu č'ič'ek muhambis.*
mara gamocdileba-k do-Ø-o-k'or-d-u
but experience-ERG PRV-IO_3-VER_L-lack-INCH-S_3SG.PST
č'ič'e-**k**muhambi-s
a.bit-**ERG** Muhambi-DAT
'But Muhambi came short of experience **a little bit**'. (Xubua 1937: 31.18)

In addition to the ergative encoding of the stimulus in these constructions, the verb in the AOR Series triggers the appearance of the ERG marker *-k* on the adverb (which occurs neither in other TAME Series in Megrelian nor elsewhere in Kartvelian), cf. *ʒalam-**k*** and *č'ič'e-**k*** in the above examples (on these phenomena, see Uridia 1960, K'ač'arava 1969, and K'iziria 1978, in Georgian, and Harris 1991: 375, in English).

On the other hand, in Series I that features the experiencer in the dative and the stimulus in the nominative case, the situation is exactly the opposite: it is the experiencer, whose case-marking is mirrored by the adverb, not the stimulus:

(3) *ate baɣana ʒalams uʔorc te vezirc.*
ate baɣana-Ø ʒalam-**s** Ø-u-ʔor-s
DEM.PROX child-NOM very.much-**DAT** IO_3-VER_O-love-S_3SG
te vezir-s
DEM.PROX vizier-DAT
'This vizier loves this child **very much**'. (Xubua 1937: 87.28)

This talk aims to provide a comprehensive account of the syntactic properties of Megrelian adverbs of manner, which include sensitivity to case-marking alignment in affective constructions, and explain what factors determine their inflection-like behavior.

Abbreviations

3 – 3rd person; ADD – additive; AFF – affirmative; ALL – allative; AOR – aorist; AUG – augment; DAT – dative; DEM.PROX – proximal demonstrative; ERG – ergative; EVID – evidential; FUT – future; INCH – inchoative; IO – indirect object; NOM – nominative; PERF – perfect; PRS – present; PRV – preverb; PST – past; S – subject person marker; SG – singular; VER_L – locative versionizer; VER_O – objective versionizer; VER_R – relative versionizer.

References

Harris, Alice (1991), Mingrelian, in Harris, Alice (ed.), (1991), *The Indigenous Languages of the Caucasus, vol. 1, The Kartvelian Languages*, Delmar/New York: Caravan Books, 313–394.

K'iziria, Ant'on (1978), Vitarebis garemoebis gamoxat'vis sašualebani kartvelur enebši [Possibilities of the Expression of Adverbial Modifiers of Manner in the Kartvelian Languages], *Iberiul-k'avk'asiuri enatmecniereba* XX, 37–60.

K'ač'arava, Giorgi (1969), Ergat'ivis erti ucnobi punkciisatvis zanurši [On an Unfamiliar Function of the Ergative in Zan], *Bulletin of the Georgian Academy of Sciences* 53 (2), 505–507.

Uridia, Otar (1960), Megrulis sint'aksuri taviseburebani kartultan mimartebit [Syntactic Specificities of Megrelian in Comparison with Georgian], *Works of the Stalin Tbilisi State University* 93, 167–178.

Xubua, Mak'ar (1937), *Megruli t'ekst'ebi* [Megrelian Texts], T'pilisi: Academy of Sciences.

WORKSHOP 6

Schedule: We 11.00 – Thu 12.25 (Room 5)

Extravagant Morphology

Matthias Eitelmann & Dagmar Haumann
(University of Mainz & University of Bergen)

Keywords: morphology, morphosyntax, language variation, language change, expressiveness

The workshop explores extravagance in morphological processes and investigates its impact on linguistic variation and change. The maxim of extravagance, as coined by Haspelmath (1999: 1055) in his account of the unidirectionality typical of grammaticalization processes, essentially picks up one of Keller's dynamic maxims, namely "Talk in such a way that you are noticed", which is active in the workings of the 'invisible hand' in language change (1994: 101). Thus, speakers, in their intent to be "socially successful with their speech" (Haspelmath 1999: 1057f.), may not only be particularly expressive but deviate noticeably and considerably from established language norms by using an expression in an innovative sense, in an "imaginative and vivid" way (ibid.), or any other clearly attention-attracting fashion. Extravagance may therefore be regarded as a trigger for language variation and change, with such ostensibly deviating and non-conforming language use ultimately initiating the formation of new patterns. It also shows considerable overlap with the notion of linguistic creativity, i.e. "the native speaker's ability to extend the language system in a motivated, but unpredictable (non-rule-governed) way" (Bauer 1983: 63). At the same time, extravagance goes beyond creativity in that it is conceptualized as an integral part of language change processes.

Against this backdrop, the workshop seeks to re-address the notion of extravagance in an attempt to operationalize the concept to a larger degree, investigate extravagant phenomena empirically and shed further theoretical light on the role of extravagance in language variation and change. To a certain extent, extravagance shares features of what Zwicky & Pullum (1983) have referred to as 'expressive morphology', a notion that predominantly refers to ad-hoc formations with characteristics such as a "specific pragmatic effect", "imperfect control" (i.e. speakers lacking the ability to apply the word-formation pattern productively) or the restricted occurrence of such expressive formations in particular syntactic contexts (Zwicky & Pullum 1983: 335-8). However, in contrast to expressive morphology, our re-conceptualization of extravagance seeks to go beyond the individual speaker's innovative language use and takes into account the eventual effect that innovations have on the language system and its structures, i.e. by discussing the concomitant effect of extravagance on linguistic variation and change and its theoretical ramifications.

Understood literally in the sense of the underlying Latin etymon *extra-vagans* 'wandering outside, out of bounds', extravagance in the present context refers to morphological phenomena that display divergent tendencies, with a specific interest in the following:

- word-formation processes that straddle boundaries and turn extravagant in that innovative formations violate alleged or actual constraints, such as the pervasive spread of affixes, as witnessed by
  - *-ish* developing from a derivational suffix deriving adjectives from an ever growing set of bases into a 'detachable' degree operator surfacing as a phrasal affix/clitic or as a free lexical item (e.g. *British, awkwardish, body-spray-ish, light-at-the-end-of-the-tunnel-ish, everybody-ish, Ish!*, see

e.g. Kuzmack 2007, Traugott & Trousdale 2013, Oltra-Massuet 2017, Eitelmann, Haugland & Haumann 2018a, b)
    - *-ly* undergoing secondary grammaticalization from a derivational suffix deriving manner adverbs form adjectives to a generalized 'adverbializer' deriving various subclasses of adverbs (e.g. epistemic *possibly*, illocutionary *briefly*, or subject-oriented *willingly*) from a variety of bases (e.g. *purposely, seldomly, shiveringly, wetly*; see Nevalainen 1994, Killie 1998, 2007)
- phenomena situated at the interface between morphology and syntax or morphology and semantics/pragmatics, thus extravagantly straying over various linguistic levels, such as
    - lexical structures involving phrasal embeddings (e.g. *stick-it-in-your-ear-attitude*; see Pascual et al. 2013 on what they call 'direct speech compounds')
    - contrastive reduplication with a variety of lexical(ized) items (e.g. *An evening bag or a BAG-bag?, I like it quite a lot but not A-LOT-a-lot*) indicating that the "prototypical meaning of the [reduplicated] lexical item is intended" (Ghomeshi et al. 2004: 312; examples gleaned from the *Corpus of English Contrastive Focus Reduplications*)
- borderline phenomena that are not easily reconcilable with traditional postulates of morphological accounts, such as
    - the cross-categoriality and multifunctionality of aforementioned *-ish-* and *-ly*-derivation, calling into question the validity of the Unitary Base Hypothesis (Aronoff 1976)
    - particle-verb nominalizations that involve a twofold attachment of the *–er* suffix (*crowd-warmer-upper, trash-picker-upper*, cf. Cappelle 2010, Lensch 2018), which challenge the assumption of a general paucity of reduplication in English and other European languages (Schwaiger 2015: 478)

While the above illustrations are all taken from English, extravagance can also be observed in other languages (see, for instance, (i) spreading suffixes such as Spanish emotive *-azo* or -ón (Lang 1990), Dutch adverbial *–erwijs* (Diepeveen & Van de Velde 2010) or German pejorative *Ge-x-e/x-erei* (Dammel & Quindt 2016); (ii) phenomena transgressing the borders of linguistic levels as in external degree modification (e.g. German *sau die gute Party*, cf. Gutzmann & Turgay 2015) or (iii) challenging or supporting evidence for more or less controversial universal tenets of morphological theorizing as instantiated by languages other than English, e.g. expressive diminutives in (non)European languages, which sheds new light on the Split Morphology Hypothesis, as discussed by Fortin 2011).

The aim of this workshop is to bring together experts in variational morphology and/or morphological theory to provide a forum to share and advance knowledge of the workings and impact of extravagance, i.e. speakers' deliberate deviation from established language norms, in morphological variation and change. The workshop also seeks to re-assess the empirical adequacy of established or alleged morphological principles and constraints and challenge their relative robustness or rather violability. Of particular interest are therefore the following research questions and their more general implications for linguistic variation and change:

- how much deviation is necessary to give rise to extravagance (and satisfy speakers' needs), and how much deviation does the system tolerate so that the speakers' needs can be satisfied (without there being any backfiring consequences doubting a speaker's linguistic competence)?
- where is the divide between innovation/change and violation/retention?
- what are the limits to extravagance, or in other words, what constrains the trespassing of constraints?
- are there specific domains that abound with morphological extravagance, as e.g. registers?

- which role do de-/re-categorization and reanalysis play? does extravagance feed these processes? how does the notion of extravagance tie in with context-sensitive approaches to language change that similarly observe deviation from established language use in the switch context (Heine 2002) or critical context (Diewald 2002)?
- are derivational processes, per se, more accommodating to speakers' creativity than inflectional morphology? is derivation-induced variation and change more pervasive than inflection-induced variation and change?
- are periphrastic constructions more expressive and extravagant than synthetic ones? does verbosity play a role?
- which (other) sociolinguistic factors 'promote' speakers' deviation from established norms? who are the 'movers'?
- what, finally, distinguishes extravagance from expressiveness or other related concepts such as 'extragrammatical morphology' or 'marginal morphology' (cf. Doneschal & Thornton 2000, Mattiello 2013)?

References

Aronoff, Mark (1976), Word *Formation in Generative Grammar*, Cambridge: MIT Press.

Bauer, Laurie (1983), *English Word-Formation*, Cambridge: Cambridge University Press.

Cappelle, Bert (2010), Doubler-upper nouns: A challenge for usage-based models of language?, in A. Onysko & S. Michel (eds.), (2010), *Cognitive Perspective on Word-Formation*, Berlin/New York: Mouton de Gruyter, 335-374.

Dammel, Antje & Olga Quindt (2016), How do evaluative derivational meanings arise? A bit of *Geforsche* and *Forscherei*, in R. Finkbeiner, J. Meibauer & H. Wiese (eds.), (2016), *Pejoration*, Amsterdam: Benjamins, 41-73.

Diewald, Gabriele (2002), A model for relevant types of context in grammaticalization, in I. Wischer & G. Diewald (eds.), (2002), *New Reflections on Grammaticalization*, Amsterdam: Benjamins, 103-120.

Diepeveen, Janneke & Freek van de Velde (2010), Adverbial morphology: How Dutch and German move away from English, *Journal of Germanic Linguistics* 22, 381–402.

Doleschal, Ursula & Anna M. Thornton (eds.), (2000), *Extragrammatical and marginal morphology*, München: Lincom Europa.

Eitelmann, Matthias, Kari Haugland & Dagmar Haumann (forthc. a), From *engl-isc* to *whatever-ish*: A corpus-based investigation of *-ish* derivation in the history of English.

Eitelmann, Matthias, Kari Haugland & Dagmar Haumann (forthc. b), *Teletubbyish paintings*, *youngerish sisters* and *ish eyes*: A corpus-based investigation of the extravagance of English *-ish*-derivatives.

Fortin, Antonio (2011), *The Morphology and Semantics of Expressive Affixes*, PhD thesis University of Oxford. Online available at http://ling.auf.net/lingbuzz/001365/v1.pdf (accessed 07/09/2018).

Geurts, Bart (2000), Explaing grammaticalization (the standard way), *Linguistics* 38(4), 781-788.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russel (2004), Contrastive focus reduplication in English (the Salad-Salad paper), *Natural Language and Linguistic Theory* 22(2), 307-357.

Gutzmann, Daniel & Katharina Turgay (2015), Expressive intensifiers and external degree modification, *The Journal of Comparative Germanic Linguistics* 17(3), 185-228.

Haspelmath, Martin (1999), Why is grammaticalization irreversible?, *Linguistics* 37(6), 1043-1068.

Heine, Bernd (2002), On the role of context in grammaticalization, in. I. Wischer & G. Diewald (eds.), (2002), *New Reflections on Grammaticalization*, Amsterdam: Benjamins, 83-101.

Keller, Rudi (1994), *Language Change: The Invisible Hand in Language*, London: Routledge.

Killie, Kristin (1998), The spread of *-ly* to present participles, in J. Fisiak & M. Krygier (eds.), (1998), *Advances in English Historical Linguistics*, Berlin: Mouton de Gruyter, 119-134.

Killie, Kristin (2007), On the development and use of appearance/attribute adverbs in English, *Diachronica* 24 (2), 327-371.

Kuzmack, Stefanie (2007), *Ish: A New Case of Antigrammaticalization?*, Paper presented at the meeting of the Linguistic Society of America (LSA), Anaheim, January 4-7.

Lang, Mervyn F. (1990), *Spanish Word Formation. Productive Derivational Morphology in the Modern Lexis*, London/New York: Routledge.

Lensch, Anke (2018), *Fixer-uppers*. Reduplication in the derivation of phrasal verbs, in R. Finkbeiner & U. Freywald (eds.), (2018), *Exact Repetition in Grammar and Discourse*, Berlin/Boston: Mouton de Gruyter, 158-181.

Mattiello, Elisa (2013), *Extra-grammatical Morphology in English: Abbreviations, Blends, Reduplicatives, and Related Phenomena*, Berlin/Boston: De Gruyter.

Nevalaien, Terttu (1994), Aspects of adverbial change in Early Modern English, in D. Kastovsky (ed.), (1994), *Studies in Early Modern English*, Berlin: Mouton de Gruyter, 243–259.

Oltra-Massuet, Isabel (2017), Towards a morphosyntactic analysis of *-ish*,*Word Structure* 10(1), 54-78.

Pascual, Esther, Emilia Królak & Theo A.J.M. Janssen (2013), Direct speech compounds: Evoking socio-cultural scenarios through fictive interaction, *Cognitive Linguistics* 24(2), 345-366.

Schwaiger, Thomas (2015), Reduplication, in P.O. v. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds.), (2015), *Word-Formation*, Berlin: de Gruyter Mouton, 467-484.

Traugott, Elizabeth Closs & Graeme Trousdale (2013), *Constructionalization and Constructional Changes*, Oxford: Oxford University Press.

Zwicky, Arnold M. & Geoffrey K. Pullum (1987), Plain morphology and expressive morphology, in J. Aske, N. Beery, L. Michaelis & H. Filip (eds.), (1987), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistic Society, 330-340.

Corpus of English Contrastive Focus Reduplications (2014)
https://home.cc.umanitoba.ca/~krussll/redup-corpus.html

What‘s extravagant about *be-sandaled feet*? The morphology and pragmatics of German pseudo-participles

Luise Kempf & Stefan Hartmann
(University of Mainz, University of Bamberg)

Keywords: extravagance, pseudo-participles, morphology, creativity, web corpora

The phenomenon of extravagance can be conceived of as a (more or less) deliberate deviation from established norms that evokes surprise or attention. In this talk, we present an empirical case study of a rare but arguably quite salient morphological pattern, namely German pseudo-participles. Pseudo-participles are forms that look like past participles but lack a verbal counterpart (see Bernstein 1992; also see Plag 2005 and Van Haeringen 1949 for pseudo-participles in English and Dutch). Attested examples include *be-brill-t* ‘be-glassed’ (derived from the noun *Brille* ‘glasses’) or *birken-be-stock-t*

(derived from the proper name *Birkenstock*, a sandal manufacturer, i.e. 'wearing Birkenstock sandals').

While not all formations of this pattern are extravagant (there are quite a few lexicalized types, e.g. *be-nachbar-t*, lit. "be-neighbour-ed" 'adjacent'), we will argue that the pattern lends itself to achieving extravagant effects in novel formations. The extravagance can be analyzed as an effect of a mismatch between the formal structure of a participle and the functional properties of a denominal adjective ("salience by surprisal", Schmid & Günther 2016). Thus, we are dealing with an interface phenomenon transgressing the boundaries of inflection and word formation. Formations with a complex base often serve the function of condensing information given in the previous discourse (e.g., *be-apfel-baum-t* lit. "be-apple-tree-d" refers back to an apple tree owner established in the preceding text). In this respect, the construction also straddles the boundaries between morphology and syntax.

Pseudo-participles derived from complex nouns often exhibit synonymous variants of the type (i) *be-sonnen-brill-t* "be-sun-glass-ed" vs. (ii) *sonnen-be-brill-t* "sun-be-glass-ed". As opposed to simplex-based pseudo-participles, they both intensify the mismatch. With type (i), the empty slot within the circumfix is filled with a polymorphemic structure, which would not occur with verbal stems, – thus creating formal deviation. Type (ii) formations mimic participles of separable verbs (*auf-ge-druck-t* past participle of *auf-druck-en* 'to print s.th. on s.th.') or participle compounds (*sonnen-be-strahl-t* 'sunlit'). This means they formally "camouflage" as being deverbal, yet semantically, they are not – thus increasing the mismatch and the effect of surprisal.

In our study, we investigate the extravagance of the pattern using data from the webcorpus DECOW16AX (Schäfer & Bildhauer 2012). We combine a detailed qualitative analysis of singleton coinages with explorative quantitative analyses, taking into account the complexity of the formations and their anchoring in the context. In addition, we examine metalinguistic indicators of extravagance such as double quotes, as exemplified in (1), and meta-comments, as exemplified in (2).

(1) Da findet mein Körper wohl endlich in den normal "bemuskelten" Zustand zurück. 'Then my body will probably finally find its way back into its normal 'be-muscled' state.' (http://www.deutschlands-dicke-seiten.de/forum/archive/index.php/t-9682.html, DECOW16AX)

(2) Ich finde die Wortwahl wie, beschwipst - *bekoffert - beanzugt* (selbstkreiert von Herbert) einfach genial. 'I find the choice of words, as in *tipsy – *be-suitcased – be-suit-ed* (created by Herbert himself), simply ingenious.' (http://www.letzte-version.de/songbuch/luxus/hartgeld, DECOW16A-NANO)

We argue that as a prime example of an "extravagant" phenomenon at the interface of morphology and pragmatics, the case of pseudo-participles can also be instructive for our understanding of extravagant morphology on a general level.

References

Bernstein, Wolf Z. 1992. *Pseudopartizipien im deutschen Sprachgebrauch. Ein Nachschlage- und Übungsbuch*. Heidelberg: Groos.

Handler, Peter. 2009. Rhetorik und Stilistik / Rhetoric and Stylistics. In Ulla Fix, Andreas Gardt & Joachim Knape (eds.), *Rhetorik und Stilistik / Rhetoric and Stylistics*, vol. 2, 1563–1575. (HSK 31.2). Berlin, New York: Mouton de Gruyter

Plag, Ingo. 2005. *Word-Formation in English*. Cambridge: Cambridge University Press.

Schäfer, Roland & Felix Bildhauer. 2012. Building Large Corpora from the Web Using a New Efficient Tool Chain. In Cicoletta Calzolari, Khalid Choukri, Terry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of LREC 2012*, 486–493.

Schmid, Hans-Jörg & Franziska Günther. 2016. Toward a Unified Socio-Cognitive Framework for Salience in Language. *Frontiers in Psychology* 7. doi:10.3389/fpsyg.2016.01110
Van Haeringen, Conrad B. 1949. Participia praeverbalia. *De Nieuwe Taalgids* 42. 187–191.

Recycling and extravagance: The case of English *-ster*

Graeme Trousdale
(University of Edinburgh)

Keywords: construction grammar, exaptation, extravagance, morphology, network

The English suffix *-ster* is remarkable for the variety of functions it has displayed in the course of its history. It began as an agentive suffix attaching to nouns with typically female referents (e.g. Old English *hoppestre* 'female dancer'); it was therefore closely aligned with the more productive agentive suffix *-ere* (Jespersen 1927, Kastovsky 1992). In certain varieties of Middle English, the *-ere* and *-estre* forms came to be used interchangeably with some lexical bases for the denotation of occupations (e.g *demere, demestre* '(male) judge'), with the outcome that agentive nouns with female referents required a new (borrowed) suffix to be added to the historically derived form (e.g. *songstress*). More recently, the *-ster* suffix has been used to create nouns that refer to people engaged in a particular (usually nefarious) activity, such as *mobster* and *gangster*. However, not all of the Late Modern English neologisms are pejorative (e.g. *punster* [first recorded 1699; *OED*] and *funster* [first recorded 1788; *OED*]). Furthermore, the form has very recently come to be used as a hypocoristic, often co-occurring with the definite article (e.g. *Robert* > (*the*) *Robster*).

This paper charts the history of *-ster* in detail, and uses principles of Construction Morphology (Booij 2010) to explain the relationship between morphological extravagance and competition in the constructional network. In particular, it seeks to address the following two research questions:

a. How does the recent development of a hypocoristic function compare to earlier changes in function in terms of extravagance?

b. How can the various types of *-ster* constructions be unified in a Construction Morphology framework?

The paper makes use of COHA (Davies 2010-) to provide frequency counts both of tokens of particular constructions, and of types/sub-schemas, to chart quantitative changes over time, for the non-hypocoristic uses. The discussion of the hypocoristic material is qualitative only.

The loss of particular features associated with a constructional schema (e.g. association with female referents for OE V.*estre*) is a semantic change akin to bleaching in grammaticalisation. Two outcomes are possible: univerbation and loss of the schema, or a recycling of the form of the schema with a new meaning (in a way similar to exaptation). Extravagance has no role to play in the first of these developments, but may have some role in the second. More than one cycle of change is possible (as in the case of hypocoristic *-ster*), and the recycling may straddle the contentful/procedural divide in terms of the function of the newly created morphological schema. The development of hypocoristic *-ster* is argued to be a clear example of extravagant morphology, due to the relatively frequent renewal of hypocoristic patterns in English (e.g. *Robert* > *Rob* ~ *Bob* ~ *Robin* ~ *Robbie* ~ *Robs* ~ *Robster* etc.). The paper therefore recasts extravagant morphology as a particular kind of development of certain

constructional schemas, typically in cases of low productivity and competition between associated subconstructions.

References

Booij, Geert (2010), *Construction Morphology*, Oxford: OUP.

Davies, Mark (2010-), *The Corpus of Historical American English* (COHA): 400 million words, 1810-2009. Available online at https://corpus.byu.edu/coha/.

Jespersen, Otto (1927), The ending *–ster*, *Modern Language Review* 22(2), 129-136.

Kastovsky, Dieter (1992), Semantics and vocabulary, in R. Hogg (ed.), *The Cambridge History of the English Language Volume 1: The Beginnings to 1066*, Cambridge: CUP, 290-407.

On *loners, easy-peelers, diggers-out, leaf clearer-uppers* and s*tayer-onner-for-nowers*: Extravagance of *-er* nominalizations in English

Anke Lensch
(Johannes Gutenberg-University Mainz)

Keywords: *-er* derivation, constructional changes, phrasal and prepositional verbs, morphosyntax, diachronic corpus study

This paper seeks to shed light on the ever-increasing extravagance of *-er* nominalizations in two respects. Firstly, it accounts for the constructional changes (Traugott & Trousdale 2013) which resulted in the diversification of the constructional network of *-er* suffixation. Throughout the history of English, the process of adding the suffix *-er* to mostly durative and habitual verbal bases to form new nouns as in *learn* → *learn**er*** has been highly productive (Dixon 2015:306). By Present-day English, the range of possible bases has been considerably extended, with adjectives (*lon**er***), place names (*London**er***), prepositions (*down**er***), cardinal numbers (*fiv**er***) and complex lexemes (*easy-peel**er***) forming, among others, part of the suffix's constructional network. During this process of host-class expansion, the suffix also widened its semantic range in that today, *-er* nominalizations still always denote nouns but no longer always denote agents (Ryder 1999), as exemplified by *fiver* denoting 'a five pound note', but also 'anything that counts as five' (OED 2019).

Secondly, this paper zooms in on multi-word verbs undergoing *-er* derivation as these verbs display extraordinarily rule-bending derivational patterns. As can be observed in (1), *-er* prototypically attaches to the verbal base, while at the same time violating the Right-hand Head rule. (2), on the other hand, exemplifies the two-fold attachment of the *-er* affix, thus instantiating a pattern that has been aptly called "doubler upper nouns" (Cappelle 2010; see also Chapman 2008). The occurrence of particularly extravagant formations involving multiple attachment of *-er* as in (3) and (4) shows that speakers are able to use the *-er* suffix in ever more creative ways (sensu Bauer 1983:63), deviating from common assumptions about English word-formation.

(1) a formidable *digg**er**-out* of facts (The Guardian 1993)
'someone very successful at unearthing the truth'

(2) the leaf *clear**er**-upp**er**s* (St Andrew's Roundhay Quarterly Magazine 2018)
'someone who professionally tidies away leaves'

(3) *stay**er**-onn**er**-for-now**er*** (The Guardian 2005)
'a politician who managed to have just enough votes to retain their seat in parliament'

(4) *<u>blanket</u>-putt**er**-onn**er**-and-off**er*** (The Guardian 2005)
'a device to maintain an infant's body temperature at the right level'

Moreover, the nominalizations in (1) to (4) show that the argument structure of the verbal base is still reflected in the derivational output and its syntactic environment. Often, the direct object of the underlying verb phrase either follows the derivative in an *of*-phrase, cf. (1), or it serves as the determinant of a compound where the *-er* nominalization of the phrasal verb is the determinatum as illustrated in (2). In this regard, this paper provides new insights into how *-er* nominalizations of multi-word verbs blur the boundaries between morphology and syntax (see also Los et al. 2012:134).

References

Bauer, Laurie (1983), *English word-formation*, Cambridge: Cambridge University Press.

Cappelle, Bert (2010), Doubler-upper nouns: A challenge for usage-based models of language?, in Onysko, Alexander and Michel, Sascha (eds), (2010), *Cognitive Perspectives on Word-formation*, Berlin/New York: Mouton de Gruyter, 335-374.

Chapman, Don (2008), *Fixer-uppers* and *passers-by*: Nominalization of verb-particle constructions, in v. Fitzmaurice, Susan M. & Minkova, Donka (eds), *Studies in the History of Language IV*, Berlin/New York: Mouton de Gruyter, 265-299.

Dixon, Richard (2015), *Making new words*, Oxford: Oxford University Press.

Los, Bettelou; Blom, Corrien; Booij, Geert; Elenbaas, Marion; Kemenade, van Ans (2012), *Morphosyntactic Change. A Comparative Study of Particles and Prefixes*. Cambridge: Cambridge University Press.

Oxford English Dictionary (2019)

Ryder, Mary Ellen (1999), *Banker*s and *blue-chippers*: An account of *-er* formations in Present-day English, *Language and Linguistics* 3(2), 269-297.

Traugott, Elizabeth & Trousdale, Graeme (2013), *Constructionalization and Constructional Changes*, Cambridge: Cambridge University Press.

Phrasal compounds with adjectival heads: Extravagant but following well-known patterns

Christine Günther, Sven Kotowski & Ingo Plag
(University of Siegen, University of Düsseldorf & University of Düsseldorf)

Keywords: word formation, adjectival compounds, phrasal compounds, compound semantics

Phrasal compounds, as in (1), are word-level structures that combine a lexical head with a phrasal non-head (Bauer et al. 2013: Ch.19). As such, they constitute a relatively marginal phenomenon that straddles the alleged syntax-morphology boundary:

(1) this "Steffi is Great" attitude (from Trips 2014: 44)

While still underresearched, phrasal compounds have received quite some attention in recent years (e.g., Meibauer 2007; Trips & Kornfilt 2015). Several claims have emerged from the pertinent literature, for example: they usually/only have nominal heads, can host a variety of syntactic structures in non-head position, are determinative (or subkind-establishing), and are cases of expressive morphology. In particular, the possibility of adjectives heading phrasal compounds is either explicitly

denied (e.g., Trips 2016) or considered a very marginal phenomenon at best (e.g., Lawrenz 2006 for German). This paper addresses the question of whether adjectival phrasal compounds exist in English and, if so, what their basic formal and semantic properties are.

Presenting corpus data from COCA, such as (2), we show that adjectival phrasal compounds exist, and that they are more common than assumed:

(2) a. one-skier-at-a-time narrow
b. make-your-stomach-hurt difficult

While problematic for restrictively modular accounts to the grammatical architecture, our general claim is that adjectival phrasal compounds follow the same patterns as more common compound types. We demonstrate that they are formally similar to non-phrasal adjectival compounds as well as to nominal phrasal compounds. At the same time, they crucially differ from their nominal cousins in the prototypical semantic relations between head and non-head: adjectival phrasal compounds are mostly similative-intensifying as opposed to determinative (see, for example, the 'difficult to a high degree'-reading in (2b)). We argue that this property is also found in NA-compounds and follows naturally from the semantics of the head category in question, viz., adjectives as mostly gradable predicates.

References

Bauer, Laurie, Rochelle Lieber & Ingo Plag (2013), *The Oxford reference guide to English morphology*, Oxford: Oxford University Press.

Lawrenz, Birgit (2006), *Moderne deutsche Wortbildung. Phrasale Wortbildung im Deutschen: Linguistische Untersuchung und sprachdidaktische Behandlung*. Hamburg: Verlag Dr. Kovač.

Meibauer, Jörg (2007), How marginal are phrasal compounds. Generalized insertion, expressivity,and I/Q-interaction, *Morphology* 17, 233-259.

Trips, Carola (2014), How to account for the expressive nature of phrasal compounds in a conceptual-semantic framework, *SKASE Journal of Theoretical Linguistics*, 11(1), 33-61.

Trips, Carola (2016), An analysis of phrasal compounds in the model of parallel architecture, in ten Hacken Pius (ed.), *The semantics of compounding*, Cambridge: Cambridge University Press, 153-177.

Trips, Carola and Jaklin Kornfilt (eds.) (2015), *Phrasal compounds from a theoretical and typological perspective* (= Special issue of Language Typology and Universals), Berlin: de Gruyter.

Compounds and idioms - a neo-constructionist approach to morphological marking

Leah S Bauke
(University of Wuppertal)

Keywords: compounding, idioms, morphosyntax, extravagance, English vs. German

This paper investigates an extravagant construction in English that straddles the line between idiom and compound in that it shows non-compositional content and categorial unfaithfulness, while it allows some modification that seems to presuppose functional morphological structure. The constructions is exemplified in (1):

(1) a. mind-blow*ing* - ‘unbelievable’
b. free-wheel*ing* - ‘blithe’
c. back-break*ing* - ‘hard, difficult, tough’

These constructions are problematic for a neo-constructionist content matching analysis (see Borer 2013) because they all involve morphological marking, indicated by italics - which clearly indicates the presence of functional structure - and thus should lead to a strictly compositional interpretation at the syntax-LF-interface. What is striking about these forms though is that their occurrence is strictly limited to *-ing* suffixation. I present an analysis that involves content matching for these forms that can accommodate *-ing* suffixation, thus alining the derivation of these idioms to a compound interpretation in a compound frame (see Borer 2013 for details). I further show how this analysis can be extended to other idiom constructions that also show morphological marking of functional structure as in (2):

(2) a. peaches *and* cream - ‘perfect’
b. small potatoes - ‘insignificant’
c. piece *of* cake - ‘easy’

Furthermore, I show that the analysis presented here can be transferred to German nominalized infinitives with *zum* (resembling nominal gerunds in English):

(3) a. Die Situation ist *zum Mäuse melken*
The situation is to mouse milking
‘The situation is frustrating’
b. Das Wetter ist *zum Eier legen*
The weather is to egg laying
‘The weather is wonderful’
c. Er fühlt sich *zum Bäume ausreißen*
He feels to tree pulling
‘He feels vigorous’
d. Der Vortrag war *zum junge Hunde kriegen*
The talk was to young dogs having
‘The talk was very exasperating’

This is reminiscent of another quite substantial set of *zum* + infinitive constructions, which also have an idiomatic interpretation:

(4) *zum Wiehern/Brüllen/Kugeln/Schreien/Kringeln/Quieken,…*
to neighing/yelling/rolling/screaming/curling/squeaking
‘very funny’

These forms can be interpreted straightforwardly under a content matching analysis, where content matching must take place when the roots are merged with the nominalizer *zum* at the latest. (Alternatively content matching takes place at the root, leading to the literal interpretations). For the forms in (3) this means that content matching can likewise be delayed until merger of *zum,* where the merger of the roots e.g. *Maus* and *melken* is an instantiation of a compound frame, thus alining the derivation of these idioms to a compound interpretation in a compound frame. Plural assignment can be derived from remerging one of the constituents (*Maus* in this case) under the relevant functional structure, i.e. adjoined to *zum.*

With this much in place, the system can then be extended to further idiom types like those in (6):

(6)	*zum*	*Greifen*	*nahe*	*zum*	*Affen*	*machen*	*zum*	*Zuge*	*kommen*
	to	grab	close	to	mokey	make	to	move	get
	'within easy reach'			'to disgrace oneself'			'become active'		

References
Borer, H. 2013. *Taking Form.* Oxford: OUP.

ꕥ

'Fake' morphemes in Dutch: Morphological productivity as a predictor of categorial flexibility?

Kristel Van Goethem & Muriel Norde
(F.R.S.-FNRS; Université catholique de Louvain & Humboldt University Berlin)

Keywords: extravagant morphology; privative semantics; morphological productivity; category change; corpus-based analysis

Dutch features several morphemes with 'privative' uses (Cappelle et al. 2018) that allow the proposition 'a PRIVATIVE X is not an X' and occur as left-hand members in compounds. Examples are: *imitatieleer* 'imitation leather', *fopspin* 'trick/joke spider', *kunstgras* 'lit. art-grass; artificial grass', *lokeend* 'duck decoy', *namaakbont* 'imitation fur', *nepjuwelen* 'fake jewels', *schijnhuwelijk* 'lit. appearance-marriage; marriage of convenience'. Moreover, the English loan *fake* is found in Dutch compound-like sequences, e.g. *fake(-)bericht* 'fake message'.

Some of these 'fake' morphemes are extravagant in that they display great categorial flexibility and innovative adjectival uses. *Nep*, for instance, is synchronically attested as free noun (e.g. *die keiharde nep* 'this rock-hard imitation'), compound member (e.g. *nepdrankjes* 'fake drinks'), but also as an inflected (example 1a) or graded (example 1b) adjective. *Fake* displays a similar degree of categorial flexibility, including comparative forms (example 2).

(1) a. *Dat is geen echte cupcake maar ik vond dat deze **neppe** cupcake, toch wel op de site mocht.*
'That's not a real cupcake but I thought this fake cupcake was fine for the site.'
b. *(...) hoe donkerder je gaat hoe **nepper** het er vaak uit gaat zien.*
'(…) the darker you go, the more fake it often looks.'

(2) *Kim lijkt wel een steeds strakker en **faker** gezicht te krijgen.*
'Kim looks like she is getting an ever tighter and faker face.'

Namaak is found in constructions that are ambiguous between noun and adjective (e.g. *een namaak rieten dak* 'a fake thatched roof'), whereas *kunst, fop, lok, imitatie* and *schijn* only seem to be used as nominal/verbal stems and compound members.

In this paper, we present an extensive analysis of 500 token samples of the 8 'fake' morphemes mentioned above, taken from the nlTenTen2014 webcorpus (Kilgarriff et al. 2014). First, we present the semantic-distributional profile of each 'fake' morpheme, based on a multiple distinctive collexeme analysis, and compare their morphological profiles and productivity. Second, we explore the question whether high type productivity of the morph correlates with debonding and its emergence as a free adjective.

Our results will be related to Barðdal's (2008) claim that there is an inverse correlation between type frequency and semantic coherence of a construction. Thus, we expect that high type frequency of the 'fake' compounds correlates with low semantic coherence between the compound members and with a higher degree of bleaching of the 'fake' morpheme. Moreover, we hypothesize that there is a correlation between semantic bleaching of the 'fake' morpheme and debonding from the compound, as evidenced by its usage as an (inflected) adjective (cf. Norde & Van Goethem 2018, among others). Conversely, 'fake' compounds with low type frequency are expected to show more semantic coherence, and would therefore less easily be reanalyzed as free adjectives.

References

Barðdal, Jóhanna (2008), *Productivity: Evidence from case and argument structure in Icelandic*, Amsterdam / Philadelphia: John Benjamins.

Cappelle, Bert, Pascal Denis, and Manuela Keller (2018), Facing the facts of *fake*: a distributional semantics and corpus annotation approach, *Yearbook of the German Cognitive Linguistics Association* 6(1), 9-42.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel (2014), The Sketch Engine: ten years on. *Lexicography* 1, 7-36.

Norde, Muriel & Kristel Van Goethem (2018), Debonding and Clipping of Prefixoids in Germanic: Constructionalization or Constructional Change? in G. Booij, (2018), *The Construction of Words* (*Studies in Morphology* 4), Cham etc.: Springer, 475-518.

Sketch Engine: http://www.sketchengine.eu

They're proing it up hardcore: An analysis of the V it up construction

Ulrike Schneider

The V *it up* construction is a subclass of the wide range of English particle constructions. It is what Cappelle (2005:9) terms a "specific template", as two of its elements are fixed: The object slot is filled by non-referential dummy *it* (cf. e.g. Postal and Pullum 1988: 650), which is followed by the particle *up*. The construction has been described as expressing an exuberant or excessive performance of the activity denoted by the verb (cf. Cappelle 2005: 418).

Despite the fact that the construction has been in use for well over half a century – as evidenced by the 1954 film *Living it up* – it has received little scholarly attention. This paper uses data from web corpora and the Urban Dictionary to analyse current uses of V *it up*. It focusses on two extravagant properties of the construction, namely the range of verbal heads it can take and the way it marks these as 'verby'.

Most of the higher frequency instances of the construction have verbs as their heads, such as (1). The low-frequency types and hapaxes, however, reveal the extravagant creativity the construction permits. The verbal slot can actually take proper names, as in (2), as well as nominal compounds and even phrasal compounds – see (3). This otherwise unusual attraction to "new or weakly established" verbs is typical for dummy *it* (Mondorf 2016: 94). The paper explores whether the productivity and morphosyntactic flexibility of the construction has increased in recent decades and whether verbal heads have to share certain features with well-established types.

(1) I'm gonna go home and sleep it up (getting rest)
(2) Boystars: you are PUT ON NOTICE. You are expected to Willis it the Fuck UP* the next time you're poledancing for a movie […]
(3) We're ugly Christmas sweater partyin' it up tonight!

A possible explanation for this attraction to un-verby heads is that both dummy *it* and the particle *up* have been described as "assist[ing] in turning non-verbs into verbs" (Cappelle 2005: 313, see Mondorf 2016: 94, 97 for dummy *it*). Thus a construction combining both of these elements should be an excellent 'verbaliser'. Yet there are indications that the 'verbiness support' does not stop there and that the construction tends to employ the progressive as a third marker. Therefore, this study furthermore tests whether there is evidence that V *it up* a) generally attracts the progressive and/or b) whether newly converted de-nominal heads are particularly prone to occur in the progressive. Thus the second question addressed is whether the construction balances out the extravagance of its (potential) heads by employing an extravagant number of syntactic and inflectional markers of verbiness.

References

Cappelle, Bert (2005): Particle Patterns in English. A Comprehensive Coverage. Thesis, Katholieke Universiteit Leuven.

Mondorf, Britta (2016): "'Snake legs it to freedom': Dummy *it* as Pseudo-Object." *Corpus Linguistics and Linguistic Theory* 12 (1). 73-102.

Postal, Paul M. and Geoffrey K. Pullum (1988): "Expletive noun phrases in subcategorized positions." *Linguistic Inquiry* 19 (4). 635-70.

Extravagant morphology: attaching -ly to present participles

Kristin Killie

<pdf>

Extravagant Harry Potter adverbs in American English: Synchronic and diachronic considerations

Cristiano Broccias & Enrico Torre
(University of Genoa)

Keywords: -*ingly* adverbs; subjective adverb; objective adverb; extravagant morphology; VVing.

Broccias (2012) discusses 'Harry Potter adverbs' (HPAs), illustrated in (1) below.

(1) a. Neville looked *pleadingly* at Harry, Ron and Hermione. (J.K. Rowling)
b. [H]e stared *unmovingly* at the sea. (Gerald Hensley)

HPAs are subject-oriented *–ingly* adjuncts, typically found in fiction (as in the Harry Potter series). They are formed out of verbs that can be predicated of the matrix subject either subjectively or objectively. In the former case, the verbal base (*plead*) pertains to the speaker's subjective assessment of the process profiled by the matrix verb (*look*) as a clue to the emotional state the clausal trajector is or, rather, seems to be in, see (2a). In the latter case, the verbal base ((*not*) *move*) describes a process unfolding simultaneously with that profiled by the matrix verb and predicated of the clausal trajector, as is in (2b).

(2) a. Neville {seemed to be /*was} pleading Harry, Ron and Hermione. [subjective HPA type]
b. He {*seemed not to be/was not} moving (as he stared at the sea). [objective HPA type]

The former case is a typical manner usage while the latter case is purely temporal. Nevertheless, both may be in parallel distribution with the VV*ing* pattern (Broccias and Torre 2018), as is shown in (3). Thus, HPAs can be characterised as an instance of 'extravagant' morphology as a morphologically simpler option is available (see Haspelmath 1999: 1055).

(3) a. He now looked *pleading* at both Tyler and Tanya. (C.J. Long)
b. Lambent stared *unmoving* at the boot for a few seconds. (Frances Hardinge)

Broccias (2012) shows that HPAs have been on the rise since the 19th Century. In order to investigate the extent of the phenomenon in more depth, the present contribution reports on ongoing research into the distribution of HPAs in American English, using data from *COCA* and *COHA* (Davies 2010, 2012), as they provide larger data sets than those used in Broccias (2012). In particular, we shall address the following questions:

- are there any diachronic differences in the distribution of the two types of HPA between the 19th and the 20th centuries?
- Are there any synchronic and/or diachronic differences between HPAs and the VV*ing* pattern? That is, why is an 'extravagant' *–ingly* form used instead of its non-extravagant counterpart?

References

Broccias, Cristiano (2012), Oriented *–ingly* adjuncts in Late Modern English, in H. Sauer, and G. Waxenberger (eds.), *English Historical Linguistics 2008. Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24-30 August 2008* [CILT 324]. Amsterdam / Philadelphia: John Benjamins, 147-164.

Broccias, Cristiano, and Enrico Torre (2018), From the VVingPP construction to the VVing pattern, *Lingue e Linguaggi*, 26, 81-99.

Davies, Mark (2010), The Corpus of Contemporary American English as the first reliable monitor corpus of English, *Literary and linguistic computing*, *25*(4), 447-464.

Davies, Mark (2012), Expanding horizons in historical linguistics with the 400-million-word Corpus of Historical American English, *Corpora*, *7*(2), 121-157.

Haspelmath, Martin (1999), Why is grammaticalization irreversible? *Linguistics*, *37*(6), 1043-1068.

Extravagant expressions denoting quite normal entities: Contrastive focus reduplication in German

Maximilian Frankowsky
(Leipzig University)

Keywords: morphology, German, reduplication, compounding, corpus linguistics

Nominal compounding in German usually combines different nouns. Compounds consisting of identical nouns are often considered impossible (e.g. Erben 1981). However, recent research has shown that compounds with identical constituents (ICCs) exist in German (e.g. Finkbeiner 2014; Freywald 2015; Kentner 2017). Apart from being formed 'accidentally' as ordinary determinative compounds (*Kindeskind* 'grandchild'), they are also used systematically as a special case of total reduplication, namely contrastive reduplication. As in English, German ICCs can indicate that the "prototypical meaning of the [reduplicated] lexical item is intended" (Ghomeshi et al. 2004: 312), e.g. *ich bin ein richtiges Mädchenmädchen und liebe Kitsch* 'I am a real girl girl and love kitsch'. This phenomenon violates the aforementioned dissimilarity constraint as well as the assumption that German is an extraordinary reduplication avoider (Stolz et al. 2011). In this respect, ICCs are extravagant constructions. They deviate from word formation norms and, in doing so, can be an effective way to highlight specific semantic features of nouns. In order to attract attention, speakers choose a synthetic (morphological) construction over the more common periphrastic one, because the former is more extravagant and expressive than the latter. They "choose a new way of saying old things" (Haspelmath 1999: 1057).

So far, research on German ICCs mostly based on small sets of examples. Hence, the lack of reliable data results in contradictory descriptions of the formal features and the nature of the phenomenon. It is still unclear, for instance, which properties of ICCs are responsible for the prototype interpretation. On the one hand, it is said that the context determines the usage and interpretation of ICCs (Hohenhaus 2004: 314); on the other hand, formal properties like the (presumptive) lack of linking elements are said to be decisive (Finkbeiner 2014: 187).

The current investigation aims to address these problems by examining large-scale corpus data. The chosen corpora DECOW16 (Schäfer & Bildhauer 2012) and deTenTen13 (Jakubíček et al. 2013) provide thousands of instances of German ICCs which help to further explore their formal properties, semantic restrictions, discourse function and domain. Preliminary results support the view that there are concrete formal and semantic features of the nouns which, alongside pragmatic enrichment, influence the use of German ICCs. Moreover, the data reveal the strategies that speakers use to keep the new word understandable and at the same time strange enough to make people notice it. Therefore, the study sheds some light on the balancing act between the need for and the limits of deviation. Finally, it turns out that ICCs became a powerful tool used by speakers of different social backgrounds. This way, a former taboo construction initiates the formation of an entirely new word formation pattern.

References

Erben, Johannes (1981), Neologismen im Spannungsfeld von System und Norm, in B. Schlieben-Lange (ed.), *Geschichte und Architektur der Sprachen*. Berlin / New York: de Gruyter, 35–43.

Finkbeiner, Rita (2014), Identical constituent compounds in German, *Word Structure* 7(2), 182–213.

Freywald, Ulrike (2015), Total Reduplication as a productive process in German, *Studies in Language* 39 (4), 905–945.

Ghomeshi, Jila; Jackendoff, Ray; Rosen, Nicole; Russell, Kevin (2004), Contrastive focus reduplication in English (The salad-salad paper), *Natural Language & Linguistic Theory* 22(2), 307–357.

Haspelmath, Martin (1999), Why is grammaticalization irreversible? *Linguistics* 37(6), 1043–1068.

Hohenhaus, Peter (2004), Identical constituent compounding – a corpus-based study, *Folia Linguistica* 38(3–4), 297–331.

Jakubíček, Miloš; Kilgarriff, Adam; Kovář, Vojtěch; Rychlý, Pavel; Suchomel, Vít (2013), The TenTen Corpus Family, *Proceedings of the 7th International Corpus Linguistics Conference*, 125–127.

Kentner, Gerrit (2017), On the emergence of reduplication in German morphophonology, *Zeitschrift für Sprachwissenschaft* 36(2), 233–277.

Schäfer, Roland; Bildhauer, Felix (2012), Building Large Corpora from the Web Using a New Efficient Tool Chain, *LREC*, 486–493.

Stolz, Thomas; Stroh, Cornelia; Urdze, Ania (2011), *Total reduplication the areal linguistics of a potential universal*. Berlin: Akademie-Verlag.

Phrasal compounds in German as a case of extravagant morphology

Katrin Hein
(Leibniz Institute for the German Language (IDS), Mannheim)

Keywords: German word-formation; compounding; construction grammar; usage-based; corpus linguistics

Phrasal compounds (PCs) like *»Man-muss-doch-über-alles-reden-können«-Credo* ‘one-should-be-able-to-talk-about-everything motto’ display an interesting case of extravagant morphology. Those „complex words with phrases in modifier position“ (Meibauer 2003: 153) can be considered as morphologically ‚extravagant‘ in the sense of Haspelmath (1999) in two respects: First, due to the integration of a syntactic unit into a complex word, we have to do with a borderline phenomenon that straddles the boundaries between morphology and syntax. Second, PCs violate morphological constraints like the ‚No phrase constraint‘ (Roeper/Siegel 1978) which does not allow for any interaction of word formation rules with phrasal elements as well as the ‚Principle of Lexical Integrity‘.

As a result of their morphological extravagance, the study of PCs is worthwhile in theoretical terms alone. Nevertheless, they have been largely ignored in the research literature for a long time, labeling them as a marginal phenomenon (but cf. Trips/Kornfilt 2017 for a recent collective volume about phrasal compounding in different languages).

While the generation of PCs poses a problem for ‚traditional‘ generative approaches which assume a modular architecture of grammar and do not allow for ‚syntax in morphology‘ (e.g. Lieber 1992, cf. Hein 2011), I will argue in my talk that a usage-based constructional model (e.g. Langacker 1987; Goldberg 2006) provides a suitable and fruitful approach to the phenomenon. For this purpose, I will sketch out the results obtained in Hein (2015): On the basis of the first corpus-based investigation of phrasal compounding in German, a large inventory of nominal PCs is modelled as pairings of form and meaning at different levels of specifity and abstractness within a bottom-up process (cf. Hein

2017). Besides the corpus-based carving out of central patterns of phrasal compounding, I will also attempt to relate the posited constructions within a ‚constructicon' (cf. Ziem/Lasch 2013: 95).

Moreover, extravagance can be understood in the sense of ‚expressiveness', i.e. it is connected with the assumption that PCs „display an expressive flavour typical of marginal morphology" (Meibauer 2007: 233). With reference to Meibauer (2007), this expressivity can be explained theoretically by adopting Levinson's (2000) "Theory of Generalized Conversational Implicatures". Moreover, the relationship of PCs to prototypical determinative compounds like *Baumhaus* 'tree house' is crucial in this context (cf. Hein 2015: 35-54).

All in all, I will present a usage-based construction grammatical view on phrasal compounding which is based on large corpus-data and takes into consideration aspects of frequency. My investigation clearly points out that phrasal compounds are by no means a marginal phenomenon in German, but that they are used productively. The approach at hand can also be seen as an attempt to applying the notion of construction to (extravagant) morphological phenomena (cf. Booij 2010).

References

Booij, Geert E. (2010), *Construction morphology*, Oxford et al.: Oxford University Press.

Goldberg, Adele E. (2006), *Constructions at work: The nature of generalization in language*, Oxford: Oxford University Press.

Haspelmath, Martin (1999), Why is grammaticalization irreversible?, *Linguistics* 37(6), 185-228.

Hein, Katrin (2011), Phrasenkomposita – ein wortbildungsfremdes Randphänomen zwischen Morphologie und Syntax?, *Deutsche Sprache* 39(4), 331–361.

Hein, Katrin (2015), *Phrasenkomposita im Deutschen. Empirische Untersuchung und konstruktionsgrammatische Modellierung*, Tübingen: Narr.

Hein, Katrin (2017), Modeling the properties of German phrasal compounds within a usage-based constructional approach, in C. Trips, and J. Kornfilt (eds), *Further investigations into the nature of phrasal compounding*, Berlin: Language Science Press, 119-148.

Langacker, Ronald W. (1987), *Foundations of cognitive grammar*, Stanford: Stanford University Press.

Levinson, Stephen C. (2000), *Presumptive meanings. The theory of generalized conversational implicature*, Cambridge, MA: MIT Press.

Lieber, Rochelle (1992), *Deconstructing Morphology. Word Formation in Syntactic Theory*, Chicago / London: The University of Chicago Press.

Meibauer, Jörg (2003), Phrasenkomposita zwischen Wortsyntax und Lexikon, *Zeitschrift für Sprachwissenschaft* 22(2), 153–188.

Meibauer, Jörg (2007), How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction, *Morphology* 17(2), 233-259.

Roeper, Thomas/Siegel, Muffy E. A. (1978), A lexical transformation for verbal compounds, *Linguistic Inquiry* 9(2), 199-260.

Trips, Carola/Kornfilt, Jaklin (eds) (2017), *Further investigations into the nature of phrasal compounding*, Berlin: Language Science Press (http://langsci-press.org/catalog/book/156).

Ziem, Alexander/Lasch, Alexander (2013), *Konstruktionsgrammatik. Konzepte und Grundlagen gebrauchsbasierter Ansätze*, Berlin et al.: de Gruyter.

Extravagant compounding

Barbara Schlücker
(University of Leipzig)

Keywords: compounding, proper names, German, morphology-syntax divide

Compounding, and in particular nominal compounding, is a highly productive word-formation pattern in present-day German but also in earlier stages of German. It is also a very old pattern that has already been attested in Old High German and before. Interestingly, though, there is one specific subpattern, viz. proper name compounds (consisting of a proper name modifier and a common noun head, e.g., *Europapolitik*, lit. Europa politics, 'European policy'; *Röntgenstrahlen*, lit. Röntgen rays, 'X-Rays') that is different from the other subpatterns of nominal compounding. Although a few sporadic examples of proper name compounds have been attested in Early New High German and before, based on data from the DTA corpus (1600-1900) it can be shown that this pattern is being systematically used only from 1600 on, and with a particular strong increase in the 19th century.

Thus, speakers at the beginning of this period are innovative in that they make use of an existing but largely dormant word-formation pattern. However, what is even more innovative is the functional development of the pattern that can be observed: whereas in the beginning, proper name compounds are predominantly names themselves, and in particular river names (e.g., *Rheinstrom*, lit. Rhine river), the use of proper name compounds as common nouns expands over the period investigated. This includes an expansion of the lexical class of the head nouns involved, such as the development from local nouns only to non-local ones, and from concrete to abstract nouns, among other things.

In addition, a second innovation (or, in other words, another extravagant use of the pattern) can be observed, namely the increasing use of proper name compounds as a competitor construction to the genitive (e.g., *Humboldt-Besuch*, lit. Humboldt visit ≈ 'Humboldt's visit'). Importantly, the competitor construction poses a challenge to traditional ideas about morphology and the morphology-syntax divide because these compounds seem to have properties typical of syntax rather than morphology, such as the referential interpretation of the modifier (cf. Di Sciullo & Williams 1987, Zifonun 2010, for instance). Moreover, the competitor use is also interesting regarding the question of the degree of deviation from the language system accepted by the speaker community and the norm since this use has already been remarked and criticized by grammarians at the turn of the 19th century, cf., e.g., Wustmann (1891). Research in present-day German corpora such as DECOW2012 and DWDS shows that the pattern continues to increase in the 20th century and is still frequently used. Interestingly, negative assessments of the competitor use are still found in linguistic and lay linguistic literature of the 20th and 21th century (e.g., Carstensen & Galinsky 1963, Zimmer 2006).

Thus, proper name compounding in German is innovative and morphologically extravagant in several ways and the paper discusses the function, the prerequisites and the limits of extravagant morphology on the basis of this morphological pattern.

References

Carstensen, Broder & Hans Galinsky. 1963. *Amerikanismen der deutschen Gegenwartssprache. Entlehnungsvorgänge und ihre stilistischen Aspekte*. Heidelberg: Carl Winter Universitätsverlag.

Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, Mass.: The MIT Press.

DTA corpus. 2018. Deutsches Textarchiv. Grundlage für ein Referenzkorpus der neuhochdeutschen Sprache. Edited by Berlin-Brandenburgische Akademie der Wissenschaften, Berlin. URL: http://www.deutschestextarchiv.de/

DWDS corpus. 2018. Das Digitale Wörterbuch der deutschen Sprache (Kernkorpus). Edited by Berlin-Brandenburgische Akademie der Wissenschaften, Berlin. URL: http://www.dwds.de/

Schäfer, Roland & Felix Bildhauer. 2012. Building Large Corpora from the Web Using a New Efficient Tool Chain. In Nicoletta Calzolari (ed.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, 486–493. European Language Resources Association.

Wustmann, Gustav. 1891. *Allerlei Sprachdummheiten. Kleine deutsche Grammatik des Zweifelhaften, des Falschen und des Häßlichen.* Leipzig: Grunow.

Zifonun, Gisela. 2010. Von *Bush administration* zu *Kohl-Regierung*: Englische Einflüsse auf deutsche Nominalkonstruktionen? In Carmen Scherer & Anke Holler (eds.), *Strategien der Integration und Isolation nicht-nativer Einheiten und Strukturen*, 165–182. Berlin, New York: de Gruyter.

Zimmer, Dieter E. 2006. *Die Wortlupe: Beobachtungen am Deutsch der Gegenwart*. Hamburg: Hoffmann und Campe.

Morphologically embedded general extenders: Consequences for morphological theory

Francesca Masini & Simone Mattiola
(University of Bologna)

Keywords: word-formation, general extender, list, pragmatics, Italian

General extenders (GEs) are expressions such as *and stuff*, (*or*) *whatever*, *etcetera* which indicate "additional members of a list, set, or category [and combine] with a named exemplar (or exemplars)" (Overstreet 1999:11). See (1), where *and everything* refers to everything related to a higher-level category that is inferred from the list of items preceding it (*name*, *address*), in this case something like "(any) useful contact information".

(1) *Make sure your bag has a tag with your name and address **and everything*** (Overstreet 1999:5)

In this paper we show that GEs, typically found in syntax/discourse – often at the end of "lists" (Masini, Mauri & Pietrandrea 2018), like in (1) – can interact with morphological schemas, giving rise to highly context-dependent expressions.

In Italian, GEs (like *eccetera* 'etcetera') and GE-related elements (like *tutto* 'all/everything', which is involved in a number of GEs: *e tutto* 'and everything', *e tutto il resto* 'and all the rest') are occasionally found embedded into word-formation schemas, typically at the end of a list of complex words formed according to the same schema:

(2) *Fini, D'Alema, post-fascisti, post-comunisti, **post-tutto***
'Fini, D'Alema, post-fascists, post-communists, **post-everything**'

(3) [...] *quello che ha il paraspalle paragomiti e **paratutto***
'that [back protector] that has the shoulder protector elbow protector and **everything protector**'

In (2), we have a list of prefixed words with *post-* 'post-' that culminates with a nonce-formation composed of *post-* and *tutto*; the same applies to (3), where two VN compounds with the verb *para(re)* 'protect' are followed by an item formed by *para-* and *tutto*. The resemblance with the syntactic list in (1) is striking. Note that *post-tutto* in (2) and *paratutto* in (3) are not stored items with a conventionalized semantics; rather, they acquire meaning by virtue of being inserted in a given context, a meaning that changes if the same item is embedded in a different environment: compare (2) with (4), where *post-tutto* refers not to the frame of political systems but to non-mainstream styles/fashions.

(4) *la casa di Emilia è tutta post-grunge, post-alternative, **post-tutto***
'Emilia's home is all post-grunge, post-alternative, **post-everything**'

Given this, nonce-formations with GEs can be regarded as "contextuals" in the sense of Clark & Clark (1979) and as a strategy to create "ad hoc categories" as intended in Mauri & Sansò (2017).

In addition to cases like (2)-(4), GEs may be found within coordinative compounds:

(5) *pregiudizi maschilisti-patriarcal-**eccetera***
'chauvinist-patriarchal-**etcetera** prejudices'

In this case, reference is made to prejudices that are chauvinist, patriarchal, *and things like that.*
This phenomenon is "extravagant" in more than one way:

(i) it straddles the boundary between morphology and syntax: GEs are normally claimed to be used in syntax, yet they are also found within morphological structures;
(ii) it raises questions on the role of pragmatics in word-formation, since the interpretation of words containing GEs is strictly dependent on (the wider) context.

The paper addresses both points, by describing how GEs are used within morphology, by offering a typology of relevant patterns, and by discussing the consequences for morphological theory.

References

Clark, Eve V. and Herbert H. Clark (1979), When nouns surface as verbs, *Language* 55(4), 767-811.

Masini, Francesca, Caterina Mauri and Paola Pietrandrea (2018), List constructions: Towards a unified account, *Italian Journal of Linguistics* 30(1), 49-94.

Mauri, Caterina & Andrea Sansò (2017), Strategie linguistiche per la costruzione online di categorie: un quadro tipologico, in Giuseppe Brincat & Sandro Caruana (eds), *Tipologia e 'dintorni'. Il metodo tipologico alla intersezione di piani d'analisi*, Roma: Bulzoni, 209-232.

Overstreet, Maryann (1999), *Whales, candlelight, and stuff like that: General extenders in English discourse*, New York: Oxford University Press.

WORKSHOP 7

Schedule: Thu 9.00 – Thu 16.55 (Room 10)

Gestures, facial expressions and body posture in CxG accounts of language: Core or periphery?

Claudia Lehmann & Meike Pentrel
(Universität Osnabrück)

Keywords: Construction Grammar, Multimodal Construction Grammar, gestures, body language, constructions

This workshop aims at investigating the relationship between verbal expressions and gestures, facial expressions and body posture using Construction Grammar (e.g. Goldberg 2006, Hoffmann & Trousdale 2013), or CxG, as the theoretical background. As Langacker (2008: 250) points out: "Manual gestures [...], facial expressions, actions performed more globally (e.g. a shrug), and even factors like body language [...] may all be closely bound up with linguistic expressions, in which case they can hardly be excluded from 'language' on an a priori basis." CxG provides a powerful framework for accounting for this very relationship. Its units of analysis, constructions, are defined as form-meaning pairings (e.g. Croft 2001: 18) and besides verbal expressions, gestures and the like can be such forms.

Recently, research on multimodal CxG has started to address these issues (e.g. Steen & Turner 2013, Zima & Bergs 2017). Quite a few publications report on notable co-occurrences of constructions and gestures (Schoonjans, Brône & Feyaerts 2015, Zima 2014a, Hsu, Brône & Feyaerts 2018). In the light of these findings, some researchers suggest that certain "constructions may not only have verbal but also gestural structure" (Zima 2014b: 27) and that gestures are thus integral parts of specific constructions. In these, gestural and verbal parts combine to arrive at a holistic meaning that is more than the sum of its parts (Schoonjans 2018). While multimodal CxG so far mainly focuses on gestures, the present workshop likes to include facial expressions and other bodily expressions, e.g. posture, in the discussion. Like gestures, body language and facial expressions also recurrently co-occur with certain expressions/constructions. An informal observation suggests, for example, that the construction *Tell me about it*, which is associated with an ironic meaning by convention, is often accompanied by a facial expression known as "blank face" whereas its nonconstructional counterpart (the sincere request) is not. However, the ironic construction is likewise accompanied by a flat intonation contour and so the question remains, if and to what extent "the body" can be seen as part of a construction. On the one hand, empirical studies were able to show frequency effects for morpheme sequences (e.g. Arnon & Snider 2010, Blumenthal-Drame 2012), which they interpret as a sign of their holistic entrenchment. Such a finding can easily be extended to multimodal constructions. On the other hand, it has often been argued thatprobably not all facets of a construction will be stored in the constructicon, i.e. the mental repository of constructs (cf. Goldberg 2013, Hoffmann 2017).

Yet, the notion of multimodal constructions is controversial. Since such constructions seem to also work in non-face-to-face situations, there are voices which analyse gestures not as integral but "as constructions in their own right that enter into crossmodal collostructions with linguistic items" (Uhrig 2018) while others argue that truly multimodal constructions (as described by Schoonjans (2018) above), which are stored in the long-term memory, are rare phenomena (Hoffmann 2017). Given the

variability of gestures and their potential infrequent co-occurrence with verbal expressions, the entire idea of multimodal constructions can easily be challenged (cf. Cienki 2018).

Although opinions remain divergent as to whether multimodal constructions truly exist (Ningelgen & Auer 2017; Schoonjans 2017), the fact that embedded depictions serve as constituents of canonical verbal utterances argues for multimodal semiotic signals being integral parts of constructions.

Following from these different approaches this workshop addresses among others the following questions and issues:

- Are constructions multimodal or is each bodily expression a construction in its own right?
- Are gestures/facial expressions part of the constructions stored in the constructicon or are they idiosyncratic occurrences that cannot be generalized?
- Are there constructions that are always accompanied by gestures, facial expressions etc., even in non-face-to-face conversations? Does the lack thereof lead to miscomprehension?
- Are facial expressions universal or cultural specific? Even if they differ, are there certain constructions that are more prone to be accompanied by facial expressions than others?
- How can we incorporate bodily actions in a CxG model?
- The entrenchment of gestures
- The acquisition of gestures and facial expressions (as part of constructions)
- Is there a special type of gesture that is especially prone to become part of a construction (e.g. metaphoric gestures, iconic, etc.)?

References

Arnon, I., & Snider, N. (2010) More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62: 67-82.

Blumenthal-Drame, A. (2012). *Entrenchment in usage-based theories. What Corpus Data Do and Do Not Reveal About The Mind.* De Gruyter.

Cienki, A. (2018). The multimodality issue: taking 'utterance' as a starting point. Paper presented at the International Conference on Construction Grammar 10, Paris, July 18th 2018.

Croft, W. (2001). Radical construction grammar: Syntactic theory in typological perspective. Oxford University Press on Demand.

Goldberg, A. E. (2013). Constructionist approaches. In Hoffmann & Trousdale (eds), 15-31.

Goldberg, A. E. (2006). Constructions at work: The nature of generalization in language. Oxford University Press on Demand.

Hoffmann, T. (2017). Multimodal constructs–multimodal constructions? The role of constructions in the working memory. Linguistics Vanguard, 3(s1).

Hoffmann, T., & Trousdale, G. (Eds.). (2013). The Oxford handbook of construction grammar. Oxford University Press.

Hsu, H., Brône, G. & Feyaerts, K. (2018). Embedded depictions as multimodal constructions. Paper presented at the International Conference on Construction Grammar 10, Paris, July 18th 2018.

Langacker, R. W. (2008). Cognitive grammar: A basic introduction. OUP USA.

Ningelgen, J. & Auer, P. (2017). Is there a multimodal construction based on non-deictic so in German? Linguistics Vanguard, 3(s1).

Schoonjans, S. (2018). Modalpartikeln als multimodale Konstruktionen. Berlin: De Gruyter.

Schoonjans, S. (2017). Multimodal Construction Grammar issues are Construction Grammar issues. Linguistics Vanguard, 3(s1).

Schoonjans, S., Brône, G., & Feyaerts, K. (2015). Multimodalität in der Konstruktionsgrammatik: Eine kritische Betrachtung illustriert anhand einer Gestikanalyse der Partikel einfach. In Bücker,

J., Imo, W. & Günthner, S. (eds.), Konstruktionsgrammatik V. Tübingen: Stauffenburg, 291–308.
Steen, F. & Turner, M. (2013). Multimodal Construction Grammar. In Borkent, M., Dancygier, B. & Hinnell, J. (eds.), Language and the Creative Mind. Stanford, CA: CSLI Publications, 255-274.
Uhrig, P. (2018). Multimodal Constructions or Crossmodal Collostructions? Paper presented at the International Conference on Construction Grammar 10, Paris, July 18th 2018.
Zima, E. (2014a). Gibt es multimodale Konstruktionen? Eine Studie zu [V(motion) in circles] und [all the way from X PREP Y]. Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion (15), 1-48
Zima, E. (2014b). English multimodal motion constructions. A construction grammar perspective. Linguistic Society of Belgium, 8, 14-29.
Zima, E., & Bergs, A. (2017). Multimodality and construction grammar. Linguistics Vanguard, 3(s1).

Come on! How much gesture does a construction need?

Alexander Bergs
(University of Osnabrück)

Keywords: construction grammar, multimodal construction grammar, extralinguistic information, gestures, corpus approach

Spoken, face-to-face language obviously does not occur in isolation. It is accompanied by body posture, gesture, facial expressions and a million other contextual factors. These provide a very rich input for hearers, who have to integrate many or even all of these factors into their parsing. But what needs to be taken into account and what is merely accidental or epiphenomenal and probably irrelevant for a successful parse (cf. Ladewig 2014a, b, Müller 2014, and Steen & Turner 2013)?

While the role of gesturing, body posture and facial expression have been generally acknowledged (see, e.g., Bressem 2013 and Zima & Bergs 2017) most linguistic analyses do not systematically include these in their approaches. Linguistic and gesture studies seem to happily coexist without any real interaction. Usage-based construction grammar (CxG) with its holistic approach to form-meaning pairings and its explicit acknowledgement of rich contextual input now offers a chance of systematically including context into the analyses (as in Zima 2014, 2017, and Págan Cánovas and Valenzuela 2017, for example).

However, this leads to some very fundamental questions that will be addressed in this paper. What is the relationship between the "traditional (verbal) linguistic sign", consisting of a sound image and a mental image, and relevant contextual clues, such as facial expressions and gestures? Are the two fully integrated, or does one 'only' support the other? What is the logical relationship between language and communication? One litmus test for this is to describe a linguistic construction that crucially needs extralinguistic information (such as gesture, posture, or facial expression) in order to be easily and unambiguously parsable.

This paper discusses one particular construction that might qualify in this respect. The "come on" construction, exemplified in (1), is inherently ambiguous with at least four different readings: support (1) and (2), hortative (3), incredulity (4), sexual innuendo (5)…. These are disambiguated by intonation, context, and probably also posture, facial expression and gesture.

(1) A few phone calls later, they said, “Come on, Ramin. We love you so much. We have to make this work.“ (COCA; 2017; Hollywood Reporter)
(2) HERNANDEZ# No. no He has only been in office four days. Come on! Give him a chance (COCA 2017, Fox Tucker Carlson Tonight)
(3) Your weapon! Give it to me! Come on! (COCA Fic 2016)
(4) No, no, That’s just – I mean, come on! (COCA Fic 2017)
(5) Come on, come on, come on, come on Do it to me now, baby, yeah (The Moments, Sexy Mama 1974)

On the basis of empirical data from the Red Hen database (www.redhenlab.org) and other corpora this paper discusses the pros and cons, and technicalities, of including extralinguistic information into the grammatical description of (verbal) linguistic constructions within a usage-based construction grammar framework.

References

Bressem, Jana (2013), A linguistic perspective on the notation of form features in gestures, in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (eds), (2013), *Body - Language - Communication: An International Handbook on Multimodality in Human Interaction*, Berlin / Boston: De Gruyter Mouton, 1079-1097.

Ladewig, Silva (2014b), Recurrent gestures, in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (eds), (2014), *Body - Language - Communication: An International Handbook on Multimodality in Human Interaction*, Berlin / Boston: De Gruyter Mouton, 1558–1574.

Ladewig, Silva (2014a), Creating multimodal utterances: The linear integration of gestures into speech, in C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (eds), (2014), *Body - Language - Communication: An International Handbook on Multimodality in Human Interaction*, Berlin / Boston: De Gruyter Mouton, 1662–1677.

Müller, Cornelia (2014), Gesture as ‘deliberate expressive movement’, in M. Seyfeddinipur and M. Gullberg (eds), (2014), *From gesture in conversation to visible action as utterance*, Amsterdam: John Benjamins, 127–151.

Pagán Cánovas, Cristóbal and Javier Valenzuela (2017), Timelines and multimodal constructions: Facing new challenges, *Linguistics Vanguard* 3(s1).

Steen, Francis and Mark B. Turner (2013), Multimodal Construction Grammar, in M. Borkent, B. Dancygier and J. Hinnell (eds), (2013), *Language and the Creative Mind*. Stanford: CSLI Publications, 255-274.

Zima, Elisabeth (2014), English multimodal motion constructions. A construction grammar perspective. *Linguistic Society of Belgium* 8, 14–29.

Zima, Elisabeth (2017), Multimodal constructional resemblance. The case of English circular motion constructions, in F. J. Ruiz de Mendoza Ibáñez, A. Luzondo Oyón, and P. Pérez Sobrino (eds), (2017), *Constructing families of constructions*, Amsterdam: John Benjamins, 301-337.

Zima, Elisabeth and Alexander Bergs (eds), (2017), Special issue: Towards a multimodal construction grammar. *Linguistics Vanguard* 3(s1).

Crossmodal Collostructions – A Usage-Based Approach to Multimodal Communication in a Construction Grammar Framework

Peter Uhrig

<pdf>

Constructions without words: Speech-embedded non-verbal depictions

Hui-Chieh Hsu, Geert Brône & Kurt Feyaerts
(University of Leuven)

Keywords: depiction, speech-embedded non-verbal depiction, multimodal construction, Construction Grammar, multimodality

Depictions, in the sense put forward by Clark (2016), are scenes people create and display, using sets of physical actions, to facilitate the addressee in imagining the depicted scenes. Iconic in nature, depictions serve as the common denominator of linguistic phenomena where the form of the sign bears physical resemblance to the meaning. While depictions are ubiquitous and come in various types (ibid.), of particular relevance to the recent multimodal CxG discussion (e.g. Ruth-Hirrel & Wilcox, 2018; Zima & Bergs, 2017) are tokens of speech-embedded non-verbal depictions; i.e. depictions that are embedded in speech, but that themselves contain no verbal elements. Communicating meaning non-verbally and without temporally co-occurring speech, such depictions are potential evidence of non-speech modalities being an integral part of language use. However, except for the few studies investigating restricted subsets of such depictions (e.g. Ladewig, forthc.), systematic research has yet to be carried out.

Given the gap, and to unravel the patterns and systematicity of speech-embedded non-verbal depictions, we tapped into ca. 10 hours of video data of American TV talk shows. Over 200 tokens of such depictions were identified, analyzed along two dimensions: Syntagmatically, a wide range of depiction-speech relations are observed, with embedding found on the levels of the word, phrase, clause, sentence, and beyond. This shows not only the possibility of sequential gesture-speech integration across syntactic levels, but also non-verbal signals' capability of "substituting" for structurally diverse verbal constituents, both in form and function. Paradigmatically, the tokens themselves exhibit internal structural complexities. In particular, series of consecutive depictions are observed building up complex composite meanings. For example, to demonstrate how he would not be able to hold back from actually enjoying the food if eating on scene were required of his role, Conan O'Brien staged the following non-verbal depictions consecutively, all of which share a common posture and arm configuration (annotation simplified):

(1) [*bites into a burger*] — [*sucks fingers*] — [*sticks out R index, signaling others to wait*]

Essentially, such series of depictions comprise a base-depiction, the common thread shared across the series; staged on top of the base are several elaboration-depictions, each of which either contributes to

the composite meaning of the series individually, or builds and elaborates on prior elaborations in the same series (as in 1).

Delving into the gap in the literature that is speech-embedded non-verbal depictions, the present study yields results with significant implications: The empirical investigation attests to non-verbal signals being able to communicate meaning without co-occurring speech, countering views dismissing them as merely para-linguistic, pointing instead to non-verbal depictions being form-meaning pairings just like speech. The versatility in syntagmatic cross-modal integration and complexity in paradigmatic variation of the tokens further render irrelevant skepticism that reduces such multimodal phenomena as idiosyncratic occurrences lacking systematicity. In addition to contributing to the multimodal CxG debate, the present study puts forward a strong case for non-speech modalities being an integral part of language use.

References

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review*, 123(3), 324–347.

Ladewig, S. (forthc.). *Integrating Gestures: Cognitive Grammar Multimodal*. Berlin: De Gruyter Mouton.

Ruth-Hirrel, L., & Wilcox, S. (2018). Speech-gesture constructions in cognitive grammar: The case of beats and points. *Cognitive Linguistics*, 29(3), 453–493.

Zima, E., & Bergs, A. (2017). Multimodality and construction grammar. *Linguistics Vanguard*, 3(s1), 20161006.

Toward a multimodal CxG analysis of Japanese mimetic expressions

Masaru Kanetani
(University of Tsukuba)

Keywords: Japanese, mimetic expressions, gestures, multimodality, construction grammar

This study considers the extent to which certain paralinguistic features are integrated with Japanese mimetic words. First, Japanese mimetic words are *constructions* in that they have certain constructional properties such as unpredictable morphophonological structures specific for them (e.g., Akita 2009, 2017). Second, they are often *multimodal* as accompanied by some paralinguistic features (e.g., Kita 1997; cf. also Dingemanse and Akita 2016). Treating thus Japanese mimetic words as constructions, this study conducts an experiment (i) to consider whether the "paralinguistic features" are stored as part of the mimetic word construction (i.e., part of the speaker's language knowledge), and (ii) to identify what – if any – paralinguistic feature (gesture, posture, and/or facial kinetic) it is that is stored for each mimetic word.

The experiment is designed as follows: (i) 20 psychomimes are extracted from a dictionary (Ono 2007) and grouped into four categories according to the type of emotion they represent (i.e., 5 expressions for each type of emotion: joy/happiness (e.g., *ukiuki* 'buoynat$_{mim}$'), anger (e.g., *mukaQ* 'huff$_{mim}$'), sadness (e.g., t*ohoho* 'woe$_{mim}$'), and other (e.g., *uttori* 'rapt$_{mim}$')); (ii) each word is shown to the subjects (10 native speakers of Japanese), who are then asked to make a gesture that may be accompanied by the word when uttering it. The subjects are also shown a specific situation in which

the expression at issue is used before making a gesture. It should be noted that the present study focuses only on psychomimes (or *gizyoogo*) although not only *gizyoogo* but also *giongo* 'phonomimes' (i.e., mimetic words for sound; e.g. *wanwan* 'bow-wow$_{mim}$') and *gitaigo* 'phenomimes' (i.e., mimetic words for appearances or states; e.g. *kirakira* 'twinkling$_{mim}$') are included in Japanese mimetic words (cf. Martin 1975). Of these three categories, psychomimes are used specifically to express the speaker's internal feelings or emotions. Given the embodiment of conceptual metaphors related to emotions such as HAPPY IS UP, SAD IS DOWN, etc. (Lakoff and Johnson 1980, Johnson 2017), I assume that psychomimes, as expressions of the speaker's emotions or feelings, are most readily – though not exclusively – accompanied by certain gestures, postures, and/or facial kinetics, and hence are suitable to the experiment. The parent study ignores particular tones or pitches, even if they are paralinguistic features that may be accompanied by phonomimes and phenomimes as well.

The data collected will first be compared across the subjects for each word (i) to identify what the most central paralinguistic feature is for a particular expression, and (ii) to see the extent to which the specific paralinguistic feature is integrated with that word. The most common gestures will be considered central for the expression in question; others may be variations, resulting from either extensions from the central one (via constructional extension) or effects of other similar psychomimes (via constructional network). After identifying the central paralinguistic feature for each word, I will also compare them within the same type of emotion and consider whether some common features may be abstracted within the same type of emotion.

Acknowledgements

This research is supported partially by JSPS KAKENHI Grant Numbers 18K00637 and 19K00681 and partially by Tsukuba Basic Research Support Program Type A.

References

Akita, Kimi (2009), *A grammar of sound-symbolic words in Japanese: Theoretical approaches to iconic and lexical properties of mimetics*, Doctoral dissertation, Kobe University.

Akita, Kimi (2017), Grammatical and functional properties of mimetics in Japanese, in N. Iwasaki, P. Sells, and K. Akita (eds.), (2009), *The Grammar of Japanese Mimetics: Perspectives from Structure, Acquisition, and Translation*, London / New York: Routledge, 20-34.

Dingemanse, Mark and Kimi Akita (2016), An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics* 53, 501-532.

Johnson, Mark (2017), *Embodied mind, meaning, and reason: How our bodies give rise to understanding*, Chicago: University of Chicago Press.

Kita, Sotaro (1997), Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35, 379-415.

Lakoff, George and Mark Johnson (1980), *Metaphors we live by*, Chicago: University of Chicago Press.

Martin, Samuel E. (1975), *A reference grammar of Japanese*, New Heaven: Yale University Press.

Ono, Masahiro (Ed.) (2007), *Giongo / gitaigo 4500: Nihongo onomatope ziten* [Phonomimes / Phenomimes 4500: Japanese onomatopoeia dictionary], Tokyo: Shogakkan.

Variable multimodality analyzed in terms of Utterance Construction Grammar: The case of "from beginning to end"

Alan Cienki & Agni-Veatriki Stanopoulou
(VU Amsterdam and Moscow State Linguistic U.; & VU Amsterdam)

Keywords: constructions; gesture; English; semantics; pragmatics

Expressions in English of the type [all the way from X PREP Y] have been shown to frequently occur with gestures (Zima 2017). They encompass not only spatial uses, but also temporal reference and listing functions (ibid.). However, some instantiations (specific constructs, to use Fried's [2015] term) concerning processes are more highly schematic semantically. The sub-construction in English "from X to Y", without articles before the reference points X and Y, is considered here as a sub-category (e.g., from start to finish, from soup to nuts, from top to bottom). We focus here on the specific construct "from beginning to end" as a case in point. It can refer to the spatial or temporal extent of an event as well as to the two extremes of a list. How are speakers gesturing with this construct, and what can this reveal about their mental simulation (semantics as imagery) of the events being described? Are gestures in fact being used to highlight semantic content of the utterance or to serve other, discourse or pragmatic functions? What can this tell us about the nature of the construct and, more broadly, the parent construction?

The research involved over 100 uses of the construct "from beginning to end" which occurred with manual gestures in various American television shows archived in The UCLA Library Broadcast NewsScape, accessed via tools developed within the Distributed Little Red Hen Lab. An analysis of form parameters (Bressem 2013) and functions (representation, discourse-structuring, or pragmatic functions) revealed that while a majority of the gestures were representational in some way (in terms of Müller's [1998] modes of representation), this combined with pragmatic functions in many of the cases. The great variety of depictive forms of the gesture share the property of providing imagistic grounding for the pragmatic function/meaning, namely that of endpoints having to do with absoluteness in a more abstract sense (totality, completeness). This correlated with particular qualities of the verbal construct "from beginning to end" as uttered, such as the use of a single intonation contour and the rapid speed of utterance, and with semantic characteristics of the preceeding speech relating to an evaluation of absoluteness.

The presentation will conclude by considering the theoretical implications of the empirical results. We will argue that the given constructional sub-type can be fruitfully analyzed within a framework such as that of Utterance Construction Grammar (Cienki 2017). This allows for treating form-meaning expression as variable in nature by admitting possibilities for flexibility in expression; in the case of this particular framework, this is possible in terms of various options for metonymic realizations of prototype structures.

References

Bressem, J. 2013. A linguistic perspective on the notation of form features in gestures. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill and S. Teßendorf (eds) *Body – Language – Communication. Volume 1*, 1079–1098. Berlin: De Gruyter Mouton.

Cienki, A. 2017. Utterance Construction Grammar (UCxG) and the variable multimodality of constructions. *Linguistics Vanguard* 3(s1) DOI 10.1515/lingvan-2016-0048.

Fried, Mirjam. 2015. Construction grammar. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (vol. 2), 974–1003. Berlin: De Gruyter Mouton.

Müller, C. 1998. *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich.* Berlin: Berlin Verlag Arno Spitz.

Zima, E. 2017. On the multimodality of [all the way from X PREP Y]. *Linguistics Vanguard* 3(s1) DOI 10.1515/lingvan-2016-0055

“Und dann” (*and then*) in German spoken narrative interactions

Elisabeth Zima
Albert-Ludwigs-Universität Freiburg

Keywords: multimodal constructions, Construction Grammar, German “und dann”, narrative connector, turn-taking

This paper sheds light on the functional diversity of “und dann” (*and then*) as used in jointly construed narrative sequences of two co-tellers in triadic interactions. Previous studies have described “und dann” as a narrative connector (Quasthoff 1979, Helmer & Deppermann 2013, Ziem & Lasch 2018), narrative structuring device (Gülich 1976, Schütze 1983) or element that serves to condensate a narrative passage (*Technik der Verdichtung*, Günthner 2007). Drawing on a video corpus of approximately five hours of triadic interactions that couples a third-eye perspective on the interaction with participants’ eye gaze recordings made by mobile eye tracking glasses (SMI, cf. author 2018 for a detailed description), our study, however, reveals a significantly broader spectrum of functions. These include issues of turn allocation management, resumption of a storytelling activity after some temporary digression, emphatic construal and projection of a sequence’s climax as well as a means for recipients to invite tellers to resume a storytelling activity or to give more detail.

While the function of “und dann” as a narrative connector, that links narrated elements of the story to be told to each other, is prevalent in both spoken and written discourse (cf. Ziem & Lasch 2018) and can be instantiated and accessed monomodally, the interactional functions that serve to both forward the storytelling activity as well as to manage turn allocation are crucially dependent on non-verbal meaning cues, such as particular eye gaze patterns or para-verbal features like particular prosodic contour. We discuss these findings in terms of the ongoing debate on the potential multimodality of constructions as modelled in Construction Grammar (see author & Bergs 2017 for a state-of-the-art) and suggest that “und dann” is not a multimodal construction *in se* but some of its functions are multimodal, as they rely on the joint instantiation of non-verbal meaning cues. We draw on prototype theory to explain this (mono-/multimodal) functional diversity.

References:

Gülich, Elisabeth. 1976. Ansätze zu einer kommunikationsorientierten Erzähltextanalyse (am Beispiel mündlicher und schriftlicher Erzähltexte). Zeitschrift für Literaturwissenschaft und Linguistik – Beihefte (4): 224-256.

Günthner, Susanne. 2007. Techniken der ‚Verdichttung‘ in alltäglichen Narrationen – Kondensierungsverfahren in Beschwerdegeschichten. In: Bär, J. & Roelcke, Thorsten (Hrsg.).

Sprachliche Kürze. Konzeptuelle, strukturelle und pragmatische Aspekte. Berlin/New York: De Gruyter. 391-412.

Deppermann, Arnulf & Helmer, Henrike. 2013. Zur Grammatik des Verstehens im Gespräch: Inferenzen anzeigen und Handlungskonsequenzen ziehen mit also und dann. Zeitschrift für Sprachwissenschaft, 1-40

Quasthoff, Uta M. 1979. Verzögerungsphänomene, Verknüpfungs- und Gliederungssignale in Alltagsargumentationen und Alltagserzählungen. In: Weydt, H. (Hg.): Die Partikeln der deutschen Sprache. Berlin: De Gruyter. 39-57.

Schütze, Fritz. 1983. Biographieforschung und narratives Interview. Neue Praxis 13(3), 283-293.

Ziem, Alexander & Lasch, Alexander. 2018. Konstruktionsgrammatische Zugänge zu narrativen Texten. Ausgangspunkte und Perspektiven. Zeitschrift für Literaturwissenschaft und Linguistik 48: 389–410.

The role of facial gestures in constructing interactional meaning

Karin Birkner & Alexandra Groß
(University of Bayreuth)

Keywords: facial gestures, multimodal constructions

This contribution focuses on the contribution of facial gestures (Bavelas & Chovil 2018) for the multimodal construction of meaning in conversation. It takes Langacker's (2008) view on gestures and language as a starting point that conventional gestures might bear a systematic relation to the expressions they occur in. By focusing on relatively fixed minimal verbal responses we will address the question of the significance of facial resources for co-constructing meaning within an interaction based C&G approach (e. g. Günthner & Imo 2006; Bücker, Günthner & Imo 2015).

Facial gestures that systematically co-occur with German "joa" will be described regarding their form. Taking also into account prosodic characteristics we aim at showing that accompanying facial gestures serve to constitute the pragmatic function of the response token in question. It will be argued that while the verbal component serves to comply with the conditional relevance set by a first action in terms of alignment, the accompanying facial gesture has the potential to communicate the speaker's evaluative stance. Since the context of usage is a significant third angle of the form-meaning-connection of (multimodal) constructions (e. g. Schoonjans 2018), we will particularly take into account how the embedding of "joa" comes into play in constructing its interactional meaning.

References

Bavelas, Janet / Chovil, Nicole (2018), *Some pragmatic functions of conversational facial gestures*, Gesture 17/1, 98-127.

Bücker, Jörg / Günthner, Susanne / Imo, Wolfgang (2015), *Konstruktionsgrammatik V. Konstruktionen im Spannungsfeld von sequenziellen Mustern, kommunikativen Gattungen und Textsorten*, Tübingen: Stauffenberg.

Günthner, Susanne / Imo. Wolfgang (2006), *Konstruktionen in der Interaktion*, Berlin/New York: Mouton de Gruyter.

Langacker, Ronald (2008), *Cognitive grammar – A basic introduction*. Oxford: Oxford University Press.

Schoonjans, Steven (2018), *Modalpartikeln als multimodale Konstruktionen – Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*, Berlin / New York: de Gruyter.

Frequency issues in multimodal construction grammar: gesture and beyond

Steven Schoonjans & Beatrix Schönherr
(KU Leuven & Leopold-Franzens-Universität Innsbruck)

Keywords: Construction Grammar, frequency, gesture, prosody, Conversation Analysis

The inclusion of non-verbal layers of expression into the notion of 'construction' is not uncontroversial and poses a number of theoretical issues. One of these issues is related to frequency as a yet vaguely defined criterion leading to the entrenchment of constructions. For instance, whereas Goldberg (2006) argues that form-meaning pairs may get entrenched when they are used and/or encountered with "sufficient frequency", Langacker (2008) speaks of "recurrent commonalities". Clear-cut cases of multimodal constructions would hence include para- and/or non-verbal elements co-occurring with every single instance of a particular verbal construction. However, there is as yet no study reporting such a 100% match. For gestures, for instance, previous analyses have revealed co-occurrence rates of gestures and verbal structures of around 60% (e.g. Sambre & Brône 2013, Zima 2013, Schoonjans 2018). Accordingly, the question has been raised whether it is justified to include gesture as a part of the construction if it is absent in no less than a third of the occurrences of the construction. The same is true for other para- and non-verbal layers of expression, where co-occurrence rates are also below 100% (e.g. Lanwer 2017 and Schönherr 1997 for prosody, albeit the latter not explicitly in CxG terms).

In this talk, we will address this frequency issue. The goal is not to argue for or against the constructional status of particular multimodal patterns. Rather, starting from the results of several multimodal analyses (including the ones mentioned above), we will argue that the role of frequency as a criterion for deciding upon construction status has to be rethought.

Our arguments will stem both from Construction Grammar (CxG) theory and from Conversation Analysis (CA) as a framework that is compatible with and an important feeding ground for CxG analyses (cf. Fischer 2007, among others). When it comes to CxG, we will refer, among other things, to the prototypical nature of constructions with optional form elements and to the granularity of constructions, strong correlations not always showing up at every level of specificity or generality (cf. Pijpops et al. [forthc.]). Related to this, from a CA perspective, we will refer, among other things, to the importance of the function in the discourse: if the position in the interaction is different, this may show up in a different multimodal behavior, impacting upon the multimodal co-occurrence rates. Finally, we will also address the particular role of external factors for the realization of non-verbal elements: more than verbal and para-verbal layers of expression, the realization of non-verbal layers (e.g. gesture) may be hampered by external factors such as restricted space to move, also reducing frequency of co-occurrence.

References

Fischer, Kerstin (2007), Konstruktionsgrammatik und Interaktion, in K. Fischer, and A. Stefanowitsch (eds), (2007), *Konstruktionsgrammatik: Von der Anwendung zur Theorie*, Tubingen: Stauffenburg, 133–150.

Goldberg, Adele Eva (2006), *Constructions at Work: The nature of generalization in language*, Oxford: Oxford University Press.

Langacker, Ronald W. (2008), *Cognitive Grammar: A basic introduction*, Oxford: Oxford University Press.

Lanwer, Jens (2017), Multimodal constructions: Explorations in the framework of interactional Construction Grammar (talk at MaMuD4, Frankfurt/Oder, 24.11.2017).

Pijpops, Dirk, Dirk Speelman, Stefan Grondelaers, and Freek Van de Velde (forthcoming), A theory-steered, data-driven approach to incorporate the multi-level nature of the constructicon into corpus research.

Sambre, Paul, and Geert Brône (2013) Cut and break. The multimodal expression of instrumentality and causality (talk at MaMuD1, Aachen, 22.11.2013).

Schönherr, Beatrix (1997), *Syntax – Prosodie – nonverbale Kommunikation: Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch*, Tübingen: Max Niemeyer.

Schoonjans, Steven (2018), *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*, Berlin: De Gruyter.

Zima, Elisabeth (2013), English multimodal motion constructions. A construction grammar perspective, *Studies van de BKL* 8, 14-28.

WORKSHOP 8

Schedule: Fri 9.00 – Sat 12.25 (Room 10)

Ideophones and interjections

Aimée Lahaussois & Yvonne Treis
(CNRS-HTL & CNRS-LLACAN)

Keywords: ideophone, interjection, typology, methodology, areality

Our workshop focusses on two word classes, ideophones and interjections, that have been faced, throughout their history, with definitional problems of a rather different nature than other word classes, in part because they are situated at the boundary of arbitrary vs. iconic language. In this workshop we would like to bring together scholars who study the semantics, morphology, syntax and pragmatics of ideophones and interjections from a typological, diachronic, areal and/or multimodal perspective.

Interjections, invented by Latin scholars to fill a gap in the 8-part word classification after the dismissal of articles (which existed in Greek but not in Latin), are on the whole defined, throughout history, in negative terms on account of their phonological marginality, their nonparticipation in morphological processes and their syntactic autonomy. Ideophones are also difficult to define notionally, on account of the great variety of ideophonic lexemes in the world's languages, and many definitions, as with interjections, highlight the non-standard phonology found with this word class. We have decided to consider these two word classes in the same workshop with the hope that setting them up in opposition to each other will make us more efficient in searching for consistent (and contrasting) definitions and in discussing issues of data collection, methodology and analysis that are common to both. Considering that in many languages, ideophones and interjections do not participate in morphological processes, it is reasonable to question whether they should, or even can be, topics in grammatical descriptions, and if so, how they can be described (and not only listed). They express emotions or reactions (interjections) and sensory imagery (ideophones), in other words subjective notions with great cultural variability. The difficulty in translating them and accurately capturing their meaning makes describing them difficult (for non-native linguists), marginalizing them even more compared to other word classes. An important question to be discussed in our workshop is that of the type of data which is needed to study these word classes: interjections and ideophones are more frequent in corpora of spontaneous speech, most often interactive, in other words types of linguistic production that are significantly different from those commonly used for research into nominal or verbal morphology. Additionally, they will often only be interpretable to the fieldworker after a relatively long period in the field, adding to the complexity of their description.

Ideophones have been given cross-linguistic consideration, through work on sound symbolism (Hinton et al. 1994), and in typology and descriptive work (see especially Voeltz and Kilian-Hatz 2001, Dingemanse 2011 and Reiter 2011). This work on ideophones has resulted in the following implicational hierarchy: sound < movement < visual patterns < other sensory perceptions < inner feelings and cognitive states (languages with ideophones covering a semantic field to the right will also have ideophones in semantic domains to the left of that point). In contrast, work on interjections has tended to be situated within the field of pragmatics (Ameka 1992, Cram 2008, Poggi 2009) rather than being carried out from the perspective of typology.

The workshop will address the following questions from the perspective of language-specific and cross-linguistic analysis.

- TYPOLOGY: How can one proceed from language-specific to cross-linguistic definitions of "interjection" and "ideophone"? What formal, semantic and pragmatic criteria can be used to compare interjections and ideophones across languages?
- CATEGORIZATION: Where are the boundaries between ideophones and adverbs, interjections and ideophones, interjections and fixed expressions, interjections and "imperativa tanta" etc.?
- MORPHOLOGY: In which languages do we find productive processes for the formation of ideophones on the basis of elements from other word classes? What derivational processes can interjections and ideophones be the input for?
- PROSODY: What are characteristic features of the word prosody and the prosodic integration of ideophones and interjections in individual languages and cross-linguistically?
- SEMANTICS: What are the semantic domains expressed by interjections and ideophones? When interjections and ideophones occur in grammars, it is often in the form of lists, divided into semantic sub-classes: are other configurations for their description possible?
- DIACHRONY: What are the lexical or syntagmatic origins of interjections across languages? What are possible origins of ideophones (onomatopoeia, loans etc.)?
- AREALITY: How do interjections and ideophones spread across language boundaries or within a linguistic area? Are there phono-symbolic patterns that are characteristic of specific linguistic areas? In which linguistic areas do we find similarly elaborated systems of interjections (e.g. interjections for different types of work, domestic animals)?
- MULTIMODALITY: Which co-verbal gestures are associated with ideophones and interjections?
- METHODOLOGY / TOOLS FOR DATA COLLECTION AND ANALYSIS: What types of linguistic data most frequently yield interjections and ideophones? Are there differences in frequency between certain linguistic genres (narratives, poetry, prayers, eulogies…) and everyday language? Which (non-)verbal stimuli can be used to trigger the use of interjections and ideophones and to help us capture their meaning?
- HISTORIOGRAPHY: How have interjections and ideophones typically been defined and described in research traditions of certain areas, language branches, families?

References

Ameka, Felix (1992), Interjections: The universal yet neglected part of speech, *Journal of Pragmatics* 18(2–3), 101–118.

Cram, David (ed.) (2008), *The Henry Sweet Society Bulletin* 50 [Special issue on Interjections].

Dingemanse, Mark (2011), *The Meaning and Use of Ideophones in Siwu*, Nijmegen: Radboud University. (Doctoral dissertation)

Hinton, Leanne, Johanna Nichols and John J. Ohala (eds) (1994), *Sound Symbolism*, Cambridge: CUP.

Poggi, Isabella (2009), The language of interjections, in A. Esposito, A. Hussain, M. Marinaro, and R. Martone (eds), *Multimodal Signals: Cognitive and Algorithmic Issues* (Lecture Notes in Artificial Intelligence 5398), Berlin / Heidelberg: Springer, 170-186.

Reiter, Sabine (2011), *Ideophones in Awetí*, Kiel: Christian-Albrechts-Universität zu Kiel. (Doctoral dissertation)

Voeltz, F.K. Erhard and Christa Kilian-Hatz (eds) (2001), *Ideophones*, Amsterdam: Benjamins.

Infixes, clicks, and the Devil: Scandinavian interjections from a Construction Grammar perspective

Steffen Höder
(Kiel University)

Keywords: interjections, Scandinavian languages, phonological schematicity, Construction Grammar

In traditional Scandinavian grammaticography, interjections are rather loosely defined as a word class and often understood as lying on the fringe of the language system, sometimes being described as para-linguistic phenomena rather than full-fledged linguistic units (e.g. Hansen & Heltoft 2011: vol. 1, 214; Teleman, Hellberg & Andersson 1999: vol. 2, 761–768). The main reason is that, while constituting a very heterogeneous group, interjections typically share some kind of exceptional formal features as compared to more prototypical word classes, a property they have in common with ideophones and other cases of sound symbolism (Dingemanse 2017). For example, they may occur independently of clausal structures (extraclausally or parenthetically) and have unusual phonological forms (such as Swedish [|] expressing, inter alia, indignation; Engstrand 2012: 71).

On the other hand, Scandinavian interjections do have formal properties that can be described in terms of well-established grammatical concepts and terminology. Examples are the morphological compositionality of many interjections (as in Danish *hovsa* 'whoops' < *hov* 'whoops' + *-sa* [colloquial marker]), their involvement in productive morphological processes (such as suffixation by Danish *-da* [colloquial marker] in *uhada*, *nåda*, *øvda*, all expressing different shades of surprise as well as additional meanings) and the fact that some interjections have an argument structure (e.g. Danish *skål for NP* 'here's to NP [toasting]'). However, the morphological or syntactic processes found in interjections often differ from those used in other word classes. For example, Swedish interjections may be formed productively by infixation (as in *jaha* 'ah, yes' < *ja* 'yes' + *-h-* 'understanding'; similarly *nehej* 'ah, no', *njaha* 'ah, well') or on the basis of phonological schemas (such as Swedish *ˈfa-X* 'Devil (swearing)'), a strategy employed primarily for purposes of taboo avoidance (Stroh-Wollin 2008: 105–107).

This talk approaches the grammar of Scandinavian interjections from a usage-based Construction Grammar perspective (Goldberg 2013), which seems well-suited for overcoming the traditionally exceptionalist view on interjections. In line with basic constructionist assumptions, it is claimed that Scandinavian interjections can be analysed in terms of constructions (i.e. form-meaning pairs) exhibiting different degrees of schematicity, thus bridging the continuum between more and less exceptional formal structures by using the same analytical devices. While lexical schematicity is widely used in Construction Grammar analyses across the board, this talk argues that it is crucial to also take the importance of phonological schemas into account (as proposed for a range of different phenomena by Höder 2014: 205–215). Drawing on earlier research on, amongst other things, the cognitive reality of phonaesthemes (Bergen 2004), it is claimed that phonological schematicity is needed in order to arrive at a cognitively realistic constructional representation of interjectional form as well as the organization of interjections into constructional networks.

References

Bergen, B. K. 2004. The psychological reality of phonaesthemes. *Language* 80, 290–311.

Dingemanse, M. 2017. On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In N. J. Enfield (ed.), *Dependencies in language. On the causal ontology of linguistic systems*, 195–203. Berlin: Language Science Press.

Engstrand, O. 2012. *Hur låter svenskan, ejengklien?* Stockholm: Norstedt.

Goldberg, A. E. 2013. Constructionist approaches. In T. Hoffmann & G. Trousdale (eds.), *The Oxford handbook of construction grammar*, 15–31. Oxford: Oxford University Press.

Hansen, E. & L. Heltoft. 2011. *Grammatik over det danske sprog.* [Copenhagen]: Det Danske Sprog- og Litteraturselskab [3 vols.].

Höder, S. 2014. Phonological elements and Diasystematic Construction Grammar. *Constructions and Frames* 6, 202–231.

Stroh-Wollin, U. 2008. *Dramernas svordomar – en lexikal och grammatisk studie i 300 års svensk dramatik*. Uppsala: Institutionen för nordiska språk, Uppsala universitet.

Teleman, U., S. Hellberg & E. Andersson (eds.). 1999. *Svenska Akademiens grammatik.* [Stockholm]: Norstedts Ordbok [4 vols.].

Comment définir la catégorie des interjections, le cas du Gbaya
Paulette Roulon-Doko

<pdf>

Ideophones and Interjections in Paiwan and Seediq

Amy Pei-jung Lee
(National Dong Hwa University)

Keywords: ideophones, interjections, Paiwan, Seediq, Formosan languages

Ideophones and interjections are two distinct categories of sound lexicon. The former is associated with iconicity resembling sensations and perceptions, while the latter with indexicality expressing sensations and emotions. This distinction can be further demonstrated by examining two languages—Paiwan and Seediq, which are Austronesian languages spoken in Taiwan (i.e. Formosan languages). In terms of subgrouping, Seediq belongs to the Atayalic branch and Paiwan constitutes a branch itself (Blust 1999). This paper compares ideophones and interjections of the two languages in terms of morphological processes and semantic categorization.

It is observed that interjections form a word class on its own, but ideophones tend to be nominal or precategorial. Ideophones in both languages take parts in derivation and inflection, while interjections are uninflected. In terms of structure, Seediq ideophones are characteristic of a CVC syllable with optional reduplication, while those in Paiwan are mostly manifested by lexicalized reduplication. Unlike Seediq, in Pawian a nonproductive infix <aL> (L as a cover letter for liquids) with <ar>, <al>, and <a×> variants is found in ideophonic lexicon to denote the meaning of 'to have sound or quality of' (Ferrell 1982, Li and Tsuchida 2009). Interjections are usually short words, where syntactic repetition may apply depending on contexts.

As for syntactic distribution, ideophones and interjections in both languages occur in quotative constructions, and are usually followed or preceded by the verb 'say'. Ideophones may function as verbal predicates or adverbials, while interjections as independent or part of exclamations.

In the two languages ideophones can be categorized semantically as onomatopoeic and depictive. The former mimic sounds from animals, actions, or the surrounding environment; the latter depict actions, states, and entities. Interjections are of three kinds based on their communicative functions:

expressive, conative, and phatic, following Ameka (1992). Expressive interjections express emotions and cognition; conative ones include those for interacting with animals; and phatic ones are used mostly in conversations for back-channeling and feedback signaling.

Examples:

(1) Seediq ideophones (Lee 2017)
   (a) ras ‘sound of water flowing’
   (b) ngiaw ‘sound of a cat; a cat’
   (c) wixwix ‘state of tiredness and exhaustion’

(2) Paiwan ideophones (Ferrell 1982)
   (a) busbus ‘very light drizzling rain’ > b<ar>usbus ‘to chatter on and on’
   (b) kingking ‘ring’ > k<al>ingking ‘to have ringing sound’
   (c) gemgem ‘fist’ > g<al>emgem ‘be furious’
   (d) zangzang ‘body heat’ > z<al>angzang ‘to perspire’

(3) Seediq interjections (Lee 2017: 199)

bah! bah! mesa		sagi=ta		ka	babuy.
INTJ	INTJ	AV.say	<AV>call.animal=1PL.INCL.NOM	NOM	pig

‘*Bah! bah!* is how we call pigs to approach.’

(4) Paiwan interjections

ai=aya	ti		kama	tjanusun.
INTJ.=say	NOM.person	father	2SG.OBL	

My father sends you his regards.

References

Ameka, Felix K. (1992), Interjections: The Universal yet Neglected Part of Speech, *Journal of Pragmatics*, 18(2-3): 101–118.

Blust, Robert (1999), Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics, in E. Zeitoun and P. J. Li, (eds)(1999), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, Taipei: Academia Sinica, 31-94.

Ferrell, Raleigh (1982), *Paiwan dictionary*. Pacific Linguistics Series C, no. 73. Canberra: The Australian National University.

Lee, Amy Pei-jung (2017), Ideophones, Interjections, and Sound symbolism in Seediq. *Oceanic Linguistics* 56(1): 181–209.

Li, Paul Jen-kuei, and Shigeru Tsuchida (2009), Yet more Proto-Austronesian infixes, in Bethwyn Evans, (ed) (2009) *Discovering History through Language: Papers in Honour of Malcolm Ross*, Pacific Linguistics 605. Canberra: Pacific Linguistics, 345-362.

Colour ideophones in African languages: A typological approach

Martine Vanhove & Guillaume Segerer
(Université Sorbonne Paris-Cité; LLACAN – CNRS; INALCO)

Keywords: ideophones, typology, colour terms, African languages, lexicon

We present the results of a large-scale typological survey of colour-related meanings expressed with ideophones in a sample of 106 African languages from various families and genetic stocks: Chadic, Cushitic (Afroasiatic), Khoe-Kwadi (Khoisan), Adamawa, Atlantic, Bantu, Dogon, Gur, Kordofanian, Kru, Kwa, Mande, Mel, Ubangi (Niger-Congo), Central Sudanic, Eastern Sudanic (Nilo-Saharan), and Songhai (varia). The data was mostly extracted from the RefLex online lexical database (Segerer & Flavier 2011-2018) by searching colours terms via translations together with grammatical categories.

We will first discuss methodological issues related to the use of a lexical database for the purpose of discovering ideophones since there is a great diversity of terminology in the sources. We will discuss the decisions we took for retrieval purposes, specifically the issue of selecting items not labelled as ideophones. Such items were included insofar as they were formally and/or semantically akin to Ameka's (2001: 26) working definition of ideophones. That includes:

a) terms labelled as ideophones by the authors: *pééb* ideo. 'very white' (Akoose); *jɔŋ* ideo. 'brown' (Mende)
b) terms labelled differently but fitting in with the working definition: *pələk-pələkà* adj. 'very black' (Bade, Chadic); *córí* adv.expr. 'very red, bright red' (Bambara, Mande); *pál* interjection 'pure white' (Bedik, atlantic-North); *ná tápatápa* onom. 'dark black' (Duala, Bantu); *kàrù* adverb (degree modifier) 'really red' (Ma'di, Central Sudanic)
c) any lexical item lacking the mention of its word category, and which combines specifically with a colour term, e.g. coy 'in intensive expressions yeeq coy – yaxig coy to be very red; to be bright red' (Wolof, Atlantic).

We will then present the genetic and areal patterns of the some 1,000 ideophones (leaving out 'multicoloured', 'spotted' and the like) found in our sample, as well as their skewed distribution across the spectrum, and their semantic properties, especially their intensifying or attenuative function of "basic" colour terms.

Since ideophones have been long claimed as being difficult to borrow (e.g. Childs 1994), but recently disclaimed as potentially borrowable because of the low degree of their morphosyntactic integration (Dingemanse 2017), we will end our presentation by the discussion of a specific case of lexical diffusion via borrowings from Fula within the Macro Sudan Belt area (Güldemann 2017) concerning the colours 'black', 'white' and 'red'. It is possible to hypothesize the origin of the diffusion: in each case, only one language, namely Fula (Atlantic-North), in fact a cluster of languages spoken by nomadic people, are or have been in contact with the other 25 languages belonging to four branches of Niger-Congo (Atlantic-North, Mande, Dogon and Adamawa), to one Songhay (Varia) variety, and to one branch of Afro-Asiatic (Chadic).

Acknowledgments: the research was funded by the ANR preoject Reflex and the Labex EFL.

References

Ameka, Felix. 2001. Ideophones and the nature of the adjective word class in Ewe. In Voeltz, E. F. K. & C. Kilian-Hatz (eds), Ideophones, Amsterdam & Philadelphia : Benjamins, pp. 25-48.

Childs, Tucker G. 1994. Expressiveness in contact situations: The fate of African ideophones. Journal of Pidgin and Creole Languages 9(2): 257-282.

Dingemanse, Mark. 2017. Expressiveness and system integration: On the typology of ideophones, with special reference to Siwu. STUF 70(2): 363–384.

Ferry, Marie-Paule. 1991. Thesaurus tenda: dictionnaire ethnolinguistique de langues sénégalo-guinnéennes (bassari, bedik, konyagi). Paris : Société des Etudes Linguistiques et Anthropologiques de France (SELAF). 1280 pp.

Güldemann, Tom. 2007. The Macro-Sudan belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In B. Heine & D. Nurse (Eds.), A Linguistic Geography of Africa, pp. 151-185. Cambridge: Cambridge University Press.

Segerer, Guillaume & Sébastien Flavier. 2011-2018. RefLex: Reference Lexicon of Africa, Version 1.1. Paris, Lyon. www.reflex.cnrs.fr (accessed 26 September 2018).Cennamo, Michela (2009), Argument structure and alignment variations and changes in Late Latin, in J. Barðdal, and S. L. Chelliah (eds), (2009), *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, Amsterdam / Philadelphia: John Benjamins, 307-346.

The emergence of ideophones in Atlantic Portuguese creoles: The case of Guinea-Bissau Kriyol and Santome

Chiara Truppi & Patrícia Costa
(Universidade de Lisboa)

Keywords: Portuguese-based creoles, substrate, Niger-Congo *Sprachbund*

The present paper focuses on ideophones in two Atlantic Portuguese-based creoles, i.e. Guinea-Bissau Kriyol (GBK; Upper Guinea creole) and Santome (ST; Gulf of Guinea creole). The aim of the study is twofold. First, we will compare formal and semantic properties of ideophones in both creoles. Moreover, the study aims to take part in the discussion on the role of substrate languages in the emergence of the category of ideophones in creoles. The presence of ideophones in these languages is, in fact, usually ascribed to substrate influence, especially with regard to Atlantic creoles (Childs 1994a; Parkvall 2000).

The paper aims to verify this hypothesis with regard to GBK and ST. In order to do that, we will look at the substrate languages of both creoles, i.e. Mandinka and Wolof for GBK (e.g. Kihm 1994, 2011; Jacobs 2010) and Edo for ST (e.g. Ferraz 1979), and at other Niger-Congo languages, whose contribution to GBK and ST grammars is still debated. The paper is based on both first-hand data and literature study: more specifically, the data on GBK and ST mainly come from our fieldwork, while the data for the study of substrate and Niger-Congo languages originate from the available literature.

The comparative part of the study will highlight that ideophones in GBK and ST show certain prototypical features found cross-linguistically such as idiosyncratic morphophonological patterns, reduplication, occurrence in multiword expressions, iconic motivation and (supra-)sensorial domains (Dingemanse 2011; Voeltz & Kilian-Hatz 2001; Childs 1994b). Furthermore, we will show that GBK and ST ideophonic inventories resemble, to a certain extent, the formal patterns found in the respective substrate languages (see (1)-(2) for examples of ideophones in GBK and ST). Yet, the source of some ideophones in the creoles under study can be traced back to Niger-Congo languages that are not traditionally considered as their contributing languages. Moreover, it is not always possible to determine the origin of certain ideophones. This could be explained in terms of internal innovation of these creoles with respect to the contributing languages. The situation described so far seems to suggest that the emergence of the category of ideophones in these creoles could be an effect of the Niger-Congo *Sprachbund*, i.e. it could be areal-induced rather than substrate-induced (as claimed by Parkvall 2000: 140).

(1) GBK a. (*pretu)*$_{\text{ADJ}}$ *nok* ‘very black’

b. *(intchi)*$_{V}$ *tep* 'to be full/to fill completely' < Mandinka (*fáa/fándí) tép* 'to be full/to fill completely' (Couto 1995; Creissels & Sambou 2013)

c. (*yentra*)$_{V}$ *fat* 'to enter all of a sudden'< Wolof *fat* 'swiftly'*;* (*dàgg*) *fatiit* 'cut in one stroke' (Torrence 2013)

d. (*limpu*)$_{ADJ}$ *pus* 'very clean' < Pepel *(fasi) pus* 'very clear' (Ndao 2011)

(2) ST: a. (*blanku)*$_{ADJ}$ *fenene* 'very white/pale'

b. (*tlêmê*)$_{V}$ *gidigidi* 'tremble uncontrollably' < Edo *(ye) gidigidi* 'big and thick or strong' (Agheyisi 1986)

c. (*fili*)$_{N}$ *petepete* 'very young, unripe'< Nembe *petepeté* 'tiny, very small' (Maduka 1988)

d. (*doxi*)$_{ADJ}$ *menemene* 'very sweet' < several Ijoid languages *menimeni* 'be(come) sweet, sugary' (Williamson 2004)

References

Agheyisi, Rebecca (1986), *An Edo-English Dictionary*, Benin: Ethiope Publishing Corporation.

Bartens, Angela (2000), *Ideophones and Sound Symbolism in Atlantic Creoles*, Saarijärvi: Gummerus Printing.

Creissels, Denis & Pierre Sambou (2013), *Le Mandinka. Phonologie, grammaire, textes*, Paris: Karthala.

Childs, Tucker (1994a), Expressiveness in contact situations: the fate of African ideophones, *Journal of Pidgin and Creole Languages,* 9:2, pp. 257-282.

Childs, Tucker (1994b), African Ideophones, in Hinton, Nichols & Holm (eds.), *Sound Symbolism*, Cambridge: Cambridge University Press.

Couto, Hildo H. do (1995), Exclusive particles (ideophones) in Guinea-Bissau Creole, in Philip Baker (ed.), *From contact to creole and beyond*, London: University of Westminster University Press, 207-215 (Westminster Creolistics Series 1).

Dingemanse, Mark (2011), *The meaning and use of ideophones in Siwu*, PhD dissertation, Nijmegen: Radboud University Nijmegen.

Ferraz, Luiz (1979), *The creole of São Tomé*, Johannesburg: Witswatersrand University Press.

Jacobs, Bart (2010), Upper Guinea creole: Evidence in favor of a Santiago birth, *Journal of Pidgin and Creole Languages* 25(2), pp. 289–343.

Kihm, Alain (1994), *Kriyol Syntax: The Portuguese-based creole language of Guinea-Bissau*, Amsterdam/Philadelphia: John Benjamins.

Kihm, Alain (2011), Substrate influence in Kriyol: Guinea-Bissau and Casamance Portuguese-related Creole, in Claire Lefebvre (ed.), *Creoles, their substrates, and language typology*, 81-103, Amsterdam: John Benjamins.

Maduka, Omen (1988), Size and shape ideophones in Nembe: A phonosemantic analysis, *Studies in African Linguistics*, Vol. 19, N.º1, pp. 93-113.

Ndao, Dame (2011), *Phonologie, morphologie et structures syntaxiques du Pepel*, PhD dissertation, Dakar/Paris: Université Cheikh Anta Diop de Dakar (UCAD)/Institut Nationaldes Langues et Civilisations Orientales (INALCO).

Parkvall, Mikael (2000), *Out of Africa: African influences in Atlantic Creoles*, London: Battlebridge Publications.

Peace Corps The Gambia (1995), *Wollof-English Dictionary*, Banjul, Gambia.

Scantamburlo, Luigi (1981), *Gramática e dicionário da língua Criol da Guiné-Bissau*, Bologna: Editrice Missionaria Italiana.

Torrence, Harold (2013), A promotion analysis of Wolof clefts, *Syntax* 16:2, pp. 176–215.

Voeltz, F. K. Erhard & Christa Kilian-Hatz (eds.), (2001), *Ideophones*, Amsterdam: John Benjamins.

Williamson, Kay (2004), *Comparative Ijoid Word List*.

ꕥ

Derived ideophones in Amharic and beyond

Ronny Meyer

<pdf>

ꕥ

Ideophones in Upper Guinea Creoles

Quint Nicolas & Noël Bernard Biagui
(LLACAN – UMR8135 – CNRS/INALCO/USPC (Paris/Villejuif –France) & CLAD – Université Cheikh Anta Diop (Dakar – Sénégal))

Keywords: Afro-Portuguese, Comparative Linguistics, Ideophones, Niger-Congo, Upper Guinea Creoles

Like many Niger-Congo languages, all Upper Guinea Creoles (UGCs) spoken in Africa (i.e. except Papiamentu) make use of a specific set of adverbs which specifically intensify various verbs of the language (see Table 1).

Table 1. Some intensive ideophones in three African UGs

	Meaning		
Language	*(be)* ***very*** *white*	*(be)* ***very*** *black*	*remain* ***completely*** *silent*
Capeverdean (Santiago)	bránku **álbu**	prétu **finu**	kála **ien**
Casamance	braŋku **fandaŋ**	pretu **nɔk**	kalá **mik**
Guinea-Bissau	branku **fandan**	Pretu **nok**	kala **yem**

These ideophones have already been studied in some details for Guinea-Bissau Creole (Couto 1994, 1995, Doneux & Rougé 1988, Scantamburlo 1981, 1999, Kihm 1980, 1994) and briefly mentioned for Capeverdean (Quint 2000, 2008) and Casamance (Biagui 2018). However, quite often, the material published mainly consists of lists of items accompanied by rather short comments. In this presentation we intend to provide a broader overview of intensive ideophones across Upper Guinea Creoles, basing ourselves on first hand and published sources.

First, drawing on typological references (Creissels 2002, 2006, Dingemanse 2012, 2018, Samarin 2001) and available data (in particular two unpublished lists of respectively 90 Casamance and 15 Capeverdean ideophones), we will endeavour to define the main phonological, syntactic and semantic characteristics of UGC ideophones. Second, we will develop a comparative approach, insisting both on what the different varieties have in common (see Table 1 **ien** [jẽ] and **yem** [jem]) or not (see Table 1 **álbu, finu, mik**, each idiosyncratic of a specific variety). Third, we will provide some insights about the origin of this interesting category which, like many other components of UGCs, has a double African and Portuguese provenience (e.g. Capverdean **álbu** < Portuguese *alvo* 'white' (variant of

branco) and **finu** < Mandinka *fĩŋ* 'black'). Finally, we will delve into the functional relationship between UGC ideophones and their Niger-Congo counterparts, with special emphasis on the substrate languages of UGCs, ie Mandinka, Wolof and other Atlantic languages.

References

Biagui, Noël Bernard (2018), *Description générale du créole afro-portugais de Ziguinchor (Sénégal)*, Paris: Karthala.

Couto, Hildo Honório (1994), *O Crioulo Português da Guiné-Bissao*, Hamburg: Helmut Buske.

Couto, Hildo Honório (1995), Exclusive particles (ideophones) in Guinea-Bissau creole, in Ph. Baker (ed.) *From contact to creole and beyond*. 1 ed., London: University of Westminster Press, v. 1, 207-215.

Creissels, Denis (2002), Adjectifs et adverbes dans les langues africaines, in *Typologie des langues d'Afrique subsahariennes et universaux de la grammaire* Vol. 1, Paris: l'Harmattan.

Creissels, Denis (2006), *Syntaxe structurale, une introduction typologique 1. Catégories et Construction*, Paris: Lavoisier.

Dingemanse, Mark (2012), Advances in the Cross-Linguistic Study of Ideophones, *Language and Linguistics Compass* 6/10, 654–672.

Dingemanse, Mark (2018), Redrawing the margins of language: Lessons from research on ideophones, *Glossa: a journal of general linguistics* 3(1): 4. 1–30.

Doneux, Jean-Léonce & Rougé, Jean Louis (1988), *En apprenant le créole à Bissau ou Ziguinchor*, Paris: L'Harmattan.

Kihm, Alain (1980), *Aspects d'une syntaxe historique. Études sur le créole portugais de Guinée-Bissau* [PhD thesis], Paris: Sorbonne Nouvelle.

Kihm, Alain (1994), *Kriyol Syntax: the Portuguese-based creole language of Guinea-Bissau*, Amsterdam/Philadelphia: John Benjamins.

Quint, Nicolas (2000), *Grammaire de la langue capverdienne (description diachronique et synchronique du créole afro-portugais des îles du Cap-Vert)*, Paris: L'Harmattan.

Quint, Nicolas (2008), *L'élément africain dans la langue capverdienne / Africanismos na língua caboverdiana*, Paris: L'Harmattan.

Samarin, William J. (2001), Testing hypotheses about African ideophones, in F. K. Erhard Voeltz and C. Kilian-Hatz (eds), *Ideophones*, Amsterdam: John Benjamins, 321-337.

Scantamburlo, Luigi (1981), *Gramática e dicionário da língua criol da Guiné-Bissau (GCr)*, Bologne: Editrice Missionaria Italiana.

Scantamburlo, Luigi (1999), *Dicionário guineense*, Volume I: Introdução e notas Gramaticais, Lisbon: Edições Colibri, Faspebi.

The prosodic features of Teko ideophones

Françoise Rose, Mao-Line Fillon, Jennifer Krzonowski
(Dynamique Du Langage, CNRS/Université Lyon 2)

parts of speech, intonation, stress, pause, vowel harmony, syllabic pattern, reduplication, intonation, Tupi-Guarani

The aim of this paper is to present a prosodic description of ideophones in Teko, contributing to the recent development of research on ideophones (Childs 1994, Voeltz and Kilian-Hatz 2001,

Dingemanse 2012). The prosody of ideophones, in particular, is rather little investigated (Dingemanse et al. 2016, Reiter 2011). Teko (aka Emérillon) is a Tupi-Guarani language spoken by 400 persons in French Guiana (Glottocode emer1243), and has been recently described (Rose 2011). A lexical database of 1327 lexemes (Rose 2018) contains 103 ideophones, "marked words that depict sensory imagery" (Dingemanse 2011:25). The corpus of 38 texts shows 388 occurrences of ideophones on a total of 2000 sentences. The phonological, morphological, syntactic, semantic and discourse-pragmatic features of Teko ideophones have been previously described by Rose (2011:400-409). They set aside ideophones as a distinct part of speech, found in both narratives and conversation to depict situations (actions or states), with different degrees of syntactic integration: ideophones can either be an optional addition to a clause, be introduced by a 'do/say' verb, or be used as the unique predicate of the clause.

The paper investigates various aspects of the prosody of ideophones. First, we will show that a syllabic pattern ending with a final closed syllable (such as in *moŋ* 'dark' and *dɨɾɨg* 'look') is statistically favoured in ideophones (66%) in comparison with other parts-of-speech (45%), $\chi^2(1) = 15.61$, $p < .001$. Second, we will highlight the fact that two thirds of Teko ideophones show vowel harmony across the entire word, a process that is not usual in the rest of the lexicon (see also Akita et al. 2013). Third, acoustic measures in Praat will confirm the perception of ideophones's prosodic saliency due to vowel lengthening, high pitch and intensity, as opposed to "plain" prosody (Gordon and Rose 2006). Studying the prosodic variation of the same ideophone in different contexts also helps explain how prosody can add semantic information to the lexeme, for example vowel lengthening and a salient pitch drop in the ideophone *tou* 'fall' indicates that the object falls from a very high location. Fourth, the length of pauses around ideophones will be evaluated, as was done for direct speech (Rose and Vanhove 2007). The third and fourth tasks are based on one hour and forty minutes of audio-recorded texts, and enable the exploration of the relation between prosodic and syntactic integration, following the methodologies of Malibert and Vanhove (2015) and Dingemanse and Akita (2017).

To summarize, this paper will first complement the existing description of Teko ideophones, and second contribute to the emerging typology of ideophones, especially their prosody, a promising under-investigated domain for ideophones. It provides answers to three general questions: Is the prosody of ideophones in general distinct from that of other parts-of-speech? How does the prosody of ideophones participate in their semantic interpretation in context? Does the prosodic integration of ideophones match their syntactic integration?

References

Akita, Kimi, Mutsumi Imai, Noburo Saji, Katarina Kantartzis, and Sotaro Kita, 2013, Mimetic vowel harmony. Japanese/Korean Linguistics, ed. by Bjarke Frellesvig and Peter Sells. Vol. 20. Stanford, CA: CSLI Publications.

Childs, G.T, 1994, "African ideophones", in *Sound symbolism*, Vol. Chapter 13, L. Hinton, J. Nichols and J. Ohala (eds): Cambridge University Press, 178-204.

Dingemanse, Mark, 2012, "Advances in the Cross-Linguistic Study of Ideophones", in *Language and Linguistics Compass* 6.10, 654–672.

Dingemanse, Mark, Schuerman, Will, Reinisch, Eva, Tufvesson, Sylvia and Mitterer, Holger, 2016, "What sound symbolism can and cannot do: testing the iconicity of ideophones from five languages", in *Language*: 92.2, 117-133.

Dingemanse, Mark and Akita, Kimi, 2017, "An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese", in *Journal of Linguistics*: 53.3, 501-532.

Gordon, Matthew and Rose, Françoise, 2006, "Emérillon stress: a phonetic and phonological study", in *Anthropological Linguistics*: 48.2, 132-168.

Malibert, Il-Il and Vanhove, Martin, 2015, "Quotative constructions and prosody in some Afroasiatic languages: Towards a typology", in *Corpus-based Studies of lesser-described Languages: the CorpAfroAs Corpus of spoken AfroAsiatic*, Amina Mettouchi, Martine Vanhove and Dominique Caubet (eds), Amsterdam, Philadelphia: John Benjamins, 117-169.

Reiter, Sabine, 2011, Ideophones in Awetí, PhD dissertation, Christian Albrechts University, Kiel.

Rose, Françoise, 2011, *Grammaire de l'émérillon Teko, Une Langue Tupi-Guarani de Guyane Française*. Langues et Sociétés d'Amérique Traditionnelle 10. Louvain: Peeters.

Rose, Françoise. 2018. *Toolbox lexical and textual database*.

Rose, Françoise, and Martine Vanhove, 2007, "Discours Rapporté Direct et Prosodie En Émérillon et En Bedja." presented at the Colloque CerLiCO Grammaire et Prosodie, Nantes.

Voeltz, F.K.E. & C. Kilian-Hatz (eds.), 2001, *Ideophones.* Amsterdam/Philadelphia: Benjamins.

WORKSHOP 9

Schedule: Thu 9.00-16.55 (Room 13)

Knowing in interaction: Fieldwork on epistemicity and intersubjectivity

Karolina Grzech, Henrik Bergqvist & Eva Schultze-Berndt
(Stockholm University, Stockholm University & University of Manchester)

Keywords: epistemicity, intersubjectivity, epistemic marking, pragmatic fieldwork, experimental methods

Epistemicity in language concerns expressions of knowing, including belief, attention, perceptual accessibility, attitude, and rights to knowledge, as well as their distribution among the speech-act participants. The distribution of knowledge of events has been investigated in terms of 'intersubjectivity', and more recently under the term 'engagement' (Evans et al. 2018). While the term epistemicity has a very wide applicability in linguistics and philosophy, the proposed workshop will mainly be concerned with investigating grammatical expressions reflecting this notion in the form of inflections, clitics, auxiliaries and particles. Lexical resources and supra-segmental phenomena (e.g. intonation) will be left aside.

Epistemic marking, as an instance of epistemicity in language, reflects how knowledge is distributed in interaction, including how knowledge states are expressed and tracked by discourse participants. It encompasses multiple functional categories, such as evidentiality, epistemic authority ('right to know or claim', cf. Heritage & Raymond 2005; Stivers et al. 2011), stance (e.g. Du Bois 2007) and the newly-proposed category of 'engagement' (Evans et al. 2018), among others. Despite the growing number of described languages where complex epistemic marking systems are attested (e.g. Gipper 2011; Bergqvist 2012; Zariquiey 2015; Grzech 2016), no systematic methodology exists for studying the use of such expressions (for a notable exception see Zeisler 2016). Moreover, methodological tools used for studying evidentiality concentrate on the truth-conditionality of the markers and their morphosyntactic properties in terms of scope and embeddability (cf. e.g. Peterson 2010), and are as such not sufficient for adequate description of intersubjective epistemic marking systems (e.g. Bergqvist 2016).

Bergqvist (submitted, cf. Bergqvist 2017) discusses the categorical overlap between epistemic modality, evidentiality, egophoricity, and engagement in terms of how these distinct categories reflect the allocation of epistemic authority. He argues that the qualification of the speaker's belief, perceptual access, and involvement constitutes ways of either claiming (direct, sensory access), or deferring epistemic authority by assigning it to someone else (e.g. reported speech), or by signalling reduced accessibility to the event in question (e.g. non-sensory access/uncertainty). Whether epistemic authority is indeed a central notion for the use of epistemic marking remains to be confirmed by empirical investigation into the use of such systems in discourse.

Semantic research based on fieldwork on lesser-spoken languages has been developing dynamically since Matthewson's (2004) seminal methodological paper. However, field semantics is still mostly concerned with truth-conditional phenomena (cf. Bochnak & Matthewson 2015), and, despite some attention being afforded to pragmatics in field methods volumes such as Chelliah and de Reusse (2011), studying language-in-use has yet to receive systematic and detailed treatment in the

literature for linguistic fieldworkers. Thus, the methodological advances in the field lag behind its technological development: with widespread access to affordable, good-quality video-recording devices and a rise in the number of digital archives, linguistic documentation can now provide accessible and transparent records of all kinds of communicative practices, from ceremonial speech to private conversations. One part of this field of research in which the mismatch between thriving theoretical interest and lack of methodological progress is particularly apparent is the study of epistemicity.

Moreover, the use of corpus data alone is not sufficient to investigate forms of epistemic marking. In order to test any hypothesis emerging from patterns attested in natural language use, we need elicitation materials/stimuli that target the relevant components of epistemic authoritativeness in terms of privileged access as relying on perception/experience, involvement, attitude, and expertise (see Kamio 1997 for a discussion of these components), as well as strategies related to epistemic stance construction (cf. e.g. Mushin 2001). Matcher-director tasks (e.g. the Map Task or the Diff-task, Enfield & De Ruiter 2003), collaborative problem solving using board games (e.g. *Mastermind*, Silva & AnderBois 2016), and the collective production of narratives (e.g. FPPT, San Roque et al. 2012), have all proved useful for eliciting epistemics, but more so for some forms of epistemic marking than others. One challenge that appears especially pertinent in the development of such methods and tools is how to consider the influence of social factors in the analysis of epistemic marking systems/strategies, since these may impact on rights to knowledge. When accounting for the distribution and meaning of epistemic marking, how can we situate socio-cognitive considerations against perceptual and spatio-temporally grounded accessibility?

This panel aims to bring together field workers and experimental linguists with an interest in describing epistemic marking systems, in order to discuss field methods and tools that can be used to document such systems. The main questions the panel will seek to answer are the following:

(1) What kind of data is needed to ground our analysis and understanding of epistemic marking systems in their communicative function as seen in everyday language use?
(2) What kinds of experimental stimuli should be developed to elicit epistemic paradigms in a fieldwork situation?
(3) What methods can be used to test the socio-cognitive relevance of our analyses of epistemic systems?
(4) How can we ensure that the methods we use in the field deliver results that can be used for comparative studies of epistemic marking systems?

In line with the questions listed above, we invite contributions centred around, but not limited to, the following topics:

- Experimental methods for eliciting forms of epistemic marking, such as epistemic modals and evidentials;
- Adaptation/use of existing tools and stimuli for work on epistemic marking systems;
- Methods for tracking knowledge states in experimental tasks and in natural discourse;
- Methods for establishing semantic distinctions within epistemic marking paradigms, in particular semantic components of intersubjectivity;
- Empirically testing the relationship between related functional categories, such as epistemic modality, evidentiality, egophoricity and engagement.

We believe that, in a field of research as complex as this one, it is important to learn from our errors, and therefore we would like to encourage contributors to talk not only about their successes, but also to discuss their failures and reflect on possible reasons behind them.

References

Bergqvist, Henrik (2012), Complex epistemic perspective in Kogi (Arawako-Chibchan). *International Journal of American Linguistics* 82 (1), 1–34.

Bochnak, M. Ryan and Lisa Matthewson (eds.) (2015), *Methodologies in semantic fieldwork*. Oxford/New York: Oxford University Press.

Chelliah, Shobhana Lakshmi & Willem Joseph de Reuse (2011), *Handbook of descriptive linguistic fieldwork*. Dordrecht/New York: Springer.

Du Bois, John W. (2007), The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: subjectivity, evaluation, interaction* (Pragmatics & Beyond New Series), vol. 164, 139–182. Amsterdam: John Benjamins.

Evans, Nicholas, Henrik Bergqvist and Lila San Roque (2018), The grammar of engagement I: framework and initial exemplification. *Language and Cognition* 10(1), 110–140. doi:10.1017/langcog.2017.21.

Gipper, Sonja (2011), *Evidentiality and Intersubjectivity in Yurakaré: an Interactional Account*, Nijmegen: MPI.

Grzech, Karolina (2016), *Discourse enclitics in Tena Kichwa: A corpus-based account of information structure and epistemic meaning*. SOAS, University of London, https://eprints.soas.ac.uk/24336/ (10 October, 2018).

Heritage, John and Geoffrey Raymond (2005), The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. *Social Psychology Quarterly* 68(1), 15–38.

Matthewson, Lisa (2004), On the methodology of semantic fieldwork. *International Journal of American Linguisitics* 70(4), 369–415.

Peterson, Tyler (2010), *Epistemic Modality and Evidentiality in Gitskan at the Semantics-Pragmatics Interface*. Vancouver: University of British Columbia.

Stivers, Tanya, Lorenza Mondada and Jakob Steensig (2011), Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada and J. Steensig (eds.), *The Morality of Knowledge in Conversation*, 3–24. Cambridge: Cambridge University Press.

Zariquiey, Roberto (2015), The encoding of addressee's perspective in Kakataibo (Panoan, Peru). *STUF - Language Typology and Universals* 68(2), 143–162.

Zeisler, Bettina (2016), Evidentiality, inferentiality, and speaker's attitude Questionnaire or exemplary set.
www.academia.edu/26073108/2016_Evidentiality_inferentiality_and_speakers_attitude_--_Questionnaire_or_exemplary_set (10 October, 2018).

The many paths to epistemic access

Eva Schultze-Berndt
(University of Manchester)

Keywords: epistemic authority; intersubjectivity; egophoricity; fieldwork methodology; language documentation

In this paper, I will reflect on the use of a range of corpus and elicited data in order to establish the functions of two clitics marking primary and shared epistemic authority, respectively, in

Jaminjung/Ngaliwurru, a Mirndi language of Australia (Schultze-Berndt 2017). These clitics are not obligatory, with most utterances unmarked for epistemic authority, and have no equivalent in speakers' spontaneous Kriol/English translations of any utterance containing them, which reflects their non-truth-conditional nature. Moreover, they proved virtually inaccessible to metalinguistic judgment: the interactional impact of both clitics alike was described by speakers (if at all) as 'letting someone know about something', and correspondingly, acceptability judgments outside a fully naturalistic context did not turn out to be reliable.

Since the language is rarely used in everyday interaction, there is also no large body of naturalistic conversational data to draw on. However, a general documentation corpus of around 50,000 words compiled over 25 years – consisting of personal and mythical narratives, staged and (some) naturalistic dialogues, and elicitations by means of visual stimuli or verbal prompts – proved sufficient to establish the function of the clitics by providing clear examples of different constellations of epistemic access. For example, personal narratives based on first-hand experiences almost exclusively featured the primary epistemic authority marker as illustrated in (1), while responses elicited by means of visual stimuli unfamiliar to the speaker, such as (2), frequently triggered the shared authority marker.

(1) *nenigot=biyang* *yirrurra-wardagarra-nyi=**ngardi*** *garrb*
goat=then 1PL.EXCL>3PL-follow-IPFV=EGO gather
'we used to follow goats around, and gather (them) [I'm in a position to tell you because I was there and you were not]' (beginning of a personal narrative about the speaker's work on cattle stations in her youth)

(2) *digirrij=jung* *ga-rdba-ny=**mindi***
die=RESTR 3SG-fall-PST=EGO+TU
'(The owl frightened the boy), and he fell down as if dead [or so it appears – you have access to the same evidence as me, so correct me if I am wrong]' (from a Frog Story narrative, based on a picture book)

Staged dialogues (role-play involving multiple speakers), and fictitious dialogues (conversations made up and enacted by a single speaker) complemented actual attested interactions in confirming the above analysis, as well as a further restriction of the shared epistemic authority marker to those contexts where a newly arising situation is not yet fully integrated into the common ground. In other words, this marker is not generally used for knowledge shared within the community, but only where the issue of the interpretation of a new situation arises and when the interlocutors share direct access to this situation. In the paper, various examples of the use of the clitics in their communicative contexts will be analysed to illustrate the conclusions that could be drawn from the different types of data. Overall, the results demonstrate the potential usefulness of staged communicative events in the sense of Himmelmann (1998) for investigations of this kind.

References

Himmelmann, Nikolaus P. (1998), Documentary and descriptive linguistics, *Linguistics* (36), 161-195.
Schultze-Berndt, Eva (2017), Shared vs. Primary Epistemic Authority in Jaminjung/Ngaliwurru, *Open Linguistics* (3), 178–218.

Conversational structure as evidence for regularity and variability in the use of epistemic markers

Sonja Gipper
(University of Cologne and Australian National University)

Keywords: conversation, epistemic marking, sequence organization, variation, Yurakaré

For any investigation of epistemic markers, there are at least two questions to answer: What is the invariable meaning of the markers, and what variable functions do they fulfill in conversation? This paper explores the usefulness of evidence from conversational structure for answering these questions. Three types of conversational evidence are considered: The sequential placement of the utterance containing the epistemic marker, the action performed by the utterance, and the treatment of the utterance by the interlocutor. The methodology is illustrated with data from Yurakaré (isolate, Bolivia) which possesses a complex system of epistemic markers. Findings from two different video-recorded corpora are reported: A corpus of everyday conversation (around 5 hours), and a corpus collected using the Family Problems Picture Task (San Roque et al. 2012, and Barth & Evans 2017), a problem-solving task designed to encourage speakers' use of expressions of social cognition, among them epistemic marking (around 6 hours). While there is enough regularity across speakers to identify the meanings of the epistemic markers, the data also reveal that speakers vary considerably regarding the epistemic markers they prefer. For instance, some speakers favor hedging expressions, while others prefer the use of boosters. Based on the conversational evidence, it is argued that this variation indexes the roles speakers take in interaction and the way they position themselves with respect to different aspects of the context. Speakers use epistemic markers to claim or grant epistemic authority, or to take a leading or following position in the discussion. Conversational structure is demonstrated to provide evidence for the regular and the variable aspects of epistemic markers in terms of their invariable meanings and the diverse uses to which they are put by speakers to position themselves in interaction.

Acknowledgments: This research is supported by DFG (Deutsche Forschungsgemeinschaft), grant number GI 1110/1-1. The data collections were funded by Volkswagen Foundation (grant numbers 81821 and 83448) and Nicholas Evans through his Anneliese Maier Research Award (Alexander von Humboldt Foundation). I am also very grateful to all Yurakaré speakers who participated in the data collections.

References

Barth, Danielle & Evans, Nicholas. 2017. The social cognition parallax corpus (SCOPIC): Design and overview. *Language Documentation & Conservation Special Publication No. 12: Social Cognition Parallax Corpus (SCOPIC)*, pp. 1-21.

San Roque, Lila, Gawne, Lauren, Hoenigman, Darja, Miller, Julia Colleen, Rumsey, Alan, Spronck, Stef, Carroll, Alice & Evans, Nicholas, 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation* 6, pp. 135-174.

Egophoric marking in interaction in Wutun: evidence from action combinations in video-recorded everyday conversations

Erika Sandman & Seppo Kittilä
University of Helsinki

In our talk we will discuss how studying video-recorded everyday conversations may shed light on the use of egophoric marking in social interaction in Wutun. We will focus on the pairings of verbal speech acts and physical actions, referred to as *action combinations* (Ford 2002). Wutun is a variety of Northwest Mandarin spoken by ca. 4000 people in Qinghai Province, Western China. Due to long term language contact, Wutun morphosyntax has been heavily influenced by Amdo Tibetan (Sandman 2016), and one of its most prominent Tibetan features is egophoric marking. In egophoric marking languages, one grammatical form (EGO) is associated with high degree of accessibility to the event in question, while one or several grammatical morphemes (NON-EGO) are associated with reduced accessibility to the event in question (Floyd et al. 2018).

Most of the earlier literature on egophoric marking languages states that ego form is typically used in first person statements and second person questions, while non-ego form is used elsewhere (see e.g. Hale 1980; DeLancey 1992; Creissels 2008). Example 1) illustrates a textbook example of egophoric marking in Wutun. The ego marker *-yek* is used in first person statement and second person question, while the non-ego marker *-li* is used in non-first person statement:

1) a. *ngu* *huan* *xhe-di-**yek***
1SG food drink-PROGR-**EGO**
'I am eating.'
b. *ni/gu* *huan* *xhe-di-**li***
2SG/3SG food drink-PROGR-**NON.EGO**
'You are eating/(S)he is eating.'
c. *ni* *ma-ge* *nian-di-**yek***
2SG what-REF read-PROGR-**EGO**
'What are you reading?'

The basic egophoric marking pattern can be explained as being due to the fact that speakers usually have privileged access to their own actions, and in second person questions the privileged access shifts from the speaker to the addressee. However, Wutun system is extremely flexible and any marker can be used with any person both in statements and in questions. Our goal is to investigate the possibility of combining the observations of verbal and physical actions in studying the interactional principles that underlie the choice of egophoric marking in naturally-occurring data. The data for the study consists of video-recorded everyday conversations collected among the Wutun speech community in autumn 2018. A striking feature of Wutun spoken interaction is the frequent use of ego marker *-yek* in third person statements in which the speaker describes the actions of others. Looking at the action combinations reveals that the statements with ego marker *-yek* often accompany the speaker's physical action directed towards the other participant(s) in a speech situation. For example, an adult can perform an action that is intended to prevent a baby from crying, and at the same time, switch to an ego marker while making a statement about baby's crying. Our preliminary results suggest that ego marker *-yek* in spoken interaction is often performative in nature; it can be used when the speaker has influenced to physical or mental states of others and therefore is in a privileged position of making statements concerning other person.

References

Creissels, Denis 2008. Remarks on so-called "conjunct/disjunct" systems. *Proceedings of Syntax of the World's Languages* III. Berlin, 25-28 September 2008.

DeLancey, Scott 1992. The historical status of the conjunct/disjunct pattern in Tibeto- Burman. *Acta Linguistica Hafniensia* 25, 39-62.

Ford, Cecilia E. 2002. Denial and the construction of conversational turns. In Joan L. Bybee (ed.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*. 61-78. Amsterdam: John Benjamin's.

Floyd, Simeon E., Elisabeth Norcliffe and Lila San Roque (eds.), *Egophoricity*. Typological Studies in Language 118. Amsterdam: John Benjamin's.

Hale, Austin 1980. Person markers: finite conjunct and disjunct forms in Newari. In R. Trail et al. (eds.), *Papers in Southeast Asian Linguistics* 7, 95-106.

Janhunen, Juha 2007. Typological interaction in the Qinghai linguistic complex. *Studia Orientalia* 101, 85-103.

Sandman, Erika 2016. *A Grammar of Wutun*. PhD dissertation, University of Helsinki.

ꕥ

Exploring Kogi epistemic marking in interactional elicitation tasks: A report from the field

Dominique Knuchel
(University of Bern)

Keywords: Epistemic marking, engagement, demonstratives, fieldwork methods, South American languages

This paper details different elicitation tasks that have been conducted with speakers of Kogi, an indigenous language of Northern Colombia, and discusses them with regard to their suitability for investigating epistemic marking. The expression of epistemicity, more precisely the recently proposed category of 'engagement', can be considered a distinctive characteristic of Kogi grammar. 'Engagement' refers to the encoding of shared/non-shared access in terms of knowledge, expectation (i.e. epistemic access) or attention (i.e. perceptual access) to a discourse object on part of the speech act participants (Evans et al. 2018).

In Kogi, engagement is manifested in a set of (ad)nominal demonstratives in the use of which shared attention of speaker and addressee towards a referent plays a crucial role (Knuchel, forthcoming). Moreover, engagement is encoded in the verbal domain in four auxiliary prefixes which express (a)symmetries between speech act participants in epistemic access to a state of affairs, and reflect two different parameters: epistemic authority (i.e. whose knowledge is targeted) and (non)-shared access (i.e. whether it is exclusive to one of the speech act participants, or shared by both) (Bergqvist 2016).

In order to study the use of these forms, a number of interactional stimuli-based tasks were carried out in which epistemic/perceptual access to a referent or state of affairs can, to a certain degree, be determined by the task design, and which allow for the observation of the participants' attentional states. A matching game based on the 'Shape Classifier Task' (Seifart 2003) provided particularly revealing data about the use of a demonstrative associated with joint attention. Epistemic prefixes

were observed in the Family Problems Picture Task, an interactional story-telling task (FPPT, San Roque et al. 2012), where the markers are frequently used in the enactment of conversations between characters. The talk discusses the employed methods and obtained data, and furthermore introduces a set of picture stimuli currently under development. These stimuli are inspired by parts of the FPPT and are designed to elicit enacted conversations between characters that are involved in situations with assumed shared or non-shared access to knowledge.

References

Bergqvist, Henrik. 2016. Complex epistemic perspective in Kogi (Arawako-Chibchan). *International Journal of American Linguistics* 82(1). 1–34.

Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018. The grammar of engagement I: framework and initial exemplification. *Language and Cognition* 10(1). 110–140. doi:10.1017/langcog.2017.21.

Knuchel, Dominique. forthcoming. Kogi Demonstratives and Engagement. *Open Linguistics.*

San Roque, Lila, L. Gawne, D. Hoenigman, J. Miller, S. Spronck, A. Rumsey, A. Carroll and N. Evans. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation* 6, 135–174.

Seifart, Frank. 2003. Encoding shape: Formal means and semantic distinctions. In Nick Enfield (ed.), *Field research manual 2003 part I: Multimodal interaction, space, event representation*. Nijmegen: Max Planck Institute for Psycholinguistics. 57–59.

Blackmail, Biblical hearsay and fortunetelling: How thought experiments and translation analysis clarify Aymara evidential/epistemic marking

Matt Coler, Nicholas Q. Emlen & Edwin Banegas-Flores
(University Gronignen – Campus Fryslân, University of Tübingen – DFG Center for Advanced Studies, Peruvian Ministry of Education)

Keywords: Aymara, evidentiality, evidential/epistemic marking, fieldwork methods, translation analysis

Aymara is a highly agglutinative Andean language with rich morphological resources for evidential/epistemic marking. However, the contextually embedded nature of such phenomena makes them difficult to capture through standard elicitation methods. In this contribution, we describe two methods for analyzing evidential/epistemic marking: thought experiments and the analysis of translated narratives that include the voices of multiple characters as reported speech. We show that these methods, taken together, illuminate evidential/epistemic marking in the language.

The first method involves thought experiments based on local practices. In these thought experiments, native Aymara speaking participants consider imaginary situations, described by the fieldworker in regional Spanish which were designed to require the participant to report on events which would be impossible to elicit naturalistically (Coler, 2014). These thought experiments include cases of blackmail, mistaken identities, corrupt fortunetellers, and a variety of situations in which firsthand knowledge contradicts secondhand knowledge and vice-versa. As Aymara cannot relay indirect speech, thought experiments proved to be a reliable way to get a quotation directly from a

speaker. For example, in (1) participants speculated whether Akhawana paid his electric bill. Although the lights in Akhawana's house are on (had he not paid, the electricity would have been cut), no one witnessed him paying the collector. In (2), participants imagine that the lights of the house are off and though no one witnessed Akhawana paying the collector, he is known to be responsible, so he likely paid. In (1), the conjectural suffixes *–chi* and *–jalla* are used, whereas in (2) no conjectural suffixes are used.

1. paga-ni-wj-ch-i-s-jalla
   pay-CISLOCATIVE-BUFFER-CONJECTURAL-3SUBJ.3OBJ.SIMPLE.TENSE-ADDITIVE-CONJECTURAL2
   'He must have paid him.'
2. paga-ni-wjama-wj-rak-ta-s
   pay-CISLOCATIVE-VERBAL.COMPARATIVE-BUFFER-ADDITIVE-2SUBJ.3OBJ.SIMPLE.TENSE-ADDITIVE
   'It seems you also paid him up.'

The second method that we describe draws on a 600-page text about the life of Jesus Christ, called the *Vita Christi*, written by a native Aymara speaker named Martín de Sancta Cruz the early 1600s. Sancta Cruz created this Aymara text by translating an earlier Spanish book, and he filled his translation with the direct reported speech of countless characters as they interacted with each other. This reported speech features a wide range of markers that attribute different kinds of evidential and epistemic statuses to the information presented by the characters. For instance, in one typical passage, the narrator quotes King Herod, who in turn quotes the three kings of the Magi; each of these embedded voices has different epistemic and evidential markers that create a dense intersubjective network regarding who knew what, and how. By comparing the Aymara translation with the Spanish original, we show that a systematic analysis of the narrative sheds light on how evidential and epistemic marking interacts with the reported speech of the narrator, the shared knowledge of the story hearers, and between the characters within the narratives themselves. This corpus of data, therefore, represents a sort of historically available thought experiment similar to that described above, and offers an invaluable source for interpreting Aymara epistemic/evidential marking, as well as its history.

References

Coler, M. (2014). *A grammar of Muylaq' Aymara: Aymara as spoken in Southern Peru*. Leiden: Brill.

Testing the hypothesis that Engagement is grammaticalized in Kari'nja

Berend Hoff, Spike Gildea & Racquel-Maria Sapién

<pdf>

The frequency and distribution of modal particles in spoken Swedish

Henrik Bergqvist
(Stockholm University)

Keywords: modal particles, Swedish, engagement, distribution, frequency

The paper investigates the frequency and distribution of modal particles in spoken Swedish in order to argue for their status as epistemic markers managing the distribution of knowledge in interaction, i.e. "engagement" (see Evans et al. 2018a). Engagement is defined as "a grammatical system for encoding the relative accessibility of an entity or state of affairs to the speaker and addressee" (Evans et al. 2018a, 118). As such, engagement targets (a)symmetries in the speaker-hearer dyad with respect to knowledge of events. The Swedish modal particles *ju* ('as you know'/'of course') and *väl* ('right?'/'isn't it?') are used to manage inter-personal access to knowledge and may therefore be viewed as markers of engagement according to the above definition. They are also paradigmatically contrastive in expressing "shared access" to an event from the point of view of one of the speech-act participants. *Ju* signals shared access to an event and at the same time places the epistemic authority with the speaker, whereas *väl* signals shared access to events placing the epistemic authority with the addressee (Bergqvist, in prep., cf. Bergqvist 2017). Both *ju* and *väl* are frequent in spoken Swedish, comparable to some pronouns forms where *ju* is as frequent in spoken Swedish as the second person pronoun *du* (GSLC, cf. Allwood 1999a). Given that Swedish grammar features obligatory subject marking, this fact is worth noting. With respect to the distribution of both forms, *väl* occurs significantly more often with the second person pronoun *du*, and significantly less with the first person pronoun *jag*, when compared to *ju*. Taken together, these observations of frequency and distribution align with the notion of knowledge (a)symmetry in engagement marking where the speaker and/or the addressee may be targeted as the epistemic authority in making reference to an event (see Bergqvist 2016, for Kogi). The paper argues that taking into account the distribution and frequency of epistemic markers in various forms of spoken interaction constitutes an important aspect of the analysis of such forms.

References

Allwood, Jens (1999a), The Swedish Spoken Corpus at Göteborg University, In *Fonetik 99: Proceedings from the 12th Swedish Phonetics Conference,. Gothenburgh Papers in Theoretical Linguistics 81*, Göteborg University: Department of Linguistics.

Bergqvist, Henrik. (in preparation), Assigning epistemic authority using modal particles in spoken Swedish, MS.

Bergqvist, Henrik (2017), The role of 'perspective' in epistemic marking, *Lingua*, 186-187, 5-20.

Bergqvist, Henrik (2016), Complex epistemic perspective in Kogi (Arwako), *International Journal of American Linguistics*, 82(1), 1-34

Evans, Nicholas R., Henrik Bergqvist, and Lila San Roque (2018a), The grammar of engagement I: framework and initial exemplification, *Language and Cognition*, 110–140.

Evans, Nicholas R., Henrik Bergqvist, and Lila San Roque (2019b), The grammar of engagement II: typology and diachrony, *Language and Cognition*, 141–170.

Göteborg Spoken Language Corpus (GSLC), https://data.flov.gu.se/gslc/. University of Gothenburgh, Department of Philosophy, Linguistics and Theory of Science.

Assessing acquisition of epistemic markers and its interplay with sociocognitive development

Ditte Boeg Thomsen, Birsu Kandemirci & Silke Brandt
(Lancaster University)

Keywords: intersubjective markers, discourse particles, evidential morphology, social cognition, L1 acquisition

Children's acquisition of epistemic marking is intimately related to their developing sociocognitive skills for monitoring and reasoning about mental states (e.g. Aksu-Koç et al. 2009). To understand to which degree acquiring language-specific epistemic markers *depends on* and/or *supports* social cognition, it is crucial to be able both to pinpoint which distinctions are relevant in a specific epistemic system and to test children's mastery of them. This talk discusses three tools for assessing proficiency with epistemic markers as well as results from comparisons with sociocognitive development.

First, we present a gap-filling test targeting Danish double-perspective particles that mark constellations of knowledge and attitude between the interlocutors (or a third person), as in *jo* (shared knowledge), *da* (shared knowledge, opposing viewpoints) and *vel* (speaker uncertainty, privileged interlocutor access) (Davidsen-Nielsen 1996). To validate semantic analyses of these intersubjective meanings, the gap-filling test checked adult agreement with small vignettes building up contrasting configurations of shared knowledge, disagreement and differential access to knowledge (Boeg Thomsen 2012). While constructing unequivocal contexts excluding different ascriptions of mental states to the story characters is nontrivial, the final 15-item version showed solid agreement among 101 adults, demonstrating the usefulness of the format for pinpointing particle meaning. Responses from schoolchildren (N = 167) further demonstrated adult-like sensitivity to the relevant intersubjective dimensions.

Second, we present results from an ongoing crosslinguistic study examining the relationship between children's acquisition of evidential morphology and their ability to monitor and remember the sources of their knowledge. Children acquiring Turkish must learn to mark knowledge source with evidential suffixes when talking about the past, and we examine whether children's use of *-DI* (direct evidence) and *-mIş* (inference) in two production tasks (Ögel 2007) is related to their performance on a source-monitoring task that tests ability to monitor mode of knowledge access: seeing, inferring or being told (Gopnik & Graf 1988, Ögel 2007). We compare 50 Turkish-speaking three- and four-year-olds with 50 English-speaking three- and four-year-olds to test whether the Turkish-speaking children are more likely to mark knowledge source in the linguistic production tasks and to attend to mode of knowledge access in the source-monitoring task, and we examine the relationships between these specific skills and children's short-term memory and mental-state-reasoning skills.

Third, we briefly discuss a comprehension task, the Hidden Objects task, that is often used for assessing children's appreciation of the certainty distinctions marked by epistemic modals and evidential expressions (e.g. Moore et al. 1990). We point out pragmatic confounds and suggest ways to improve the validity of the task.

Finally, we discuss the possibilities for and the challenges in adapting the linguistic and sociocognitive tests to other languages to compare developments crosslinguistically and investigate to which degree language-specific epistemic marking influences and depends on specific aspects of sociocognitive development.

References

Aksu-Koç, Ayhan, Hale Ögel-Balaban, and I. Ercan Alp (2009), Evidentials and source knowledge in Turkish, in S. A. Fitneva, and T. Matsui (eds), (2009), *Evidentiality: A Window into Language and Cognitive Development, New Directions for Child and Adolescent Development*, San Francisco: Jossey-Bass, 13–28.

Boeg Thomsen, Ditte (2012), Viet til vinkler, viernes vinkler: typisk udviklede danske børns beherskelse af synsvinkelmarkerende diskurspartikler [Typically Developing Danish Children's Mastery of Perspective-Marking Discourse Particles]. Master's thesis. Copenhagen: University of Copenhagen.

Davidsen-Nielsen, Niels (1996), Discourse particles in Danish, in E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft, and L. Falster Jakobsen (eds), (1996), *Content, Expression and Structure: Studies in Danish Functional Grammar*, Amsterdam / Philadelphia: John Benjamins Publishing Company, 283–314.

Gopnik, Alison, and Peter Graf. (1988), Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development* 59: 1366–1371.

Ögel, Hale (2007), Developments in Source Monitoring and Linguistic Encoding of Source. Master's thesis. Istanbul: Bogazici University.

Moore, Chris, Kiran Pure, and David Furrow (1990), Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development* 61: 722–730.

Consciousness, context, and epistemicity

Marianne Mithun
(University of California, Santa Barbara)

Keywords: evidentials, grammaticalization, documentation, interaction, metalinguistic knowledge

The more we learn about the grammatical expression of epistemicity across languages and cultures, the clearer it becomes that multiple considerations must enter into our techniques for documentation. Two factors discussed here are a) the variety of contexts in which markers are used, and b) our understanding of what is in the minds of the speakers we work with. Material cited here comes from nine years of work with speakers of Central Pomo, indigenous to California. None had ever read or written the language, nor thought about its structure.

One important variable is context. Evidentials may be completely absent from isolated elicited sentences. Hearsay evidentials may be pervasive in legends but rare in procedural texts. Non-verbal sensory evidence may never be mentioned in either. And claims of certainty may vary widely within conversations depending on the topics under discussions and relations among participants. Furthermore, context can be crucial to a full understanding of the larger discourse and social functions of the markers. Many Central Pomo evidentials, such as the hearsay and factual markers, for example, play important roles in structuring discourse. Quotatives interact with logophoric pronouns.

A second major variable is the degree of grammaticalization of the markers. Central Pomo evidentials range from particles still close in form and meaning to their lexical sources, through loosely bound second-position clausal enclitics, to final markers tightly bound to their verbal hosts. Free particles still close to their origins were easily recognized. Clausal enclitics went largely unnoticed, but became more accessible once pointed out. Verb endings were another matter. As we

transcribed recordings, sentences appeared like *ʟo: šʔú:č'ka* and *ʟo: šʔúčʰya*, both translated by speakers as 'I forgot'. Asked about the difference, the speakers typically answered that there was none. A speaker translated *dačéya* 'he caught it'. Asked how this differed from *dačéw*, seen earlier, she responded that *dačéya* is 'he just caught it'. Tense? We later encountered *dačéwla* 'I just caught it'. Person? In fact *=ka* marks inference, *-ya* direct observation, *-w* perfective aspect, and *=la* the certainty of speakers who have themselves acted voluntarily. The evidentials are simply more appropriate for sudden discoveries than established situations. The speakers used these markers flawlessly in speech, but were unaware of their existence.

Such variables suggest that a single experimental procedure may not give us everything. Documentation of extensive unscripted speech, in a variety of genres, much of it interactive, continues to be important, not only for spotting markers, but also for discovering subtle functions beyond indicating source and certainty of information. Of course knowledge on the part of the linguist of the epistemic distinctions that occur cross-linguistically can facilitate identification of them in a new language and point the way to further issues to investigate. And consideration of the metalinguistic knowledge of speakers can be crucial. Evidentials are often below the level of consciousness and may pass by unnoticed, but once speakers are alerted, their insights can greatly enrich our understanding of their roles in the language and culture.

WORKSHOP 10

Schedule: Thu 9.00-16.55 (Room 10)

Language as Network: How to conceive of a usage-based constructicon

Or: What's in the constructicon? The nature of the syntactic/constructional networks in language development and theory

Heike Behrens & Karin Madlener
(University of Basel)

Keywords: construction grammar, network models, usage-based models, language acquisition, formulaic language

Modern structuralist linguistics has been looking for rules and generalizations. The units of analyses are defined for different structural levels (like morphology, syntax, phonology), and the lexicon that feeds into these structures. This led to the postulations of different units of representations on different layers, and— depending on the model—interfaces to mediate between them. An implication of the search for rules and compositionality is that irregular, non-compositional phenomena cannot easily be accounted for and are subject to additional mechanisms (e.g., holistic storage rather than compositionality). A further, and perhaps more fundamental assumption is that the speakers of a language share and access the same grammar.

Construction grammar questions the existence of separate linguistic layers of representation as well as the divide between lexicon and grammar, and instead works with form-function units (or constructions) of different size, compositionality, and complexity (Ellis 2002, Goldberg 2009). The advantage of this account is that it can integrate regular and irregular phenomena by analyzing the degrees of freedom in the range of a particular construction, and that it can explain linguistic change (both in the historical dimension as well as in the individual dimension in language learning) by analyzing the factors that determine the intake and uptake of linguistic forms and the changing form-function relationships. On a theoretical level, this ideally leads to a small descriptive apparatus that accounts for a variety of phenomena. The first aim of this workshop is to delineate more clearly which theoretical constructs are needed to account for the grammatical and functional relations within and between constructions.

A further advantage lies in the usage-based nature of this theory: Learners learn language from the actual-usage events they are involved in, and can generalize more abstract knowledge by noticing analogies or relations between constructions (Abbot-Smith & Tomasello 2006). This entails that speakers can, but need not represent or activate all possible grammatical relations. In terms of processing, the access to formulaic or chunked units can speed up access and lead to more reliable performance (e.g., Wray 2017). This also entails that there are individual differences. In sum, the assumption of a constructicon with network relations is attractive because it possibly minimizes the descriptive apparatus we need for language and because this apparatus is based on empirical findings from language learning, use, and processing. In addition, the concept of form-function units with flexible, probabilistic relations means that change (historical, developmental) and variation (individual, regional, social) can be accounted for by shifts in the distributional patterns in the speaker's experience and production.

In this workshop we want to flesh out such a model of language by looking at (1) the theoretical specifications of the constructicon, (2) the online access of constructions and their representations, and

(3) individual differences resulting from the usage-based nature of linguistic representations. We will elaborate on the notion of linguistic networks by looking at the following issues:

1. Linguistic Theory: What is the nature of linguistic relations between form-function units?
As a more general model of language, a so-called "constructicon" as an assembly of all constructions in a language has been proposed (Goldberg 2009) and it is assumed that the constructions are related in a network. However, it is less clear how to think of this network both in theoretical terms (inheritance links, structural classes?) and in empirical research (how can we show that users process exemplars/chunks vs. more abstract schemas, that they make use of these networks, what is their mental representation). Finally, what would a network-model of language look like?

2. Online Access of Units: Chunking and Dechunking
It seems that the same utterance can be produced with different mental representations: as pre-specified chunk, or generated from scratch (Bybee & Scheibman 1999, Wray 2017). Though chunks can be analyzed compositionally, the question is whether and when speakers access the holistic or the compositional representation. Chunking has attested advantages in cognitive processing (increased speed in retrieval/processing and increased reliability). If that is so, what exactly is the added value of "abstract" linguistic knowledge in processing? What are the conditions when we access it (e.g., communicative situations that require pattern extension and creativity)? What is the evidence for and advantage of language processing with low-level schemas?

3. Individual differences in the representation of linguistics knowledge
If linguistic knowledge is derived from language use, it follows that speakers of the same language may differ in their mental representation. Recent research has pointed out that individual differences in language acquisition, language representations, and language attainment exist (Kidd, Donelly & Christiansen 2017; Dabrowska 2015). What is the nature and extent of individual representation and what do individual differences entail for linguistic theory? How can we access the level and nature of online processing experimentally, given that there is growing evidence that linguistic knowledge is speaker-dependent both in the extent of linguistic knowledge and in the analytical depth of processing the linguistic signal?

References

Abbot-Smith, Kirsten & Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23, 275–290.

Ellis, Nick C. 2002. Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143–188.

Goldberg, Adele E. 2009. The nature of generalization in language. *Cognitive Linguistics* 20 (1), 93–127.

Wray, Alison. 2017. Formulaic Sequences as a Regulatory Mechanism for Cognitive Perturbations During the Achievement of Social Goals. *Topics in Cognitive Science* 9, 569–587.

Filler-slot relations and the cognitive organization of syntactic categories

Holger Diessel
<pdf>

Constructions as usage patterns, constructions as cognitive routines

Ewa Dąbrowska
(Friedrich-Alexander-Universität Erlangen-Nürnberg and University of Birmingham)

Should we see constructions as patterns of language use, or as cognitive routines for producing or comprehending utterances? Since patterns in language use arise from individual speakers producing and comprehending utterances, most cognitive linguists assume that there is a fairly direct correspondence between the two: that is to say, usage patterns exist because speakers have acquired certain generalizations about their language and apply them when producing utterances.

In this presentation, I argue that this is not always the case, and therefore, it is important to distinguish between the two perspectives. I begin with a discussion of two instances of patterns which are arguably present in the language but are not explicitly represented in speakers minds: a statistical generalization about the distribution of Polish masculine genitive endings (Dąbrowska 2008) and constraints on questions with long distance dependencies in English (Dąbrowska 2013). I argue that such discrepancies arise partly because linguistic knowledge, including grammatical knowledge, is distributed: that is to say, individual speakers know only a subset of the constructions that are arguably present in the language and partly because different speakers may represent "the same" knowledge at different levels of abstraction. In addition, patterns may arise not because speakers "know" something, but simply because they do what is easiest for them.

Thus, when we study units of language in cognitive terms, we need to distinguish between constructions in the sense of recurring form-meaning pairings and constructions in the sense of the cognitive routines that underlie language use. This has interesting consequences. Different speakers (or even the same speaker on different occasions) may produce the same utterance in different ways. Furthermore, the cognitive routines used in production are almost certainly different from those used in comprehension.

Constructions as attractors in neural networks - a modelling perspective

Lars Konieczny
(University of Freiburg)

Keywords: Construction grammar, local syntactic coherence, recurrent networks, sentence processing, context effects

The *constructicon* as a network is a promising hypothesis about how linguistic knowledge is represented and organized in the mind. The cognitive nature of these representations is however far from clear. How precisely is construction knowledge acquired and put to use in comprehension and production? From a connectionist point of view (Elman, 2009), constructions can be construed as locations (attractors) in a large dimensional state space. In this view, the *constructicon* can be conceived of as the multi-dimensional landscape of constructions emerging from and being shaped by exposure to language. Comprehension and production amounts to passing through the landscape, where each new input pulls into a new direction.

We will present a series of neural network simulations to illustrate the connection between construction acquisition and processing. Recurrent networks trained with the task to predict the next event in a sequence of linguistic inputs will not only implicitly acquire construction-like knowledge representations, they will also make systematic prediction errors based on local, i.e. intra-constructional, transitional probabilities. These *local syntactic coherences* (LSC; Tabor et al. 2004, Konieczny 2005, Konieczny et al. 2009, Konieczny et al. 2010) can be shown to affect human sentence processing, which will be substantiated by results from a series of eye-tracking experiments.

German verb-final subordinate clauses like (1) were presented. In that clause, an ambiguous adverb ("erfreut"/ *delightful*) succeeds a dative pronoun and a nominative NP. Form-wise, the adverb could also be a finite verb (*delights*). As a finite verb, it could locally be combined with the proper noun *Daniela* and the subsequent NP *her boyfriend* to form a short transitive sentence (*Daniela delights her boyfriend).* This combination however should be ruled out by the preceding dative pronoun *him*, which cannot be integrated in such a construction (*ihm Daniela erfreut ihren Freund*/ him Daniela delights her boyfriend*).

(1) Als ihm **Daniela erfreut ihren Freund** vorstellt, lächelt Mark über das ganze Gesicht.
*When him **Daniela** delightful/ **delights her boyfriend** introduces, smiles Mark over the all face*.
(German word order)
When Daniela delightful introduces her boyfriend to him, Mark smiles all over the face.

Experimental target sentences like (1) containing a LSC were contrasted with controls, where the ambiguous adverb ("erfreut") was replaced by an unambiguous synonymous adverb („heiter"/ *lively*). The results of our experiments consistently suggest that the sequence of words *Daniela delights her boyfriend* activates a transitive construction – even when it is embedded in a sentence where it would be rendered ungrammatical – hence interfering with the processing of global grammatical dependencies. Moreover, this effect was elevated the sentences were preceded by short contexts (2) containing propositions that entailed the local transitive construction meaning.

(2) Danielas new boyfriend **ate nothing better** than Tiramisu. To **make him happy**, Daniela had made some herself. Mark was also invited to dinner. (translated from German)

This result suggests that a construction can become more likely to falsely interfere with syntactic processing when its meaning has been primed by the context. As the result is compatible with a construction grammarian view on sentence processing, we will finish by discussing the implications for a neural network approach to constructions and the *constructicon*.

Acknowledgments
This work was supported by the German Research Foundation (DFG, grant KO 1932/3-1)

References

Elman, J. L. (2009). On the Meaning of Words and Dinosaur Bones: Lexical Knowledge Without a Lexicon. *Cognitive Science*, *33*, 1-36.

Konieczny, L. (2005). The psychological reality of local coherences in sentence processing. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, pages 1178–1183.

Konieczny, L., Müller, D., Hachmann, W., Schwarzkopf, S., and Wolfer, S. A. (2009). Local syntactic coherence interpretation. evidence from a visual world study. In Taatgen, N. and van Rijn, H., editors, *Proceedings of the the 31st Annual Conference of the Cognitive Science Society*, pages 1133– 1138, Austin, TX.

Konieczny, L., Weldle, H., Wolfer, S. A., Baumann, P., and Mu¨ller, D. (2010). Anaphors and local coherences. In Ohlsson, S. and Catrambone, R., editors, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pages 1204–1209, Austin, TX. Cognitive Science Society.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of Merely Local Syntactic Coherence on Sentence Processing. *Journal of Memory and Language, 50 (4)*, 355-370.

(Rethinking) Representations of case

Maciej Borowski, Adnane Ez-zizi, Petar Milin & Dagmar Divjak
(University of Birmigham)

Keywords: Polish genitive, computational simulation, behavioural experiment, Naive Discriminative Learner, usage-based approach

Constructions are units of language defined as form–meaning pairings, connected by inheritance links. Goldberg (1995) argued that the linguist's task should be to distinguish all and only those constructions that speakers use. Unfortunately, this principle was not supported with a reliable methodology that would yield replicable results. The genitive case in Polish is the perfect illustration of a phenomenon that would benefit from such methodological advancement. For example, Swan (2002) describes 10 main 'uses' of the genitive with various meanings that can be divided further. This raises the question of where to draw the line: how many possible subgroups should be treated as distinct constructions and are all of them linked to the same overarching genitive category? Since the genitive case has so many manifestations and each of those manifestations can fulfil so many different functions, naive speakers may never unite all of them into one abstract category.

We address these questions by combining the results of computational simulations and behavioural experiments. Using a computational simulation study, we determine what level of generalization speakers could arrive at, given the usage data we have from the corpus. Our computational approach uses a simplified version of the Delta learning rule (cf. Widrow & Hoff 1960; Rescorla & Wagner 1972), as implemented in the Naive Discriminative Learner (NDL: Baayen et al. 2011; Milin et al. 2017). The model is trained on the 1.2 billion-word Araneum corpus (Benko 2014) to predict a noun in the genitive case as it occurs in naturalistic sentences, using other words in that same context as discriminative cues. The weight matrix generated by NDL is used to group similar nouns together. An in-depth analysis of the most discriminative contextual cues for each of the clusters of Polish genitive nouns sheds light on the functions that could be discriminated by naive speakers based on exposure to language.

The results from the computational simulation study are then corroborated with experimental evidence. We present data from a categorization study conducted to determine if and when abstract linguistics concepts such as case become important (Divjak et al. 2015). In the experiment, native speakers of Polish sort sentences from the corpus into groups according to similarity. Groupings are possible along a number of dimensions (lexical, functional, grammatical) and regression modelling shows whether the grammar-based functions or the usage-based clusters best predict the groupings the participants detect in the data.

We argue that combining computational simulations and experiments is a step towards a more principled and less subjective way of modelling language – one which, as illustrated by the case of

Polish genitive, provides valuable insights into the kinds of language units that might lay claim to cognitive plausibility.

This research was supported by a Leverhulme Trust Research Leadership Award (RL-2016-001) to Dagmar Divjak.

References

Baayen, Harald R., Milin, Petar, Filipovic Durdevic, Dusica., Hendrix, Peter & Marelli, Marco, (2011), An amorphous model for morphological processing in visual comprehension based on naïve discriminative learning, *Psychological Review* 118, 438-482.

Benko, Vladimir (2014), Aranea: Yet Another Family of (Comparable) Web Corpora, in P. Sojka, A. Horák, I. Kopeček, and K. Pala, K (eds), (2014), *Text, Speech and Dialogue. TSD 2014. Lecture Notes in Computer Science, vol 8655*, Springer, Cham: 247-256.

Divjak, Dagmar, Szymor, Nina, Socha-Michalik, Anna, (2015), Less is More: possibility and necessity as centres of gravity in a usage-based classification of core modals in Polish, *Russian Linguistics. International journal for the study of Russian and other Slavonic Languages*, 39 (3), 327-349.

Goldberg, Adele. E., (1995*), A Construction Grammar Approach to Argument Structure*, Chicago: University of Chicago Press.

Milin Petar., Feldman, Laurie B., Ramscar Michael, Hendrix Peter, Baayen Harald R., (2017), Discrimination in lexical decision. *PLOS ONE* 12(2): e0171935.

Rescorla, Robert A. and Wagner, Allan R., (1972), A theory of Pavlovian conditioning:

Variations in the effectiveness of reinforcement and nonreinforcement, in A.H. Black, and W.F. Prokasy (eds), (1972), *Classical Conditioning II: Current Research and Theory*, New York: Apple-Century-Crofts, 64-99.

Swan, Oscar E., (2002), *A Grammar of Contemporary Polish*, Bloomington: Slavica Publisher.

Widrow, Bernard and Hoff, Marcian E., (1960), Adaptive Switching Circuits, *IRE WESCON Convention Record*, 4, 96-104.

A Dynamic Network Analysis of Emergent Grammar

Paul Ibbotson
(The Open University, UK)

Network Structure; Grammar; Emergence; Dynamics; Modelling

We explored the extent to which grammatical constructions can be an emergent property of how simpler linguistic elements interact with one another. To do this we instantiated a corpus of early child directed speech into a dynamic network. We allowed the network to grow word by word as the mother uses her language in a corpus of naturalistic speech. We measured whether organizational properties of the network, specifically *community structure*, map onto grammatical patterns in any way that a child could plausibly capitalize on when constructing their language (Figure 1). Unlike other network approaches to learning the grammatical clusters that emerged did so in an unsupervised fashion and with no a priori constraint on the number of hidden layers relevant for the particular learning task.

Neither did the approach call for any specific learning biases of word learning models other than the general capacity to represent words, the transitions between them, and cluster frequently co-occurring words together. We found network communities emerged with distinct grammatical pathways when the network itself was blind to grammatical information. Because grammatical patterns in the input disassociated by community structure, they provide a potentially valuable foothold into the who-did-what-to-whom of grammatical relationships for the learner. We believe the incremental growth of the network captures something fundamentally developmental and complex (in the sense of many interacting parts) about the process of language acquisition, that neither batch-processing of corpus data nor non-dynamic models of development can. Networks also offer a highly plausible psychological medium in which to simulate cognitive processes because, like language, the brain itself is a complex dynamic network. Furthermore, this approach helps to visualize what a structured inventory of constructions – conceptualized as some sort of organized network of linguistic form and function – might actually look like. If network communities show distinct grammatical characteristics then the dynamic network approach suggests some of language's complexity (grammar) can be an emergent property of how simpler elements (words) interact with one another. It would also suggest that early grammatical patterns can be represented at a level that is grounded in the distributed properties of the network. We present these initial findings as proof-of-concept in the hope that other researchers will explore the possibilities and limitations of this approach on a larger scale and with more languages.

Figure 1. *An example of within module trigram grammatical patterns, showing a close-up of a community identified in the CDS network. From these trigram maps we characterized some of the most typical grammatical patterns for trigrams within communities, for example, the preposition→determiner(article) →noun for the pathway highlighted in red below.*

Form-meaning relations in acquisition: the case of polysemous constructions

Anna Theakston, Heather Lemen, Kimberley Bell, Silke Brandt & Elena Lieven
(University of Manchester, University of Manchester, University of Manchester, Lancaster University, & University of Manchester)

Keywords: acquisition, modals, causal/conditionals, polysemy, usage-based

Acknowledgements: This work was funded by the UK Economic and Social Research Council (ESRC ref. ES/L008955/1) and an ESRC PhD studentship to K. Bell (ref. 10054247)

Many constructions are associated with a variety of pragmatic meanings, posing a challenge in acquisition. How do children learn the meanings of these constructions, and what factors influence their production and comprehension of the various forms? Here we consider two construction types; modals, and causal/conditional adverbial sentences. Individual modals (e.g. *can, must*) can convey a variety of deontic (e.g. ability – *He can reach the book*, permission – *He can go to the park now*) and epistemic meanings (e.g. *He can read French*) (Papafragou, 1998), while causal and conditional adverbial sentences explain real-world causality (Content, e.g. *You're hungry because you didn't eat*); justify/restrict conclusions (Epistemic, e.g. *She must not have eaten, if there is food left*); or

justify/restrict speech acts (Speech-Act, e.g. *Eat your dinner, because you're hungry*) (Sweetser, 1990). We ask to what extent the frequency distribution of form-meaning mappings in the input children hear predicts (i) the form-meaning mappings they produce and (ii) their comprehension of these forms with different meanings. Specifically, do children start out by producing and comprehending those mappings which they hear most frequently?

To address (i), we conducted two corpus studies. Utterances containing modals or the adverbials *because/if* were extracted from dense naturalistic corpora of mothers interacting with their 3-5yr-old English-speaking children. These utterances were coded for their form and pragmatic meaning, and the distribution of form-meaning mappings in the input and in the children's speech was compared. We found that for all speakers, different modal/causal/conditional forms were associated with a different distribution of form-meaning mappings (e.g. *because* was most frequently produced in Speech-Acts, while *if* most frequently encoded Content). Some modals (e.g. *may*) appeared with only one or two meanings, whereas others (e.g. *can*) took many different meanings. In general, the children's patterns of usage mirrored those in their input (e.g. frequency of modal form-meaning pairings in the input predicted use in the children's speech); children did not start out producing only the most frequently encountered form-meaning mapping for each form. On the other hand, there were also some differences (e.g. in the use of *because* and acquisition of epistemic modal meanings) which suggests that immediate discourse priorities and the children's cognitive capabilities interact with distributional information to determine the patterns of acquisition.

To address (ii), we conducted two comprehension experiments to explore whether 3-5yr-old children's interpretation of modal *can/must* and adverbial *because/if* constructions varies as a function of their frequency of exposure to different form-meaning mappings in the input. Both studies utilize a forced-choice picture selection task, gathering picture choice and reaction time data. Bayesian analysis suggests that for conditional constructions, in line with the corpus data where we observed frequency asymmetries in form-meaning associations, children's performance was better for higher frequency mappings, although this pattern was not observed for causal constructions. For modal constructions, data analysis is underway.

Both sets of results will be discussed in the context of the role of input frequency in determining the comprehension and production of polysemous constructions across acquisition.

References

Papafragou, A. (1998). The Acquisition of Modality: Implications for Theories of Semantic Representation. *Mind & Language*, 13(3), 370-399.

Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

Using structural priming to test links between constructions: Caused-motion sentences prime resultatives

Tobias Ungerer
(University of Edinburgh)

Keywords: structural priming, Construction Grammar, linguistic network, caused-motion, resultative

Cognitive theories of grammar (e.g. Croft 2001, Goldberg 1995, Langacker 1987) view speakers' linguistic knowledge as a hierarchically structured network of form-meaning pairings, or constructions. Different kinds of links have been suggested to relate constructions in this network with one another, but few empirical attempts have been made to test the psychological reality of those relations. Extending Branigan and Pickering's (2017) recent arguments for using structural priming to investigate linguistic representations, this paper reports on a psycholinguistic study which tested similarities and differences between constructions to derive insights about their underlying relationships.

Many constraint-based theories assume that the linguistic network is organised by one central linking mechanism: 'inheritance', the taxonomic relation between superordinate and subordinate constructions. Goldberg (1995) suggests a further sub-classification of inheritance links into four types: instance, subpart, polysemy and metaphorical extension links. For example, she argues that the caused-motion construction in (1) and the resultative construction in (2) are related via a metaphorical extension link.

(1) Bill rolled the ball down the hill.
(2) Herman hammered the metal flat.

Goldberg's account, however, relies mainly on theoretical arguments and lacks empirical corroboration. Psycholinguistic evidence is needed to address the following questions: (a) Are caused-motion and resultative sentences instances of distinct but related constructions?; (b) Which psycholinguistic correlates can be used to identify distinct types of links between constructions?

I present the results of an online experiment testing structural priming effects between caused-motion and resultative sentences (see (1) and (2) above). The study used a novel experimental design combining self-paced reading with speeded acceptability judgments. 159 native speakers of English were presented with 24 prime-target pairs: prime sentences were either resultative, caused-motion or unrelated constructions; targets were always marginally acceptable resultative sentences (to provoke variation among judgment scores so that a priming effect would become visible). A linear mixed effects model with random effects for items and subjects revealed that participants read resultative sentences on average 18.8 ms faster after having been primed with caused-motion sentences than after reading unrelated constructions (β = -0.02, SE = 0.006, t(3411) = -3.24, p = .001 for log-transformed reading times). This suggests that resultative and caused-motion are different but related constructions (question (a) above). Surprisingly, however, no priming effect was found between resultative primes and resultative targets, pointing to limitations of the experimental design.

Based on these mixed results, I discuss some of the challenges for creating structural priming experiments of this sort, for example regarding priming modality (production vs. comprehension), choice of priming method and 'lexical boost' (i.e. verb repetition) between prime and target. Moreover, I argue that further structural priming experiments will reveal whether differences in the size of priming effects can be reliably used to distinguish between types of constructional links (question (b) above). Once these challenges are addressed, structural priming promises to provide a powerful tool for advancing our models of the structure of the linguistic network.

References

Branigan, Holly P., and Martin J. Pickering (2017), An experimental approach to linguistic representation, *Behavioral and Brain Sciences, 40*, 1-61. https://doi.org/10.1017/S0140525X16002028

Croft, William (2001), *Radical Construction Grammar: Syntactic theory in typological perspective*, Oxford: Oxford University Press.

Goldberg, Adele E. (1995), *Constructions: A Construction Grammar approach to argument structure*, Chicago, IL: The University of Chicago Press.
Langacker, Ronald W. (1987), *Foundations of Cognitive Grammar*, vol. 1, *Theoretical prerequisites*, Stanford, CA: Stanford University Press.

Constructional meaning in the English Constructicon

Florent Perek and Amanda Patten

<pdf>

Libfixes and the organisation of constructional networks

Muriel Norde & Sarah Sippach
(Humboldt University of Berlin & Humboldt University of Berlin)

Keywords: Construction Morphology, Constructional networks, Word formation, English, Corpus Linguistics

One of the basic tenets in Construction Morphology is that the minimal construction is the word, which implies that bound morphemes, which are both morphologically and semantically dependent on the morphological schema they appear in, are no constructions, that is, no independent pairings of form and meaning (Booij & Audring 2018: 61). On this view, words containing the same bound morpheme are considered to be interparadigmatically linked in a constructional network (Norde & Morris 2018: 54). From a diachronic perspective, this means that free morphemes that become bound (e.g. Old English Noun *had* -> derivational suffix *-hood*; Traugott & Trousdale 2014) cease to be independent constructions, whereas bound morphemes that become free (e.g. English *ish* as a free adverb; Bochnak & Csipak 2014) come to form a morphological schema of their own. However, there also exists a type of bound morpheme that has never been a free morpheme, but arose out of language users' ability for "global pattern matching" (Kemmer 2003: 77). These are what Zwicky (2010) charmingly termed "libfixes": "word-forming elements that are semantically like the elements of compounds but are affix-like in that they are typically bound." Examples include *-tainment* (*infotainment, traveltainment, eco-tainment*), *-licious* (*tacolicious, hunk-o-licious*, *diamondalicious*) and *-meter* (*hoaxmeter*, *baloney-meter, WTFometer*).
Libfixes are particularly interesting for a network analysis for a number of reasons. First of all, they do not always have discrete morpheme boundaries. For example, *ea**ter**tainment* and *shopp**er**tainment* include more phonological material from the model exemplar word (*entertainment*), a property they have in common with blends. Secondly, they may include linking syllables (*war-**o**-tainment*, *nerd**uh**licious*), a property they have in common with compounds. Thirdly, they can be combined with a wide range of bases, including phrases, as in *horrible-job-o-meter*. Finally, they may be the source of back formations such as *infotain* or *edutain*.

In this talk, we will present a corpus study of eight English libfixes, six of which (*-fection, -flation, -gasm, -tainment, -licious* and *-tastic*) are based on a single exemplar word, and which (at least initially) emerged as blends, whereas two others (*-cracy* and *-meter*) developed out of a set of similar neo-classical compounds. Data are drawn from the ENCOW16A corpus (Schäfer 2015, and Schäfer & Bildhauer 2012), a web-based corpus of world Englishes, built up from random sentences crawled from internet sources, many of them reflecting informal usage. For each libfix, we analysed a random selection of 500 types (or all types when type frequency was < 500), which were annotated for the following morphological properties: spelling (presence or absence of hyphenation); form of base (single word or phrase); modification of base (e.g. clipping, linking syllables); prosodic similarity to the exemplar word (number of syllables, stress). Preliminary results suggest that libfix constructions are most fruitfully analysed from an associative network model as advanced in the works of Bybee (e.g. 2010, 2013), in which constructions are linked on the basis of phonological similarity, which allows for both fully compositional and non-compositional constructions to be linked without an exhaustive analysis into morphemes.

References

Bochnak, M. Ryan & Eva Csipak (2014), A new metalinguistic degree morpheme, in T. Snider, S. D'Antonio and M. Weigand (eds), (2014), *Proceedings of Semantics and Linguistic Theory (SALT)* 24, 432- 452.

Booij, Geert & Jenny Audring (2018), Partial motivation, multiple motivation: The role of output schemas in morphology, in G. Booij (ed), (2018), *The Construction of Words: Advances in Construction Morphology* (= *Studies in Morphology* 4), Dordrecht: Springer, 59-80.

Bybee, Joan (2010), *Language, usage and cognition*, Cambridge: Cambridge University Press.

Bybee, Joan (2013), Usage-based theory and exemplar representations of constructions, in Th. Hoffmann & G. Trousdale (eds), (2013), *The Oxford handbook of construction grammar*, Oxford: Oxford University Press, 49-69.

Kemmer, Suzanne (2003), Schemas and lexical blends, in Th. Berg et al. (eds), (2003), *Motivation in Language: From Case Grammar to Cognitive Linguistics. A Festschrift for Günter Radden*, Amsterdam / Philadelphia: John Benjamins, 69–97.

Norde, Muriel & Caroline Morris (2018), Derivation without category change: a network-based analysis of diminutive prefixoids in Dutch, in K. Van Goethem, M. Norde, E. Coussé and G. Vanderbauwhede (eds), (2018), *Category change from a constructional perspective*, Amsterdam / Philadelphia: John Benjamins, 47-90.

Schäfer, Roland (2015), Processing and querying large web corpora with the COW 14 architecture, in *Proceedings of Challenges in the Management of Large Corpora (CMLC-3, July 20, 2015, Lancaster)*.

Schäfer, Roland, & Felix Bildhauer (2012), Building large corpora from the Web using a new efficient tool chain, in N. Calzolari et al. (eds), (2012) *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul: ELRA, 486-493.

Traugott, Elizabeth C. & Graeme Trousdale (2014), Contentful constructionalization. *Historical Linguistics* 4(2), 256-283.

Zwicky, Arnold (2010), Libfixes. https://arnoldzwicky.org/2010/01/23/libfixes/

WORKSHOP 11

Schedule: Fri 9.00-16.25 (Room 9)

Learners and variationist theory: New speakers and new sources of variation

Jonathan R. Kasstan & Itxaso Rodríguez-Ordóñez
(Queen Mary University of London & Southern Illinois University Carbondale)

Key words: sociolinguistic variation, new speakers, minoritized languages, language contact, language change

Workshop description

A recent focus on 'new speakers' of minoritized languages in Europe has shaped a novel trajectory of sociolinguistic research that questions naturalized assumptions about what constitutes a 'competent' speaker of a language, and promises novel understandings surrounding the plasticity of the grammar in adulthood. Although the notion of 'new speakers' is not 'new', it has recently been conceived of as an 'emic' category (Jaffe 2015) to describe the lived experiences and sociolinguistic realities of individuals who acquire a minority language not through traditional transmission contexts (i.e. home, family), but as adults through language revitalisation initiatives (O'Rourke Pujolar and Ramallo 2015). This largely qualitative body of work has focused on questions surrounding the perceived 'legitimacy' and 'authenticity' (Woolard 2008) of new speakers in traditional minoritized speech communities. While this work has repeatedly shown these social actors to be perceived of as 'illegitimate' speakers of their chosen target language in most cases, there is nonetheless significant variation in the way these individuals mobilize within these new sociolinguistic regimes (cf. Jaffe 1999, 2015, Costa et al. 2018, O'Rourke and Ramallo 2013, Gal 2018, Ó Murchadha and Flynn 2018, Ó Murchadha and Ó hIfearnáin 2018; Sallabank and Marquis 2018, Puigdevall et al. 2018, Ortega et al. 2014, 2015, Urla et al. 2018).

However, a heavily understudied dimension to the research on 'new speakers' concerns their linguistic systems, especially the means by which they exploit their linguistic repertories and the various linguistic resources that they deploy to navigate the sociolinguistic field. Yet, this is beginning to change. Variationist sociolinguistic work on 'new speakers' has shown that newly enregistered varieties of minoritized languages may be emerging within new speaker communities (Nance 2015, 2018, Nance et al. 2016), the implication being that new speakers can operate as agents of change in the diffusion of new variants into their communities, even in cases of severe language endangerment (Kasstan 2017; Kasstan and Müller 2018). This variation, in turn, can also represent a linguistic manifestation of 'new speakerness' (cf. Nance et al. 2016, Kennard 2018). Researchers have also begun to focus on the role that language and dialect contact plays in new speaker practices, raising questions concerning the nature of code-switching and vernacularisation of standard norms (Lantto 2014, 2018a, 2018b) or the role that linguistic ideologies play in shaping mechanisms of contact-induced processes among learners (Rodríguez-Ordóñez 2016, 2018).

The results from this more recent quantitative line of work shows that some of the patterns of language change cannot be explained within the scope of existing (variationist) sociolinguistic theory. For instance, Nance et al. (2016) compared variation in rhoticity among traditional and new speakers of Scottish Gaelic in Edinburgh and Glasgow, where they found that the variability that emerges among new speakers is better explained in terms of a Type III variation model, i.e. a type of variation

that that accounts for identity construction among L2 speakers, *contra* traditional speaker types that remain the staple of variationist research. Similarly, Rodríguez-Ordóñez (2018) examines the social meaning and use of Differential Object Marking in Basque, a contact-induced phenomenon that is highly salient and stigmatized within the Basque community. Her results show that 'new speakers of Basque' adopt and diffuse stigmatized variants once they have claimed some level of 'authority' and 'legitimacy' as Basque speakers, further arguing that salient and stigmatized features are prone to change inviting more nuanced explanations that go beyond Labov's (1994) principles of change (i.e. 'change from below' and 'change from above'). From a stylistic point of view, Kasstan (2018) examines the variable use of /l/-palatalization in traditional and new speakers of Francoprovençal, and shows that previously obligatory rules can become variable, and that this underspecification renders variants available for social work as dialect icons. As such, Kasstan (2018) argues that new forms of style variation can emerge in situations of severe language endangerment, questioning the notion that language obsolescence necessarily leads to stylistic shrinkage, as is commonly argued in the literature (see e.g. Dorian 1994).

These new advancements in our understanding of the linguistic variation among new speakers calls for more comparative sociolinguistic work (per Stanford 2016) that situates the study of linguistic variation at the center of linguistic inquiry in minoritized contexts. As such, this workshop welcomes contributions that engage with questions pertaining to linguistic variation among new speakers of minoritized languages. In particular, we aim to address the following research questions:

- In what ways can 'new speaker' variation be modelled?
- What principles of linguistic diffusion and change stand out in these contexts?
- In what ways are 'new varieties' of minoritized languages emerging?
- How and why are new speakers diffusing new vernacular forms in their communities?
- What are the processes by which contact features become part of 'new speakers'' linguistics systems?
- What aspects of the language are more vulnerable to cross-linguistic influence?
- How do aspects of social meaning interact in contact-induced language change?
- How do contact features become enregistered in minoritized contact situations?
- What kind of stylistic practices do new speakers engage in?
- Why do certain features become enregistered while others don't?
- How can existing analytical domains such as social networks and communities of practice aid us in understanding the emergence and change of new variants among new speakers?
- What is the role of dialect and language contact in the use of new variants in minoritized contexts?
- What are methodological challenges that arise from studying the variation among 'new speakers of minoritized languages?

The current workshop proposal is accompanied by nine abstracts that seek to address the proposed research questions. These original studies, representing a wealth of new research on six minoritized languages (Basque, Scottish Gaelic, Irish, Breton, Francoprovençal, Belarusian and Diné bizaad (Navajo)), allow us to move forward in our theorizations of linguistic variation in minoritized contexts. The workshop welcomes further studies that addresses questions of linguistic variation in minoritized contexts.

References

Costa, James, De Korne, Hayley Jean, and Lane, Pia (2018), Standardising Minority Languages Reinventing Peripheral Languages in the 21st Century, in P. Lane, J. Costa, and H. J. De Korne (eds), (2018), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*, New York: Routledge, 1-23.

Dorian, Nancy C (1994), Stylistic variation in a language restricted to private-sphere use, in D. Biber, and E. Finegan (eds), (1994), *Sociolinguistic Perspectives on Register*, Oxford: Oxford University Press, 217-232.

Jaffe, Alexandra (1999) *Ideologies in action: language politics on Corsica*. Berlin / New York: Mouton de Gruyter.

Jaffe, Alexandra (2015), Defining the new speaker: theoretical perspectives and learner trajectories, *International Journal of the Sociology of Language* 231, 21-44.

Gal, Susan (2018), Visions and Revisions of Minority Languages: Standardization and Its Dilemmas, in P. Lane, J. Costa, and H. J. De Korne (eds), (2018), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*, New York: Routledge, 222-242.

Kasstan, Jonathan R. (2017), New speakers: Challenges and opportunities for variationist sociolinguistics, *Language and Linguistic Compass* 11(8), e12249.

Kasstan, Jonathan R (2018), Maintaining style in language death, Paper presented at *New Ways of Analyzing Variation 47*, New York University, New York City, United States.

Kasstan, Jonathan R., and Müller, Daniela. (2018). (l) as a sociolinguistic variable in Francoprovençal, *International Journal of the Sociology of Language* 249, 99-118.

Kennard, Holly J., (2018), Verbal Lenition Among Young Speakers of Breton: Acquisition and Maintenance, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 231-252.

Labov, William (1994), *Principles of Linguistic Change: Volume 1: Internal Factors*, Cambridge/ Oxford: Blackwell.

Lantto, Hanna (2014), Code-switching, swearing and slang: The colloquial register of Basque in Greater Bilbao, *International Journal of Bilingualism* 18(6), 633-648.

Lantto, Hanna (2018a), New Basques and Code-Switching: Purist Tendencies, Social Pressures, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 165-187.

Lantto, Hanna (2018b), Language contact and variation in new Basque speakers' individual styles, Paper presented at *The 51st Annual Meeting of the Societas Linguistica Europaea*, Tallin, Estonia.

Nance, Claire (2015), 'New' Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalization, *Language in Society* 44(4), 553-579.

Nance, Claire, McLeod, Wilson, O'Rourke, Bernadette, and Dunmore, Stuart (2016), Identity, accent aim, and motivation in second language users: New Scottish Gaelic speakers' use of phonetic variation, *Journal of Sociolinguistics* 20(2), 164–191.

Nance, Claire (2018), Linguistic Innovation Among Glasgow Gaelic New Speakers, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 213-230.

Ó Murchadha, Noel, and Flynn, Colin (2018), Language educators' regard for variation in late modernity: Perceptions of linguistic variation in minority contexts, *Journal of Sociolinguistics* 22(3), 288-311.

Ó Murchadha, Noel, and Ó hIfearnáin, Tadhg (2018), Converging and diverging stances on target revival varieties in collateral languages: the ideologies of linguistic variation in Irish and Manx Gaelic, *Journal of Multilingual and Multicultural Development* 39(5), 458–469.

O'Rourke Bernadette, and Ramallo, Fernando (2013), Competing ideologies of linguistic authority amongst new speakers in contemporary Galicia, *Language in Society* 42(3), 287-305.

O'Rourke, Bernadette, Pujolar, Joan, and Ramallo, Fernando (2015), New speakers of minority languages: the challenging opportunity - foreword, *International Journal of the Sociology of Language* 231, 1–20.

Ortega, Ane, Amorrortu, Estibalitz, Goirigolzarri, Jone, Urla, Jacqueline, and Uranga, Belen (2014), New Basque speakers: Linguistic identity and legitimacy, *Digithum* 16, 47-58.

Ortega, Ane, Urla, Jacqueline, Amorrortu, Estibalitz, Goirigolzarri, Jone, and Uranga, Belen (2015), Linguistic identity among new speakers of Basque. *International Journal of the Sociology of Language* 231, 85-105.

Puigdevall, Maite, Walsh, John, Amorrortu, Estibalitz, and Ortega, Ane (2018) 'I'll be one of them': linguistic mudes and new speakers in three minority language contexts, *Journal of Multilingual and Multicultural Development* 39(5), 445-457.

Rodríguez-Ordóñez, Itxaso (2016), *Differential Object Marking in Basque: grammaticalization, attitudes and ideological representations*, Unpublished Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

Rodríguez-Ordóñez, Itxaso (2018), Language contact and social meaning from the perspective of 'new speakers' of Basque, Paper presented at *New Ways of Analyzing Variation 47*, New York University, New York City, United States.

Sallabank, Julia, and Marquis, Yan (2018) 'We Don't Say It Like That': Language Ownership and (De)Legitimising the New Speaker. in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 67-90.

Stanford, James (2016), A call for more diverse sources of data: Variationist approaches in non-English contexts, *Journal of Sociolinguistics* 20(4), 525-541.

Urla, Jacqueline, Amorrortu, Estibalitz, Ortega, Ane, and Goirigolzarri, Jone (2018). Basque Standardization and the New Speaker: Political Praxis and the Shifting Dynamics of Authority and Value. in P. Lane, J. Costa, and H. J. De Korne (eds), (2018), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*, New York: Routledge, 25-46.

Woolard, Kathryn (2008), Language and Identity Choice in Catalonia: The Interplay of Contrasting Ideologies of Linguistic Authority, In K. Süselbeck, U. Mühlschlegel, and P. Masson (eds), (2008), *Lengua, nación e identidad. La regulación del plurilingüismo en España y América Latina*, Madrid: Iberoamericana, 303–323.

New speakers as agents of linguistic creativity: the case of Basque ergativity

Itxaso Rodríguez-Ordóñez

Keywords: new speakers, ergativity, language contact, social meaning, language variation

Recent work has demonstrated that 'new speakers' of Basque do not enjoy the same legitimacy as those who speak a 'local' variety (Ortega et al. 2014, 2015, Urla et al. 2018). However, there is a dearth of research focusing on the linguistic system of these speakers and an understanding of how the social meaning behind their linguistic repertoires shape new variations that may emerge in revitalization contexts. In filling this gap, the present study focuses on the variable use and social meaning behind nominal ergative case-marking among 42 adult 'new speakers' (*euskaldunberriak*) and 44 'traditional' speakers (*euskaldunzaharrak*) of Basque, considering their proficiency, learning trajectories and social identities.

Following variationist approaches to sociolinguistics and language acquisition, production data was collected through sociolinguistic interviews and their use of subject ergative marking was coded for a number of linguistic factors such as verb type (transitive, unergative, unaccusative), phonological context (pre-consonantal, pre-vowel, and final position) and grammatical person (first, second and third). The social meaning of ergative case marking was gathered by means of a matched-guise experiment whereby participants listened to 4 guises, controlling for the presence or absence of ergative case marking in two varieties (a regional variety vs. Standard Basque). For each guise, participants responded to 18 questions such as "How Basque is this person?" or "Does this person make mistakes?" using a 1-7 Likert-Scale. Results from the MGE were analyzed using a number of mixed-effects models and correlations in *R*. These results were combined with overt metapragmatic commentary.

Production results show that all speakers omit the ergative case marker to some extent, but those in the lowest proficiency omitted the most. Results with respect to self-identification show that they pattern similarly to those reported for proficiency, indicating that self-identification and proficiency are co-linear with respect to ergative case marking. In terms of linguistic factors, new speakers show a mastery of the contexts in which ergative case-marking is prescriptively expected to be marked (transitive and unergative subjects) and largely show a sensitive towards a phonological rule in that ergative case-marking is more likely to be omitted when the following phonological segment is a consonant. Additionally, new speakers show another variable rule in that they extend ergative subjects into unaccusative contexts, mainly for topicalization purposes. Although there is no direct link between their attitudes towards ergative case-marking and use, MGE results show that lack of ergative-case marking is strongly associated with 'new speakers' and 'lack of competence', but goes unnoticed in the vernacular guises.

These results corroborate recent findings on the enregisterment of *Batua* as an illegitimate variety (Urla et al., 2018). Furthermore, I discuss these results along similar findings that situate new speakers' linguistic repertoires as 'different' rather than 'deficient' (Nance 2018, Lantto, 2014, 2018) providing further support to the argument that 'new speakers' play a role in linguistic creativity and potential agents of language change (Kasstan 2017).

References

Lantto, Hanna (2014), Code-switching, swearing and slang: The colloquial register of Basque in Greater Bilbao, *International Journal of Bilingualism* 18(6), 633-648.

Lantto, Hanna (2018), New Basques and Code-Switching: Purist Tendencies, Social Pressures, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 165-187.

Kasstan, Jonathan (2017), New speakers: Challenges and opportunities for variationist sociolinguistics, *Language and Linguistic Compass* 11(8), e12249.

Nance, Claire (2018), Linguistic Innovation Among Glasgow Gaelic New Speakers, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 213, 230.

Ortega, Ane, Amorrortu, Estibaliz, Goirigolzarri, Jone, Urla, Jacqueline, and Uranga, Belen (2014), New Basque speakers: Linguistic identity and legitimacy, *Digithum* 16, 47-58.

Ortega, Ane, Urla, Jacqueline, Amorrortu, Estibaliz, Goirigolzarri, Jone, and Uranga, Belen (2015), Linguistic identity among new speakers of Basque, *International Journal of the Sociology of Language* 231, 85-105.

Urla, Jacqueline, Amorrortu, Estibaliz, Ortega, Ane, and Goirigolzarri, Jone (2018), Basque Standardization and the New Speaker: Political Praxis and the Shifting Dynamics of Authority and Value. In P. Lane, J. Costa, and H. De Korne (eds), (2018), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*, New York: Routledge, 24-46.

Incipient morphological change among different types of modern Chukchi speakers

Jessica Kantarovich
(University of Chicago)

Keywords: <language shift, attrition, Chukotko-Kamchatkan, variation, morphological leveling>

Chukchi is an endangered language spoken in Siberia: there are few proficient speakers under the age of 50, and virtually all speakers are bilingual in Russian. This work analyzes the different types of speakers that emerge in situations of language shift (attriting L1 speakers, semi-speakers/heritage speakers, and L2 learners). Can we identify morphological patterns that are unique to their circumstances, and how do modern interspeaker differences fit with what is known about existing variation in Chukchi?

Recent fieldwork with speakers of these different categories reveals striking morphosyntactic differences from published grammars. Of particular interest are changes to the polysynthetic verb, which have the potential to drive other changes in the language's overall morphological organization. Chukchi verbal agreement is also relevant for studies of cross-linguistic realizations of ergativity, as it displays a typologically unusual kind of positional split ergativity (Bobaljik 1998): verbal agreement prefixes encode the subject in transitive and intransitive verbs (a nominative pattern), but suffixes encode the subject in intransitive verbs and the direct object in transitive verbs (an absolutive pattern). In intransitive verbs, both affixes are present and agree with the subject.

(1) Standard Chukchi aorist agreement (Dunn 1999)

	S Prefix	S Suffix	A Prefix	O Suffix
1sg	*tə-*	*-gˀek*	*tə-*	*-gəm*
1pl	*mət-*	*-mək*	*mət-*	*-mək*
2sg	*Ø-*	*-gˀi*	*Ø-*	*-gət*
2pl	*Ø-*	*-tək*	*Ø-*	*-tək*

3sg	*Ø-*	*-gˀi*	*(ne)-*	*-(gˀe)n, -nin*
3pl	*Ø-*	*-gˀet*	*ne*	*-(ni)net*

(2) a. *tə-lˀu-net*
1sg.SBJ-see-3pl.OBJ
'I saw them'
b. *Ø-pˀeŋiwe-gˀet*
3pl.SBJ-become.tired-3pl.SBJ
'They became tired'

The speech of some modern speakers evidences a shift away from this pattern. In transitive verbs, both agreement slots are preserved. However, morphological leveling among object suffixes has neutralized some of the information conveyed by the suffix slot. Specifically, 3sg *-gˀen* and 3pl *-(ni)net* have been adopted in place of the expected 1st and 2nd person object suffixes throughout the paradigm. These unexpected forms are bolded below:

(3) Consultant's full transitive aorist agreement system

	1sg.OBJ	1pl.OBJ	2sg.OBJ	2pl.OBJ	3sg.OBJ	3pl.OBJ
1sg.SBJ	–	–	*tə-* ***-gˀen***	*tə- -rək*	*tə- -gˀen*	*tə- -net*
1pl.SBJ	–	–	*mət-* ***-gˀen***	*mət- -net*	*mət- -gPen*	*mət- -net*
2sg.SBJ	*ine- -gˀi*	*ine-* ***-gˀen***	–	–	*-gˀen*	*-net*
2pl.SBJ	*ine- -tək*	*ine- -tək*	–	–	*-tək*	*-tək*
3sg.SBJ	*ine- -gˀi*	*ine-* ***-ninet***	*ne- -gˀet*	*ine-* ***-ninet***	*-nin*	*-ninet*
3pl.SBJ	*ne-* ***-gˀen***	*ne- -mək*	*ge- -gət*	*ne- -tək*	*ne- -gˀen*	*ne- -net*

Because the prefixes are largely unchanged, this leveling has in some places produced syncretism between entire verb forms for the same subject with different objects—nominative agreement.

This pattern is noteworthy because regional varieties of Chukchi display few differences in inflectional morphology. We might expect a loss of object agreement due to contact-induced change from Russian, which encodes only subject agreement. Another possibility is linguistic loss due to incomplete acquisition or attrition. To further complicate matters, regional variation in Chukchi is on the whole underdescribed. The speakers consulted for this study reside in the Sakha Republic; most documentation of Chukchi has been conducted in Chukotka. Thus, it is possible this morphological variation was acquired in childhood.

Nevertheless, these data indicate that there is some systematic variation that is specific to modern speakers of an obsolescing language. Although distinguishing among regional differences, attrition, and change in Chukchi requires further research, these initial findings inform our understanding of the linguistic systems of shifting speakers.

References

Bobaljik, Jonathan David. 1998. Pseudo-ergativity in Chukotko-Kamchatkan Agreement Systems. *Recherches Linguistiques de Vincennes* 27. 21-44.

Dunn, Michael. 1999. A Grammar of Chukchi. Australian National University dissertation: Canberra.

Subject Pronoun Expression in Basque and Spanish: comparing 'traditional' and 'new speakers'

Eider Etxebarria-Zuluaga, Itxaso Rodríguez-Ordóñez and Lorena Sainzmaza-Lecanda

Key words: subjects, sociolinguistic variation, new speakers, Basque, language contact

The variable use of Spanish Subject Pronoun Expression (SPE) has rendered increasing attention to the effects of linguistic contact by either showing convergence effects or lack thereof (Otheguy and Zentella 2012, Travis and Torres-Cacoullos 2011, Carvalho et al 2015; Michnowicz 2015). Whereas most scholars have focused on the influence of English onto Spanish SPE, very little is known about the outcomes of contact between languages that systematically allow a null subject variant. In this study, we follow a variationist approach to the study of SPE in Basque and Spanish in order to quantitatively assess potential cross-linguistic effects between these two null-subject languages and compare the role that 'new speakers' vs. 'traditional' speakers of Basque play in shaping such variation.

The spontaneous speech of 10 'traditional' Basque and 10 'new speakers' was analyzed both in Basque and Spanish. Additionally, we examined the spontaneous speech of two control groups of monolingual speakers of Spanish (n=10 from Bilbao and n=10 from Madrid). Extracting an average of 100 variable uses of SPE per speaker, 4,000 tokens have been coded for the following linguistic factors: PERSON & NUMBER, PRIMING, SWITCH REFERENCE, VERB SEMANTICS, and TENSE. Following statistical modeling principles in variationist sociolinguistics, data was analyzed using hierarchical constraint analyses and mixed-effects models in *R* (Bates et al. 2015; Walker 2014).

Results indicate that 'traditional speakers' use less SPE both in their Basque and Spanish (12.1%, 13.4%, respectively) than 'new speakers' (19.6% Basque; 15.6% Spanish) but both groups use less SPE than Spanish monolingual controls (~21%). With respect to linguistic factors, the 'traditional Basque' group shows that their SPE systems are mainly conditioned by Person & Number and Switch Reference, in that over subject pronouns are more likely to occur when the subject is produced in the first and second and when the previous subject is different. 'New speakers' of Basque show that their SPE systems both in Basque and Spanish are conditioned by the same factors as 'traditional speakers' with the addition of a priming effect: an overt subject is more likely to occur if the previous subject was also overtly mentioned. The results of 'new speakers' are consistent with the results obtained from the two monolingual control groups.

These results suggest that the influence of language contact is bidirectional in the realization of SPE. Specifically, the Spanish system of 'traditional Basque' bilinguals is influenced by their Basque, whereas new speakers' Basque SPE system is influenced by their Spanish. These results lead us to conclude that the Basque SPE system of 'new speakers' is different from 'traditional speakers', rendering these findings consistent with previous phonological findings in other 'new speaker populations' (Nance 2015, 2018). Finally, it is argued that languages do not necessarily need to be typologically similar to convergence in long-standing contact situations (Muysken 2006), but such convergence is facilitated through 'pivot-matching' (Matras 2009) in that typologically similar linguistic features (i.e. null subjects) may influence each other in long-term contact situations.

References

Bates, Douglas, Maechler, Martin, Bolker, Ben, and Walker, Steve (2015), Fitting Linear Mixed-Effects Models Using lme4, *Journal of Statistical Software* 67(1), 1-48.

Carvalho, Ana María, Orozco, Rafael, and Shin, Naomi (2015), *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective,* Washington DC: Georgetown University Press.

Matras, Yaron (2009), *Language Contact*, Cambridge: Cambridge University Press.

Michnowicz, Jim (2015), Subject Pronoun Expression in Contact with Maya in Yucatan Spanish, in Carvalho, Ana María, Orozco, Rafael, and Shin, Naomi (eds), (2015), *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, Washington DC: Georgetown University Press, pp.101-119.

Muysken, Pieter (2006), Two linguistic systems in contact: Grammar, phonology and lexicon. In Tej Bhatia and William Ritchie (eds), *The Handbook of Bilingualism*, Oxford: Wiley-Blackwell, 147-168.

Nance, Claire (2015), 'New' Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalization, *Language in Society* 44(4), 553-579.

Nance, Claire (2018), Linguistic Innovation Among Glasgow Gaelic New Speakers, in C. Smith-Christmas, N. Ó Murchadha, M. Hornsby, and M. Moriarty (eds), (2018), *New Speakers of Minority Languages: Linguistic Ideologies and Practices*, London: Palgrave Macmillan, 213-230.

Otheguy, Ricardo and Zentella, Ana Celia (2012), *Spanish in New York: Language Contact, Dialect Leveling and Structural Continuity*, Oxford: Oxford University Press.

Torres-Cacoullos, Rena, and Travis, Catherine (2011), Testing convergence via code-switching: Priming and the structure of variable subject expression, *International Journal of Bilingualism* 15 (3), 241–267.

Walker, James (2014), Variation analysis, in Robert Podesva and Devyani Sharma (eds), (2014), *Research Methods in Linguistics*, Cambridge: Cambridge University Press, 440-459.

Analogy and similarity as a resource in new speaker varieties of Basque

Hanna Lantto
(University of Turku)

Keywords: New speakers, Basque, language contact, similarity

We do not learn "languages" as complete systems with clear boundaries, but as constructions, form-meaning pairings, that are associated with different languages. A linguistic repertoire consists of constructions that the speaker has accumulated in interaction with other speakers during their life trajectories (Blommaert and Backus 2011). Speakers' individual varieties, however, are not merely products of these accumulated interactions, but speakers actively create new language forms and registers. Bi- and multilinguals, which new speakers are by definition, have the cognitive representation of the constructions related to different languages in their repertoire. They can establish equivalence through identification of both interlingual and intralingual links between these constructions.

This paper examines the different ways in which new speakers of Basque use analogy and similarity in the creation of their colloquial styles. The data used for this study was collected in the city of Bilbao in the years 2016 and 2017 via ethnographic observation and semi-structured metalinguistic interviews. Twenty-five new Basque speakers were interviewed and observed for the study. New speakers of Basque learn the standard *Euskara Batua* in classroom contexts. The standard variety of Basque, however, is often perceived as artificial and too formal outside of the classroom (Urla et al. 2016). Therefore, most new Basque speakers adapt their own speech to the requirements of

the more informal situations, often by drawing on similarity and analogy between the constructions of Batua, Spanish and Basque vernacular dialects in register construction. Depending on the degree of their multilingual awareness, the use of analogy and similarity as a resource can take many forms, which is reflected in the variation in new Basque speakers' individual styles.

The use of similarity as a resource seems to follow hierarchical patterns: the new speakers who use some Basque in their daily life, yet do not show a high degree of engagement in the Basque-speaking culture (here labelled *non-activists*), mostly create colloquial registers by *identifying dissimilarity*, filling the gaps of their standard Basque with code-switching to Spanish at the points where they lack the Basque equivalents. The second group, labelled *activists*, use similar strategies of code-switching as a stylistic resource, but also add to their otherwise standard speech dialectal Basque constructions that are similar to their standard equivalents both in form and meaning. The third speaker group consists of new Basque speakers who show a high investment in the Basque-speaking culture. These speakers, labelled *language activists*, in addition to using code-switching to Spanish and vernacular features, create new colloquial innovations using both intralingual (Basque > Basque) and interlingual (Spanish > Basque) analogies as a resource in register construction.

References

Blommaert, Jan and Backus, Ad (2011), Repertoires revisited: 'Knowing language' in superdiversity. *Working Papers in Urban Language & Literacies.* Tillburg University.

Urla, Jacqueline, Amorrortu, Estibaliz, Ortega, Ane and Goirigolzarri, Jone (2016), *Authenticity and Linguistic Variety among New Speakers of Basque.* In LD&C Special Publication No. 9: Language Documentation and Conservation in Europe.

"New speakerness" and language variation: Constructing sociolinguistic authenticity in contemporary Belarus

Curt Woolhiser

Keywords: new speakers, authenticity, divergence, Belarusian, Russian

Studies of new speakers in minority and minoritized language contexts have generally focused on the sociolinguistic opposition between learners of standardized supra-dialectal varieties and speakers of traditional rural dialects, where sociolinguistic authenticity tends to be discursively situated in the latter (Hornsby 2005, MacCaluim 2007, O'Rourke and Ramallo 2013, Costa 2015, Costa, De Korne and Lane 2018). In the case of Belarusian, due to decades of convergence of Belarusian dialectal varieties not toward standard Belarusian, but rather toward closely-related standard Russian, the primary focus of ideological contestation with regard to the language practices of new speakers is not the standard/dialect axis, but rather the opposition between two varieties of the Belarusian standard: one based on the post-1933 Soviet codification, which is somewhat closer to Russian in lexicon, inflectional morphology and syntax, and which remains the sole standard variety of Belarusian enjoying official recognition (alongside the dominant Russian language) in contemporary Belarus, and "unofficial" varieties based on the pre-1933 standard.

A striking feature of Belarusian new speakers' linguistic repertoires is a marked preference for the "unofficial" standard variants, despite institutional support for the official standard in the sphere of education (primarily as a subject, rather than as the dominant medium of instruction), the state-owned

media, signage, and a number of other domains. Frequently, new speakers favoring the more divergent "unofficial" standard variety show a tendency toward hypercorrection, particularly in inflectional morphology and the lexicon, employing variants that are more divergent from Russian than those considered standard according to the pre-1933 norms. While many new speakers regard the pre-1933 standard as a more "authentic" form of Belarusian than the official standard, members of the Belarusian linguistic, educational and cultural establishment openly criticize the innovations associated with new speakers as "neophyte jargon," which they maintain only weaken the already tenuous position of the Belarusian language in a predominantly Russian-speaking society.

In this paper I will examine how speaker profiles, ideological stances and social networks (reflected in reported domains of Belarusian language use, speakers' views on Belarusian identity and cultural affinities, and density and multiplexity of Belarusophone social networks) correlate with the use of variables in inflectional morphology and syntax that are more or less divergent vis-à-vis the dominant language, Russian. I will argue that many of these variables are consciously manipulated by speakers in order to construct a particular version of sociolinguistic authenticity based on an avoidance of variants that are more similar to their Russian counterparts. On the basis of survey and interview data obtained during fieldwork in Belarus in 2013 and the spring and summer of 2019, I will show that the new speakers who use Belarusian in the broadest range of domains, show the strongest ideological commitment to the idea of Belarusian political and cultural independence, and have denser and more multiplex Belarusophone social networks, are also those whose language use is most divergent from their first language, Russian.

References

Costa, James (2015). New speakers, new language: on being a legitimate speaker of a minority language in Provence. *IJSL* 231: 127-145.

Costa, J. De Korne, H. & Lane, P. (2018). Standardising Minority Languages Reinventing Peripheral Languages in the 21st Century. In Lane, P., Costa, J. & De Korne, H. (Eds.), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery*, New York: Routledge, 1-23.

Hornsby, Michael (2005). *Néo-breton* and questions of authenticity, *Estudios de Sociolingüística* 6 (2), 191-218.

MacCaluim, Alasdair (2007). *Reversing Language Shift: The Social Identity and Role of Scottish Gaelic Learners.* Belfast.

O'Rourke, Bernadette and Ramallo, Fernando (2013). Competing ideologies of linguistic authority amongst new speakers in contemporary Galicia, *Language in Society* 42: 287-305.

Teachers' regard for variation in a minority language context

Noel Ó Murchadha

<pdf>

WORKSHOP 12

Schedule: Fri 9.00 – Sat 12.55 (Room 11)

Managing information structure in spoken and sign languages: Formal properties and natural discourse organization

Peter Arkadiev, Vadim Kimmelman, Nikolay Korotaev & Vera Podlesskaya
(Institute of Slavic Studies of the Russian Academy of Sciences and RSUH; RSUH and Bergen University; RSUH and RANEPA & RSUH andRANEPA)

Keywords: discourse organization, information structure, sign languages, spoken discourse, prosody

Modern linguistics deals with information structure (IS) at least since Mathesius (1929). The existence of such phenomena as theme (topic), rheme (comment, focus), categorical vs. non-categorical (thetic) utterances, seem to be generally accepted and addressed in studies on both spoken and sign modalities (Féry & Ishihara 2016; Kimmelman & Pfau 2016), although even the most fundamental notions have been recently criticized (Matić & Wedgwood 2013; Ozerov 2018, see also the workshop "Beyond Information Structure" at the 50th SLE meeting in Zürich, http://sle2017.eu/downloads/workshops/Beyond%20information%20structure.pdf)..

IS-related phenomena have been approached from different angles. They have been thoroughly analyzed in logic and semantics (Jacobs 2001; Maslova & Bernini 2006; Krifka & Musan 2012), as well as from pragmatic (Moneglia & Raso 2014), cognitive (Chafe 1976; Gundel 1988; Tomlin 1995) and grammatical (Lambrecht 1994, Kuroda 2006, Zimmermann & Féry 2010, Dalrymple & Nikolaeva 2011, Van Gijn et al 2014) perspectives. A great deal of attention has been drawn to formal properties of IS that help distinguish one category from another, both universally (Lambrecht 1994; Féry 2016) and cross-linguistically. To give but one example, in Russian, syntactic (esp. word order) and intonational properties of utterances are strongly influenced by IS factors (Kovtunova 1976; Yanko 2008; Slioussar 2007; Paducheva 2015), and this is also the case in Russian Sign Language where sign order, non-manual markers and manual prosody all interact with IS (Kimmelman 2019). When it comes to oral and sign production, prosody seems to play an extremely important role (Kodzasov 2009; Calhoun 2012; Féry 2013, Crasborn & van der Kooij 2013, Herrmann 2015).

Although the theme – rheme distinction has a great impact on the overall structure of a coherent text (see, for instance, the notion of *theme dynamics* in Daneš 1964, Enkvist 1976), the formal properties of these and other IS-related categories have been primarily studied on the basis of isolated sentences. However, recently the focus of attention has been shifted to the interplay between IS and the organization of natural spoken discourse (Raso & Melo 2014; Moneglia & Cresti 2015; Fernandez-Vest & Van Valin 2016; Kibrik et al. to appear). The goal of our workshop is to follow this new line of research integrating the data of sign languages into a broader context of IS in natural discourse production. The suggested overall approach and the pool of research questions set below are in the spirit of the ongoing project "Managing information structure in spoken and sign natural discourse" (supported by RSF grant 17-18-01184)

The scope of the workshop will include, but not be limited to, the following key questions:

(b) Does the fundamental distinction between thetic and categorical utterances on the one hand, and that between theme and rheme on the other hand, stand against data of natural discourse in spoken and sign languages? If yes, what are the main formal properties of these categories as they arise in corpora of natural discourse?

(c) What is the exact role of prosody in delimiting IS categories in natural speech? What approach(es) to describing intonational structure yields better results when analyzing the information structure of spoken discourse in spoken and sign languages?
(d) What are functional and structural parallels between prosodic means of information structure encoding in spoken languages (i.e., phrasal accents, their placement rules, tonal patterns which are associated with accents) and prosodic means of information structure encoding in sign languages (including non-manual prosody, i.e. face expressions, head and body movement as well as manual prosody – pauses, speed, size and other integral characteristics of movement in sign systems)?
(e) How does the grammatical structure and / or intonation inventory of a language affect the interplay between grammar and prosody as they contribute to encode the IS? Specifically, (a) what is the role of prosody in the languages which grammaticalize IS marking (e.g., have grammaticalized topic); (b) what is the role of the phrase-level intonation in tonal languages?
(f) What are the possible contexts for neutralization of the theme – rheme opposition in natural discourse? For instance, in Russian, clausal themes and rhemes share a great number of formal properties when combined with non-final transitional continuity (Du Bois et al. 1993; see Yanko 2008; Korotaev 2018), and in American Sign Language, topics and foci can be marked by the same non-manual markers (eyebrow raise) in some contexts (Wilbur 2012). Do such contexts differ across languages?
(g) How are sentences with different illocutionary force integrated into the complex hierarchical structure in spoken and sign languages? Specifically, how are they integrated in the contexts which are sensitive to neutralizing illocutionary force meanings, e.g. in reported speech? Which grammatical and prosodic patterns are at play? Are prosodic signals of integrating IS accompanied by such grammatical phenomena as indexical shift?
(h) How is the prosody-IS coherence maintained in natural discourse in the case of communicative breakdowns – caused by speech generation problems or by interactional problems? Are coherency restoration strategies different in spoken and sign languages?
(i) What are the best practices for tagging IS in prosodically annotated spoken and sign language corpora?
(j) How do gestures participate in packaging information in spoken and sign languages?

References

Calhoun, Sasha (2012), The theme/rheme distinction: Accent type or relative prominence? *Journal of Phonetics* 40, 329-349.

Chafe, Wallace (1976), Givenness, contrastiveness, definiteness, subjects, topics, and point of view, in Charles N. Li (ed.), *Subject and topic*, New York: Academic Press, 25-55

Crasborn, Onno & Els van der Kooij (2013), The phonology of focus in Sign Language of the Netherlands, *Journal of Linguistics* 49(3), 515-565.

Dalrymple, Mary and Nikolaeva, Irina (2011), *Objects and Information Structure*, Cambridge: Cambridge University Press

Daneš, František (1964), A three-level approach to syntax, *Travaux linguistiques de Prague* 1: 225-240.

Du Bois, John, Stephan Schuetze-Coburn, Danae Paolino, and Susanna Cumming (1993), Outline of discourse transcription, in Jane A. Edwards, Martin D. Lampert (eds.), *Talking Data: Transcription and Coding in Discourse Research*, Hillsdale, NJ: Lawrence Erlbaum. 45-89.

Enkvist, Nils Erik (1973), Theme dynamics and style: an experiment, *Studia anglica posnaniensia* 5, 127-135

Fernandez-Vest, M. M. Jocelyne and Robert D. Van Valin Jr. (eds.) (2016), *Information Structuring of Spoken Language from a Cross-linguistic Perspective*, Berlin/Boston: de Gruyter Mouton.

Féry, Caroline (2013), Focus as prosodic alignment. *Natural Language and Linguistic Theory*, 31(3), 683-734.

Féry, Caroline (2016), *Intonation and Prosodic Structure*, Cambridge: Cambridge University Press.

Féry, Caroline and Shinichiro Ishihara (eds.) (2016), *The Oxford Handbook of Information Structure*, Oxford: Oxford University Press.

Gundel, Jeanette K. (1988), Universals of topic-comments structure, in M. Hammond, E. A. Moravcsik, and J. R. Wirth (eds.), *Studies in Syntactic Typology*, Amsterdam/Philadelphia: John Benjamins, 209-242.

Herrmann, Annika (2015), The marking of information structure in German Sign Language, *Lingua* 165, 277-297.

Jacobs, Joachim (2001), The dimensions of topic-comment, *Linguistics* 39(4), 641-681.

Kibrik, Andrej A., Nikolay A. Korotaev and Vera I. Podlesskaya. To appear. Russian spoken discourse: Local structure and prosody.

Kimmelman, Vadim (2019), *Information Structure in Sign Languages*, Berlin/Boston: de Gruyter Mouton.

Kimmelman, Vadim and Roland Pfau (2016), Information structure in sign languages, in Caroline Féry and Shinichiro Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford: Oxford University Press, 814-834.

Kodzasov, Sandro V. (2009), *Issledovanija v oblasti russkoj prosodii* [Studies in Russian Prosody], Moscow: Jazyki Slavjanskix Kul'tur.

Korotaev, Nikolay (2018). Topic, focus, and transitional continuity in Russian spoken narratives, in *51st Annual Meeting of the Societas Linguistica Europaea. Book of Abstracts*, 137-138.

Kovtunova, Irina I. (1976), *Sovremennyj russkij jazyk: porjadok slov i aktual'noe členenie predloženija* [Word order and information structure in modern Russian], Moscow: Prosveščenie.

Krifka, Manfred and Renate Musan (eds) (2012), *The Expression of Information Structure*, Berlin/Boston: De Gruyter Mouton.

Kuroda, Shige-Yuki (2005), Focusing on the matter of topic: a study of *wa* and *ga* in Japanese, *Journal of East Asian Linguistics*, 14(1), 1-58.

Lambrecht, Knud (1994), *Information Structure and Sentence Form*, Cambridge: Cambridge University Press.

Maslova, Elena and Giuliano Bernini (2006), Sentence topics in the languages of Europe and beyond, in Giuliano Bernini and Marcia L. Schwartz (eds.), *Pragmatic Organization of Discourse in the Languages of Europe*, Berlin/New York: Mouton de Gruyter, 67-120.

Matić, Dejan, and Daniel Wedgwood (2013). The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis, *Journal of Linguistics*, 49(1), 127-163.

Mathesius, Vilém (1929), Zur Satzperspektive im modernen Englisch, *Archiv für das Studium der modernen Sprachen und Literaturen*, 155, 200-210.

Moneglia, Massimi and Tommaso Raso (2014), Notes on the Language into Act Theory, in Tommaso Raso and Heliana Mello (eds), *Spoken Corpora and Linguistics Studies*. Amsterdam/Philadelphia: John Benjamins, 468-494.

Ozerov, Pavel (2018). Tracing the sources of information structure: Towards the study of interactional management of information, *Journal of Pragmatics* 138, 77-97.

Paducheva, Elena V. (2015), Kommunikativnaja struktura predloženija. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki [Communicative structure of the sentence. Materials for a corpus-based Russian grammar], Moscow, http://rusgram.ru/.

Raso, Tommaso and Heliana Melo (eds) (2014), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins.

Sloussar, Natalia (2007), *Grammar and Information Structure. A Study with Reference to Russian*, Utrecht: LOT Publications.

Tomlin, Russel. S. (1995), Focal attention, voice, and word order, in P. Downing and M. Noonan (eds.), *Word Order In Discourse*. Amsterdam/Philadelphia: John Benjamins, 517-552.

Van Gijn, Rik, Jeremy Hammond, Dejan Matić, Saskia Van Putten & Ana Vilacy Galucio (eds.) (2014), *Information Structure and Reference Tracking in Complex Sentences*. Amsterdam/Philadelphia: John Benjamins.

Wilbur, Ronnie B. (2012), Information structure, in Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An International Handbook*. Berlin/Boston: De Gruyter Mouton, 462-489.

Yanko, Tatiana (2008), *Intonacionnye strategii russkoj reči v tipologičeskom aspekte* [Intonational strategies in spoken Russian from a comparative perspective], Moscow: Jazyki Slavjanskix Kul'tur.

Zimmermann, Malte and Caroline Féry (eds) (2010), *Information Structure: Theoretical, Typological, and Experimental Perspectives*, Oxford University Press.

Aboutness topics and referential hierarchy in Italian Sign Language (LIS)

Chiara Calderone
(University of Venice)

Keywords: Aboutness Topic, Referential Hierarchy, Accessibility, Italian Sign Language, Retrievability

The topicality in spoken languages has been addressed from a pragmatic, syntactic and prosodic perspective (Chafe 1976, Givón 1983, Gundel 1985, Frascarelli & Hinterhölzl 2007, Reinhart 1981). According to previous studies, the **aboutness topic**(AbT) is a given referent which represents what the sentence is about; cross-linguistically it may be realized with different referential categories(Ariel 1988, 1991) and marked by specific intonation contours.

As both syntactic and prosodic cross-linguistic differences have been detected in sign languages (SLs) (Aarons 1994, Brunelli 2011, Wilbur 2012, Sze 2013, Kimmelman 2014, Kimmelman and Pfau 2016), the purpose of this study is threefold: (i) to investigate the **prosodic features** of AbTs in LIS, considering manual(MMs) and non-manual markers (NMMs), (ii) to account for the **accessibility** of a referent in relation to its **informative status** (continued/shifted), to the **sentential distance** between the same referents and to the number of **competitors** (iii) to sketch a first **referential hierarchy** in LIS, by considering the distribution of syntactic categories such as full DPs, pronouns and null arguments with different verb types (Meir 2002, Benedicto & Brentari 2004, Frederiksen & Mayberry 2016, Koulidobrova 2016, Czubec 2017, Ahn 2019).

The data (1654 items), collected from nine native-signers, and analysed through ELAN and "R" (Crasborn & Sloetjes 2008) consists of two types: spontaneous storytelling between couple of signers and single monologues.

In LIS three tendencies of NMMs accompany AbT(ex. 1): (i) **raised-eyebrow** (re) (30%); (ii) **squinted-eyebrows** (sq) (32%) (iii) **head-tilt back** accompanying 13% of pronominal AbT subjects; prosodic markers (eye-blink and/or head-nod) separate AbT (20%) from the sentence. Similarly to

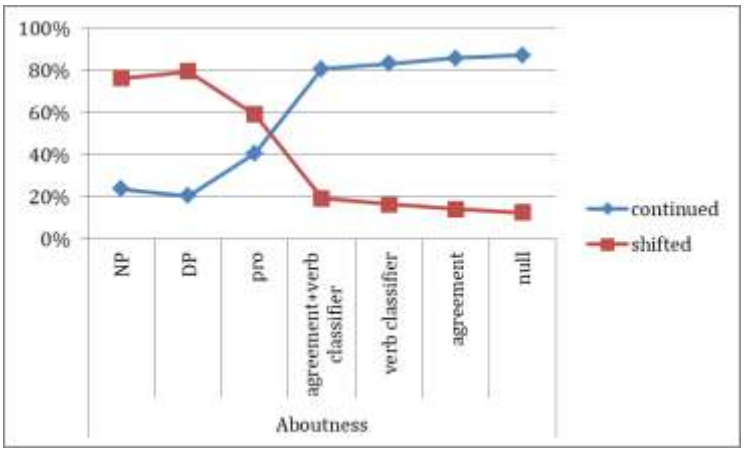
(Dachkovsky & Sandler 2009), in LIS squinted-eyebrows fulfils a pragmatic function by signalling the low accessibility of a referent, its occurrence with DPs is statistically significant (*odds ratio* 6.06, $p<0.001$).

re
sq sq hn eb
(1) **IX-loc, IX-3 WIFE DAUGHTER TABLE SIT.**

'Tomorrow, in an estate agency John buys a house'. [Fi_5st_60]

In line with (Ariel 1988, 1991), which state that the more salient the information is, the less linguistic material the speaker needs to codify it, in LIS 84% of continued AbT are realized as null arguments, on the contrary 78% of shifted AbTs (20%) are realized as NP/DPs, and the correlation is significant (*odds ratio* 9.29, $p<0.001$), while pronouns occupy an intermediate position, as can be seen in the Chart (1).

1.



Sometimes, null arguments can also be realized as shifted AbTs. Indeed, language-specific strategies make the omitted referent retrievable even when shifted, such as predicative classifiers and agreeing verbs. By conveying anaphoric functions, the predicative classifier (JUMP) from the example (1) allows the drop of the Ab referent, in fact, it may only refers to the dog.

(1) sq

DOG BARK, BEAR$_i$ IX-3$_{i-a}$ FEAR, CL-FALL$_a$, **CL-JUMP**…
'The dog barks, the bear is scared, (he) falls and (the dog) jumps over him'. [Mi_Mo_53-56]

For a more comprehensive picture of the accessibility of AbT, also sentential distance between antecedent-anaphora and the occurrence of competitors within this distance have been taken into account. The statistical analysis confirms that the interaction between squinted-eyebrows and DPs is emphasized by sentential distance (*odds ratio* 1.027, $p<0.01$) and competitors (*odds ratio* 1.41, $p<0.01$).

Such a research represents a first study on referentiality in LIS and an important litmus test for validating theories on the pragmatic, syntax and prosody of information structure proposed for spoken languages(Ariel 1991).

I would like to express my gratitude to the prof. Vadim Kimmelman and the prof. Davide Bernasconi for their crucial help in statistics and to the signers who have been involved in the collection of data. Part of this research would not have been possible without the SIGN-HUB project,

which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693349.

References

Aarons, Debra (1994), *Aspect of the Syntax of American Sign Language*, Doctoral dissertation, University of Boston.

Ahn, Dorothy (2019), Anaphoric expressions in ASL, Doctoral Dissertation, Harvard University.

Ariel, Mira (1988), Referring and accessibility, *Journal of Linguistics*, 24: 65–87.

Ariel, Mira (1991), The function of accessibility in a theory of grammar, *Journal of Pragmatics*, 443-463.

Benedicto, Elena & Diane Brentari (2004), Where Did All the Arguments Go?: Argument-Changing Properties of Classifiers in ASL, *Natural Language & Linguistic Theory* 22(4),743-810.

Brunelli, Michele (2011), *Antisymmetry and sign languages: A comparison between NGT and LIS*, Doctoral dissertation, University of Amsterdam & Ca' Foscari University.

Chafe, Wallace (1976), Givenness, contrastiveness, definiteness, subject, topics, and point of view. In Charles Li (ed.), *Subject and Topic*, New York, Academic Press, 25–56.

Crasborn, Onno & Han Sloetjes (2008), Enhanced ELAN functionality for sign language corpora. *Proceedings of LREC 2008*, Sixth International Conference on Language Resources and Evaluation, Marrakech, Morocco.

Czubek, Todd A (2017), A comprehensive study of referring expressions in ASL, Doctoral Dissertation, Boston University.

Dachkovsky Svetlana & Wendy Sandler (2009), Visual Intonation in the Prosody of a Sign Language, *Lang. Speech*, 52, 287.

Frascarelli, Mara & Roland Hinterhölzl (2007), Types of Topics in German and Italian, Schwabe, Kurt; Winkler, Susanne (eds.), *On Information Structure, Meaning and Form*, Amsterdam: John Benjamins, 87-116.

Frederiksen, Anne Therese & Mayberry, Rachel I.. 2016. Who's First? Investigating the referential hierarchy in simple native ASL narratives, *Lingua*, 180, 49-68.

Givón, Talmy (1983), *Topic continuity in discourse: A quantitative cross-language study*, Philadelphia, PA: John Benjamins.

Gundel, Jeanette (1985), 'Shared Knowledge' and topicality, *Journal of Pragmatics*, 9, 83-107, North-Holland.

Kimmelman, Vadim (2014), *Information Structure in Russian Sign Language and Sign Language of the Netherlands*, Doctoral dissertation, University of Amsterdam.

Kimmelman, Vadim & Roland Pfau. (2016), Information Structure, In C. Féry & S. Ishihara (eds.) *The Oxford handbook on information structure*, Oxford: Oxford University Press, 814–833.

Lambrecht, Knud (1994), *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*, Cambridge: Cambridge University Press.

Meir, Irit (2002), A cross-modality perspective on verb agreement, *Natural Language and Linguistic Theory*, 20, 413-450.

Koulidobrova, Elena (2016), Elide me bare, Null argument in American Sign Language (ASL), *Nat Lang Linguistic Theory*, 35, 397-446.

Reinhart, Tanya (1981), Pragmatics and Linguistics: an analysis of sentence topics. *Philosophica*, 27, 53-94 Sze, Felix (2013), *Nonmanual marking for topic constructions in Hong Kong Sign Language*, Doctoral Dissertation, University of Bristol.

Wilbur, Ronnie B (2012), Information Structure, In Pfau, Roland, Markus Steinbach & Bencie Woll (eds.), *Sign language. An international handbook*, 462-489, Berlin: Mouton de Gruyter.

Focus in Russian Sign Language: A quantitative study

Vadim Kimmelman
(University of Bergen; Russian State University for the Humanities)

Keywords: focus, contrast, prosody, doubling, Russian Sign Language

Focus has been claimed to be a universal category (Zimmermann & Onea 2011), at least in spoken languages. However, any universal claims have to be tested against sign languages (SLs). While some research on SLs exists (Wilbur 2012), most SLs have not been studied in this respect. In this paper we present the results of a research on expression of focus in Russian Sign Language (RSL).

We collected data from 10 native RSL signers. We conducted a quasi-experimental task in which the signers were presented questions in RSL about pictures; they had to repeat the question and answer it using the information from the pictures. The data was statistically analyzed using mixed effects modeling (Baayen et al. 2008) to study the effect of focus type, syntactic scope of focus, and movement type of the focused signs on potential focus markers.

The research has shown that focus in RSL is expressed by a variety of means, namely word order, non-manual markers and manual prosodic markers. Word order marking primarily concerns (optional) doubling whereby a constituent appears in the sentence twice to mark focus (1), furthermore, the non-focused part of the sentence is often elided (2). Non-manual markers of focus in RSL are used less often than in other sign languages studied before (Crasborn & van der Kooij 20013), e.g. eyebrow raise and head tilts are barely used. However, corrective and selective focus is sometimes marked by head nods and forward body leans (3).

(1) GIRL [KISS BOY KISS]$_{FOC}$
(What did the girl do?) 'The girl kissed the boy.'

(2) INDEX GIRL INDEX
(Who is hitting the boy?) 'A girl.'
nods

(3) [BOY]$_{FOC}$ WALK
(Who is walking, a boy or a girl?) 'A boy is walking.'

Finally, focused signs are generally different in manual prosody from their non-focused counterparts. In particular, focused signs are longer, larger, slower, contain more repetitions, and are articulated higher. All types of focus (information, selective and corrective), and all scopes of focus (subject, object, verb) are marked this way, although information focus is marked significantly more often by size and height than selective or corrective focus (for size, the percentage of marked cases is 26% for information focus, 5% for selective and 12% for corrective focus, and for height, the percentages are 18%, 7% and 5% for the same categories).

To conclude, we find that RSL uses a variety of syntactic and prosodic means to mark focus as expected if focus is a universal category. One surprising finding is that, contrary what has been found in spoken languages (Zimmermann & Onea 2011), contrastive focus in RSL seems to be less marked than non-contrastive focus, at least by manual prosody. A possible explanation is that non-manual prosody is used as the primary marker of contrast in this language, as in (3).

Acknowledgments: This research has been supported by the RSF grant 17-18-01184.

References:

Baayen, Rolf H., Douglas J. Davidson & Douglas M. Bates (2008), Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4), 390–412.

Crasborn, Onno & Els van der Kooij (2013), The phonology of focus in Sign Language of the Netherlands. *Journal of Linguistics* 49(03), 515–565.

Wilbur, Ronnie B. (2012), Information Structure, in R. Pfau, M. Steinbach & B. Woll (eds), (2012), *Sign Language: An International Handbook*. Berlin: De Gruyter Mouton, 462–489.

Zimmermann, Malte & Edgar Onea (2011), Focus marking and focus interpretation. *Lingua* 121(11), 1651–1670.

ꕥ

The marking and interpretation of contrast in Catalan Sign Language

Alexandra Navarrete-González
(Universitat Pompeu Fabra, Barcelona)

Keywords: contrast, selection, correction, expectations, exhaustivity

In Catalan Sign Language (LSC) foci and topics seem to display the same marking in order to express contrast (Barberà 2012, Navarrete-González 2016 and Zorzi 2018). In this research we aim at describing the different syntactic and prosodic non-manual markers (NMMs) involved in the expression of contrast in LSC. Moreover, we offer an analysis of the expression of contrast from a semantic-pragmatic perspective following Vallduví & Vilkuna (1998).

The data for this research have been collected from two deaf native signers of LSC through different elicitation tasks: question-answer pairs, dialogues, translations, and felicity judgments. Moreover, the LSC corpus from the *Institut d'Estudis Catalans* has been consulted in order to contrast the elicited data with natural discourse data.

Contrast in LSC is mainly expressed through a combination of prosodic NMMs: left and right body leans (bl), and head tilts (ht), and morphophonological marking: the use of the opposite sides of the signing space. These markers are always present when there are salient contextually contrasted alternatives and can spread across different types of constituents (1).

(1) left_bl right_bl
[[GIORGIA]$_T$ [LINGUIST]$_F$]$_x$ [[RAQUEL]$_T$ [INTERPRETER]$_F$]$_y$.
'Giorgia is a linguist and Raquel is an interpreter.'

When the context requires selecting one alternative and excluding the other, a repeated head nod (hn) is commonly added, which is directed towards the selected alternative (2).

(2) What is the woman doing: riding a bike or riding a horse?
right bl+ht+hn
[BIKE RIDE]$_x$
'Riding a bike.'

In order to express a correction, body leans and head tilts are again displayed, but, in addition, a strong head thrust (hthr) is found in the correction (3).

		o		hthr
	left_bl+ht		right_bl+ht	
(3)	[MARY PIZZA-EAT NOTHING]$_x$, [OTHER BURGER]$_y$.			
	Mary didn't eat a pizza, but a burger.'			

In LSC forward-backward body leans can be found in combination to left-right body leans, especially in corrections. If the correction includes more than one element, a left-right body lean expresses the contrast between those elements, and the contrast between the correction and the corrected element is expressed with a forward-backward body lean —similar findings for Sign Language of the Netherlands (Crasborn & van der Kooij 2013)—.

Following Umbach (2004) and inspired by the classification used in Kimmelman (2014), we argue that in LSC three different types of contrast can be distinguished: i) *parallel contrast* (1), which involves contrast due to similarity plus dissimilarity; ii) *selective contrast* (2), which involves contrast due to similarity plus dissimilarity, plus contrast due to exclusion; and iii) *corrective contrast* (3), which involves contrast due to similarity plus dissimilarity, due to exclusion, and a violation of expectations (Destruel & Velleman 2014). All these types of contrast are expressed through the combination of markers mentioned before. However, in ii) an additional head nod is found towards the selection, and in iii) a head thrust is added to mark the correction. Our analysis proposes that the types of contrast identified before correlate with different interpretations in terms of exhaustivity, related to the selection of an alternative, and expectations, related to the correction of an alternative.

This research is possible thanks to the SIGN-HUB project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693349. Thanks also to the Spanish Ministry of Economy and Competitiveness and FEDER Funds (FFI2012-36238 "ClauseCombi2 "GramRefLSC FFI2015-68594-P, BES2013062848) and the Generalitat de Catalunya (2017 SGR 1478).

References

Barberà, Gemma (2012), *The meaning of space in Catalan Sign Language (LSC): Reference, specificity and structure in signed discourse*. PhD Dissertation. Universitat Pompeu Fabra.

Crasborn, Onno & Els van der Kooij (2013), The phonology of focus in Sign Language of the Netherlands. *Journal of Linguistics* 49, 515-565.

Destruel, Emilie & Leah Velleman (2014), Refining contrast: Empirical evidence from the English it-cleft. *Empirical Issues in Syntax and Semantics* 10, 197-214.

Kimmelman, Vadim (2014), *Information Structure in Russian Sign Language and Sign Language of the Netherlands*. PhD Dissertation. University of Amsterdam.

Navarrete-González, Alexandra (2016), *Focus constructions in Catalan Sign Language*. MA thesis. Universitat Pompeu Fabra.

Umbach, Carla (2004), On the notion of contrast in information structure and discourse structure. *Journal of Semantics* 21, 155-175.

Vallduví, Enric & Maria Vilkuna (1998), On rheme and kontrast. *The Limits of Syntax*. New York: Academic Press. 79-108.

Zorzi, Giorgia (2018), *Coordination and gapping in Catalan Sign Language (LSC)*. PhD Dissertation. Universitat Pompeu Fabra.

On the lack of integration of reported speech in Russian Sign Language

Evgeniia Khristoforova
(Russian State University for the Humanities)

Keywords: Russian Sign Language; information structure; reported speech; non-manual markers; sign language discourse

Reported speech is an interesting phenomenon on the discourse level because while being by definition recursive (speech embedded within speech) it is often not syntactically embedded or prosodically integrated, especially when direct speech is involved (Brendel et al. 2011). We wanted to find out whether this simultaneous integration and lack thereof is also present in the visual modality: in sign languages. We thus conducted a corpus study of reported speech in Russian Sign Language (RSL) in naturalistic discourse.

We used the Moscow narrative sub-corpus of the online RSL corpus (http://rsl.nstu.ru/) containing narratives and detected and annotated contexts involving reported speech and thought (341 tokens). We found that, similar to what has been described for other sign languages (Quer 2011), reported speech and thought in RSL is often marked by non-manual markers (head and body turns, eye gaze direction change), as in (1). Consequently, on the discourse level, reported speech is distinguished from the main narrative, albeit not obligatorily (approximately 90% of tokens are marked by one or more non-manual markers).

eyes left, head turn left
(1) I GOOD FINISH
'I say: good, that's it.'

At the same time, we found very little signs of syntactic embedding or prosodic integration. A handful of examples demonstrate one clear sign of syntactic subordination (the use of a conjunction THAT, which cannot be used in a main clause) (2), but they might be instances of interference from spoken Russian; other examples show no evidence of syntactic embedding. PU refers to a palms up sign, a multi-functional discourse marker.

eyes right, head turn right, body turn right
(2) TELL THAT I PU LAMP INDEX$_1$ NO
'I told her that my lamp was missing.'

Prosodically, reported speech can form full-fledged utterances. This is evidenced, for instance, by examples with imperative illocutionary force. For instance, (3) contains the manual imperative sign IMP also marked by imperative non-manuals; note that the whole example is accompanied with head turn marking reported speech. We also find by examples with reported questions also non-manually marked as regular questions, as in (4): note that eyebrow raise marks the quote as a question, and the head turn as reported speech.

head turn left — head turn left + eyes wide open, head nod
(3) WOMAN PU TICKET CHANGE CAN PAST, IMP!
'The woman says: you could change the tickets, go!'

brow raised, head turn
(4) I PU I WHAT THIS?
'I'm like: what is this?'

So, similarly to spoken languages, reported speech in RSL shows embedding on discourse level, but, frequently, a lack of embedding syntactically or prosodically. This duality is even more transparent than in spoken languages because reported speech is often marked by specific non-manuals in RSL. Due to the affordances of the visual modality (simultaneous use of manual and several non-manual markers), it is possible both to preserve prosodic independence of reported speech and also clearly mark it as reported speech by using different markers in parallel.

Acknowledgment: This research has been funded by the RSF grant 17-18-01184.

References:

Brendel, Elke, Meibauer, Jörg and Markus Steinbach (eds.) (2011), *Understanding quotation*. Berlin: Mouton de Gruyter.

Quer, Josep (2011), Reporting and quoting in signed discourse, *Understanding quotation* (pp. 277–302).

Theticity and sentence structure: What can we learn from colloquial Israeli Hebrew?

Shlomo Izre'el
(Tel Aviv University)

Keywords: Theticity, sentence structure, information structure, Israeli Hebrew

In constructions expressing categorical judgments, syntactic predication reflects semantic predication, whereas constructions expressing thetic judgments manifest different relationship between these respective predications of or suggest the lack of predication thereof, depending on theory (Sasse 2006; Cornish 2008).

The syntactic approach adopted here is functional, communicational, discursive, and information oriented. For the study of spoken language, it integrates segmental and suprasegmental features. Within this paradigm, it has been suggested that the only necessary and sufficient component constituting a clause is the predicate phrase, which need not be seen as depending on a subject, thus forming by itself a unipartite clause. Corpus research has revealed that about half of the (matrix) clauses in colloquial Hebrew do not include subjects. Most significantly, in Hebrew, any part of speech can function as a predicative nucleus. Given the above, a new perspective of what consists of a predicate is in order. The predicate phrase is defined as the component that carries the informational load of the clause within the discourse context, including by default a "new" element in the discourse, carries modality (broadly defined), and focused (Izre'el 2018a). This framework opens a path to converge information structure and syntax in the analysis of thetic constructions. Thus, the approach proffered here predicts that (the majority of) thetic constructions merge syntactic and semantic structure, being essentially unipartite clauses.

E.g., in presentative-existential sentences, usually regarded as typical thetic constructions, the existential component is analyzed as a modal phrase, which can include either a single particle (1) or consist of a verbal complex (2), whereas the following NP ("pivot") is analyzed as the lexical component of the predicate phrase.

	EXT	**PIVOT**
(1)	*jef*	*nivχer-et=ʃait* ‖
	EXT	team-F=sailing
	‘There is a sailing team.’	
(2)	*haja-ø*	*tmun-a* (…) ‖
	be\PFV-3SGM	picture-F
	‘There was a picture (…).’	

This V1 order (which contrasts the default SV order of Modern Hebrew) does not necessarily suggest that the pivot be regarded as subject. It will be noted that the person marker of the verb is non-referential, which takes the form of the 3SGM person. So are other thetic constructions, like the one in (3).

(3)	*niʃar-ø*	*od*	*ʧips-im* /
	remained-3SGM	more	chip-PL
	‘Are there any (potato-)chips left?’		

An alternative configuration shows agreement between the two referential components of the sentence, as in (4).

(4)	*afilu*	*χaser-a*	*aruχ-a*	‖
	even	missing-F	meal-	F
	‘There is even a meal missing.’			

In many of these constructions, the anticipatory agreement morpheme can be analyzed as a focus marker (cf. Izre’el 2018b).

In my talk, I will look at segmental and suprasegmental configurations of various types of thetic constructions and examine their functions, trying to show that the syntax of thetic constructions reflects their function and meaning in a straightforward manner. The constraints on both types of configurations will be looked at, suggesting a continuum of thetic-like constructions shown in their variant structures between unipartite and bipartite clauses and the interrelations between information structure and syntax.

The research takes a corpus-driven approach, using mainly *The Corpus of Spoken Israeli Hebrew*.

References

Cornish, Francis (2008), L’absence de prédication, le topique et le focus : le cas des phrases "thétiques", in Jean-Marie Merle (ed), *La predication*, (Faîtes de Langues, 31-32), Paris: Ophrys, 121-131.

Izre’el, Shlomo (2018a), Syntax, prosody, discourse and information structure: The case for unipartite clauses — A view from spoken Israeli Hebrew, *Revista de Estudos da Linguagem* 28(4), 1675-1726.

Izre’el, Shlomo (2018b), Presentative-existential sentences as unipartite sentences: A view from spoken Israeli Hebrew, in S. Masalóva, V. Polyakov, and V. Solovyev (eds), *Cognitive Modeling: Proceedings of the Fourth International Forum on Cognitive Modeling* (30 September – 7 October, 2018, Tel Aviv, Israel), Rostov-on-Don: Science and Studies Foundation, 79-88.

Sasse, Hans-Jürgen (2006), Theticity, in Giuliano Bernini and Marcia L. Schwartz (eds), *Pragmatic Organization of Discourse in the Languages of Europe*, (Empirical Approaches to Language Typology; EUROTYP, 20-8), Berlin: Mouton de Gruyter, 255-308.
The Corpus of Spoken Israeli Hebrew (CoSIH). <cosih.com>.

Sentence focus and topicality in focus constituents

Vydrina Alexandra
(LACITO-CNRS)

Keywords: Sentence Focus, topic, Mande languages, pragmatics, intersubjectivity

There are two common beliefs about Sentence Focus (SF), or thetic utterances, in the literature: 1) if the whole sentence is in focus, the whole of its content, in particular, the subject, is discourse-new; 2) sentence focus is implemented through the focalization of the subject or, at least, though its de-topicalization (e.g. Lambrecht and Polinsky 1997). These beliefs are partly due to the fact that the previous studies on SF (Sasse 1987; Lambrecht 2000; Lambrecht and Polinsky 1997) were centered on constructions with a new subject and an intransitive predicate as in (1), whereas the transitive SF constructions as well as SF constructions containing non-new elements are usually left out in the discussion.

The present study fills in this gap, and provides an account of how the discourse status of the arguments included in a SF constituent influences the expression of the focus. Using the example of SF constructions in Kakabe, a Western Mande language, I will show the two points concerning SF that are mentioned above, need reconsideration.

First, SF articulation are often used by speakers in a way that its arguments are already activated, in particular, in SF inferentials, illustrated in (2); see Delahunty (1995); Bearth (1997, 1999) on inferentials. The inferential construction highlights a new aspect of the intersubjectivity aspect of focus, in terms of Traugott (2010). It reveals the speaker's consciousness of the inferences that the listener makes (hence the term): it shows that the speaker monitors the interpretations that the listener makes of his speech.

Second, and this is related to first, the discourse status of the arguments within the SF influences the implementation of this type of focus. Thus, in Kakabe, the focus marker is preferably hosted by the subject (appearing to its right), unless the latter is activated. For example, in (3) from a tale, *lè* is placed on the DO but the focus is not on the direct object nor on VP. Instead, the whole proposition 'I took my hoes' is the answer to the question 'what happened' and the utterance is therefore SF.

In line with e.g. Lambrecht and Polinsky (1997), when the scope of focus ranges over the whole clause, the subject is prioritized for hosting the focus marker. A new point about the SF articulation that will be brought to light in my talk is that a SF utterance can contain a subject that is activated and therefore cannot host the focus marker. In this case the focus marker is passed on to other, non-activated, argument, or, in lack of the latter, to the verb. The existence of such mechanism highlights the continuity between sentence focus and the pragmatic articulations of narrower scope such as VP and argument focus: in all these cases, the focus marker tends to be placed on the more informative referent.

(1) a. *(What's new?) RAEGAN died*
b. *(What about Raegan?) Raegan DIED.*

(2) Kakabe: Sentence focus, focus particle on subject, inferential value

jé˖é	*=lè*	*báànù*	*tó-lén*	*ɲɔ́ɔ̀*	*là*
water.ART-FOC	be.3PL	leave-PC.ST	that	OBL	

[The crocodile asked me to carry him here] It is that the water had left him there [so that he was stuck here].

(3) Kakabe: Sentence focus, focus particle on DO

ǹ	*bí*	*ǹ*	*nà*	*kérè-nú=lè*	*jìgàlèn*
1SG	be	1SG	POSS	hoe.ART-PL-FOC	take-PC.ST

kà	*á*	*dòní*	*nùméè*	*bàtà*
INF	3SG	send	smith.ART	at

[Birds, listen to me what happened]: I took my hoes and went with them to the smith's.

References

Bearth, Thomas. 1997. "Inferential and Counter-Inferential Markers in Swahili Dialogue." *Afrikanistische Arbeitspapiere* 51.

———. 1999. "The Inferential Gap Condition." *Pragmatics* 9 (2).

Delahunty, Gerald P. 1995. "The Inferential Construction." *Pragmatics* 5 (3): 341–64. https://doi.org/10.1075/prag.5.3.03del.

Lambrecht, Knud. 2000. "When Subjects Behave like Objects: An Analysis of the Merging of S and O in Sentence-Focus Constructions across Languages." *Studies in Language* 24 (3): 611–82.

Lambrecht, Knud, and Maria Polinsky. 1997. "Papers from the Regional Meetings of the Chicago Linguistic Society." *Papers from the Regional Meetings of the Chicago Linguistic Society* 33 (2): 189–206.

Sasse, Hans-Jürgen. 1987. "The Thetic/Categorical Distinction Revisited." *Linsuistics* 25: 511–80.

Traugott, Elizabeth. 2010. "(Inter)Subjectivity and (Inter)Subjectification: A Reassessment." In *Subjectification, Intersubjectification and Grammaticalization*, edited by Kristin Davidse, Lieven Vandelanotte, and Hubert Cuyckens, 66:29–74. Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110226102.1.29.

Left Dislocation in spontaneous Hebrew – it is neither topicalising nor a construction

Pavel Ozerov
(University of Münster)

Keywords: topic, information structure (IS), Hebrew, interaction, projection

Left Dislocation (LD) is a classical IS-construction that exhibits a clear-cut topic-comment partition (e.g. Kerr 2014). However, interactional analyses question this view, analyzing LD as a routinised resource for interaction management (Pekarek-Doehler et al. 2015 for French). Using the theoretical framework of incremental-projecting online syntax (Auer 2005; Hopper 2011), I examine bare NP's

that occupy their own Intonation Unit (IU) without complementing a previously initiated construction (i.e. not *projected* syntactically (Auer 2005)). The data come from two corpora of spontaneous Hebrew.

From the perspective of online syntax, producing a stand-alone NP the speaker faces various potential continuation trajectories. This NP can form a self-sufficient contribution, and/or to be continued by subsequent IU's. In the latter case it can be cross-referenced within the following IU's (hence producing a LD) or left unmentioned. While around 20% of the examined ca. 600 tokens of stand-alone NP's can be classified as LD-like relative to the subsequently produced material, this would be an *a posteriori* analysis of the dynamically evolving structure. The contribution of the NP fulfills in each case **a local action** shared with other occurrences of stand-alone NP's which do not evolve into a LD.

For instance, a prominent role typically attributed to LD is to introduce a new referent, with the proposition being construed as "about it". However, corpus data show that activation of a referent is a separate interactional move, structurally unrelated to the yet unplanned continuation. It can be aimed at the negotiation of its identification with the interlocutor or shift their attention towards it. The NP can also achieve local goals such as update or stance-expression. Even if the structure eventually develops into a LD, the introduction of the new referents fulfills local discourse goals and is independent of the subsequent material. This content-managing function of stand-alone NP's accounts for ~30% of LD-like constructions.

Another prominent source for LD outcomes is NP's used as **starting points** and a subsequent **modification** of the chosen trajectory into constructions. The online trajectory modification is evident in these cases from numerous production difficulties. The apparent frequency of LD here is the outcome of the routinised usage of bare NP's as **turn-taking devices**, due to their very vague syntactic *projection* (i.e. multiple continuation possibilities). In 36% of LD-like examples the NP serves for turn-taking, while 46% involve trajectory modification with evidence for discontinuity. Moreover, the function of NP's and pronouns in this role is identical, yet the latter examples are disregarded as false starts in IS-oriented studies.

Hence, this study demonstrates that what is perceived as LD is a discourse collocation of an initial stand-alone NP and a follow-up material. The usage of this NP is a discourse move with local content-managing and interaction-regulating roles. It is not related structurally to a possible, but not obligatory and yet unplanned continuation. Topicality is not required for its analysis and is a vague, optional, theory-driven interpretation of the actual devices used for interactional information management (Matić and Wedgwood 2013; Ozerov 2018).

References

Auer, Peter. 2005. "Projection in Interaction and Projection in Grammar." *Text* 25 (1): 7–36.

Hopper, Paul. 2011. "Emergent Grammar and Temporality in Interactional Linguistics." In *Constructions: Emerging and Emergent*, edited by Peter Auer and Stefan Pfänder, 22–44. Berlin/Boston: De Gruyter Mouton.

Kerr, Betsy. 2014. "Left Dislocation in French: Information Structure vs. (?) Interactional Linguistics." In *Perspectives on Linguistic Structure and Context: Studies in Honor of Knud Lambrecht*, edited by Stacey Katz Bourns and Lindsy L. Myers, 223–40. Pragmatics & Beyond 224. Amsterdam/Philadelphia: John Benjamins.

Matić, Dejan, and Daniel Wedgwood. 2013. "The Meanings of Focus: The Significance of an Interpretation-Based Category in Cross-Linguistic Analysis." *Journal of Linguistics* 49 (1): 127–63. https://doi.org/10.1017/S0022226712000345.

Ozerov, Pavel. 2018. "Tracing the Sources of Information Structure: Towards the Study of Interactional Management of Information." *Journal of Pragmatics* 138: 77–97. https://doi.org/10.1016/j.pragma.2018.08.017.

Pekarek-Doehler, Simona, Elwys De Stefani, and Anne-Sylvie Horlacher. 2015. *Time and Emergence in Grammar: Dislocation, Topicalization and Hanging Topic in French Talk-in-Interaction*. Amsterdam/Philadelphia: John Benjamins.

Intonation patterns of complex utterances in Russian spoken discourse

Nikolay Korotaev
(RSUH, Moscow)

Keywords: discourse prosody, transitional continuity, pitch adaptation, intonation and syntax, spoken narration

In this talk, I analyze the intonation structure of complex utterances (CUs) in Russian spoken narratives. I define Cus as comprising at least two clauses and/or other discourse units such that only the last one exhibits formal properties of completion (or, in terms of Du Bois et al. 1993, final transitional continuity), whereas the preceding unit(s) is/are intonationally non-final. CUs provide a regular context for neutralization of the core IS distinction of topic vs. focus (see, e.g., Yanko 2008 for a Russian discussion). While the topic-focus opposition is most saliently manifested in isolated sentences (see Krifka & Musan 2012 for a cross-linguistic perspective), in natural spoken production discourse chunks tend to combine in stretches of various semantic, syntactic, and intonational organization (Chafe 1988; Kibrik 2011). Basing on the data from the prosodically annotated collection "Stories about presents and skiing" (Spokencorpora 2013), I claim that the analysis in terms of final vs. non-final transitional continuity yields useful insights on the internal structure of Cus regardless of whether the topic-focus distinction may be consistently applied.

Since narratives mostly consist of statements, I concentrate on this illocutionary type. In Russian complex statements, final and non-final elements are primarily contrasted on prosodic grounds. The base intonation pattern of Cus in Russian is driven by the principle of *pitch adaptation*: the word bearing the main accent in the non-final unit is pronounced with a /rising pitch that mirror-images the \falling accent in the final unit (Kodzasov 2009), see example (1). However, when a CU consists of more than two accented units, this base pattern is complicated. Three strategies of complication can be distinguished: (i) *multiple parallel adaptation*, (ii) *consecutive adaptation*, and (iii) *parenthetical strategy*. All of these are present in example (2). In multiple parallel adaptation, each non-final unit has a pitch accent that independently mirror-images that in the final unit; see rising accents in constituents with blue opening brackets. The consecutive adaptation, on the contrary, presupposes that a non-final unit adapts to a subsequent non-final unit, which results in a non-final falling accent; see emu \raspisal in (2). In parenthetical strategy, a non-final unit providing a (meta-)comment wedges in between two units that constitute a mainline (see discourse marker \vo-ot in green brackets).

Strategies (i) and (ii) are the most frequent in the analyzed data and their distribution varies significantly across stretches of different internal discourse organization. Non-final units that follow the consecutive adaptation strategy are usually more closely linked together in terms of syntactic

structure and/or semantic relations (see, e.g., an indirect question in red brackets) than those following the multiple parallel adaptation strategy.

Examples

(1) Kogda on /zašël,
when he entered
tam okazalsja \ba-anket.
there turned.out party
'*When he entered that place, it turned out there was a party'*

(2) [1/Tot [2emu \raspisal,
That.one to.him painted
[3kakaja ona super-puper /↓navoročennaja,3]2]
what it super fancy
( ) [4(\vo-ot,)4]
well
( ) [5i /–skazal-l,
and said
( ) [6čto-o /stoit ona-a ( ) [7\dorogo.7]6]5]1]
that costs it expensive
'*[the salesman] described to him how fancy it [the car] was and told him that it was expensive'*

(The research is supported by the Russian Science Foundation, grant 17-18-01184.)

References

Chafe, Wallace (1988), Linking intonation units in spoken English, in J. Haiman, and S. A. Thompson (eds.) (1988), *Clause combining in grammar and discourse*, Amsterdam / Philadelphia: John Benjamins, 1-29.

Du Bois, John., Schuetze-Coburn, Stephan, Paolino, Danae, and Cumming, Susanna (1993), Outline of discourse transcription, in J. A. Edwards, and M. D. Lampert (eds.) (1993), *Talking data: Transcription and coding in discourse research*, Hillsdale, NJ: Lawrence Erlbaum Associates, 45-89.

Kibrik, Andrej A. (2011), Cognitive discourse analysis: local discourse structure, in M. Grygiel, and L.A. Janda (eds.) (2011), *Slavic Linguistics in a Cognitive Framework,* Frankfurt / New York: Peter Lang Publishing Company, 273–304.

Kodzasov, Sandro V. (2009), *Issledovanija v oblasti russkoj prosodii* [*Studies in the field of Russian prosody*], Moscow: Jazyki slavjanskix kul'tur.

Krifka, Manfred, and Musan, Renate (eds.) (2012), *The Expression of Information Structure*, Berlin / Boston: De Gruyter Mouton.

Spokencorpora (2013), *Prosodically Annotated Corpus of Spoken Russian (PrACS-Russ),* Pilot version, available at: http://spokencorpora.ru.

Yanko, Tatiana (2008), *Intonacionnye strategii russkoj reči v tipologičeskom aspekte* [*Intonational strategies in spoken Russian from a comparative perspective*]. Moscow: Jazyki slavjanskix kul'tur.

Information structure and prosody in Russian reported speech

Vera Podlesskaya
(Russian State University for the Humanities)

Keywords: information structure, reported speech, prosody, Russian, main clause phenomena

Background and research questions. The paper investigates information structure (IS) and prosodic shaping of Russian speech reporting constructions (SRC). SRC have been thoroughly studied in theoretical, typological and descriptive linguistics (cf. inter alia Coulmas (ed.) 1986; Janssen, van der Wurff (eds.) 1996; Güldemann, von Roncador (eds.) 2002; de Vries 2008; Aikhenvald 2011; Buchstaller, van Alphen (eds.) 2012; Buchstaller (ed.) 2013, Evans 2013,) The wide variety of SRC can be located on a continuous scale of integration with direct and indirect speech serving as idealized extreme points: being the least integrated (direct speech, cf. *Vera said: "I do linguistics"*) and the most integrated (indirect speech, cf. *Vera said that she does linguistics"*). In canonical indirect speech, grammatical integration of the reporting frame (*Vera said)* and the reported utterance (*I do linguistics*) is known to be performed through embedding and indexical shift; in canonical direct speech, grammatical integration can be blocked in case of the so called "main clause (or, root) phenomena" (Hooper & Thompson 1973, Padučeva 1996 Lobke et al 2012). In canonical cases, grammatical (dis)integration is supported by prosody (cf. e.g. Jansen, Gregory, & Brenier 2001; Kalmanovitch 2015; Kasimir 2008; Couper-Kuhlen 1998; Klewitz & Couper-Kuhlen 1999; Oliveira & Cunha 2004). However, further systematic studies are needed to reveal how the IS of the reporting frame and the reported utterance can be integrated and how the degree of IS integration interplays with prosodic shaping of SRC both in individual languages and cross-linguistically. The aim of this paper is to partially fill this gap bringing in Russian data.

Data. The analysis is based on data from the Prosodically Annotated Corpus of Spoken Russian (Spokencorpora 2018). Retrieved were all instances of SRC (n= 162) from 40 elicited spontaneous personal stories (in total, 10 000 words, 70 min of speech)

Analysis and Results. Four main types of SRC were detected in the corpus: Type 1 (direct SRC) and Type 2 (indirect SRC) are canonical cases where grammatical (dis)integration is supported by IS and prosody, while Types 3 and 4 deviate from the canon with grammar and prosody operating in opposite directions:

Type 3 are non-canonical direct SRC where grammatical disintegration (no complementizer, no indexical shift, main clause phenomena) is opposed by prosodic\IS integration: the reporting frame is produced with a non-terminal pitch movement (most often this is some sort of a rising pitch) while the reported utterance looses the ability to be independently marked for the illocutionary force.

Type 4 are non-canonical indirect SRC where grammatical integration (complementizer, indexical shift, no main clause phenomena) is opposed by prosodic\IS disintegration: the reporting frame is produced with a terminal pitch movement (most often, some sort of a falling pitch) making the listener believe that next comes a reporting utterance in a "direct" grammatical form, but this expectation is not fulfilled.

Deviant Types 3 and 4 together comprise 25% of the total amount of SCR retrieved from the corpus, which shows that they are by no means marginal, and thus should find their place in the system of SRC.

* The study has been supported by the Russian Science Foundation, grant # 17-18-01184

References

Aikhenvald A. Y. (2011) Speech reports: A cross-linguistic perspective. Language at large: Essays on syntax and semantics. Aikhenvald A. Y., Dixon R. M. W. (eds.). Brill: Leiden, 290–326.

Buchstaller (ed.) (2013) Quotatives: New trends and sociolinguistic implications. Hoboken (NJ): John Wiley & Sons.

Buchstaller, van Alphen (eds.) (2012) Quotatives: Cross-linguistic and cross-disciplinary perspectives. Amsterdam: John Benjamins.

Coulmas F. (ed.) (1986) Direct and indirect speech. (Trends in Linguistics. Studies and Monographs, 31.) Berlin: De Gruyter Mouton.

Couper-Kuhlen, E. (1998) Coherent voicing. On prosody in conversational reported speech. Interaction and Linguistic Structures, 1-28.

Evans N. (2013) Some problems in the typology of quotation: A canonical approach. Canonical morphology and syntax. Brown D., Chumakina M., Corbett G. G. (eds.). Oxford: Oxford Univ. Press, 66–98.

Güldemann T., von Roncador M. (eds.). Reported discourse: A meeting ground for different linguistic domains. Amsterdam: John Benjamins, 2002.

Janssen Th. A. J. M., van der Wurff W. (eds.). (1996) Reported speech: Forms and functions of the verb. Amsterdam: John Benjamins.

Jansen, W., Gregory, M. L., & Brenier, J. M. (2001) Prosodic correlates of directly reported speech: Evidence from conversational speech. Proceedings of the ISCA Workshop on Prosody in Speech Recognition and Understanding. Red Bank, NJ, 77–80.

Hooper, Joan & Thompson, Sandra (1973) On the applicability of Root Transformations.
Linguistic Inquiry, 4, 465–491.

Kalmanovitch Y. (2015) Jumping out of context — Jumping out of tone. Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, 10–14 August 2015. Available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0804.pdf .

Kasimir, E. (2008) Prosodic correlates of subclausal quotation marks. ZAS Papers in Linguistics, 49, 67–77.

Klewitz, G. & E. Couper-Kuhlen (1999) 'Quote - Unquote?' The role of prosody in the
contextualization of reported speech sequences. Interaction and Linguistic Structure, 1-34

Lobke A., Haegeman L., Nye R. (eds.). (2012) Main clause phenomena: New horizons. (Linguistik Aktuell / Linguistics Today, 190.) Amsterdam: John Benjamins.

Oliveira, M., & Cunha, D. A. C. (2004) Prosody as marker of direct reported speech boundary. Proceedings of Speech Prosody, 263–266

Paducheva E. V. (1996) Subjective modality: Illocutionary markers and parenthetical words. Paducheva E. V. Semanticheskie issledovaniya. Moscow: Shkola «Yazyki Russkoi Kul'tury», 297–320

de Vries M. (2008) The representation of language within language: A syntactico-pragmatic typology of direct speech. Studia Linguistica.. Vol. 62. No. 1, 39–77.

Spokencorpora (2018) Prosodically Annotated Corpus of Spoken Russian (PrACS-Russ). Pilot version. Available at: http://spokencorpora.ru.

ꝏ

Different means to encode IS: how they interact and when they do not

Natalia Slioussar, Natalia Tyshkevich & Sophia Popova
(HSE, Moscow, and SPbU, St.Petersburg; MSSES, Moscow; & SPbU, St.Petersburg)

Keywords: information structure, production, comprehension, syntax, prosody

It is well known that different means can be used to encode IS: prosodic, syntactic or morphological. However, we know much less how these means may interact. We address this question in three experiments on Russian and three pilot experiments on Adyghe. We chose Russian because it can use both word order and prosody to encode IS. Adyghe was chosen for the following reason. It has a special construction (FC) to encode focus. FC is very widespread, but sentences without it are frequent as well. According to the unpublished previous studies (e.g. Kimmelman, p.c.), when FC is not used, the word order is free and sentences have very similar prosodic structures independently from their word order and IS. We wanted to check the latter claim experimentally. Analyzing prosody, we used *Praat*, logistic regression was used for statistics.

In a production experiment, we provided Russian speakers with several texts. Some sentences were replaced with blanks followed by words in parentheses. Participants were supposed to put the words together and to read the texts aloud. Every text contained several dialogues in which the answers were sentences with the same subject, object and verb, but the different ISs, e.g. "Where is the soup? — Masha ate the soup", "What did Masha eat? — Masha ate the soup". These sentences were among those that participants had to put together, so both syntax and prosody could be studied.

In Adyghe, we were interested in prosody without the FC, so we asked a native speaker to translate texts using SOV without the FC in crucial sentences, if possible (studying word order would also be interesting, but requires a separate experiment). This was possible in all dialogues, which is interesting in itself. Then we asked several native speakers to read the texts aloud. In Russian, focused words always carried the main stress, while syntactic means were used much less frequently: sometimes to make the focused constituent sentence-final and sometimes to make it sentence-initial (focus-fronting is very infrequent in written texts, but was often used in our study). In Adyghe, focused words also were prosodically prominent.

Two comprehension experiments were the same in both languages. Firstly, we cut out crucial questions and answers (read by different speakers) and recombined them, so that in some pairs, they were from the same dialogue, and in the others — from different ones (with different IS). Participants were asked to rate their appropriateness (1-5). Secondly, we gave another group of participants only answers and asked them to come up with appropriate questions. Russian speakers could detect the IS of all answers well, althogh focus fronting was significantly more difficult than other constructions. Adyghe speakers did not use prosody to decode IS (only very weak tendencies were observed). Thus, the tendency to make focused information prosodically prominent may be universal (if the phonology of the language allows for this) (Szendrői 2005), but not all languages grammaticalized this tendency and use it to *encode* IS — we cannot speak of encoding if listeners do not decode this information.

References

Szendrői, Kriszta (2005), Focus movement (with special reference to Hungarian), in M. Everaert, and H. van Riemsdijk (eds), (2005), *The Syntactic Companion*, London: Blackwell, 270-335.

ꕥ

Gesture / Prosody Synchronization in Spontaneous Speech: A View from the Language into AcT Theory

Massimo Moneglia & Giorgina Cantalini
(University of Florence & Civica Scuola di Teatro Paolo Grassi, Milan)

Keywords: Gesticulation; Utterance; Prosody; Information Structure; Synchronization

Gesture shows a linear structure that can be segmented into units aligned to the verbal stream (Mc Neill 2005; Kendon 2004). The tight interrelation of gesture and speech is often remarked by the quote: "speech and movement appear together as manifestations of the same process of utterance" (Kendon 1980).

In speech, syntactic constituents are patterned into Information units, following the *flow of thought*, each one corresponding to Prosodic units (Chafe 1970). According to the Language into Act Theory-L-AcT (Cresti 2000; Moneglia & Raso 2014) Information units are structured into higher level Reference units, which are terminated from a prosodic point of view and correlate with the performance of Speech acts (Cresti & Moneglia 2005).

Gestures are structured in a configurational model. The minimal gestural linear pattern (Gesture Phrase) foresees a compulsory root (the Expressive Phase made by at least one Stroke) and possibly other Gesture Phases eventually collocated before and after it. Gestural Phrases are packaged within larger Gesture Units (Kendon 2004; Kita et al. 1998).

The paper presents the results of a corpus-based study focusing on the relations between gestural and prosodic units in spontaneous speech. The dataset is made by three heavily annotated samples of three video recorded structured interviews with actors about their professional experience.

Comparable samples of around three minutes of each interview have been extracted forming a corpus of around 10 minutes.

Gesture annotation is based on LASG (Bressem et al. 2013); Co-speech gestures have been segmented at three hierarchic levels, each one aligned to the acoustic source:

- Gesture Units: sequences of gestures between two rest positions;
- Gesture Phrases: phases of a gesture around a focus (Stroke);
- Strokes.

Gesture annotation has been replicated by a second expert annotator. The rate of overlapped units has been calculated through "Unitizing" measure (NEUROGES; Dvoretska et al. 2013; Holle et al. 2013) and shows an Average overlap/extent ratio 0.83 for Units and 0.70 for Phrases. K measure for the categorization of Strokes is over 8.5 (Cantalini, Gagliardi, Proietti, 2018).

Prosodic cues have been annotated by experts following the L-AcT methodology (Cresti 2000; Moneglia & Raso 2014):

- Speech acts are identified through their correspondence with sequences of prosodic units ending with a Terminal prosodic break;
- Prosodic units are those sequence of syllables ending with a prosodic boundary which bear a Perceptively Relevant Prosodic movement ('t Hart et al. 1990);

- Prosodic units match with information units characterized by one Information function;
- Information functions display a type and a role: 1) Textual: Comment, Topic, Appendix, Parenthesis; 2) Dialogical: Conative, Phatic, Incipit, Allocutive, Expressive.

Gesture and prosodic layers have been annotated independently one from the other and then aligned in ELAN files (Wittemburg et al. 2006). Results show strong synchronization constraints:

- Gestures Units always reset at prosodic boundaries;
- Gesture Phrases correlate with Speech acts and do not extend across terminal prosodic breaks;
- Strokes correlate with prosodic units;
- All types of Textual units can host a Stroke, but never Dialogical units.
- Co-speech gestures are characterized as being continuous. All recorded Gesture Units start at a Gesture Unit end with a minimum rest interval.

References

Bressem, J., Ladwig, S. H., & Müller, C. (2013). Linguistic Annotation System for Gestures (LASG). In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.1)* (Vol. 1, p. 1098-1125). Berlin: De Gruyter Mouton.

Cantalini, G., Moneglia, M., Gagliardi, G., & Proietti, M. (2018). La correlazione gesto/prosodia e la sua variabilità: il parlato spontaneo di contro alla performance attorale. Atti del convegno: *GSCP – La comunicazione parlata 2018* (forthcoming).

Chafe, W. (1970). *Meaning and the structure of language.* Chicago: University of Chicago Press.

Cresti, E. (2000). *Corpus di italiano parlato.* Firenze: Accademia della Crusca.

Cresti, E., & Moneglia, M. (eds.). (2005). *C-ORAL-ROM. Integrated reference corpora for spoken romance languages.* Amsterdam: Benjamins.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (ed.), *Non verbal Communication and Language* (p. 207-227). The Hague: Mouton.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth, & M. Fröhlich (eds.), *Gesture and Sign Language in Human-Computer Interaction* (p. 23-35). Berlin: Springer.

Ladewig, S. H., & Bressem, J. (2013). A linguistic perspective on the notation of gesture phases. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.1)* (Vol. 1, p. 1060-1079). Berlin: De Gruyter Mouton.

Loehr, D. (2014). Gesture and prosody. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction (Handbooks of Linguistics and Communication Science 38.2)* (Vol. 2, p. 1381-1391). Berlin: De Gruyter Mouton.

McNeill, D. (2005). *Gesture and Thought.* Chicago: University of Chicago Press.

McNeill, D. (2015). *Why We Gesture.* Cambridge: Cambridge University Press.

Moneglia, M., & Raso, T. (2014). Notes on the Language into Act Theory. In T. Raso, & H. Mello (eds.), *Spoken corpora and linguistics studies* (p. 468–494). Amsterdam: Benjamins.

NEUROGES: http://neuroges.neuroges-bast.info/

't Hart, J., Collier, R., & Cohen, A. (1990). *A Perceptual Study of Intonation. An Experimental-Phonetic Approach to Speech Melody.* Cambridge: Cambridge University Press.

Wittenburg, P., Brugman, H., Russal, A., Klassman, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)*, (p. 1556-1559).

Dvoretska, D., Petermann, K., & Skomroch, H. (2013). Calculating Temporal Interrater Agreement for Binary Movement Categories. In H. Lausberg (Ed.), *Understanding Body Movement. A Guide to Empirical Research on Nonverbal Behaviour. With an Introduction to the NEUROGES Coding System*. Frankfurt/Main: Peter Lang. 253-260

Holle, H. & Rein, R. (2013). The modified Cohen's kappa: Calculating interrater agreement for segmentation and annotation. In H. Lausberg (Ed.), Understanding body movement: A guide to empirical research on non verbal behaviour (With an introduction to the NEUROGES coding system). Frankfurt: Peter Lang Verlag. 261-275.

DB-IPIC: A database for information patterning interlinguistic comparison

Alessandro Panunzi & Valentina Saccone
(University of Florence & University of Basel)

Keywords: information patterning; prosodic patterning; speech analysis; Language into Act Theory; language database;

This proposal aims to present the DB-IPIC spoken language database, an online XML-based language resource for the study of linear relations among information units (IU) in spoken corpora. Beside this, we will present data on some specific IUs in the framework of Language into Act Theory (L-ACT, Cresti 2000; Moneglia & Raso 2014), which is the theoretical basis that underpins the DB-IPIC resource. According to L-ACT, information structure is based on pragmatic principles (Cresti and Moneglia 2018), and it is encoded by prosody. Prosodic units (PU) mostly relate to IU. Moreover, the Comment IU corresponds to a core (root) PU that carries the illocutionary force of an Utterance.

The data collected in DB-IPIC derives from 3 larger spoken corpora: C-ORAL-ROM (Cresti and Moneglia 2005), C-ORAL-BRASIL (Raso and Mello 2012) and Cor-DIAL (Nicolás 2012). The composition of the different resources facilitates comparative studies on how information structure varies across languages (Moneglia and Cresti 2015). After giving a general and comparative picture of the tree corpora, the contribution will focus on the analysis of some specific IU, mostly within the Italian corpus.

First, we will present the informative and prosodic comparison between two specific types of complex illocutionary units, i.e. the Multiple Comments (CMM) and the Bound Comments (COB), corresponding to almost 20% of total Comment units. The two structures follow two opposite prosodic encoding, which reflects different information strategies.

CMMs create a compositional IU formed by two or more PU linked together in a unitary prosodic model, enabling conventional meta-illocutive structures as alternatives (1), comparisons (2) and lists (3).

(1) perché c'è chi vende /^{SCA} dieci /^{CMM} e chi vende cento ?^{CMM}
[*why some sell / ten / and other sell a hundred?*]

(2) noi la nostra /^{CMM} e loro la loro //^{CMM}
[*we have ours / and they have theirs //*]

(3) pattina /^{CMM} quadrante /^{CMM} fianchi /^{CMM} e maniglia // ^{CMM}
[*flap / quadrant / sides / and grip //*]

Conversely, COBs form a chain of weakened illocutive units using a progressive adjunction strategy (4).

(4) la mi' mamma era stata malata /^{COB} era &st [//2]^{EMP} come al solito /^{PAR} era stata all' ospedale /^{COB} e fu proprio il periodo /^{TOP} in cui /^{SCA} mio marito prese /^{SCA} l'azienda /^{SCA} col mi' babbo //^{COM}
[my mom was sick / she was / as usual / she went to the hospital / and it was right around the time / during which / my husband took over / the business / with my dad //]

The second case is the comparison between two types of secondary but widespread units (occurring in about 10% of total Utterances), which share a similar prosodic cues (tonal lowering and decrease of intensity) but again different information properties: Parentheticals (PAR), which correspond to meta-textual insertions on a different discourse level, and Appendixes (APC, APT), which have the value of textual integrations of Comment or Topic units. We will present the differential features of these units from distributional, prosodic and informational perspectives.

References

Cresti, Emanuela (2000), *Corpus di italiano parlato*. Firenze: Accademia della Crusca.

Cresti, Emanuela and Massimo Moneglia (2018), The illocutionary basis of information structure. The Language into Act Theory (L-AcT), in E. Adamou, K. Haude, and M. Vanhove (eds), (2018) *Information Structure in Lesser-described Languages. Studies in prosody and syntax*, Amsterdam / Philadelphia: John Benjamins, 360–402.

Moneglia, Massimo, and Emanuela Cresti (2015), The Cross-linguistic comparison of information patterning in spontaneous speech corpora: Data from C-ORAL- ROM ITALIAN and C-ORAL-BRASIL, in S. Diao-Klaeger, and B. Thörle (eds), (2015), *Linguistique interactionnelle contrastive. Grammaire et interaction dans les langues romanes*, Tübingen: Stauffenburg Verlag, 107–128.

Moneglia, Massimo and Tommaso Raso (2014), Notes on the Language into Act Theory, in Tommaso Raso, and Heliana Mello (eds), (2014), *Spoken corpora and linguistics studies*, Amsterdam / Philadelphia: John Benjamins, 468–495.

Nicolás, Carlota (2012), *Cor-DiAL (Corpus oral didáctico anotado lingüísticamente)*, Madrid: Liceus.

Raso, Tommaso and Heliana Mello (2012), *C-ORAL-BRASIL I: Corpus de referência de portu- guês brasileiro falado informal*, Belo Horizonte: UFMG.

The prosodic forms of Topic information unit: a crosslinguistic, statistically-based investigation in spontaneous speech corpora

Tommaso Raso, Frederico Cavalcante & Plínio Barbosa
(Federal University of Minas Gerais, Federal University of Minas Gerais & Campinas State University)

This is a statistically-based, prosodic study of *Topic* in spontaneous speech, following *Language into Act Theory* (L-AcT; Cresti 2000; Moneglia; Raso 2014), in AE, EP, BP, IT. L-AcT identifies *utterances* as the smallest stretch of speech exhibiting pragmatic autonomy, with boundaries marked by terminal prosodic breaks. Utterances necessarily feature one prosodic unit conveying illocutionary force (*Comment*). Topic is an optional unit found in 9-15% of utterances and precedes Comment, with which it maintains pragmatic aboutness. Prosodically, it has a nucleus, with three functionally equivalent forms (Raso et al. 2017), featuring longer duration, higher intensity, and specific f0 patterns. *Type-1*: rising-falling on the unit's last stressed syllable; *type-2*: rising on the last stressed syllable; *type-3*: two frequently discontinuous semi-nuclei, the first exhibiting (extra-)high values, the second with (much) lower values (classification based on spontaneous speech corpora of IT, EP, BP, AE; Cresti; Moneglia 2005; Raso; Mello 2012; Cavalcante et al. 2018).

We applied Functional Data Analysis (FDA), a technique for qualitative and quantitative investigation of dynamic phenomena (Ramsay; Silverman 2005), to a set of Topic contours to verify our classification. One advantage of FDA is that curves are analyzed as a whole, requiring no *a priori* assumptions. FDA provides time-aligned functional representations of individual curves (Fig. 1) and numerical descriptors. We used functional PCA on the FDA output to see how they cluster. We assessed the clustering using Analysis of Variance of the PC coefficients associated with each contour. Class information was only introduced at the end (Gubian et al. 2009).

The PCA analysis approximates each curve *f(t)* according to the equation $f(t) \approx \mu(t) + s1 * PC1(t) + s2 * PC2(t) + s3 * PC3(t)$, where $\mu(t)$ is the mean curve for all the curves, and *PC1(t)*, *PC2(t)*, *PC3(3)* (Fig. 2) correspond to the three first PC curves expressing the main modes of variation. The coefficients (*s1*, *s2*, *s3*) describe the PC scores for individual curves. They are weights that should be multiplied to each PC curve to approximate an individual curve. We used the median PC score of curves associated with a type to see whether the approximation resembles each type.

The emerging patterns (Fig. 3) are consistent with the described forms. By carrying ANOVA on PC1-3 scores computed for individual curves we can corroborate our classification. PC1 scores alone suffice to separate the curves of each type (Table 1).

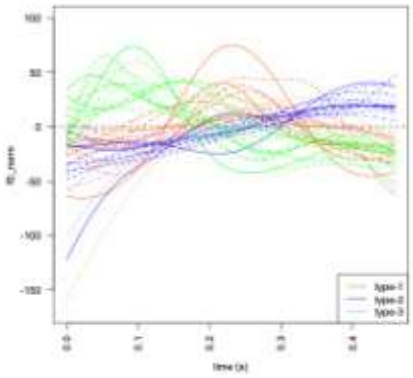


Figure 1 –Linear registration of curves.

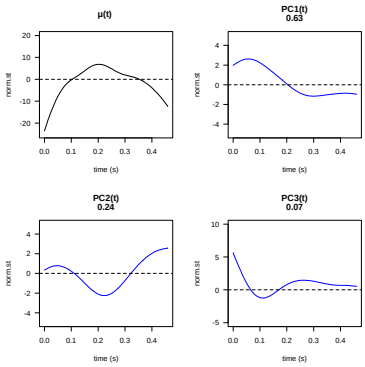


Figure 2 –Mean and PC1-3 curves expressing the main modes of variation. The first three PC curves account for 94% of variation.

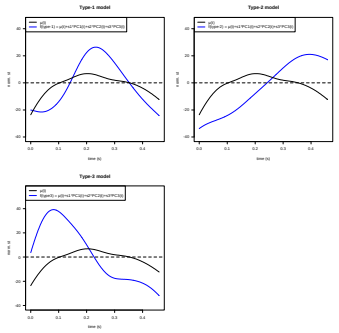


Figure 3 – Mean curve (black) and model curves (blue) obtained by plugging the median of PC scores of curves of the same type in the equation.

Table 1 – Pairwise comparisons of PC1-3 scores using Wilcox on rank sum test (P-value adjustment method: Bonferroni)

	PC1		PC2		PC3	
type	**p-value**	**diff.**	**p-value**	**diff.**	**p-value**	**diff.**
1-2	0.045	yes	3.2e-07	yes	0.6	no
1-3	7.7e-08	yes	0.0983	no	0.17	no
2-3	3.0e-07	yes	0.0013	yes	0.56	no

References

Cavalcante, Frederico; Raso, Tommaso; Ramos, Adriana (2018), *Informationally tagged American English minicorpus*, Creative Commons 4.0. http://www.c-oral-brasil.org/

Cresti, Emanuela (2000), *Corpus di italiano parlato*. Firenze: Accademia della Crusca.

Cresti, E.; Moneglia, M. (Eds.). C-ORAL-ROM. *Integrated reference corpora for spoken Romance languages*. Amsterdam: John Benjamins, 2005.

Gubian, Michele; Torreira, Francisco J.; Strik, Helmer; Boves, Lou W. J, (2009), Functional Data Analysis as a Tool for Analyzing Speech Dynamics: A Case Study on the French Word c'était. Proceedings of Interspeech, 6 September 2009, 2199-2202.

Ramsay, James; Silverman, Bernard W. *Functional Data Analysis*. New York: Springer-Verlag, 2005.

Raso, Tommaso; Cavalcante, Frederico; Mittmann, Maryualê, (2017) Prosodic forms of the Topic information unit in a cross-linguistic perspective: a first survey. Proceedings of the *SLI-GSCP International Conference*, 13-15 June, 2016. Anna de Meo & Francesca Maria Dovetto. Rome: Aracne editrice, 2017, 473-98.

Raso, Tommaso; Mello, Heliana (eds.). *C-ORAL-BRASIL I: corpus de referência de português brasileiro falado informal*. Belo Horizonte: UFMG, 2012.

WORKSHOP 13

Schedule: Fri 9.00-12.55 (Room 5)

Paradigms regained: Theoretical and empirical arguments for the reassessment of the notion of paradigm

Gabriele Diewald, Katja Politt
(Leibniz Universität Hannover, Leibniz Universität Hannover)

Keywords: paradigms, cognitive linguistics, grammaticalization, construction grammar, morphology

Workshop description

The workshop aims at assessing the cognitive reality of (grammatical) paradigms throughout various linguistic domains, thereby testing this notion for its ability to allow for "graceful integration" (Jackendoff 2011), meaning that it should be able to account for empirical findings and general cognitive mechanisms. Paradigms are seen as part of speaker knowledge and important generalizations that enable speakers to deal with previously unencountered elements. Moreover, grammatical paradigms are an important tool in describing the target structure of grammaticalization. It is assumed that grammatical categories are structured paradigmatically in this sense. The theoretical and conceptual foundations of the workshop are grammaticalization theory, implicational morphology, usage-based constructional approaches, cognitive semantics, as well as corpus-based and experimental approaches to grammatical structures in diachronic and synchronic phenomena.

Background and aims

The notion of paradigms is primarily discussed in morphological theories, where it plays a central role as a tool for describing the structures in which inflectional forms are organized. Each member of a paradigm corresponds to a cell, which can be either filled by a form or by a form-feature pair (Lieb 2005, Werner 1994, Wurzel 1984). Lately, work on relational structures in morphological paradigms (Ackerman et al. 2009, Blevins 2015, Blevins 2016) has shown that this purely instrumental conception of paradigms as nothing but a useful descriptive convention clearly underestimates its cognitive foundation and functional importance. Current morphological theory defines paradigms as "cohesive wholes" (Blevins 2015: 94) that consist of relations and associations between its cells. As such, they are part of speaker knowledge, because they are necessary generalizations that allow speakers to infer previously unencountered forms of lexical items (Ackerman et al. 2009: 54). Knowing the overall organizational structure of the forms allows for inferring forms and their functions from one another.

This inferential nature of paradigms is what can be generalized as a structuring principle to other areas of grammar. (Nørgård-Sørensen et al. 2011: XI). In addition to their prominence in morphological approaches, paradigms are an important notion in grammaticalization as well (Diewald 2017, [in preparation]), where grammatical paradigms are the target structure of grammaticalization processes. They are not just accumulations of inflectional forms but holistic semiotic structures, consisting of ordered bundles of oppositions between all members of the category in question (which in grammaticalization are modified in various ways). Take the grammatical category tense as an example: The members of the tense paradigm share a common categorical function, namely situating events relatively to the speech time. The unmarked zero in tense is the present. The specific function of all other members of the category – like past and future forms – can be described relatively to that

unmarked zero, i.e. in opposition to it and of course also in opposition to one another. In short: The oppositions between the members of a grammatical paradigm are the very essence of grammatical structures that "cannot be described without reference to the paradigmatic organisation that lies behind the syntagmatic realisations" (Nørgård-Sørensen et al. 2011: 71).

During grammaticalization processes, elements acquire a place in this structure or change their place within it. Elements form oppositional pairs with other grammatical elements that are members of the same superordinate category, e.g. "tense" (Diewald and Smirnova 2010: 4). By developing this opposition, the newly grammaticalized item becomes a member of a grammatical paradigm (Bybee et al. 1994, Lehmann 2002[1982]).

While it is undisputed that "grammar" is the target domain of grammaticalization processes, and that "paradigms" play a role in the development of Indo-European languages, the exact extend of the notion of paradigm and grammatical paradigm and its usefulness for languages with little or no inflectional morphology has been under dispute for some time by now. For once, there is fundamental criticism concerning the lack of an exact definition of "grammar", as put forth by Himmelmann: "[w]ork in grammaticalization also hardly ever makes explicit the concept of grammar underlying a given investigation" (Himmelmann 1992: 2). Furthermore, there is a lively discussion about (i.) whether the notion of paradigm should be extended to syntagmatic linguistic structures beyond bound morphology and periphrastic forms, as for example in Construction Morphology (Booij 2010, 2016), and include, for example, grammatical oppositions on the level of the whole clause, like the opposition between sentences particles and modal particles. Another hotly disputed issue is (ii.) what the benefit of such an extension might be (cf. (Haspelmath 2000: 663, Wiemer & Bisang 2004, Bisang 2014, Diewald 2017, [in preparation]).

This discussion, which arose in typological research and grammaticalization studies, meets with current questions and challenges in construction grammar. In constructional approaches, paradigms are often "marginalized or even lost" (Diewald, [in preparation]). However, construction grammar aims at describing grammatical structures in their entirety. It is therefore necessary to find an integrative approach that combines both construction grammar and paradigms as organizational structures of grammar (Diewald 2009, Diewald and Smirnova 2010, Diewald 2015). In such an approach, paradigms are, as in morphology (Blevins 2016), necessary generalizations of associative structures and can be seen as constructions "whose function and meaning is defined by the specific number and constellation of its components", which "mutually define each other's values" through their inherent indexical structure (Diewald [in preparation]).

The basic assumptions derived from this background are

1. paradigms are necessary generalizations of grammatical structures,
2. paradigms are part of the grammatical knowledge of speakers, and
3. paradigms are what makes grammaticalization processes structured processes.

This workshop intends to investigate and test these assumptions, and thus raises the following questions:

- Can research from different linguistic subdisciplines underpin the importance of the notion of paradigms?
- What are the advantages (and limitations) of such an integrative approach of describing grammatical structures as paradigmatic, i.e. as consisting of oppositions and relations?
- Is there independent evidence from neighbouring disciplines supporting the assumption of paradigms as cognitive entities?

We invite contributions from various areas of linguistics focusing on where evidence for – or against – an integrated approach of grammatical paradigms as cognitively relevant factors can be found:

cognitive linguistics, psycholinguistics, morphology, general linguistics, semantics, syntax, historical linguistics, and others. Possible research topics and research questions to be covered include

- paradigmatic structures in implicational and constructional morphology,
- paradigms between bound inflectional and non-bound constructional structures,
- the diachronic development of paradigmatic oppositions in grammatical categories,
- investigating oppositions between closely related constructions in a grammatical category,
- typological views on paradigms of various shapes,
- empirical and/or experimental evidence for paradigms as cognitive entities.

References

Ackerman, Farrell, James P. Blevins & Robert Malouf (2009), "Parts and wholes: Implicative patterns in inflectional paradigms". In: Blevins, James P. and Juliette Blevins (eds.): *Analogy in grammar. Form and acquisition*. Oxford: Oxford University Press: 54–82

Bisang, Walter (2014), "Overt and hidden complexity – two types of complexity and their implications." In: Poznań Studies in Contemporary Linguistics 50, 127–143.

Blevins, James P. (2015), "Inflectional paradigms". In: Baerman, Matthew (ed.): *Oxford handbook of inflection*. Oxford: Oxford University Press: 87–111.

Blevins, James P. (2016), *Word and Paradigm Morphology*. Oxford: Oxford University Press.

Booij, Geert (2010), *Construction Morphology*. Oxford: OUP.

Booij, Geert (2016), "Construction Morphology." In: Andrew Hippisley & Gregory Stump (eds.), *The Cambridge Handbook of Morphology*. Cambridge: CUP, 424-448.

Bybee, Joan, Revere Dale Perkins & William Pagliuca (1994), *The evolution of grammar. Tense, aspect, and modality in the languages of the world*. Chicago: Chicago University Press. (=196).

Diewald, Gabriele (2009), "Konstruktionen und Paradigmen". *Zeitschrift für germanistische Linguistik* 3/37: 445–468.

Diewald, Gabriele & Elena Smirnova (2010), "Introduction". In: Diewald, Gabriele and Elena Smirnova (eds.): *Paradigmaticity and Obligatoriness*. Special Isssue: 1–10.

Diewald, Gabriele (2017), „Auf- und Abbau linguistischer Komplexität in Grammatikalisierungsprozessen." In: Mathilde Hennig (ed.): Linguistische Komplexität – ein Phantom? Tübingen: Stauffenburg, 223-252.

Diewald, Gabriele [in preparation], „Paradigms lost – paradigms regained. Paradigms as hyper-constructions." In: Lotte Sommerer & Elena Smirnova (eds.), Nodes and Links in the Network: Open questions in Diachronic Construction Grammar. Amsterdam: Benjamins.

Himmelmann, Nikolaus P. (1992), "Grammaticalization and Grammar". *Arbeitspapiere des Instituts für Sprachwissenschaft der Universität zu Köln* 16.

Jackendoff, Ray (2011), "What is the human language faculty? Two views". *Language* 87(3), 586-624.

Lehmann, Christian (2002), *Thoughts on grammaticalization*. Erfurt: Seminar für Sprachwissenschaft der Universität Erfurt. (=*Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt* 9).

Lieb, Hans-Heinrich (2005), "Notions of paradigm in grammar". In: Cruse, Alan D., Michael Job Hundsnurscher & Peter Lutzeier (eds.): *Lexikologie / Lexicology. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen / An international handbook on the nature and structure of words and vocabularies*. Berlin: de Gruyter: 1613–1646.

Nørgård-Sørensen, Jens, Heltoft, Lars, & Schøsler, Lene (2017), Connecting grammaticalisation. Amsterdam: Benjamins.

Werner, Otmar (1994), "Auch Paradigmen entstehen und vergehen". In: Köpcke, Klaus-Michael (ed.): *Funktionale Untersuchungen zur deutschen Nominal- und Verbalmorphologie*. Tübingen: Niemeyer: 5–28.

Wiemer, Björn & Walter Bisang (2004), "What makes grammaticalization? An appraisal of its components and its fringes?". In: Walter Bisang, Nikolaus Himmelmann & Björn Wiemer (eds.), What makes Grammaticalization? A Look from its Fringes and its Components. Berlin, New York: Mouton de Gruyter. 3-20.

Wurzel, Wolfgang U. (1984), *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie-Verlag.

Inflectional paradigms and second-order schemas

Jenny Audring & Ray Jackendoff
(Leiden University & Tufts University, MIT)

Keywords: Construction Morphology, paradigms, second-order schemas, paradigmatic relations, sister relations

The heartland of paradigms is inflectional morphology. Yet, paradigmatic relations in a wider sense can also be found in derivational morphology, in syntax, and even in phonology. From a usage-based, construction-theoretic perspective (e.g. Goldberg 1995, Booij 2010, Hilpert 2014, Jackendoff & Audring 2019), paradigmatic relations are best understood as *sister relations* between constructions. (1) shows a pair of morphological sisters and (2) a pair of phrasal sisters. In each case the sisters form oppositional pairs: they share parts of their structure, while differing in other parts.

(1) *reiziger* 'traveler' ≈ *reizigster* 'female traveler' (**reizig*)
(Dutch, van Marle 1985)
(2) *tennis da tavolo* 'table tennis' ≈ *tennista da tavolo* 'table tennis player'
(Italian, Booij & Masini 2015: 56)

Paradigmatic relations that transcend individual words and form grammatical patterns can be modeled by means of *second-order schemas*. For example, the relations exemplified in (1) and (2) can be systematized as in (3) and (4). The parts coindexed as x, y and z are shared between the sisters.

(3) $[X_x + er]_N \approx [X_x + ster]_N$
(4) $[N_x\ P_y\ N_z]_{NP} \approx [[N_x + ista]_N\ P_y\ N_z]_{NP}$
(after Booij & Masini 2015: 57)

This type of representation reflects the understanding that paradigmatic relations rely on links between identical subparts of constructions (Hilpert 2014: 83). Such *subpart links* (Goldberg 1995: 78) can be expressed by coindexation, as shown in (3) and (4).

Second-order schemas have proved insightful for the representation of various morphological relations, and it has been argued that they are the key to modeling inflectional paradigms in a construction-theoretic framework (Masini & Audring 2019, van der Spuy 2017). However, they have not yet been applied to full inflectional paradigms. The present talk takes up this challenge, showing how the coindexing formalism proposed in Booij (2010) and elaborated in Jackendoff & Audring

(2019, forthcoming) can indeed be useful in modeling inflectional paradigms, including configurations like inflectional classes.

However, inflectional paradigms are restricted in ways that other paradigmatic relations are not, and we argue that second-order schemas are not expressive enough to state these restrictions. For example, the two German verb forms in (5) are paradigm mates not only because they share the same stem but also because the non-shared parts, the suffixes, express the same grammatical features of person, number and tense.

(5) *lieb-e* 'love-1SG.PRS' ≈ *lieb-tet* 'love-2PL.PST'

Thus inflectional paradigms are more tightly constrained in that they additionally demand parallelism between non-shared parts of structure. This parallelism cannot be expressed by coindexation; this requires a morphosyntactic template stating the feature matrix. Inflectional classes can then be understood as alternative families of sisters that correlate with the template in different ways.

By working out the details of the representation we hope to contribute to a better integration of paradigms in a construction-theoretic framework, and to a better understanding of the ways inflectional paradigms differ from other paradigmatic relations.

References

Booij, Geert (2010), *Construction Morphology*. Oxford: Oxford University Press.

Booij, Geert & Francesca Masini (2015), The Role of Second Order Schemas in the Construction of Complex Words. In: Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds.), *Semantics of Complex Words*, vol. 3. Cham: Springer International Publishing, 47-66.

Goldberg, Adele E. (1995), *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hilpert, Martin (2014), *Construction Grammar and its application to English*. Edinburgh: Edinburg University Press.

Jackendoff, Ray. S. & Audring, Jenny (2019), Relational Morphology in the Parallel Architecture. In: Jenny Audring & Francesca Masini (eds.), *The Oxford Handbook of Morphological Theory*, 390-408.

Jackendoff, Ray. S. & Audring, Jenny. forthcoming. *The Texture of the Lexicon*. Oxford: Oxford University Press.

Masini, Francesca & Audring, Jenny (2019), Construction Morphology. In: Jenny Audring & Francesca Masini (eds.), The Oxford Handbook of Morphological Theory, 365-389.

Van Marle, Jaap (1985) *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.

Van der Spuy, Andrew (2017), Construction Morphology and inflection. *Lingua* 199, 60–71. doi:10.1016/j.lingua.2017.07.010.

Applying the notion of a paradigm to word order

Lars Heltoft

The points of the present contribution is that the notion of a paradigm can be extended from morphology to word order (Nørgård-Sørensen et al. 2011). Once the necessary tools are available, we

can describe adequately the synchronic input and output stages of word order changes, exemplified here by the change from Middle Danish through Renaissance Danish to Modern Danish.

Andersen (2010) handles the inferential nature of paradigmatic organization in morphology by applying Peircean semiotics, esp. symbols, indices and icons; he analyses Latin *cucurristi* (Matthews 1972) in terms of symbolic and indexical relations. Elsewhere, he has sketchily argued that word order can index the semantic character of NPs. Danish and English have an indirect object position that indexes the NP2's Receiver role and thus, positions can index the valency of predicate stems.

Constituent order can also express symbolic oppositions with no bearing on valency, for instance the opposition focus vs. background (compare Mithun 1992). Single constituents cannot have both values at the same time, and in this sense, this opposition is a paradigmatic one. Middle Danish word order distinguishes between focused and backgrounded constituents. In relation to the *non-finite verb*, arguments and adverbs could occur in both orders (VX and XV). In the Middle Field between finite and non-finite verb, XV occur as either backgrounded or in focus, see (1). Constituents preceding a focus operator (here the negation *ei*) are backgrounded; those immediately following this operator must be (part of) the focus.

			Middle Field					Postfield
		Finite V	Background		Focus	V		
(1)	*tha*	*mughæ*	*men*	*ei*	*børn*	*fra hennæ*	*takæ*	
	then	may	men	not	children	from her	take	

'Then people may not take the children away from her' (Jutish Law, DgL II 71, 6, app. 1300)

The paradigmatic opposition between focus and background is a symbolic one, projected onto the syntagmatic axis in the form of a diagram (iconic relations: backgrounded constituents precede focused constituents). If present, a focus operator will index the positions. It is essential that paradigmatic alternations can be mapped onto syntagmatic sequences, a point also highlighted for constructional paradigms by Diewald & Smirnova (2010).

Such systems are not mentioned by Sun & Traugott (2012), claiming that word order is not grammaticalised per se. To them, word order is the emergent result of other processes such as bleaching.

Modern Danish allows only subjects and free adverbials in the Middle Field. Renaissance Danish (1500-appr. 1700) retains the system of (1), but restricts the syntactic domain of the paradigm to pronominal arguments and pronominal adverbials (irrespective of stress and cliticisation). This transitional system is exemplified in (2ab):

			Focus		
(2a)	du skalt	icke	saadant	tencke	
	thou shalt	not	such	imagine'	

'You must not imagine such things'(Syv Mestre 131, 25-26)

				Focus	
(2b)	aldrig	haffuer ieg		nogen anden	hafft /
	never	have I		anybody else	had
	aldrig	uil ieg	oc	nogen anden	haffue
	never	will I	either	anybody else	have

'Never have I had anybody else, nor will I ever have anybody else' (Griseldis 116, 18-23)

References

Andersen, Henning. 2010. From Morphologization to Demorphologization. In Silvia Luraghi & Vi't Bubeni'k (eds.), 117-146.

Diewald, Gabriele & Elena Smirnova. 2010. Introduction. In Gabriele Diewald & Elena Smirnova (eds.): *Acta Linguistica Hafniensia* 42(1), Special Issue: Paradigmaticity and Obligatoriness.1–10.

Luraghi, Silvia & Vi't Bubeni'k (eds.). 2010. *The Bloomsbury Companion to Historical Linguistics*. The Bloomsbury Publishing Company PLC.

Matthews, P.H. 1972. *Inflectional morphology: a theoretical study of based on aspects of Latin verb conjugations*. Cambridge: Cambridge University Press.

Mithun, Marianne. 1992. Is Basic Word Order Universal? In Doris L. Payne (ed.), *Pragmatics of Word Order Flexibility*, 15-61. Amsterdam/Philadelphia: John Benjamins.

Narrog, Heiko & Bernd Heine (eds.). 2011. *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.

Nørgård-Sørensen, Jens, Lars Heltoft & Lene Schøsler. 2011. *Connecting Grammaticalisation*. John Benjamins.

Sun, Chaofen & Elizabeth Closs Traugott. 2011. Grammaticalization and Word Order Change. In Narrog & Heine (eds.), 378-388.

Sources:

Jutish Law DgL II = *Danmarks gamle Landskabslove* II, ed. by Peter Skautrup. Copenhagen: Gyldendalske Boghandel – Nordisk Forlag. 1933.

Griseldis = *Danske Folkebøger* VIII, ed. by R. Paulli, 79-118. Copenhagen: Gyldendalske Boghandel – Nordisk Forlag. 1920.

Syv Mestre = *Danske Folkebøger* II, ed. by J. P. Jacobsen, 1-200. Copenhagen: Gyldendalske Boghandel – Nordisk Forlag. 1916.

ꕥ

The importance of paradigmatic analyses. From one lexical input into multiple grammatical paradigms

Kirsten Jeppesen Kragh & Lene Schøsler
(University of Copenhagen)

Keywords:
Diachrony, French verb of perception *voir* 'to see', grammaticalization, paradigm, synchrony.

We take as our starting point that when lexical entities grammaticalize, they enter preexisting paradigms. Therefore, grammatical paradigms are important for the understanding of the reanalyses leading to grammaticalization.

In the line of Henning Andersen's thinking we propose to conceive of grammar as composed of sets of paradigms (Nørgård-Sørensen et al. 2011). We here use the term *paradigm* not in the narrow sense of inflectional paradigm, nor entirely in the line of the 'classical' grammaticalization approach of Lehmann (1985), but in the more general sense of a selectional set, composed of marked or

unmarked members (Andersen 2008: 19). In previous studies we have shown that the notion of a paradigm is useful for the understanding of grammatical structure.

The lexical input that we will use to illustrate our point is the French verb *voir* 'to see', in order to show the pathway of multifunctional lexical item into grammar, i.e. into a number of individual paradigms. When aiming at analyzing a polysemous and multifunctional lexical entity like this verb, the researcher can chose between a polysemic or a monosemic approach, see Waltereit (2002, 2006). However, when studying how a lexical item grammaticalizes, we do not consider this discussion to be the essential one. Rather, we propose that the interesting point is how the lexical item enters different grammatical paradigms.

Our approach combines synchronic and diachronic investigations on electronic corpora. Each paradigm presents the actual synchronic status of diachronic grammaticalization processes. It is based on synchronic paradigmatic analyses of functions that seem to be very ambiguous and diverse. The paradigms comprise the following grammatical categories: tense, aspect and mood, i.e. progression (*je le vois qui arrive* 'I see him arriving'), voice (*il se voit refuser l'accès* 'he is denied access'), and pragmatics: presentatives (*voilà le bateau* 'here is the boat'), focus constructions (*voici le bateau qui arrive* 'here is the boat arriving'), and discourse markers (*le bateau arrive, tu vois* 'the boat is arriving, in fact'). By distinguishing the different contexts (labelled *domains*) in which the forms appear, and state which semantic fields they cover (labelled *frames*), we can generate synchronic paradigms of which the grammatical entities are members. We are convinced that synchronic paradigms provide a precise and relatively simple presentation of what otherwise would seem utterly diverse usages of a lexical entity.

References

Andersen, Henning. 2008. Grammaticalization in a speaker-oriented theory of change. In Thorhallur Eythórsson (ed.) *Grammatical change and linguistic theory: The Rosendal Papers*, 11-44. Amsterdam: John Benjamins Publishing Company.

Lehmann, Christian. 1985. Grammaticalization: synchronic varation and diachronic change. *Lingua e Stile* 3. 303-318.

Nørgård-Sørensen, Jens, Lars Heltoft & Lene Schøsler. 2011. *Connecting grammaticalization. The role of paradigmatic structure*. Amsterdam: John Benjamins.

Waltereit, Richard. 2002. Imperatives, interruption in conversation, and the rise of discourse markers: a study of Italian guarda. *Linguistics* 40. 987-1010.

Waltereit, Richard. 2006. Comparer la polysémie des marqueurs discursifs. Les marqueurs discursifs dans les langues romanes: staff.ncl.ac.uk.

Formalizing paradigms in Construction Grammar

Jaakko Leino
(University of Helsinki)

Keywords: construction grammar, paradigms, syntax, morphology, analogy

Construction Grammar sees the language system to consist solely of conventional pairings of form and meaning, i.e. constructions. Constructions may be of any size and complexity, and they may be

abstract (or schematic) to any degree. They may be templates for sentences, lexical items, inflectional morphemes, discourse patterns (Östman 2005) that organize whole texts or even genres, etc.

However, the notion of construction seems incapable to capture patterns found within the grammar: systematic similarities between constructions and, notably, paradigms of different sorts. For instance, inflection paradigms consist of sets of constructions, but nothing in common varieties of Construction Grammar explains how those constructions join together to form a paradigm.

The paper argues that in addition to constructions, the language system must also include specifiable relations which hold between the constructions of a language and which organize them into a functional system. Crucially such relations are necessary for the organization of paradigms, be they of morphological, syntactic or other nature.

Relations between constructions within the grammar can be—and have previously been—described in terms of inheritance (e.g. Goldberg 1995), taxonomic and meronomic links (Croft 2001), and the like. However, such very abstract links can only capture simple relations between constructions. Yet, more complex relations, notably of an analogical nature, exist widely within the grammar of apparently all human languages.

To capture such analogical relations, the paper uses the notion of *metaconstruction*, previously introduced briefly in Leino & Östman (2005). Metaconstructions may be thought of as generalizations of constructions, partly in the same sense as constructions may be seen as generalizations of actual expressions. It will be argued that such analogical relations, formalizable as metaconstructions, hold paradigms together and also facilitate both producing and interpreting complex expressions.

It will also be shown that while analogical relations in morphology most often is static and merely "holds paradigms together", as it were, analogical relations between (morpho)syntactic constructions are a notable source of linguistic creativity and innovations. Systematic analogical structures often both show existing gaps in the language system and provide means of coining novel but instantly comprehensible ways of filling such gaps. The mechanism is ubiquitous in language, but it seems to lead to (conceived) paradigms only in certain parts of grammar. This, in turn, may be revelatory of the nature of paradigms.

References

Croft, William 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

Goldberg, Adele E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure.* Chicago: The University of Chicago Press.

Leino, Jaakko & Jan-Ola Östman 2005. Constructions and variability. In Fried, Mirjam & Hans Boas (eds.), *Grammatical Constructions: Back to the Roots.* Constructional Approaches to Language 4. Amsterdam: John Benjamins.

Östman, Jan-Ola 2005. Construction Discourse: A Prolegomenon. In Östman, Jan-Ola & Mirjam Fried (eds.), *Construction Grammars. Cognitive grounding and theoretical extensions.* Constructional Approaches to Language 3. Amsterdam: John Benjamins. 121–144.

ℒ

Transcategorial paradigms: specification of functional potential and grammaticalisation of oppositions across categories

Peter Juul Nielsen

<pdf>

ℒ

Usage-Based Paradigms: A bottom-up experiment in distinctive categorisation of Polish verbal prefixes

Dylan Glynn & Piotr Wyrośląk
(Université Paris VIII; Adam Mickiewicz University in Poznań & Université Paris VIII)

Keywords: Polish, paradigm, verbal prefixes, Profile-Based Approach, quantitative usage-based analysis

The notion of the grammatical paradigm is central to much of descriptive grammar. In usage-based linguistics, the paradigm can be understood as a theoretical model of paradigmatic semantic relations rather than a description *per se*. This study tests to see if a grammatical paradigm can be empirically derived from quantitative usage-based analysis (cf. e.g. Janda to appear). In other words, can the multidimensional distribution of contextual characteristics be interpreted as a set of distinctive features that map the paradigmatic relations of grammatical categories? If this proves successful, usage-based methods could be used to empirically confirm introspection-based studies of paradigms and potentially add nuance to these descriptions. In order to test this proposal, the study examines a subset of perfectivising verbal prefixes in contemporary Polish.

The study employs the Profile-Based Approach (Geeraerts *et al*. 1994, Gries 2003), which employs a combination of feature analysis and multivariate quantitative analysis to map complex usage patterns. This method begins with the detailed manual annotation of large numbers of contextualised examples for a set of characteristic features. This results in a complex set of metadata: the usage-profile of the form in question. Statistical analysis is then needed to identify patterns (usage-based structure) in the metadata. In order to assure that the analysis is based on contextual clues, the form under investigation is hidden from the analyst and in order to assure maximal repeatability, subjective features require a secondary analysis. Particular care was taken to ensure that the object of the study challenges the adopted method yet, if successful, will yield a coherent onomasiological map, interpretable in terms of a grammatical paradigm.

In this study, three verbal perfectivising prefixes (*wy-*, *za-*, *na-*) are examined as combined with four verb bases from the same lexical field (*-pchać*/*-pychać* ‘push’, *-pełnić*/*-pełniać* ‘fill’). All the included items are near-synonymous verbs of filling (cf. analyses of the analogous English classes in Levin 1993:111-121). The relations between the prefixes are paradigmatic and are traditionally mapped out as a set of distinctive semantic features. The corpus employed is the National Corpus of Polish (NKJP; Przepiórkowski et al. 2012), with an extracted dataset of over 900 occurrences (comprising comparable samples of each prefix-lexeme pairing). These observations will be subjected to feature annotation, focusing on the semantic features of the nominal elements, agency, force dynamics, as well as event characteristics and the morphosyntactic profile. The features annotated will be based on the results of previous research (Janda 1986, Dąbrowska 1996, Fabiszak et al. 2014 Grzegorczykowa et al. 1999). One specific hypothesis is that plexity will be a crucial factor in determining the choice between prefixes (Talmy 2000). We expect that the results will quantitatively map the usage using discriminative quantitative analysis to cluster the contextual features. The descriptive accuracy will then be determined through predictive modelling. It is hoped that these features will produce a non-discrete usage-based map that distinguishes the prefixes paradigmatically.

References

Dabrowska, E. (1996). The Spatial Structuring of Events: A Study of Polish Perfectivizing Prefixes. In M. Putz & R. Dirven (Eds.), *The Construal of Space in Language and Thought* (pp. 467-490). Berlin & New York: Mouton de Gruyter.

Fabiszak, M., Hebda, A., Kokorniak, I., & Krawczak, K. (2014). The semasiological structure of Polish myśleć 'to think'. D. Glynn & J. Rominson (Eds.), *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.

Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). T*he structure of lexical variation. A descriptive framework for cognitive lexicology*. Berlin: Mouton de Gruyter.

Gries, S. T. (2003). *Multifactorial Analysis in Corpus Linguistics: A study of particle placement*. London & New York: Continuum Press.

Grzegorczykowa, R., Laskowski, R., & Wróbel, H. (1999). *Gramatyka współczesnego języka polskiego: Morfologia* [Grammar of contemporary Polish: Morphology] (3rd. ed). Warsaw: Wydawnictwo Naukowe PWN.

Janda, L. A. (1986). *A Semantic Analysis of the Russian Verbal Prefixes za-, pere-, do-, and ot*. Munich: Sagner.

Janda, L.A. & Tyers, M. F. (to appear). Less is more: why all paradigms are defective, and why that is a good thing. *Corpus Linguistics and Linguistic Theory*.

Levin, B. (1993). *English Verb Classes and Alternations: A preliminary investigation*. Chicago & London: University of Chicago Press.

Przepiórkowski, A. (2012). *Narodowy Korpus Języka Polskiego* [The National Corpus of Polish].Warsaw: Wydawnictwo Naukowe PWN.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge,MA & London: MIT Press.

ø

Recursion and paradigms

Tabea Reiner
(LMU Munich)

Keywords: recursion, paradigms, inflectional vs. derivational morphology, periphrasis, reduplication

Judging from certain passages in recent literature, one might get the impression that paradigms, together with morphology more generally, are ready for the ash heap of (linguistic) history. For example, it has been argued that at least the morphology of strictly agglutinating languages like Turkish cannot be adequately described by (necessarily gigantic) paradigms but should rather be captured by principles of compositional syntax (Haspelmath 2011:59). And if agglutinating morphology is compositional syntax, then fusional morphology in languages like Latin must be conceived of as a set of constructional idioms (ibid.). So there is no room left for paradigms or morphology in any traditional sense.

However, consider the following asymmetry. Compounding, derivation, and periphrasis all allow for some degree of recursion, as witnessed by 0, 0 and 0 from German, whereas inflection (to my knowledge) does not allow any recursion, as exemplified by 0.

(1) Donau-dampfschifffahrts-gesellschafts-kapitäns-anwärter-posten
Danube-steamboating-society-shipmaster-candidate-position
'position as a candidate for Shipmaster of the Danube Steamboating Society'

(2) Ur-ur-ur-enkel
great-great-great-grandchild
'great-great-great-grandchild'

(3) Ich hab' das ge-mach-t ge-hab-t.
I AUX that PST.PTCP-do-PST.PTCP PST.PTCP-AUX-PST.PTCP
'I had done this.' (one of several readings)

(4) a. German: *mach-te-te
do-PST.1/3SG-PST.1/3SG
intended: 'had done'

b. Turkish: *yap-tı-tı-m
do-PST-PST-1SG
intended: 'had done' (p.c.)

Why? A possible explanation is that in the case of inflection, but not in the case of syntax, derivation, or compounding there is something at work that strictly delimits recursion. Something that imposes a firm grid. Conceivably: a paradigm. For example, assuming a paradigm like in table 1 below for the featural content of synthetic verbal forms in German, it does not come as a surprise that 0a. is ungrammatical: its intended content simply falls outside of the grid's boundaries.

	Narrative past	Non-past
1st person singular	…	…
2nd person singular	…	…
3rd person singular	…	…

	Narrative past	Non-past
1st person plural	…	…
2nd person plural	…	…
3rd person plural	…	…

table 1: paradigm

Thus, the notion of paradigm might be instrumental in explaining the limits of recursion. A consequence of this approach is that, conversely, periphrasis of the type in (3) is *not* to be captured by paradigms and in this sense is not to be captured morphologically, but rather syntactically. This might seem counterintuitive to some; however, it is exactly what is done in the vast formal literature on verbal clusters (e.g., Haider 2010, ch. 7; Müller 2013, ch. 15), and also from a constructionist point of view it seems perfectly fine to treat at least certain periphrases as constructions near the syntactic endpoint of a lexicon-syntax continuum. Ironically, however, a syntactic take on periphrasis is in contradiction to realizational morphology, including Paradigm Function Morphology (PFM, Stump 2001) – i.e. one of the few approaches that put paradigms on centre stage. So I will discuss how PFM can be motivated without referring to periphrasis. Beyond that, it has to be considered how reduplicative patterns (Hurch 2005ff, Haugen 2015) fit into the general picture.

References

Haider, Hubert (2010), *The syntax of German*, Cambridge: CUP.

Haspelmath, Martin (2011), The indeterminacy of word segmentation and the nature of morphology and syntax, *Folia Linguistica* 45(1), 31–80.

Haugen, Jason D. (2015), Reduplication of affixes. Handout for Workshop on Replicative Processes in Grammar. Leipzig University, 1st October 2015, available online: http://home.uni-leipzig.de/muellerg/replication/haugen.pdf.

Hurch, Bernhard (2005ff), Graz database on reduplication, available online: http://reduplication.uni-graz.at/redup/.

Müller, Stefan (2013), *Head-Driven Phrase Structure Grammar*: *Eine Einführung*. 3rd edn., Tübingen: Stauffenburg.

Stump, Gregory T. (2001), *Inflectional morphology*, Cambridge: CUP.

How systematic is the system of the Dutch modals?

Wim Caers & Jan Nuyts
(University of Antwerp)

Keywords: modals, diachrony, meaning change, analogy, grammatical paradigm

Dutch is traditionally considered to have six modal verbs: *kunnen* 'can', *mogen* 'may', *moeten* 'must', *hoeven* 'need', *zullen* 'shall/will' and *willen* 'want'. In a series of investigations (Byloo & Nuyts 2014, Nuyts 2013, Nuyts & Byloo 2015, Nuyts et al. 2018) we have addressed the question how these forms have evolved diachronically, and how their evolutions converge or diverge.

All these studies were performed in the same way. We have used corpus data from four different time periods: Old Dutch, Early Middle Dutch, Early New Dutch and Present Day Dutch. For each period a random sample of 200 instances of each verb was used (for PDD there were two separate sets of 200 instances for written and spoken language), selected from the available materials according to criteria such as representativity (e.g. in terms of text genres and regional spreading) and comparability across the periods.

The findings show that only the first four of these modal verbs really behave as a system. They do so semantically: they have all developed complementary meanings within the same range of semantic categories (including dynamic, deontic and epistemic modality, and directivity). And they do so grammatically: they have all developed an auxiliary status in the past, but in current Dutch they are all showing a tendency to regain grammatical autonomy. The four forms differ considerably in terms of when, how, and how fast they undergo these processes, but both the commonalities and differences between them can at least in part be explained with reference to processes 'aiming' to make them form a system, such as semantic and grammatical analogy, or avoidance of synonymy/semantic overlap. *Zullen* and *willen* are not participating in these developments, arguably because they do not fit well into the system (especially for semantic reasons: *zullen* has predominantly assumed the role of a tense marker, and *willen* has always been focused on one specific meaning – volition – which in itself is 'at the edge' of the modal domain).

These empirical facts and their interpretation demonstrate that the notion of a grammatical system is very real, linguistically, hence cognitively.

References

Byloo, P. & J. Nuyts (2014). Meaning change in the Dutch core modals. *Acta Linguistica Hafniensia* 46: 85-116.

Nuyts, J. (2013). De-auxiliarization without de-modalization in the Dutch core modals. *Language Sciences* 36: 124-133.

Nuyts, J. & P. Byloo (2015). Competing modals. *Diachronica* 32: 34–68.

Nuyts, J., H.-J. Goelen & W. Caers (2018). *Hoeven* diachroon. *Taal & Tongval*. In print.

Redundant indexicality and paradigmatic reorganisations in the Middle Danish case system

Bjarne Simmelkjær Sandgaard Hansen

<pdf>

Expressions of genericity in Norwegian – a case study

Anna Kurek
(Adam Mickiewicz University in Poznań)

Keywords: generics, Norwegian, paradigms, well-established kinds, corpus

This research project is funded by the National Science Center, grant number 2017/25/N/HS2/00003.

Genericity has two senses which are expressed either as kind-referring sentences or habitual sentences (Krifka et al. 1995: 2-4). The first type of generics renders reference to kinds, whereas habituals make generalisations about individiduals. What is more, none of the studied languages has an articulated generic operator, such as an article or a given noun form (Behrens 2005: 276 and Leslie 2007: 381). This proves that the phenomenon is a part of speakers' knowledge about the world. Numerous studies have shown that even children comprehend the phenomenon and are able to express it correctly before they master the use of quantifiers (Leslie 2007: 380). Genericity as a grammatical phenomenon is therefore paradigmatic.

The phenomenon has been studied with the use of different analysis methods. The most common ones include analysis of individual sentences (Lyons 1977, Oosterhof 2008), surveys (Ionin et al. 2011, Leslie and Gelman 2012) and corpus research (Carlsson 2012). This paper presents a case study combining two methods, namely survey and corpus analysis.

The survey was conducted among 599 native speakers of Norwegian and it consisted of 30 short generic texts. In each of them, the respondents had to fill a gap, similarly as in Ionin et al.'s study (2011). Survey results have shown that when it comes to uncountable nouns, bare nouns (BNs) and definite singular forms were chosen most often (92,77% and 7,07% respectively). Countable nouns have shown much more diversity with most frequent use of bare plural (BPs, 27,85%), BN (25,43%) and indefinite singular form (24,84%). Definite singular and plural forms constituted 15,82% and 6,07% of the answers.

The second part of the research consisted in building and tagging a specialised corpus of 170 texts (27 740 words) with the use of R software. Generics were tagged with respect to position in the sentence, as well as their syntactic function. Most of generic nouns were subjects (66,43%) occurring in initial position. Similarly as in the survey, BNs, BPs and definite singular forms were chosen most often (40,5%, 26,13% and 21,02% respectively), whereas definite plural and indefinite singular forms represented only 8,94% and 3,37%.

The differences were also observed in the case of well-established kinds (WEKs, Krifka et al. 1995), both in the survey and the corpus. WEKs seem to function as proper names (cf. Carlson 1997 and Carlson 1982), namely they are more often used in definite forms than non-WEKs, as illustrated in the examples below:

(1)
Gitaren*kom* *opprinnelig* *fra* *Midtøsten.* [WEK]
Guitar-DEF came originally from Middle East.
'The guitar came originally from Middle East.'

(2)
En ***aksel*** *kan* *også* *være* *en* *stillestående* *maskindel.* [non-WEK]
An axis can also be a motionless machine.part.
'An axis can also be a motionless machine part.'

Genericity in Norwegian is highly context prone and can be expressed with all noun forms. The use of particular noun forms depends not only on the sentence structure and the context, but also on the noun type (WEK, non-WEK).

References

Behrens, Leila (2005), Genericity from a cross-linguistic perspective, *Linguistics*, 43(2), pp. 275-344.

Carlson, Greg N. (1977), A Unified Analysis of the English Bare Plural, *Linguistics and Philosophy*, 1, pp. 413-457.

Carlson, Greg N. (1982), Generic terms and generic sentences, *Journal of Philosophicaf Logic*, 11, pp. 145-181.

Carlsson, Yvonne (2012), *Genericitet i text*, Stockholm: Stockholms universitet.

Ionin, Tanja et al. (2011), An experimental investigation of the expression of genericity in English, Spanish and Brazilian Portuguese, *Lingua*, 121, pp. 963-985.

Krifka, Manfred et al. (1995), Genericity: An Introduction. In: Krifka, M. et al. (eds), *The Generic Book*, Chicago and London: The University of Chicago Press.

Leslie, Sarah-Jane (2007), Generics and the Structure of the Mind, *Philosophical Perspectives – Philosophy of Mind*, 21, pp. 376-403.

Leslie, Sarah-Jane and Susan A. Gelman (2012), Quantified statements are recalled as generics: Evidence from preschool children and adults, *Cognitive Psychology*, 64, pp. 186-214.

Lyons, John (1977), *Semantics*, Cambridge: Cambridge University Press.

Oosterhof, Alfred (2008), *The Semantics of Generics in Dutch and Related Languages*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

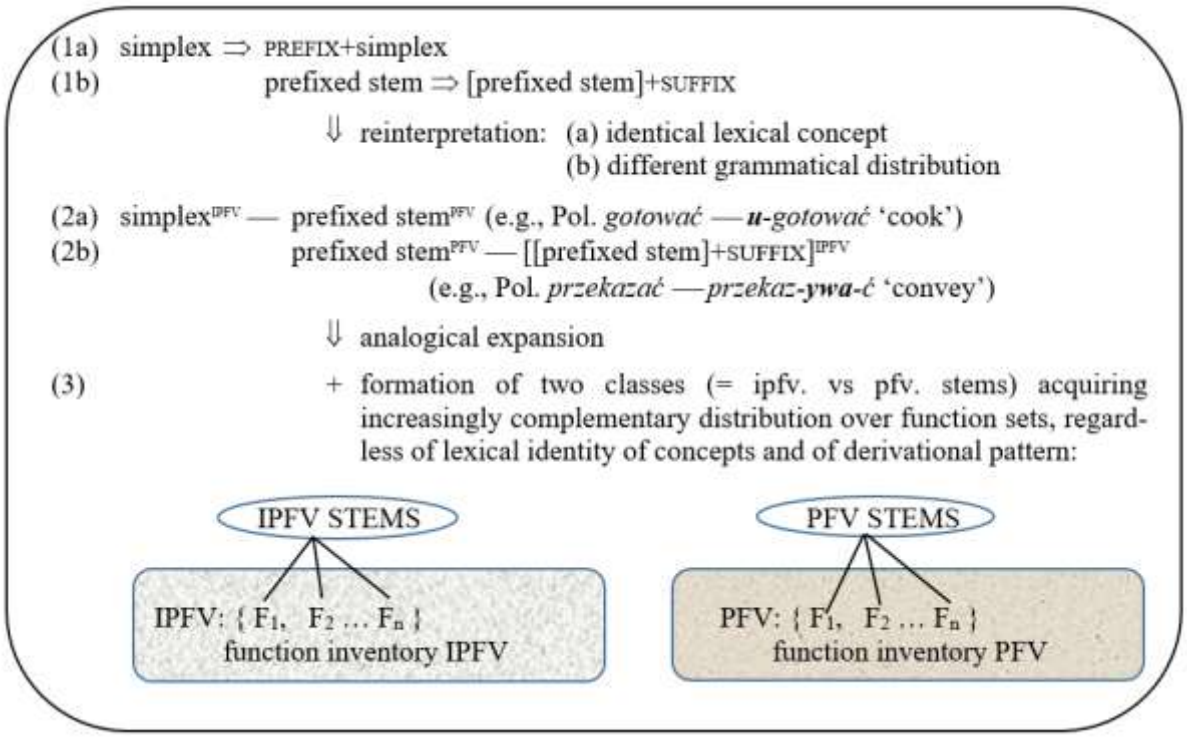
No paradigms without classification: how stem-derivation develops into grammatical aspect

Björn Wiemer, Marek Łaziński, Joanna Kwiatkowska, Ilja Egorov, Katarzyna Osior-Szot & Dorota Urban-Wójnicka
(University of Mainz, University of Warsaw, University of Mainz, University of Mainz, University of Warsaw, University of Warsaw)

Keywords: aspect, stem-derivation, classificatory categories, Slavic, paradigmatic tightening

Discussions concerning the grammatical character of Slavic aspect have usually concentrated on the derivation—inflection divide (cf. still Matasović 2018), whereas discussions about the paradigmatic structure of constructions have largely ignored aspect. Probably, the reason is that, on the one hand, the development of the Slavic perfective:imperfective (pfv.:ipfv.) opposition cannot be adequately captured by morphologization and semantic bleaching assumed for canonical grammaticalization (e.g., Nicolle 2012); on the other hand, research into paradigmatic tightening, and concomitant obligatorification, has hitherto left aside the rise and structure of classificatory categories, of which Slavic aspect is a paradigm example (Plungjan 2000: 124-126, 2011: 51-55, Wiemer 2006, 2017). Our talk aims at closing this gap.

In a nutshell, the rise of the Slavic pfv.:ipfv. opposition rests on the functional reinterpretation of productive patterns of verbal stem derivation, here schematized for the two least restricted patterns (Wiemer 2008, Wiemer/Seržant 2017):



Increasing complementary distribution affects contexts on word-form, predication, clause, and discourse level. Functions result from interaction with other categories (e.g., tense, imperative/modals±negation). They involve all, finite and non-finite, forms and apply regardless of the particular morphological and/or lexical relation between stems (as suggested in (2a-b)), since the function sets apply to pfv. vs ipfv. stems as classes (see (3)). The complementary association with these sets leads to complex paradigms which are very tight because aspect choice cannot be avoided in any, finite or non-finite, verb form.

Our theoretical contribution is to pinpoint the development of classificatory aspect and to compare its architecture with typical inflection and its paradigms. Our empirical contribution concerns the core of the system, i.e. the interaction between actionality types and aspect, which yields pairs or groups of derivationally related stems with shared lexical concepts (Lehmann 1988, Mende et al. 2013); functionally, these pairs and groups complement each other in ways comparable to inflectional AT-systems (Tatevosov 2016). In addition, as Janda/Lyashevskaya (2011) showed, the behavioral profiles of stems organized into aspect pairs according to the most frequent morphological patterns (see 2a-2b) are basically identical. While their investigation was restricted to the dominant grammatical forms of Russian verbs, we want to demonstrate how these two patterns equilibrate the system when they overlap in so-called triplets: two ipfv. stems occur as lexical equivalents of the same pfv. stem (e.g., Russ. *že-č'*^{IPFV}—***s-že-č'***^{PFV}—*s-žig-**a**-t'*^{IPFV} 'burn(TR)'). Paradigmatic coherence is maintained, or regained, either by dropping (or re-lexicalizing) one of the two ipfv. stems or by redistributing the two ipfv. stems over different functions from the function set assigned to ipfv. stems (Apresjan 1995). We will illustrate this dynamic process of paradigmatic strengthening by showing how the significance of triplets for the aspect system has changed since the mid-18th century in Russian, Polish, and Czech. Furthermore, for selected high-frequency triplets, we will present a token-based model of the distribution of both ipfv. stems over the standard inventory of functions and collocationary restrictions ascribed to ipfv. stems. The database created for this purpose comprises the period 1750-2018 and is based on corpora (listed below).

Acknowledgments: Research has been funded by German Science Foundation (WI 1286/19-1) and Polish Science Foundation (2016/23/G/HS2/00922), programme *Beethoven II*.

References

Apresjan, Jurij D. (1995), Traktovka izbytočnyx aspektual'nyx paradigm v tolkovom slovare, in Ju. D. Apresjan, (1995), *Izbrannye trudy, tom II: Integral'noe opisanie jazyka i sistemnaja leksikografija*. Moscow: Škola «Jazyki russkoj kul'tury», 102–113.

Janda, Laura A., Olga Lyashevskaya (2011), Grammatical profiles and the interaction of the lexicon with aspect, tense, and mood in Russian, *Cognitive Linguistics* 22(4), 719–763.

Lehmann, Volkmar (1988), Der Aspekt und die lexikalische Bedeutung des Verbs. *Zeitschrift für slavische Philologie* 48(1), 170–181.

Matasović, Ranko (2018), Aspect and *Aktionsart* in Slavic, Inflection and Derivation in RRG, in R. Kailuweit, L. Künkel, and E. Staudinger (eds), (2018), *Applying and Expanding Role and Reference Grammar*, Freiburg: FRIAS, 95–109.

Mende, Julia et al. (2013), *Vid i akcional'nost' russkogo glagola: Opyt slovarja*, Munich:Sagner.

Nicolle, Steve (2012), Diachrony and grammaticalization, in R. I. Binnick (ed.), (2012), *The Oxford Handbook of Tense and Aspect*, Oxford etc.: Oxford University Press, 370–397.

Plungjan, Vladimir A. (2000), *Obščaja morfologija. Vvedenie v problematiku*, Moscow: URSS.

Plungjan, Vladimir A. (2011), *Vvedenie v grammatičeskuju semantiku: grammatičeskie značenija i grammatičeskie sistemy v jazykax mira*, Moscow: Izdatel'stvo RGGU.

Tatevosov, Sergej G. (2016), *Glagol'nye klassy i tipologija akcional'nosti*. Moscow: JaSK.

Wiemer, Björn (2006), O razgraničenii grammatičeskix i leksičeskix protivopostavlenij v glagol'nom slovoobrazovanii, ili: čemu mogut naučit'sja aspektologi na primere *sja*-glagolov?, in V. Lehmann (ed.), (2006), *Semantika i struktura slavjanskogo vida* IV,Munich: Sagner, 97–123.

Wiemer, Björn (2008), Zur innerslavischen Variation bei der Aspektwahl und der Gewichtung ihrer Faktoren, in K. Gutschmidt, U. Jekutsch, S. Kempgen, and L. Udolph (eds), (2008), *Deutsche Beiträge zum 14. Internationalen Slavistenkongreß, Ohrid 2008*, Munich: Sagner, 383–409.

Wiemer, Björn (2017), O roli pristavok i suffiksov na rannix i pozdnix ėtapax istorii slavjanskogo vida, in R. Benacchio, A. Muro, and S. Slavkova (eds), (2017), *The role of prefixes in the formation of aspectuality (Issues of grammaticalization)*, Florence: Firenze University Press, 219–253.

Wiemer, Björn, and Ilja A. Seržant (2017), Diachrony and typology of Slavic aspect: What does morphology tell us?, in W. Bisang, and A. Malchukov (eds), (2017), *Unity and diversityin grammaticalization scenarios*, Berlin: Language Science Press, 230–307.

Corpora

Czech National Corpus: https://korpus.cz/

Electronic corpus of 17th and 18th century Polish texts („Korba"): http://clip.ipipan.waw.pl/KORBA

Polish National Corpus: http://nkjp.pl/

Russian National Corpus: http://www.ruscorpora.ru/

Grammaticalisation, schematisation and paradigmaticisation: How they intersect in the development of German degree modifiers

Jakob Neels & Stefan Hartmann
(Leipzig University & University of Bamberg)

Keywords: paradigmaticisation, mid-level schemas, diachronic construction grammar, analogy, degree modifiers

Grammaticalisation research has identified the tight integration of structures into paradigms as the final state of many grammaticalisation processes (e.g. Lehmann 2015 [1982], Diewald 2009). Construction grammar approaches are particularly suited to model such cases of paradigmaticisation since they invite researchers to study constructions not in isolation but in a wider network context (Traugott and Trousdale 2013). At the same time, usage-based construction grammar has called into question the psychological importance of paradigms in the sense of high-level generalisations or schemas, arguing instead that language users rely on more local generalisations and lower-level schemas (e.g. Dąbrowska 2015). From this perspective, the notion of paradigmaticisation may be theoretically more useful when it captures such lower-level schemas as well.

The present paper reflects on this idea, investigating a case of language change in which a whole family of constructions grammaticalised. Based on diachronic usage data from the *Deutsches Textarchiv* (DTA, Geyken and Gloning 2015), this paper presents quantitative and qualitative analyses on the interrelated German constructions [*ein wenig* X] ('a little'), [*ein bisschen* X] ('a bit'), [*ein Quäntchen* X] (lit. 'a quantum'), [*ein Tick* X] (lit. 'a tick') and [*eine Idee* X] (lit. 'an idea'). In particular, we investigate their gradual context expansion from more referential ("lexical") uses to more procedural ("grammatical") quantifier and degree-modifier uses. In constructionist terms, shifts towards grammatical functions translate into increasing schematicity not only at the level of constructional semantics but also at the level of the constructional slots as reflected by the types of collocates entering these slots. The picture emerging from the corpus analyses is that the older quantifier/degree-modifier constructions served as attractor sets for an increasingly strengthened mid-level schema, thereby promoting the analogically driven grammaticalisation of younger micro-constructions. This scenario is supported by the finding that younger (low-frequent) micro-constructions (i) approximate the usage patterns of more established ones, and (ii) "skip" stages on the cross-linguistically attested grammaticalisation path pre-partitive > partitive > quantifier > degree modifier (cf. Traugott 2008). However, despite similarities in this constructional family, *ein Tick* and *eine Idee* especially exhibit constraints that cannot be derived solely from a single overarching mid-level schema. Such properties specific to individual micro-constructions or subsets of constructions may be interpreted not just as persistence (Hopper 1991) but as entrenched subschemas or "local paradigms". Thus, we reason that the process of paradigmaticisation involves emergent paradigms of multiple orders in the sense of constructions at multiple levels of schematicity. We argue that change is guided mainly by associations to micro-constructions and lower-level meso-constructions (e.g. Traugott 2007); only in advanced stages of grammaticalisation, when micro-constructions become sufficiently homogeneous, do higher-level meso-constructions and macro-constructions – i.e. paradigms in the traditional understanding of the term – act as decisive organisational forces.

References

Dąbrowska, Ewa (2015), Individual differences in grammatical knowledge, in E. Dąbrowska and D. Divjak (eds.), *Handbook of cognitive linguistics*, Boston: De Gruyter, 650–668.

Diewald, Gabriele (2009), Konstruktionen und Paradigmen, *Zeitschrift für Germanistische Linguistik* 37(3), 445–468.

Geyken, Alexander and Thomas Gloning (2015), A living text archive of 15th–19th century German: Corpus strategies, technology, organization, in J. Gippert and R. Gehrke (eds.), *Historical Corpora*, Tübingen: Narr, 165–180.

Hopper, Paul J. (1991), On some principles of grammaticization, in E. C. Traugott and B. Heine (eds.), *Approaches to grammaticalization*, vol. 1, Amsterdam: John Benjamins, 17–35.

Lehmann, Christian (2015 [1982]), *Thoughts on grammaticalization*, 3rd edn. (Classics in Linguistics 1), Berlin: Language Science Press.

Traugott, Elizabeth C. (2007), The concepts of constructional mismatch and type-shifting from the perspective of grammaticalization, *Cognitive Linguistics* 18(4), 523–557.

Traugott, Elizabeth C. (2008), The grammaticalization of NP of NP patterns, in A. Bergs and G. Diewald (eds.), *Constructions and language change*, Berlin: De Gruyter, 23–45.

Traugott, Elizabeth C. and Graeme Trousdale (2013), *Constructionalization and constructional changes*, Oxford: Oxford University Press.

ꕥ

Modelling paradigmatic knowledge in DCxG: The case of Middle English *sum(e)* and its development into an indefinite article

Lotte Sommerer & Klaus Hoffmann
(University of Vienna, University of Vienna)

Keywords: indefinite article, nominal determination, Diachronic Construction Grammar, constructionalization, paradigm pressure, horizontal links

By investigating the grammaticalization of the Middle English determinative *sum(e)* into an indefinite article, this paper discusses a) if and how paradigmatic pressures can drive language change and b) if and how a speaker's knowledge of grammatical paradigms and/or (changing) paradigmatic relations can and should be modelled in usage-based, cognitive CxG (Goldberg 2006, Hilpert 2013; Barthdal et al. 2015; Sommerer & Smirnova, forthc.).

In PDE, the determinative *some* fulfills various functions: It can be used as a quantifier with a partitive/selective reading (ex.1) or expresses non-specificity/vagueness/speaker-attitude with singular nouns (ex.2). Additionally, *some* can also function as an indefinite 'near-article' (non-partitive reading) selecting plural and non-count heads (ex.3a-b) (e.g. Sahlin 1979, Israel 1999, Jacobsson 2002, and Berezowski 2009).

(1) *Some people like pineapple on their pizza*.
(2) *Some idiot must have left the fridge open!*
(3) a. *I need to buy some apples*.
b. *I need some milk*.

Two questions are when and why *some* took up the article function. English developed an article category as a result of a systemic reorganization of (in)definiteness marking. The process began in late OE when overt definiteness marking became obligatory for referential NPs (Mustanoja 1960, Rissanen

1967, and Sommerer 2018). This is equivalent to the emergence of an abstract NP-construction with a determination slot requiring to be filled. Consequently, the OE demonstrative *se* ('that') was recruited as a definite article. By analogy, the numeral *ān* ('one') was recruited with some delay as a default marker of indefiniteness for singular nouns, grammaticalizing to *a(n)*. We will show that *sum(e)* started to grammaticalize into an indefinite article before plural and mass nouns even later, namely in late ME (ex.4).

(4) […] **þ**at he schuld sende <u>*summe prestes*</u> to **þ**is lond […] (*Capgrave's Chronicle*, 15th c.)

Empirically, the paper relies on the qualitative and quantitative analysis of textual material from PPCME2 and PPCEME. We track distributional changes of determinative *some* before count and mass nouns in the singular and plural in contrast with other constructions competing for the same functional niches (e.g. indefinite *a* and 'bare' nouns). In addition to standard tests of statistical significance, we also supply measures of productivity (type/token ratio) to trace the spread of innovative functions across linguistic contexts (Gries & Ellis 2015).

In terms of theoretical perspective, we reconceptualize grammaticalization as constructionalization (Traugott & Trousdale 2013; Coussé et al. 2018) and suggest that cognitive factors (frequency, processing efficiency, analogy, …) and the influence of related constructions are primarily responsible for the development. We discuss to what extent the phenomenon can be described in terms of paradigm pressure (i.e. development of *some* into an indefinite article as a means to reduce zero marking of referentiality and to complete the English article paradigm). We also show how such a systemic change can be conceptualized in a constructional network model via node creation and the emergence of horizontal, paradigmatic links without postulating the existence of 'hyper-constructions' (Diewald, forthc.), or 'second order schemas' (Booij 2016).

References

Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), (2015), *Diachronic Construction Grammar*, Amsterdam: John Benjamins.

Booij, Geert (2016), Construction morphology, in A. Hippisley & G. Stump (eds.), *The Cambridge Handbook of Morphology*, Cambridge: CUP, 424-448.

Berezowski, Leszek (2009), *The Myth of the Zero Article*, London: continuum.

Coussé, Evie, Joel Olofsson & Peter Andersson (2018), (eds.), *Grammaticalization meets Construction Grammar*, Amsterdam: Benjamins.

Christophersen, Paul. (1939), *The Articles: A Study of their Theory and Use in English*, Oxford: OUP.

Diewald, Gabriele (forthc.), Paradigms lost – paradigms regained. Paradigms as hyper-constructions, in L. Sommerer & E. Smirnova (eds.), *Nodes and Links in the Network: Advances in Diachronic Construction Grammar*, Amsterdam: Benjamins.

Goldberg, Adele E. (2006), *Constructions at Work. The Nature of Generalization in Language*, Oxford: OUP.

Gries, Stefan Th. & Nick C. Ellis. (2015), Statistical measures for usage-based linguistics. *Language Learning* 65(S1), 228-255.

Hilpert, Martin (2013), *Constructional Change in English. Developments in Allomorphy, Word Formation, and Syntax*, Cambridge: Cambridge University Press.

Israel, Michael (1999), *Some* and the pragmatics of indefinite construal. *Annual Meeting of the Berkley Linguistics Society* 25(1), 169-182.

Jacobsson, Bengt (2002), The Indefinites *some* and *any* in linguistic theory and actual usage. *Studia Neophilologica* 74(1), 1-14.

Mustanoja, Tauno F. (1960), *A Middle English Syntax. Part I*, Helsinki: Société Néophilologique.

PPCME2 = Kroch, Anthony & Ann Taylor. 2000. *The Penn-Helsinki Parsed Corpus of Middle English*, Department of Linguistics, University of Pennsylvania.

PPCEME = Kroch, Anthony; Santorini, Beatrice & Lauren Delfs. 2004. *The Penn-Helsinki Parsed Corpus of Early Modern English*, Department of Linguistics, University of Pennsylvania.

Rissanen, Matti (1967), *The Uses of* One *in Old and Early Middle English*, Helsinki: Modern Language Society.

Sahlin, Elisabeth (1979*)*, Some *and* Any *in Spoken and Written English*, Stockholm: Almquist & Wiksell.

Sommerer, Lotte (2018), *Article Emergence in Old English. A Constructionalist Perspective*, Berlin: De Gruyter Mouton.

Sommerer, Lotte & Elena Smirnova (Eds.) (forthc.), *Nodes and Links in the Network. Advances in Diachronic Construction Grammar*, Amsterdam: John Benjamins.

Traugott, Elizabeth C. & Graeme Trousdale (2013), *Constructionalization and Constructional Changes*, Oxford: OUP.

WORKSHOP 14

Schedule: Thu 14.00 – Fri 16.55 (Room 2)

Patterns of language contact within and across phylogenies

Carlota de Benito Moreno, Rik van Gijn, Anja Hasse, Patrick Mächler & Max Wahlström
(University of Zurich; University of Zurich; University of Zurich; University of Zurich; & University of Helsinki)

Keywords: language contact; areal linguistics; linguistic phylogenies; language change; historical linguistics

A long-standing debate in historical linguistics concerns the question how language change progresses. Two competing models were proposed in the 19 th century: the tree model (Schleicher 1853) and the wave model (Schuchardt 1870; Schmidt 1872). Although there are good reasons to regard these models as complementary rather than oppositional, they represented an opposition at the time, and often have been presented as opposed, mutually exclusive models since. This has led to quite separate traditions in historical linguistics that focus on inheritance on the one hand, and contact on the other.

One consequence of this is that the empirical basis for theories of contact-induced change is skewed towards situations involving unrelated or distantly related languages. This is perhaps especially apparent in areal linguistics, which clearly focuses on areas with unrelated or distantly related languages, excluding – sometimes by definition (see e.g. Thomason 2001) – contact between closely related languages.

With this workshop we would like to focus on two related general questions:

Are there systematic differences between contact involving (closely) related languages on the one hand, and unrelated or distantly related languages on the other?
In spite of Thomason and Kaufman's (1988) famous claim that, linguistically speaking, anything is possible in language contact, general tendencies have been established, e.g. in the borrowability of types of words, or features of language (see e.g. Moravcsik 1978, Curnow 2002, Field 2002, Matras & Sakel 2007, Haspelmath & Tadmor 2009, Dediu & Cysouw 2013).

Given that related languages are generally more similar to each other than unrelated languages, transfer of features from one variety to another can b e expected to be easier than with unrelated languages. Moreover, a higher degree of mutual intelligibility may facilitate communication between speakers, increasing the intensity of contact. When it comes to dialects in contact, similarity between such closely related languages is considered to be a factor facilitating contact as well as a result of contact. Furthermore, phylogenetic and geographic distance tend to correlate (Bowern 2013). These factors can have consequences for what is possible in language contact. This raises a number of questions about potential systematic differences between contact between related versus unrelated languages, such as:

- Is borrowability influenced by relatedness and/or typological distance?
- Are differences in contact between related and unrelated languages more apparent in specific subsystems of language (phonology, morphology, syntax, lexicon)?
- Are certain features in language families inherently stable or prone to contact -induced change, regardless of the specific contact situation?

- What is the effect of typological similarity in contact situations (i.e. both in contact situations involving related languages and those involving unrelated languages)?
- What are the distributions and propensities of matter and pattern borr owing (Matras & Sakel 2007), respectively?

To what extent do we need separate methodologies for study ing contact situations between related and unrelated languages?

Detecting contact signals rests on establishing similarities between languages that are spoken closely together, which cannot be explained by parallel development due to universal pressures or genetic inheritance. Detecting contact effects between related languages is made more difficult because of the dominant presence of genetic inheritance as a source of similarity, and rests heavily on exceptions to established regular change. This also raises a number of questions:

- To what extent do we have to interpret patterns of similarity differently if the languages are related or not?
- To what extent can we use a single approach to detecting contact signals across all types of situations?
- To what extent is geographical proximity between languages (areality) a justified assumption for detecting contact signals?
- To what extent is regularity, or lack thereof, a reliable baseline for distinguishing contact from inheritance?

References

Bowern, Claire. 2013. Relatedness as a Factor in Language Contact. *Journal of Language Contact* 6: 411-432.

Dan, Dediu & Cysouw, Michael. 2013. Some Structural Aspects of Language Are More Stable than Others: A Comparison of Seven Methods. *PLoS ONE* 8(1): e55009.

Field, Fredric W. 2002. *Linguistic Borrow ing in Bilingual Contexts*. Amsterdam: John Benjamins Publishing Company.

Haspelmath, Martin & Tadmor, Uri (eds.). 2009. *Loanwords in the world's languages: a comparative handbook.* Berlin: Mouton de Gruyter.

Matras, Yaron & Sakel, Jeanette. 2007. *Grammatical Borrow ing in C ross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Moravcsik, Edith. 1978. Universals of Language Contact. In J. H. Greenberg (ed.). *Universals of Human Language*, 94-122. Stanford: Stanford University Press.

Schleicher, August. 1853. Die ersten Spaltungen des indogermanischen Urvolkes. *Allgemeine Monatssch rift für Wissenschaft und Literatur 3*, 786-787.

Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen.* Weimar, H. Böhlau.

Schuchardt, Hugo. [1870] 1900. *Über die Klassif ikation der romanischen Mundarten*. http://schuchardt.uni-graz.at/werk/jahr/1900.

Thomason, Sarah & Kaufman, Terrence. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Different mechanisms of contact-induced sound change

Michael Frotscher

Although the characteristics of sound change are nowadays rather well understood, the exact mechanisms and the origin of sound change are still debated.

One of the proposed potential sources for phonetic change is language contact. In this paper it will be shown that there exist two essentially different types of contact-induced sound change, which differ from each other regarding the phylogenetic distance of the languages/dialects involved:

(i) sound change through contact between distantly related/unrelated languages
(ii) sound change through contact between closely related idioms

The first type applies when phone(me)s of the L2, which are difficult or impossible to articulate for the L1 speaker, are substituted for by phone(me)s present in the L1. This type of contact-induced sound change is widely accepted (see for instance the review article by SANKOFF 2001). The second type, sound change induced by contact between closely related idioms, however, is not generally recognised. It applies when a dialect is acquired by speakers of another dialect of the same language. Much like the first type, it also involves substitution of sounds, which are however directly identifiable with sounds of cognate words of the native dialect. A diagnostic criterium of this type is generalisations / hypercorrection. An example for this kind of change can be found in SCHRIJVER's 2011 article on the spread of the High German consonant shift, originating in the Rhenish fan and spreading through over-generalisation to the southern areas of Germany where it is most fully established. This second type will also be illustrated by similar case studies, such as the consonant gemination by *j*, which is commonly considered a Westgermanic phenomenon. A similar sound law is, however, also operative in Northgermanic, where *j*-gemination only affected velars (cf. e.g. Old Norse *leggja* ~ Gothic *lagjan*). It will be shown that both phenomena can be regarded as two ends of a continuous process, starting in the North (probably already during the time of the North-Westgermanic language unit), where it affected initially only velars, and then spreading through the whole dialect continuum which was to become the Westgermanic branch, where the sound change involves all consonants (except *r*).

Both types, working with substitutions necessarily affect the entire lexicon and thus result in exceptionless sound change as postulated by the Neo-Grammarian model. The case of sound change shows that language contact scenarios must be carefully distinguished regarding the phylogenetic distance of the languages involved.

References

Sankoff, Gillian. 2001. Linguistic Outcomes of Language Contact. In: Peter Trudgill et al. (edd.), *Handbook of Sociolinguistics*. Oxford: Basil Blackwell, 2001, 638-668.

Schrijver, Peter. 2011. The High German Consonant Shift and Language Contact. *Studies in Slavic and General Linguistics* 38, Language Contact in Times of Globalization, 217-249.

𝒟

Schrödinger's Sound Change: the interplay of inheritance and borrowing in related languages

Danny Law

<pdf>

𝒟

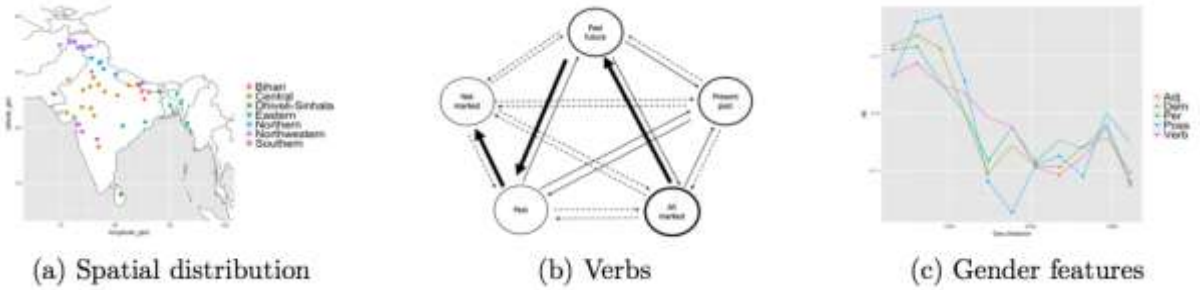
Measuring the phylogenetic signal and spatial effect of grammatical gender in Indo-Aryan languages

Marc Tang & Michael Dunn
(Uppsala University)

Keywords: Phylogenetics, Grammatical gender, Indo-Aryan, Diachrony, Contact

Phylogenetic comparative methods are used to investigate the processes of development and change of grammatical gender in Indo-Aryan languages. 44 presence-absence features pertaining to grammatical gender are compiled for 48 Indo-Aryan languages. The information relates to gender marking on the verbs, adjectives, personal pronouns, demonstrative pronouns, and possessive adjectives. For instance, gender marking on the verb is encoded with tense (past/ present/ future), number (singular/ plural), argument (subject/ object), and gender type (biological/ neuter/ animacy/ humanness). The languages included are shown in Figure 1a.

The phylogenetic signal of each gender feature is assessed according to Pagel's lambda (Pagel 1999, Ives et al 2007). On the one hand, the phylogenetic signal of animacy-based gender, humanness-based gender, and gender marking on adjectives is relatively weak. For instance, the signal on human/non-human distinction is weak since gender based on humanness is more Dravidian than Indo-Aryan. On the other hand, the phylogenetic signal is stronger on various types of pronouns, verbs, and grammatical gender systems based on biological gender and neuter gender.



(a) Spatial distribution (b) Verbs (c) Gender features

Figure 1. Overview of our dataset and its phylogenetic and spatial analysis

A Bayesian Reverse Jump Hyper Prior analysis (Green 1995, Gowri-Shankar and Rattray 2007), which infers the evolutionary dynamics of changes between feature values, is then used on the features considered to have a strong phylogenetic signal. The results are mostly consistent with historical linguistic and typological studies on gender systems in Indo-Aryan languages. Gender marking on the verb is more likely to be lost than gained while gender marking on demonstrative and personal pronouns is more likely to start from unmarked and then be first marked on the singular and finally be marked in both singular and plural. For instance, Figure 1b shows that while it is possible to acquire gender marking first in the past tense, then in the future, then in the present tense, one of the most probable evolution for verbs is to lose gender marking on the present, then on the future, and finally on the past. These transition probabilities give the same prediction as previous studies (Dahl 2004, Dediu and Cysouw 2013), as most Eastern Indo-Aryan languages have lost gender (Priestly 1983, Corbett 1991, Deo and Sharma 2007).

Furthermore, a non-phylogenetic partial Mantel correlogram controlling for phylogeny is used as an attempt to demonstrate the spatial influence (Legendre et al 2015). Fig 1c shows that the correlation of gender on possessives (blue) with geographical distance is stronger than other categories. This observation matches with our phylogenetic analysis. The phylogenetic signal of grammatical gender is weak on adjectives, which infers that adjectives in Indo-Aryan are more likely to have undergone interference, e.g., from contact. We suggest that neighbouring languages from other families should be included in this spatial analysis to further verify this speculation.

References

Corbett, G. G. (1991). *Gender*. Cambridge University Press, Cambridge.

Dahl, O. (2004). *The growth and maintenance of linguistic complexity*. John Benjamins, Amsterdam.

Dediu, D. and Cysouw, M. (2013). Some structural aspects of language are more stable than others: A comparison of seven methods. *PLoS ONE* 8(1).

Deo, A. and Sharma, D. (2007). Typological variation in the ergative morphology of Indo-Aryan languages. *Linguistic Typology* 10(3) 369–418.

Gowri-Shankar, V. and Rattray, M. (2007). A reversible jump method for bayesian phylogenetic inference with a nonhomogeneous substitution model. *Molecular Biology and Evolution* 24(6), 1286–1299.

Green, P. J. (1995). Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732.

Ives, A. R., Midford, P. E., and Garland Jr, T. (2007). Within-species variation and measurement error in phylogenetic comparative biology. *Systematic Biology* 56, 252–270.

Legendre, P., Fortin, M.-J., and Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution* 6, 1239–1247.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature* 401, 877–884.

Priestly, T. (1983). On drift in Indo-European gender systems. *Journal of Indo-European Studies* 11, 339–363.

Language change in German contact varieties

Christian Zimmer
(FU Berlin)

Keywords: German contact varieties, Namibian German, possessive *-s*, case marking, variation

Worldwide, there are contact varieties of German (traditionally called *language islands*). The contact languages include English, Russian, Brazilian Portuguese, and Afrikaans (cf. Ammon 2014). Furthermore, as a result of immigration to the German speaking area in Europe, German is also in contact with languages such as Turkish and Arabic (cf., e.g., Wiese 2013). This range of contact situations makes German contact varieties a promising topic for a study on the influence of different types of contact languages on language change: How similar are developments of German varieties in situations involving a closely related language (e.g. English) to those in which German is in contact with unrelated or distantly related languages (e.g. Turkish and Arabic)?

The starting point of my talk will be a description of the informal language use within the German-speaking community in Namibia (*Namdeutsch*). Today, approximately 20,000 speakers belong to this community. Almost all of them are also fluent in English (the official language of Namibia) and Afrikaans (the lingua franca in certain domains). Hence, we are dealing with a multilingual setting mainly involving three closely related languages (cf., e.g., Shah & Zappen-Thomson 2018, Zimmer in press).

On the basis of data taken from a systematically compiled corpus (cf. Wiese et al. 2017) and acceptability judgements of 211 speakers I will analyse grammatical innovations of Namdeutsch. Features to be examined include the possessive *-s* (cf. (1), for terminological considerations cf. Ackermann 2018: 238–322) and variation regarding case (cf. (2) vs. (3)).

(1) *mein-e* *Mutter-s* *Mutter ist* *in* *Deutschland*
my-NOM/AKK.SG mother-s mother is in Germany
geboren
born
'my mother's mother was born in Germany';

2) *und* *dann* *hab* *ich* *d-ie* *geholfen*
and then have I the-NOM/AKK.SG helped
'and then I helped her'

3) *und* *dann* *hab* *ich* *d-er* *geholfen*
and then have I the-GEN/DAT.SG helped
'and then I helped her'

My results will be compared to other German contact varieties such as Texas German and Kiezdeutsch (cf. Boas 2009, Wiese 2013), revealing striking similarities between contact settings involving closely related languages and settings involving unrelated languages (supporting, among others, findings by Rosenberg 2003 and Wiese et al. 2014). Thus, my results hint at some robust patterns of change in German contact varieties irrespective of the relatedness of their contact languages.

References

Ackermann, Tanja (2018), *Grammatik der Namen im Wandel. Diachrone Morphosyntax der Personennamen im Deutschen* (Studia Linguistica Germanica 134), Berlin & Boston: de Gruyter.

Ammon, Ulrich (2014), *Die Stellung der deutschen Sprache in der Welt*, Berlin: de Gruyter.

Boas, Hans C (2009), *The Life and Death of Texas German*. (Publications of the American Dialect Society 93), Durham, NC: Duke University Press.

Rosenberg, Peter (2003), Comparative speech island research: some results from studies in Russia and Brazil, in W. D. Keel & K. J. Mattheier (eds), *Deutsche Sprachinseln weltweit: Interne und externe Perspektiven German Language Varieties Worldwide: Internal and external Perspectives*, Frankfurt am Main: Peter Lang, 199–238.

Shah, Sheena & Marianne Zappen-Thomson (2018), German in Namibia, in Corinne A. Seals & Sheena Shah (eds), *Heritage Language Policies around the World*, Abingdon & New York: Routledge, 128–147.

Wiese, Heike (2013), What can new urban dialects tell us about internal language dynamics? The power of language diversity, in W. Abraham & E. Leiss (eds), *Dialektologie in neuem Gewand. Zu Mikro-/Varietätenlinguistik, Sprachenvergleich und Universalgrammatik* (Linguistische Berichte, Special Issue 19). Hamburg: Buske, 207–245.

Wiese, Heike, Horst J. Simon, Marianne Zappen-Thomson & Kathleen Schumann (2014), Deutsch im mehrsprachigen Kontext: Beobachtungen zu lexikalisch-grammatischen Entwicklungen im Namdeutschen und im Kiezdeutschen. *Zeitschrift für Dialektologie und Linguistik* 81, 274–307.

Wiese, Heike, Horst J. Simon, Christian Zimmer & Kathleen Schumann (2017), German in Namibia: A vital speech community and its multilingual dynamics, in P. Maitz & C. A. Volker (eds), *Language Contact in the German Colonies: Papua New Guinea and beyond* (Language & Linguistics in Melanesia, Special Issue 2017), 221–245.

Zimmer, Christian (in press), Deutsch als Minderheitensprache in Afrika, in J. Herrgen & J. E. Schmidt (eds), *Sprache und Raum – Deutsch. Ein internationales Handbuch der Sprachvariation* (Handbücher zur Sprach- und Kommunikationswissenschaft 30.4), Berlin: De Gruyter Mouton.

Dative in language contact: the Mennonite case

Neele Harlos

<pdf>

Differences in the polyglossic situtations in Walser language islands and their influence on language contact

Sandro Bachmann

<pdf>

Retracing language contact in the Ancient Central Andes by lexical comparison of trade good terminology

Hermann Sonntag
(Uni Tübingen)

Keywords: Ancient Central Andes; trade goods; lexical comparison; language contact; areal linguistics

In my talk I present the preliminary results of my current research into the linguistic situation of the (Ancient) Central Andes. Combining data from (historical) linguistics, (historical) ethnography and archaeology, I want to contribute to the reconstruction of the linguistic and socio-cultural situation in

the Ancient Central Andes, involving the examination of language contact, trade relations and even population movements.

Today, the Quechuan and Aymaran languages are the dominant language families in the Central Andes, surrounded by a multitude of smaller language families and isolates. However, it is known from numerous ethno-historical sources and descriptions that several other languages were spoken before the Spanish conquest but are extinct today (e.g., Pukina, Mochica, Culli, among many others), leaving only residual information about their grammar and vocabulary. In addition, many of the smaller present-day languages of the Central Andes are only incompletely documented and their speakers continuously exposed to more prestigious foreign languages, resulting in the danger of language loss (cf., e.g., Adelaar & Muysken 2004, Cerrón-Palomino 2015, Torero 2002).

As many of the (ancient) target languages in the Andean realm are only fragmentarily documented via several wordlists, my research intentionally focuses on the cross-linguistic comparison of terms for trade goods that are archaeologically attested in the Ancient Andes (i.e., goods found in places where they cannot originally stem from).

Putting this focus yields several advantages: On the one hand, trade good terminology typically is nominal and nominals are known to be more prone to borrowing than other word classes or bound morphology (cf., e.g., Hock 1991: 386). On the other hand, archaeologically attested (long-distance) trade often is a strong indication of contact between different populations. Furthermore, terms for trade items (being otherwise unavailable non-local goods) can plausibly be expected to be especially likely to be borrowed.

By examining related as well as presumably unrelated languages, my research addresses the workshop question how the borrowability of linguistic content is influenced by relatedness and/or typological distance. Additionally, by examining the involved languages' grammars when checking conspicuous loanword candidates, my research can contribute to the question how typological (dis)similarity might affect the outcome of contact situations.

References

Adelaar, Willem F.H. & Muysken, Pieter C. 2004. *The Languages of the Andes*. Cambridge: Cambridge University Press.

Cerrón-Palomino, Rodolfo. 2015. *The Languages of the Incas*. 39-54 in: Shimada, Izumi (ed.). 2015. *The Inka Empire: A Multidisciplinary Approach*. Austin: University of Texas Press.

Hock, Hans Henrich. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

Torero, Alfredo. 2002. *Idiomas de los Andes: lingüística e historia*. Lima: IFEA.

Diversity, stability and diffusion in the Hindu Kush region of Inner Asia

Henrik Liljegren
(Stockholm University)

The mountainous Hindu Kush, on the northwestern edge of the Indian subcontinent, offers a promising "test site" for questions relating to language contact, diffusion and stability, considering its high linguistic density and diversity, with languages from six genera: Indo-Aryan, Iranian, Nuristani, Tibetan, Turkic and Burushaski (Masica 2001: 225; Liljegren 2017: 107–108). While traces of several substrata suggest that the region in a distant past served as an important accretion zone (Tikkanen 1988: 304), akin to Caucasus (Nichols 2003: 306), the region as of today bears witness of several

waves of small- or larger scale diffusion (Bashir 2003: 823; Liljegren 2014: 162–167), mainly related to Indo-Aryan northward expansion within the last few millennia (Morgenstierne 1932: 51; 1961: 138; Strand 2001: 200), and in more recent times by superstratal influences (Bashir 2007) from a few languages of wider scope (e.g. Pashto, Urdu and Dari).

The main question asked in the present study is to what extent properties of some linguistic subsystems are more prone to diffuse than others, and whether they cluster similarly (geographically or genealogically) or significantly differently. We also ask to what extent individual properties show a higher or lower degree of intra-genealogical stability. A set of comparable first-hand data was collected from 59 varieties (representing all six genera) in a handful of collaborative workshops held in the region. The data was coded and analysed for approximately 50, mainly binary, linguistic features belonging to five different subsystems: phonology, word order syntax, grammatical categories, simple clause properties, and lexical structure, with an even distribution of features belonging to each of these five. In addition, a basic word list was cognacy-coded and used as the basis for measuring genealogical relatedness. It was primarily Indo-Aryan, due to its high representation (33 of the 59 varieties) that was analysed for feature stability.

The preliminary results, visualized with NeighborNet representations (Huson & Bryant 2006), indicate differential clustering depending on subsystem. The clearest examples of diffusion affecting substantial parts of the region were found within phonology, word order and lexical structure, while simple clause features (alignment in particular) display more limited (subareal) clustering, and grammatical categories (especially gender) a relatively high degree of intra-genealogical stability.

Acknowledgments

This study is financed by the Swedish Research Council (421-2014-631) under the project heading *Language contact and relatedness in the Hindukush region*.

References

Bashir, Elena. 2003. Dardic. In George Cardona & Danesh Jain (eds.), *The Indo-Aryan Languages*, 818–894. London: Routledge.

Bashir, Elena. 2007. Contact-induced change in Khowar. In Heather Bolton & Saeed Shafqat (eds.), *New perspectives on Pakistan: Visions for the future*, 205–238. Oxford: Oxford University Press.

Huson, Daniel H. & David Bryant. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* 23(2). 254–267. doi:10.1093/molbev/msj030.

Liljegren, Henrik. 2014. A survey of alignment features in the Greater Hindukush with special references to Indo-Aryan. In Pirkko Suihkonen & Lindsay J. Whaley (eds.), *On diversity and complexity of languages spoken in Europe and North and Central Asia* (Studies in Language Companion Series 164), 133–174. Amsterdam: John Benjamins Publishing Company.

Liljegren, Henrik. 2017. Profiling Indo-Aryan in the Hindukush-Karakoram: A preliminary study of micro-typological patterns. *Journal of South Asian Languages and Linguistics* 4(1). 107–156. doi:10.1515/jsall-2017-0004.

Masica, Colin P. 2001. The definition and significance of linguistic areas: Methods, pitfalls, and possibilities (with special reference to the validity of South Asia as a linguistic area). In Peri Bhaskararao (ed.), *The yearbook of South Asian languages and linguistics 2001*, 205–267. London: SAGE.

Morgenstierne, Georg. 1932. *Report on a linguistic mission to North-Western India* (Serie C III - 1). Oslo: Instituttet for sammenlignende kulturforskning.

Morgenstierne, Georg. 1961. Dardic and Kafir Languages. *Encyclopedia of Islam*, vol. 2, Fasc. 25, 138–139. New Edition. Leiden: E.J. Brill.

Nichols, Johanna. 2003. Diversity and stability in language. In Brian D Joseph & Richard D Janda (eds.), *The handbook of historical linguistics*, 283–310. Malden, MA: Blackwell Pub.

Strand, Richard F. 2001. The tongues of Peristân. Appendix 1. In Alberto Cacopardo & Augusto Cacopardo (eds.), *Gates of Peristan: History, Religion and Society in the Hindu Kush* (Reports and Memoirs 5), 251–257. Rome: Istituto Italiano per l'Africa e l'Oriente (IsIAO).

Tikkanen, Bertil. 1988. On Burushaski and other ancient substrata in northwestern South Asia. *Studia Orientalia* 64. 303–325.

Contact-induced changes in the multilingual area of Western Ingria

Fedor Rozhanskiy & Elena Markus
(University of Tartu; Institute for Linguistic Studies RAS; & University of Tartu)

Keywords: language contacts, Finnic languages, convergence, borrowings, mixed languages

The proposed paper offers an analysis of contact-induced changes in Finnic languages in Western Ingria (the Kingisepp district of the Leningrad oblast, Russian Federation). This multilingual area presents a challenging case since it reflects the effects of simultaneous contact between several closely-related languages (Votic, Ingrian, Ingrian Finnish, and Estonian) and further contact with the unrelated Russian language. Contact between these languages gives rise to various convergent processes (in both a strict and broad sense of the term, see Berruto 2005: 82). In fact, the whole Western Ingria is definitely a convergence area (in terms of Weinreich 1958: 379).

While particular aspects of the contact situation in Ingria have been studied (see Haarmann 1984a, 1984b, Muslimov 2005), there is currently no comprehensive research on the subject. We aim to compare the changes that arise due to contact between various Finnic languages on the one hand, and Finnic and Russian languages on the other. The analysis is based mainly on the corpus of our field materials which consist of about 1200 hours of elicitations and spontaneous speech recorded from the speakers of minor Finnic languages.

Our main conclusions are the following:

1) Generally, we find no strict correlation between the type of change and whether the contact takes place between related or unrelated languages. However, changes on the morphophonological level typically reflect the influence of related languages.

2) Contact-induced changes caused by related and unrelated languages can overlap. For instance, Votic acquired voiced geminates *bb*, *dd*, *gg* by (a) borrowing some Russian words with voiced consonants (*griba* 'mushroom' < Rus. *grib*, *parahoda* 'steamboat' < Rus. *parahod*, *kruga* 'circle' < Rus. *krug*) and (b) geminating single consonants in certain morphological forms as a pattern borrowed from Ingrian (hence the partitive forms *gribba*, *parahodda*, *krugga*).

3) Contacts between closely-related languages are more likely to result in convergent varieties. In Ingria we found at least two examples. The Lower Luga Ingrian is a convergent variety which developed through interethnic communication between the original Votic and Ingrian population living in the same area, and had some influence from Ingrian Finnish and Estonian. The mixed Votic-Ingrian variety known as "Kukkuzi Votic" is a typical mixed language (in terms of Thomason 2003) that does not have an unambiguous genetic affiliation (in line with the analysis by Noonan 2010: 61).

4) Closely-related languages can use the MAT and PAT loan (Matras & Sakel 2007) of the same phenomenon simultaneously. For instance, the Lower Luga Ingrian has parallel comitative

forms: one with the borrowed marker *ka* attached to the Ingrian genitive form (*saha-n-ka* saw-GEN-COM 'with a saw'), and the other as a borrowing of a full Votic comitative form (*saha-ka* saw.GEN-COM 'with a saw').

5) Contacts between two closely related languages can result in "soft borrowing" when nothing principally new is adopted but some forms conform to the donor language. E.g due to Ingrian influence, some Western Votic words restored the original initial *h*, cf. Finnish *hirvi* 'elk', Estonian *hirv* – Votic *irvi* – contemporary Western Votic *hirvi*).

References

Berruto, Gaetano (2005), Dialect/standard convergence, mixing, and models of language contact: the case of Italy. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.). *Dialect change. Convergence and Divergence in European Languages*. Cambridge University Press.

Haarmann, Harald (1984a), *Elemente einer Soziologie der kleinen Sprachen Europas. Band 3: Aspekte der ingrisch-russischen Sprachkontakten.* Hamburg: Buske.

Haarmann, Harald, (1984b). *Soziolinguistisch-lexikologische Studien zu den ingrisch-russischen Sprachkontakten : mit einem ingrisch-russischen Wörterverzeichnis* (Fenno-Ugrica; Bd. 7). Hamburg: Buske.

Matras, Yaron & Sakel, Jeanette (2007), Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31 (4), 829–865.

Muslimov, Mehmed Z. (2005), *Jazykovyje kontakty v Zapadnoj Ingermanlandii (nizhnee techenije reki Lugi)* [Language contacts in Western Ingria (the lower course of the Luga river)]. Ph.D. thesis from the Institute for Linguistic Studies. St. Petersburg.

Noonan, Michael (2010), Genetic Classification and Language Contact. In Raymond Hickey (ed.). *The Handbook of Language Contact*. Chichester: Blackwell, 48–65.

Thomason, Sarah G. (2003), Social factors and linguistic processes in the emergence of stable mixed languages. In Yaron Matras & Peter Bakker (eds.). *The mixed language debate.* Berlin–New-York: de Gruyter, 21–39.

Weinreich, Uriel (1958), On the Compatibility of Genetic Relationship and Convergent Development. *WORD* 14:2-3, 374–379.

Towards a typology of pattern borrowings: Case study of subject reference change in Votic and Ingrian

Evgeniya Budennaya
(Institute of Linguistics, Russian Academy of Sciences)

Keywords: null subject, personal pronoun, language contact, Russian, Votic, Ingrian

This talk focuses on the loss of referential null subjects in Votic and Ingrian, two minor Finnic languages spoken to the south-west of the Leningrad region (historical Ingria). Their long-lasting contacts with Russian and ensuing borrowings on different linguistic levels are well-documented (Hartmann 1984; Rozhanskiy 2009), yet the uncommon double pattern of expressing pronominal subjects has not been studied so far. In the talk I trace the development of this pattern with regard to language contact and discuss how the results apply to a more general typology of factors that facilitate contact-induced change.

To express the reduced subject, Votic and Ingrian employ both personal pronouns and verbal inflection in all persons:

(1) Votic (Markus, Rozhanskiy 2017):

a. *sis* *tämä* *tul-i-ø*
then 3SG come-PST-3SG
'Then he came'

b. *miä* *tätä* *e-n* *e* *kensa* *näh-nü*
1SG 3SG.PART NEG-1SG be never see-PTCP.ACT
'I have never seen him'

(2) Ingrian (non-published recording by F. I. Rozhanskiy and E. B. Markus, 2011):

a. *hā̄* *kūl-i-ø*
3SG die-PST-3SG
'She died'

b. *miä* *muišša-n* *šene-n* *hüväšt*
1SG remember.PRS-1SG it-GEN.SG good
'I remember it well'

Here Votic and Ingrian differ from other Uralic languages, where referential subject pronouns are mostly omitted (Kibrik 2013: 239-241). Taken into account the typological rarity of the double-marking pattern (Siewierska 2004: 268), its genesis in minor Finnic is of great interest. Internal scenarios dealing with the loss of agreement (Rizzi 1986) cannot be suggested since both Votic and Ingrian have preserved non-syncretic verbal inflection. However, the presence of a similar double pattern in Russian allows external influence to be viewed as a possible trigger for the rise of pronouns in Finnic.

To shed the light, I carried out a diachronic comparative study of Votic and Ingrian texts from the 19th century till today. The overall volume of analyzed finite clauses comprised 607 units. The binominal test showed that in both languages during the second half of the 20th century a significant expansion of the third person pronouns took place , with a higher rate for Votic that was always considered less prestigious in comparison to Russian and Ingrian (Markus, Rozhanskiy 2017: 13). Since contacts with Russian speakers considerably enhanced after the 1930s, the rise of third-person pronouns in minor Finnic which resulted in a regular double-marking referential pattern in all three persons, can be treated as a loss of linguistic complexity induced by a massively increased adult-only language contact (Trudgill 2011). In addition, some structural similarities between Russian and minor Finnic, such as SVO word order, might also contribute to the expansion of pronouns in Votic and Ingrian. This implication can account for a range of minor Turkic and Mongolian languages that preserved SOV order and were not affected by any rise of subject pronouns (Dryer 2013), despite the similar degree of contacts with Russian. Based on our data, typological similarity tends to play a more crucial role in facilitating the change in unrelated languages, compared to the sociolinguistic properties of the contact.

The research was supported by RFBR, grant № 17-40-00460.

References

Dryer, Matthew S. (2013), Expression of pronominal subjects, *The World Atlas of Language Structures Online*, Leipzig: Max Plank Digital Library (Available online at http://wals.info/chapter/101)

Hartmann, Harald (1984), *Soziolinguistisch-lexikologische Studien zu den ingrisch-russischen Sprachkontakten mit einem ingrisch-russischen Worterverzeichnis*. Band 7. Hamburg: Helmut Buske Verlag.

Rizzi, Luigi (1986), Null Objects in Italian and the theory of pro, *Linguistic Inquiry* 17, 501– 557.
Rozhanskiy, Fedor I. & Markus, Elena B. (2012), Zolotaja ptitsa (publikatsija izhorskoj skazki, zapisannoj v XIX veke). [The golden bird. A publication of an Ingrian tale recorded in the 19th century], *Acta Linguistica Petropolitana* VIII (1), 447–502.
Rozhanskiy, Fedor I. & Markus, Elena B. (2017), Sovremennyj vodskij jazyk [The contemporary Votic language], Saint-Petersburg: Nestor-Istorija.
Siewierska, Anna (2004), *Person*, Cambridge: Cambridge University Press.
Trudgill, Peter (2011), *Sociolinguistic typology: Social determinants of linguistic complexity*, Oxford: Oxford University Press.

ꕥ

The impact of relatedness and typological distance on pattern borrowability: A comparative study of D-marking on proper names

Iker Salaberri
(Public University of Navarre)

Keywords: contact situation, typological distance, pattern borrowability, D-marking, proper names

A central point of discussion in the literature on language contact in recent decades has addressed the question whether and to what degree genetic relationship and typological similarity favor the borrowing of various grammatical elements (Thomason & Kaufman 1988: 14, Haspelmath & Tadmor 2009: 2, Bowern 2013: 413 among others). In this sense, languages have been argued to be more prone to borrow from genetically close and typologically similar linguistic systems (Hock 1991: 388, Masica 2005: 182).

This study aims to contribute to the discussion by providing an analysis of one rarely investigated morphosyntactic property, namely the borrowing of definiteness and specificity marking —which Himmelmann (1997: 6-7) and Handschuh (2017: 491) refer to by the cover-term "D-marking"— on proper names in three contact situations: (1) Basque, Catalan and Iberian Spanish in Spain; (2) Fulacunda (Pulaar), Sereer and Wolof in Senegal; (3) Paraguayan Guaraní and Paraguayan Spanish in Paraguay. These contact areas provide valuable insights to the discussion at hand, since they encompass (i) unrelated languages and languages related to each other to varying degrees; (ii) typologically similar as well as distinct languages; (iii) different contact situations, including a more or less balanced bilingualism in Paraguay and a multilingual context in Senegal. Various onymic subclasses have been considered for the study, including anthroponyms (*Andy*, *Mary*, *Ms. Hopkins*), zoonyms (*Pickles*, *Rocky*, *Tango*), toponyms (*Morocco*, *New Orleans*, *Tuscany*, *Kilimanjaro*) and phenonyms (*Hurricane Katrina*, *The Valdivia Earthquake*), among others. The data have been drawn directly from native bilingual (i.e. Basque-Iberian Spanish, Fulacunda-Wolof, Paraguayan Guaraní-Paraguayan Spanish, etc.) speakers by means of questionnaires specifically designed to encourage the use of proper names.

First results suggest that the contact situation is a stronger predictor for the borrowability of D-marking on proper names than relatedness and typological distance. Moreover, onymic subclasses seem to differ with respect to the likelihood of developing D-marking under contact. This is illustrated by Paraguayan Guaraní (Yopará variety), a Tupian language which under considerable influence from Paraguayan Spanish has borrowed the D-markers *el*, *la*, *lo*. In the corpus of collected data these

elements are attested only with family names (1a), names that designate bodies of water (1b) and names of wars (1c), but not first names (1d) and names of countries (1e, Kallfell 2010: 219):

(1) a. *La González-kuéra iñ-akã porã*
D Gonzalez-PL first-head good
'The Gonzalezes are intelligent'

b. *La Mar Negro-pe ndaje-ko*
D Black Sea-LOC REPRT-DEM
'They say that in the Black Sea...'

c. *La Guerra Guasu-ramo guare*
D War large-SUB POSTP
'At the time of the World War'

d. *María o-gue-ro-guata imemby-míme*
Mary 3-FACT-CAUS-walk son-little
'Mary accompanies her little son'

e. *Kalo o-ho España-gui Marruecos-pe*
Charles 3-go.PST Spain-from Morocco-LOC
Gibraltar Estrecho rupi-ve
Gibraltar Strait across-LOC
'Charles went from Spain to Morocco across the Gibraltar Strait'

References

Bowern, Claire (2013), Relatedness as a factor in language contact, *Journal of Language Contact* 6, 411-432.

Handschuh, Corinna (2017), Nominal category marking on personal names: A typological study of case and definiteness, *Folia Linguistica* 51(2), 483-504.

Haspelmath, Martin & Uri Tadmor (2009), The Loanword Typology project and the World Loanword Database, in M. Haspelmath and U. Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, Berlin: Mouton de Gruyter, 1-34.

Himmelmann, Nikolaus P. (1997), *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*, Tübingen: Max Niemeyer.

Hock, Hans H. (1991), *Principles of historical linguistics*, Berlin / New York: Mouton de Gruyter.

Kallfell, Guido (2010), *Grammatik des Jopara: Gesprochenes Guaraní und Spanisch in Paraguay*, Münster: Westfälische Wilhelms-Universität dissertation.

Masica, Colin P. (2005), *Defining a linguistic area - South Asia*, New Delhi: Chronicle Books.

Thomason, Sarah G. & Terrence Kaufman (1988), *Language contact, creolization, and genetic linguistics*, Berkeley / Lons Angeles / London: University of California Press.

Language death and language contact in Iskonawa (Pano, Peru): teasing apart the causes of abrupt grammatical change in a language obsolescence situation

Roberto Zariquiey & Gabriela Tello

<pdf>

WORKSHOP 15

Schedule: We 11.00-17.55 (Room 9)

Phonological (in)stability and language evolution

Steven Moran & Eitan Grossman
University of Zürich & Hebrew University of Jerusalem

Keywords: phonology, typology, historical linguistics, language evolution, stability

The aim of this workshop is to explore the stability and instability of sound patterns, understood here as the set of phonetic and phonological properties of languages. The inherent stability of linguistic properties is a crucial component of any explanation of cross-linguistic and language-specific distributions, alongside considerations such as the number, frequency, and complexity of diachronic sources and developmental pathways (Greenberg 1978; Harris 2008) on the one hand, and the likelihood of diffusion from speaker to speaker or language to language, on the other.

In particular, the question of stability is an important one because linguists often draw inferences about human language on the basis of a sample, whether small and biased or large and balanced in terms of area and genealogy. Specifically, it would be ideal if linguists could infer the universal probability (or 'learnability') of a linguistic type from the empirical frequency of that type (Cysouw 2011). The possibility to draw valid inferences of this sort depends, however, to a large extent on some version of the uniformitarian assumption, i.e., the idea that `human languages have always been pretty much the same in terms of the typological distribution of the units that compose them' (Newmeyer 2002).

The uniformitarian assumption has been called into question in a number of ways. For example, Maslova (2000) argued that the assumption of a stationary distribution, such that the present-day distribution of linguistic properties in the world's languages, is independent of an initial state, and cannot be maintained. As a result, it becomes crucial to directly target transition probabilities between types (see also Dunn et al. 2011; Cysouw 2011; Bickel 2015). Greenberg (1978) observed that particular distributions might indicate different degrees of inherent stability. Nichols (1992, 2003) and Wichmann & Holman (2009) provide concrete measures of the relative stability of cross-linguistically comparable properties. Of the 137 properties examined, 19 deal with sound patterns, which show varying degrees of stability. For example, consonant inventories are rated as 'very unstable,' while tone is 'very stable.' While these studies provide us with a big picture of the relative stability of a number of properties, as well as some methodological foundations, we are still far from understanding the relative stability of a wide range of sound patterns. In particular, many aspects of (in)stability are potentially invisible to particular methodologies. For example, it may be the case that the phonetic precursors of, e.g., three-way length distinctions or prenasalized stops distinctions, are frequently innovated by speakers yet are not phonologized. Such frequent but evanescent innovations were envisioned already by Greenberg (1978), who predicted that they would

> be relatively frequent in languages and distributed relatively evenly among genealogical stocks.

Some proposals have been made about the inherent stability or instability of particular sound patterns. For example, Jacques (2011) argues that aspirated fricatives, despite the multiplicity of diachronic sources, are inherently unstable, due to their tendency to merge with other sounds. Dediu & Cysouw (2013) found that the feature [round] is unstable, i.e., hard to get and easy to lose. Blevins

(2008) proposes that three-way vowel nasality distinctions, as in Palantla Chinantec, or three-way length distinctions, as in Estonian, Saami, or Dinka, may be inherently unstable, and tend to be eradicated by sound change. On the other hand, coronal places of articulation for consonants seem to be especially stable, since total coronal loss is vanishingly rare (Blevins 2009). Moran & Verkerk (2018) found that consonants and vowels change at different rates, i.e., not uniformly across language families; these findings may point to broad differences between consonant inventories and vowel inventories in terms of stability.

A crucial issue in studying the stability of sound patterns is the need to tease apart the relative contributions of – and interactions between – inheritance, on the one hand, and areal effects, on the other. For example, it has recently been argued that affricate-dense inventories are inherently unstable in Eurasia, and tend to simplify unless supported areally (Nikolaev & Grossman 2018).

Sound patterns are an especially exciting domain for the study of stability, thanks to the possibility of experimental modelling (Ohala 1995, Silverman 2006). Another exciting avenue for studies of stability involve the comparison between the abundant data on reconstructed phonologies of proto-languages, on the one hand, and on present-day phonologies, on the other (see, e.g., Marsico et al. 2018; Moran & Verkerk 2018). Phonological areas provide yet another fascinating domain for research on phonological (in)stability, to the extent that they can reveal differential, areally-conditioned, patterns of innovation, loss, and retention. Furthermore, they allow a detailed examination of simple persistence vs. 'merry-go-round' areal stability, in which sound patterns diffuse from language to language.

Finally, it may be that particular types of sound pattern are 'attractors,' i.e., any state that is easier to enter or acquire than to leave or lose, and/or easier to retain than lose (Nichols 2018). This is essentially a matter of stability, although it is not clear to what extent the notion of attractor is explanatory, or requires, in turn, more primitive explanations.

We invite proposals for 20-minute talks that explore the stability of particular sound pattern types and on any of the following (or related) questions:

- How can the notion of 'stability' be defined and operationalized?
- What are the units of analysis in the study of phonological stability (phonemes, oppositions, features, etc.)?
- What are the differences between present-day distributions of sound patterns and earlier distributions, whether at a global level, at a macro-areal level, or at a micro-areal level?
- Can differential rates of change for different types of sound patterns be identified, and if so, what explanations explain these differences?
- Are different patterns of (in)stability found in different parts of the world or at different stages in the evolution of human language?
- What light can experimental phonetics and phonology shed on (in)stability?
- What light can modelling shed on (in)stability?
- What are the causal links between facts of human physiology and cognition, on the one hand, and the (in)stability of sound patterns?

References

Bickel, Balthasar. 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In: Bernd Heine and Heiko Narrog (eds.) *The Oxford handbook of linguistic analysis*, 2nd ed., 901–23. Oxford: Oxford University Press.

Blevins, Juliette. 2008. Natural and unnatural sound patterns: A pocket field guide. In Klaas Willems & Ludovic De Cuypere (eds.), *Naturalness and iconicity in language*, 121–148. Amsterdam: Benjamins.

Blevins, Juliette. 2009. Another universal bites the dust: Northwest Mekeo lacks coronal phonemes. *Oceanic Linguistics*, 48/1: 264-273.

Cysouw, Michael. 2011. Understanding transition probabilities. *Linguistic Typology*, 15: 415–431.

Dediu, Dan & Michael Cysouw. 2013. Some structural aspects of language are more stable than others: A comparison of seven methods. *PloS ONE*, 8(1), e55009.

Dunn, Michael J., Simon J. Greenhill; Stephen C. Levinson; and Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473: 79–82.

Greenberg, Joseph H. 1978. Diachrony, synchrony and language universals. In Joseph H. Greenberg, Charles A. Ferguson and Moravcsik, Edith A. (eds.), *Universals of Human Language. Volume 1: Method and Theory*, 61–92. Stanford: Stanford University Press.

Guillaume Jacques. 2011. A panchronic study of aspirated fricatives, with new evidence from Pumi. *Lingua*, 121 (9): 1518–1538.

Harris, Alice. 2008. On the explanation of cross-linguistically unusual structures, in Jeff Good (ed.), *Linguistic universals and language change*, 54–76. Oxford: Oxford University Press.

Maslova, Elena. 2000. A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4: 307–333.

Marsico, Egidio, Sebastien Flavier, Annemarie Verkerk & Steven Moran. 2018. BDPROTO: A Database of Phonological Inventories from Ancient and Reconstructed Languages. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018), 1654–1658. May 7–12, Miyazaki, Japan.

Moran, Steven & Annemarie Verkerk. 2018. Differential rates of change in consonant and vowel systems. In C. Cuskley and M. Flaherty and H. Little and Luke McCrohon and A. Ravignani and T. Verhoef (eds.), *The Evolution of Language: Proceedings of the 12th International Conference* (EVOLANGXII). April 16–19, Toruń, Poland.

Newmeyer, Frederick. 2002. Uniformitarian assumptions and language evolution research. In *The Transition to Language*, ed. by Alison Wray. Oxford: Oxford University Press.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.

Nichols, Johanna. 2003. Diversity and stability in languages. In: Brian D. Joseph, and Richard D. Janda (eds.), *The Handbook of Historical Linguistic*s, 283–310. Malden/Oxford/Melbourne/Berlin: Blackwell Publishing.

Nichols, Johanna. 2018. Non-linguistic conditions for causativization as a linguistic attractor. *Frontiers in Psychology* 8:2356. doi: 10.3389/fpsyg.2017.02356

Nikolaev, Dmitry & Eitan Grossman. 2018. Areal sound change and the distributional typology of affricate richness in Eurasia. *Studies in Language*.

Ohala, J. J. 1995. A probable case of clicks influencing the sound patterns of some European languages. *Phonetica* 52: 160-170.

Silverman, Daniel. (2006). The diachrony of labiality in Trique, and the functional relevance of gradience and variation. *Papers in Laboratory Phonology VIII*.

Wichmann, Søren & Eric Holman. 2009. *Assessing temporal stability for linguistic typological Features*. München: LINCOM Europa.

ꕥ

Diachronic phonological typology: A plea for detail

Patrick Honeybone
(University of Edinburgh)

Keywords: historical phonology; diachronic typology; phonological stability; lenition; Laryngeal Realism

Phonological stability and instability are core explananda for diachronic phonological typology, a central branch of historical phonology. How stable are series of fortis or lenis stops (p,t,k; b,d,g), for example? Is one more stable than the other? How likely are specific changes which could eliminate them? The importance of this goes beyond historical phonology: answers might explain patterns in synchronic phonological typology, and might allow us to trace back more ancient genealogical relationships than is conventionally thought possible (by focusing on slow-changing features: Greenhill *et al*. 2017).

In this paper, I argue for caution in diachronic phonological typology, because (i) the surface description of phonological states can be misleading, and (ii) we need to be sure that the changes being compared really involve the same units, in the same kinds of systems. As exemplification, I ask these questions: How stable are series of fortis and lenis stops? How likely are changes which could eliminate them? I consider three points.

Firstly, a change might be inhibited not by the inherent stability of the segments involved, but rather because of the systemic relationships that the segments contract. For example, Gurevich (2004) shows that lenition-like changes (which might turn b,d,g into fricatives, for example) are far more likely if they are non-neutralising than if they are neutralising. This shows that the other segments that exist in a system (which themselves do not change) need to be considered when we investigate the potential for change: we cannot simply talk of the likelihood of the spirantisation of b,d,g.

Secondly, it has been argued that a change of the type p,t,k > b,d,g is impossible because it would only leave voiced stops in the phonological system (a 'marked' state). A case of this change has however been reported in the literature, however, as an aspect of the binnendeutsche Konsonantenschwaechung (Lessiak 1933). Are such changes possible? A growing position in phonological theory – Laryngeal Realism – argues that this change was not, in fact, a case of p,t,k > b,d,g but was rather p^h,t^h,k^h > p,t,k, so the diachronic impossibility of p,t,k > b,d,g holds, and p,t,k remain highly stable.

Thirdly, the detailed analysis of changes requires considerable energy in the investigation of the systems that are involved: in cases of stability, this involves everything which might weigh on the segments concerned (the segmental system, its phonotactics, the prosodic system, etc); in cases of change, it involves all this for both pre-change and post-change states. This detail could preclude a 'big data' approach for some questions, and might need what something akin to what Baker (2009) calls 'Formal Generative Typology' – the detailed analysis of structure in a manageable number of unrelated languages.

My research question is therefore: how should we do diachronic phonological typology? My data is gathered from the history of several languages, extracted from previous work; the methodology involves the detailed analysis of the data, applying the insights of phonological theory. Diachronic phonological typology is an important task which deserves phonological detail.

References

Baker, Mark (2009), Formal generative typology, in B. Heine & H. Narrog (eds), *The Oxford Handbook of Linguistic Analysis*, Oxford: Oxford University Press.

Greenhill, Simon, Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen Levinson & Russell Gray (2017), Evolutionary Dynamics of Language Systems, *Proceedings of the National Academy of Sciences* 114 (42).

Gurevich, Naomi (2004), *Lenition and Contrast: the Functional Consequences of Certain Phonetically Conditioned Sound Changes*, New York / London: Routledge.

Lessiak, Primus (1933), *Beiträge zur Geschichte des deutschen Konsonantismus*, Brünn: Rohrer.

ꝺ

How stochastic is sound change?

Cormac Anderson
(Max-Planck-Institute for the Science of Human History, Jena)

Keywords: phonology, historical linguistics, phonological typology

An open question regarding sound change is to what extent it can be considered a stochastic process. With the computational turn in historical linguistics and the increased availability of cognate-coded lexical corpora (i.a. the Australian Basic Vocabulary Database: Greenhill et al. 2008; TransNewGuinea.org: Greenhill 2015), as well as large cross-linguistic databases of phoneme inventories (Phoible: Moran et al. 2014; the World Phontactic Database: Donohue et al. 2013), also of proto-languages (Marsico et al. accepted), questions such as this are becoming increasingly amenable to quantitative investigation. This paper discusses a number of constraints on the stochasticity of sound change with a view to building better quantitative models of phonological evolution.

There are two separate issues at stake here. On the one hand, there is the status of the various internal conditioning factors on sound change that have been proposed in the literature. These fall into two main categories: phonological and phonetic. On the other hand, the adequacy of the existing quantitative resources to assess these conditioning factors must be considered.

On the phonological side, since Jakobson (1929), there has been recognition that sound change is structure-dependent. The shape of a sound system conditions the type of sound change that is likely to occur (e.g. Martinet 1952 for a good example of this type of argumentation in Celtic) and phonologists frequently evoke structure preservation as an explanatory mechanism.

As well as these phonological factors, there are also language-specific phonetic factors that condition sound change. In some cases, this might be trivial, whereby a contingent phonetic element becomes phonologised due to change elsewhere in the system (e.g. Newman 1976 for such a sound change in the development of Hausa). However, synchronically non-contrastive variation can also condition the types of sound change that occur and their outcomes. This may be an explanation, alongside polymorphism in the protolanguage, for the phenomenon of Sapirean drift.

The second category of constraint lends itself more to fine-grained phonetic enquiry and is clearly less amenable to quantitative investigation than the first. Phonological constraints on the stochasticity of sound change can perhaps be investigated, however, by considering structural cohesion in synchronic systems as a proxy for the diachronic developments leading to such systems, as has been done, for example, in papers considering environmental or cognitive biases on synchronic typological distributions (i.a. Blasi et al. 2016; Everett et al. 2016).

References

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., Christiansen, M. H. 2016. “Sound-meaning association biases evidenced across thousands of languages”. In PNAS.

Donohue, M., Hetherington, R., McElvenny, J, Dawson, V. 2013. World phonotactics database. Canberra: Department of Linguistics, The Australian National University.

Everett, C., Blasí, D. E., Roberts, S. 2016. “Language evolution and climate: the case of desiccation and tone”. In Journal of Language Evolution 1.1, 33–46.

Greenhill, S.J., Blust. R, & Gray, R.D. 2008. “The Australian Basic Vocabulary Database: from bioinformatics to lexomics”. In Evolutionary Bioninformatics 4, 271–283.

Greenhill, S. J. 2015. "TransNewGuinea.org: An Online Database of New Guinea Languages". In PLoS One.

Jakobson, R. 1929. Remarques sur l'évolution phonologique du Russe comparée à celle des autres langues slaves. Prague: Travaux du cercle linguistique.

Marsico, E., Flavier, S., Verkerk, A., & Moran, S. Accepted. BDPROTO: A Database of Phonological Inventories from Ancient and Reconstructed Languages. To appear in Proceedings of the 11th Language Resources and Evaluation Conference (LREC).

Martinet, A. 1952. "Celtic lenition and Western Romance consonants" .In Language 28.2, 192-217.

Moran, S., McCloy, D., Wright, R. eds. 2014. PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Newman, P. 1976. "The origin of Hausa /h/". In Studies in African Linuistics, Supplement 6, 165–175.

ø

Eurasian consonant systems through time

Dmitry Nikolaev
(Hebrew University)

Keywords: languages of Eurasia, phonological stability, historical phonology, statistical analysis, phonological databases

Despite the lively interest that the problem of sound change has attracted over the last 150 years (see an overview by Garrett [2015]), the systemic aspect of this problem, viz. how segmental inventories evolve when viewed as a whole, received only scant attention. After the issues of 'chain shifts' and 'gap filling' were studied in the classic works by Martinet (1952) and Labov (1994), the main focus largely shifted to the study of phonetic, phonological, and cognitive underpinnings of individual sound-change processes (Garrett & Johnson 2013). The framework of *evolutionary phonology* advanced by Blevins (2004) aims to explain the 'sound patterns' of the world's languages based on the typology of sound-change processes. However, it does not trace the evolution of segmental systems and is mostly concerned with showing how particular types of sound change give rise to particular types of segments, phonological alternations, and phonotactic constraints. The only systemic topic that attracted attention recently is general diachronic stability of phonological inventories: consonantal inventories were found to be changing faster than vowel inventories (Moran & Verkerk 2018) and to be generally 'very unstable' (Wichmann & Holman 2009).

In this paper, we will provide a more nuanced overview of the evolution of phonological systems in Eurasia. As our data, we will use phonological systems of contemporary languages stemming from a historically-recorded protolanguage with a well-understood system of consonants. Several language groups from three language families provide data of this kind:

- Indo-Aryan languages < Vedic Sanskrit
- Romance languages < Latin
- Arabic varieties < Classical Arabic
- Tibetic languages < Old Tibetan
- Sinitic languages < Middle Chinese

Several other language groups, such as Slavic and Germanic languages, do not possess a recorded protolanguage, but have early attestations that can be used to confidently reconstruct large subsets of their protolanguages' phonological systems.

Based on these data, enriched with available information on sound-change processes in these language groups and on the phonological systems of circa 500 contemporary Eurasian languages, we will try to answer the following questions:

- Which subsets of phonological systems tend to be stable and what are the common patterns of loss in the segmental inventories?
- Are there common innovations in the structures of phonological systems?
- How important is the rôle of language contact in the development of consonant inventories?
- Does the uniformitarian assumption hold for Eurasian segmental inventories of the last two millenia?

References

Blevins, Juliette (2004), *Evolutionary Phonology: The Emergence of Sound Patterns*, Cambridge: Cambridge University Press.

Garrett, Andrew (2015), Sound change, in C. Bowern & B. Evans (eds.), *The Routledge Handbook of Historical Linguistics*, Abingdon-on-Thames: Routledge, 227–248.

Garrett, Andrew & Keith Johnson (2013), Phonetic bias in sound change, in Alan C. L. Yu (ed.), *Origins of Sound Change: Approaches to Phonologization*, Oxford: Oxford University Press, 51–97.

Labov, William (1994), *Principles of Linguistic Change: Internal Factors*, Malden: Blackwell.

Martinet, André (1952), Function, structure, and sound change, *Word* 8(1), 1–32.

Moran, Steven & Annemarie Verkerk (2018), Differential rates of change in consonant and vowel systems, in C. Cuskley, M. Flaherty, H. Little, L. McCrohon, A. Ravignani & T. Verhoef (eds.), *The Evolution of Language: Proceedings of the 12th international conference (EVOLANGXII). April 16–19, Toruń, Poland*. https://evolang.org/torun/proceedings/index.html

Wichmann, Søren & Eric W. Holman (2009), *Temporal Stability of Linguistic Typological Features*. München: Lincom Europa.

Sound change sources for consonant inventory structure: A paradox revealed

Shelece Easterday & Joan Bybee
(Laboratoire Dynamique du Langage UMR5596 & University of New Mexico)

Keywords: <consonant inventories, sound change, phonological stability, language evolution, phonological typology>

In investigating whether the distribution of sound change patterns can account for the structure of phoneme inventories, we uncover a paradox whereby basic stops and nasals are less often the outcome of phonetic processes than their high crosslinguistic frequency in consonant inventories (Maddieson 1984) would suggest.

Consonant inventory size predicts some of the phonological contrasts that occur in a language (Lindblom and Maddieson 1988). A set of 'basic' consonants occurs in inventories of all sizes, but as inventory size increases, particular 'elaborated' articulations additionally occur. One possibility is that elaborated consonants evolve from basic ones in specific phonetic contexts. We test this prediction in

a database of over 800 allophonic processes from 81 diverse languages. We find that phonetic changes frequently create the elaborations identified by Lindblom and Maddieson (1988), including voiced fricatives/affricates, post-aspiration, prenasalization, and consonants at various non-basic places of articulation.

However, we also find that for basic stops and nasals / p t k b d g m n ŋ /, more languages are reported to have processes starting from these articulations (76 languages) than producing them (55 languages). The distribution of processes involving basic voiceless stops / p t k / is particularly skewed in this respect (63 languages and 22 languages, respectively).

N lgs. with process(es) leading:	**N lgs.**	**Property changed in process**				
		voicing	**nasalization**	**place**	**manner**	**other**
from / **p t k** /	63	34	4	41	14	10
to / **p t k** /	22	10	—	4	5	5
from / **b d g** /	45	10	7	19	22	10
to / **b d g** /	30	23	2	5	2	3
from / **m n ŋ** /	54	5	4	39	6	19
to / **m n ŋ** /	35	—	9	26	2	5

Table 1. Number of languages in sample with allophonic processes leading from and to basic articulations, by property changed in the process.

We find that changes in voicing are the most common allophonic sources for basic voiceless stop and voiced stop articulations, and changes in place of articulation are the most common allophonic sources for basic nasal articulations (Table 1). New non-nasal stops at the three basic places of articulation stand out as being rarely created afresh and when they are it is often as a result of processes that are otherwise crosslinguistically rare: these include strengthening processes (specifically occlusivization), total assimilation, and degemination processes. Processes creating new instances of simple non-nasal labial, coronal and velar stops are also relatively restricted in terms of their conditioning contexts.

Here we attempt to reconcile the crosslinguistic prevalence of basic consonants with the apparent scarcity of their allophonic sources. In addition to being highly frequent crosslinguistically and having high type frequencies within languages (Gordon 2016), basic stops and nasals are often argued to be phonologically unmarked, phonetically natural, and acoustically/perceptually favored (Clements 2003, Maddieson 1996, Ohala 1983, Stevens 1989). The findings here thus raise an intriguing question: if basic consonants are favored in any or all of these ways, why are they more frequently the source of change than the outcome? One avenue for explaining the common prevalence of these consonants follows MacNeilage et al. (2000) in proposing a ‘deep evolutionary heritage’ for the three basic places of articulation, as they are dominant in pre-speech babbling and are also the result of simple mandibular oscillation, with the tongue in a central position for labials and in a forward or back position for coronals and dorsals respectively. That is, these consonants may have been part of the limited vocal repertoire present since the origins of speech in hominids.

References

Clements, George Nick (2003), Feature economy in sound systems, *Phonology* 20, 287-333.

Lindblom, Björn and Maddieson, Ian (1988), Phonetic universals in consonant systems, in L. M. Hyman, V. Fromkin, and C. N. Li (eds), (1988), *Language, speech, and mind: Studies in honor of Victoria A. Fromkin*, London: Taylor and Francis, 62-78.

Maddieson, Ian (1984), *Patterns of sounds. With a chapter contributed by Sandra Ferrari Disner*, Cambridge: Cambridge University Press.
Maddieson, Ian (1996), Gestural economy, *UCLA Working Papers in Phonetics* 94, 1-6.
MacNeilage, Peter F., Barbara L. Davis, Ashlynn Kinney and Christine L. Matyear (2000), The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development* 71 (1), 153-163.
Ohala, John J. (1983), The origin of sound patterns in vocal tract constraints, in P. F. MacNeilage (ed.), *The production of speech*, New York: Springer-Verlag, 189-216.
Stevens, Kenneth N. (1989), On the quantal nature of speech, *Journal of Phonetics* 17, 3-45.

Cohort-conditioned asymmetries in sound change within words

Christian D Brendel

<pdf>

The phonology of rare consonants: A typological approach

Leah Vandeveer
(Université Bordeaux Montaigne, CLLE-ERSSàB)

Keywords: phonology, typology, rare consonants, non-sibilant dental fricatives

Phoneme rarity has an essential role to play in understanding the variation and stability of different phonemic inventories. According to Ian Maddieson (Maddieson et al., 2013), rare or uncommon consonants can be divided into four classes: non-sibilant dental fricatives /θ, ð/, clicks, labiovelars, and pharyngeals. The three latter classes share important linguistic commonalities. Firstly, the languages that contain them within their phonemic inventories are compact in terms of their geographic distribution, visibly showing clusters in concentrated regions. These clusters are also reflected in the languages and language families that contain these consonants, with certain families being more representative of certain classes. For example, click consonants are found only within two language families in Africa, while labiovelars cluster geographically in languages of central Africa and Southeast Asia, with pharyngeals represented predominantly in Semitic languages.

However, the first class of non-sibilant dental fricatives distinguishes itself from these three other classes both geographically and genealogically. These consonants favor neither a geographic location nor a language family, found in distinctly distant languages worldwide with no region or language being representative of the class. Both /θ/ and /ð/ can be found contrasted in languages such as English with nearly two billion speakers worldwide as well as in Ngiyambaa, a moribund language of New South Wales with less than 10 speakers. This curious distribution raises questions as to the stability of these phonemes in the languages that contain and contrast them, namely that of the phonological conditions that give rise to them in some inventories and those that, potentially, discourage their appearance in others.

This research proposes an analysis into the phonemic distribution of non-sibilant dental fricatives with a primary focus into the languages that contain and contrast both the voiced /ð/ and non-voiced /θ/. This necessitates a diachronic analysis of instances of language contact that may have influenced

the evolution of the phonemic inventory of each language. We use comparative analysis by cross-referencing a 200-language sample provided in the World Atlas of Language Structures (Dryer et al., 2013) with phonemic inventories compiled in Lyon-Albuquerque Phonological Systems Database (Maddieson et al., 2014-2018). By examining phoneme co-frequencies between these phonemic inventories, we aim to understand the vast geographical and linguistic variation between languages that contain « th » consonants. In doing so, we aim to establish an explanation via underlying phonological patterns within these inventories and to construct a typology for the distribution of these rare consonants. This typology seeks to explain the conditions that are conducive to the appearance of dental fricatives /θ, ð/ in certain languages but not others.

Our research makes the claim that phonemic typology plays a crucial role in not only predicting phoneme frequency within inventories but also in explaining linguistic variation across languages. Understanding phoneme rarity and the conditions that restrict or facilitate its appearance in certain environments allows for a deeper understanding of phoneme stability and of the evolution of phonemic inventories within languages.

References

Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info.

Maddieson, Ian, 2013. Presence of Uncommon Consonants. In : Dryer, Matthew S. et Haspelmath, Martin (éd.), The *World Atlas of Language Structures Online*. Leipzig : Max Planck Institute for Evolutionary Anthropology.
p. Chapter 19. http://wals.info/chapter/19.

Maddieson I., Flavier S., Marsico E., Pellegrino F., 2014-2018. *LAPSyD: Lyon-Albuquerque Phonological Systems Databases*, Version 1.1.
http://www.lapsyd.ddl.cnrs.fr/lapsyd/

ø

Stable instability: How recurrent lenition across Australian languages reinforces existing phonological systems

Erich R. Round
(Max Planck Institute for the Science of Human History, Jena, Surrey Morphology Group, and University of Queensland)

Keywords: sound change, lenition, phoneme inventories, Australian languages, stability

The languages of Australia are known for their remarkably low level of phonological diversity (Busby 1980; Dixon 1980; Evans 1995; Hamilton 1996; Dixon 2002; Butcher 2006; Fletcher and Butcher 2014; Baker 2014), which raises the question of how this is possible and what diachronic processes underpin it across a continental expanse and two dozen language families. One source of evidence for diachrony is the synchronic phonological alternations created by sound change. We use this to reveal one kind of recurrent ***change***, whose effect is to ***stabilise*** the system around it: the lenition of phonemic stops into phonemic glides and liquids.

Stops	p	t	ʈ	ʧ	ʦ	k
> glides, liquids	w	r	ɻ	j		Ø

Empirical results in global context Gurevich's (2004) survey of 153 languages world-wide (including 8 Australian) showed that when a sound's phonemic category is changed by diachronic lenition, the result in over 90% of cases is a phonemic category which is novel. In Australia the reverse is true. Of 125 Australian languages with documentation of synchronic alternations, 78 reveal evidence of stop lenition. Among these, lenition into new phonemic categories is vanishingly rare: of over 200 individual alternations, just 5 cases reflect a sound change that plausibly created a new phonemic contrast. Otherwise, lenition merely shifts stop tokens into pre-existing phonemic categories /w,r,ɻ,j/ present in almost all Australian languages (Busby 1980; Round 2019a). These changes also to contribute to stability within morpho-phonological systems, where the alternations so created persist for long periods, as evidenced by their feeding or bleeding of, or counter-feeding or -bleeding by, other alternations due to later changes, such as nasal cluster dissimilation (McConvell 1988, Round 2019b). In rare instances, positive evidence exists of their role as attractors for later changes too (Round 2010).

Discussion Our data contains one important gap: representation of Cape York languages in the dataset of phonological alternations is poor, and in Cape York, within one subclade of one family, stops have often become fricatives diachronically. A priority for future work is to determine if this apparent low rate of synchronic alternations is linked to different historical processes playing out in Cape York versus elsewhere. Returning the other 99% of the continent's genealogical diversity, the question becomes: Why should lenition of stops be so common, and why to existing categories, not fricatives? Pieces of an explanation are now emerging from phonetic research. Regarding perception, Butcher (2018) highlights chronic otitis media, an inner ear infection widespread among contemporary indigenous Australians, which reduces perceptibility of high-frequency cues associated with frication noise. Regarding articulation, Ennever et al. (2017) and Carignan et al. (2019) use acoustic and ultrasound evidence to show that stops in Gurindji and Iwaidja have frequent, lenited allophones with a range of open articulatory targets, and very few are fricated (Ennever 2014). This pushes the question back further: why are these phonetic ***precursors*** so prevalent throughout Australia? Is the phonetic system itself in some ***inherently*** stable state? Or are there ***language-external factors*** holding it in place? The Australian facts raise deeply interesting questions for ultimate explanations of global phonological diversity.

Acknowledgements This research has been supported by grants to the author: *High-definition linguistic pasts* DE150101024 from the Australian Research Council and *Canonical Phonology* VF1/101602 from the British Council.

References

Baker, Brett. 2014. Word structure in Australian languages. *The Languages and Linguistics of Australia: A Comprehensive Guide*, ed. by Rachel Nordlinger and Harold Koch, 139–213. Berlin: Walter de Gruyter.

Busby, Peter A. 1980. The distribution of phonemes in Australian Aboriginal languages. *Papers in Australian Linguistics* 14, 73–139. Canberra: Pacific Linguistics.

Butcher, Andrew. 2006. Australian Aboriginal languages: consonant-salient phonologies and the "place-of-articulation imperative." *Speech production: models, phonetic processes and techniques*, ed. by J. M. Harrington and M. Tabain, 187–210. New York: Psychology Press.

Butcher, Andrew. 2018. The special nature of Australian phonologies: Why auditory constraints on the sound systems of human languages are not universal. *The Journal of the Acoustical Society of America* 144.1939.

Carignan, Christopher; Jason Shaw; Tonya Agostini; Robert Mailhammer; Mark Harvey and Donald Derrick. 2019. A temporal ultrasound study of lenition in Iwaidja. LSA annual conference, New York.

Dixon, R. M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
Dixon, R. M. W. 2002. *Australian languages: their nature and development*. Cambridge University Press.
Ennever, Thomas. 2014. Stop Lenition in Gurindji: An Acoustic Phonetic Study. University of Queensland b.a. (hons) thesis.
Ennever, Thomas; Felicity Meakins and Erich R. Round. 2017. A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology* 8.
Evans, Nicholas. 1995. Current issues in the phonology of Australian languages. *The handbook of phonological theory*, ed. by John A. Goldsmith, 723–761. Cambridge, MA: Blackwell.
Fletcher, Janet and Andrew Butcher. 2014. Sound patterns of Australian languages. *The Languages and Linguistics of Australia*, ed. by H. Koch and R. Nordlinger (Mouton de Gruyter, Berlin).91–138.
Gurevich, Naomi. 2004. *Lenition and contrast: the functional consequences of certain phonetically conditioned sound changes*. Psychology Press.
Hamilton, Philip James. 1996. Phonetic constraints and markedness in the phonotactics of Australian Aboriginal languages. University of Toronto ph.d. dissertation.
McConvell, Patrick. 1988. Nasal cluster dissimilation and constraints on phonological variables in Gurindji and related languages. *Aboriginal linguistics* 1:135–165.
Round, Erich R. 2010. Evolution(s) of some Tangkic morphophonology. Australian Linguistic Socitey annual conference, Canberra.
Round, Erich R. 2019a. forthc. Phonemic inventories of Australia. In *PHOIBLE online v.2*, ed. by Steven Moran.
Round, Erich R. 2019b. The typology of nasal cluster dissimilation in Australian languages. Ninth European Australianist Workshop, Leuven.

The phonological (in)stability of consonants is feature-based

Karolina Broś
(University of Warsaw)

Keywords: consonant lenition, phonological stability, features, language change

In this study, we analyse the behaviour of consonants in several European languages in order to elucidate the principles that govern their change patterns. We propose that consonant stability is not segment-based but feature-based, and that certain features tend to be more stable than others. Place features tend to be quite persistent: non-assimilatory place changes are rare, although different places show different behaviours (coronals vs. dorsals, Blevins 2009). Against this background, manner features and laryngeal specifications are particularly prone to change or loss in intralinguistic terms.

This is particularly acute in Romance languages. Voicing, spirantisation and gliding in Corsican (Oftedal 1985) or stop and affricate lenition in Sardinian (Hayes & White 2015) contribute to the blurring of categories and phonemic overlap. Similarly, in Spanish, the voicing contrast in stops was overridden by manner distinctions in the form of different degrees of aperture. Here, consonantal contrasts are retained, but the featural makeup of the sounds changed. In some dialects voiceless stops are voiced and often spirantised intervocalically, but voiced stops are spirantised and often lost in the same environment (Oftedal 1985). Moreover, it is argued that in the Canary Islands the feature [+/-

tense] started to be employed to mark the formerly voiced-voiceless contrast across a word boundary, after word-final *s* deletion (Almeida 1990).

Historical changes in Romance (Cravens 2002, Carvalho 2008) encompassing manner changes and voicing without any place of articulation shifts confirm that manner features and laryngeal specifications are particularly unstable across consonant inventories. At the same time, they provide abundant evidence for the relative stability of certain place features compared to others.

Further evidence comes from other language families in which laryngeal contrasts tend to blend and the division into true-voicing and aspiration languages is not sufficient to explain the observed sound changes. In Polish, 'laryngeal relativism' has been invoked to address the issue (Cyran 2014) given the spread of presonorant voicing throughout the country. In Swiss German, the distribution of *lenis, fortis* and aspirated stops calls for a redefinition of voicing/aspiration contrasts based on features other than VOT (Ladd & Schmid 2018).

All this leads to the conclusion that the relative instability of consonants (Gurevich 2004, Wichmann & Holman 2009) is governed by certain phonologically weak featural distinctions. We argue that it is not particular sounds that tend to dissipate in diachronic terms. Rather, particular features are either change-attracting or change-repelling. Among these, manner is especially blurry in terms of definition and acoustic/articulatory cues involved in consonant production. Often, there are very slight differences in obstruent aperture and muscle tenseness which make consonants difficult to distinguish based on auditory information (take voiced stops, non-strident fricatives and approximants of the same place of articulation). Similarly, given coarticulatory voicing across sonorants, the phonological distinction between voiced and voiceless obstruents has little to do with the phonetic reality in which several degrees of voicing can be identified between any two sounds. Such gradual effects lead to phonetic changes and recategorisation, which we see as phonological readjustments in diachronic terms.

References:

Almeida, M. (1990). *El habla de las Palmas de Gran Canaria. Niveles sociolingüísticos*. La Laguna: Centro de la Cultura Popular Canaria.

Blevins, J. (2009). Another universal bites the dust: Northwest Mekeo lacks coronal phonemes. *Oceanic Linguistics, 48/1*: 264-273.

Carvalho, J. Brandäo (2008). Western Romance. *Lenition and Fortition*, ed. by Joaquim Brandao de Carvalho, Tobias Scheer & Philippe Ségéral, 207-233. Berlin: Mouton de Gruyter.

Cravens, T. (2002). *Comparative Historical Dialectology: Italo-Romance clues to IberoRomance sound change*. Amsterdam: John Benjamins

Cyran, E. (2014). *Between Phonology and Phonetics. Polish Voicing.* Berlin: Mouton de Gruyter.

Gurevich, N. (2004). *Lenition and contrast.* Routledge.

Hayes, B. & J. White (2015). Saltation and the P-map. *Phonology 32(2):* 267-302.

Ladd, R. & Schmid, S. (2018). Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics 71*: 229-248.

Oftedal, M. (1985). *Lenition in Celtic and in Insular Spanish*. Universitetsforlaget Oslo.

Wichmann, S. & E. Holman (2009). *Assessing temporal stability for linguistic typological Features*. München: LINCOM Europa.

WORKSHOP 16

Schedule: We 11.00 – Thu 12.25 (Room 3)

Pragmatic markers and clause peripheries

Daniël Van Olmen & Jolanta Šinkūnienė
(University of Lancaster & University of Vilnius)

Keywords: pragmatic marker, left/right periphery, grammaticalization, pragmaticalization, (inter)subjectification

The last three decades have witnessed an increasing interest in linguistic units showing the link between different elements of discourse, expressing some non-propositional communicative stance by the speaker and/or managing the interaction with the addressee. In the literature, they are described as discourse connectives (e.g. Blakemore 1987), discourse particles (e.g. Aijmer 2002 and Fischer 2006), pragmatic markers (e.g. Brinton 1996, Aijmer & Simon-Vandenbergen 2006 and Van Olmen 2013), pragmatic particles (e.g. Fried & Östman 2005), connectives (e.g. Celle & Huart 2007) or discourse markers (e.g. Schiffrin 1987, Blakemore 2002 and Siepmann 2005). The extensive body of synchronic as well as diachronic and language-specific as well as comparative research into such linguistic units has revealed much about their functions, forms and development and it has also shed light on such general processes of language change as pragmaticalization, grammaticalization and (inter)subjectification (e.g. Traugott 2010, Fagard 2010, Lewis 2011, Prévost 2011, Van Olmen 2012 and Bolly & Degand 2013).

We choose the neutral and inclusive term "pragmatic marker" for our workshop and regard it as encompassing, among other things, discourse-structuring devices like the new topic introducer *o* 'but, and' in Lithuanian, indicators of stance such as *no doubt* in English and interactional tools like the attention-getter *zeg* 'say' in Dutch. The actual definition of such linguistic units and their inventory, functions, boundaries and status have been a matter of considerable debate in the literature (e.g. Degand et al. 2013 and Fedriani & Sansò 2017). The rise and the development of pragmatic markers is the subject of much discussion as well. On the one hand, their formation has been argued to be linked with the process of grammaticalization. On the other hand, it has been said to be the result of a process of pragmaticalization (e.g. Heine 2013 and Degand & Evers-Vermeul 2015 for an overview).

A related area of research, which has attracted growing attention in recent years, is the study not only of the role of the clause peripheries in the evolution of pragmatic markers but also of the impact of a left versus right – or, in other words, a clause-initial versus clause-final – position on the range of their semantic and pragmatic properties (e.g. Beeching & Detges 2014). As far as the development of pragmatic markers is concerned, for instance, the right periphery use of Italian *guarda* 'look' has essentially been claimed to be a by-product of the change from full-fledged imperative to left periphery attention-getter (see Waltereit 2002). In the same vein, the clause-final use of Dutch *zeg* has been said to postdate its clause-initial use (see Schermer 2007). The validity of such hypotheses merits further examination and so do the more general questions whether any directionality can be established for (different types of) forms with both left and right periphery uses and why it exists or not.

With respect to the functional features of pragmatic markers, it has been argued that they serve quite different purposes clause-initially than clause-finally (e.g. Degand 2011). The assumption is that markers in the latter position in particular – to which "far less attention has been paid" (Traugott 2016:

27) – correlate with intersubjective or, put differently, addressee-oriented functions. They include turn-yielding, hedging and stressing the illocution. In English, according to Traugott (2012), this assumption is indeed a tendency, though not a strict rule, and there is evidence from many other languages with different genealogical, areal and typological profiles for intersubjective right periphery pragmatic markers (e.g. Chor et al. 2016, Beeching 2016 and Rhee 2016).

More research is needed, though. For Japanese, for example, the case has been made that it is clause-initial position that is best for conveying intersubjectivity (e.g. Onodera 2007). The literature has also focused mainly on European and East Asian languages (e.g. Beeching & Detges 2014 cover Chinese, English, French, Italian, Japanese and Korean). Furthermore, to our knowledge, only a small number of studies have tried to map the full range of forms and functions in one of the peripheries (e.g. Van der Wouden & Foolen 2015 on the clause-final particles in Dutch) and contrast it with the range in the other periphery. This kind of comprehensive comparison could also prove useful for uncovering the (dis)similarities between languages in the (types of) textual, subjective and intersubjective meanings that they tend to express in their peripheries (e.g. are certain East Asian languages more concerned with "attitudinal" intersubjectivity than particular European languages in the right periphery?; see Ghesquière et al. 2012).

In short, the goal of this workshop is to revisit the relationship between the clause-initial or clause-final position of pragmatic markers and their functions, forms and evolutions. We seek to shed new light on, among other things, the meanings associated with the left and right peripheries within a language and across languages and the historical developments into either or both peripheries that pragmatic markers can undergo. As such, the workshop also aims to contribute to the ongoing debates about such phenomena as grammaticalization, pragmaticalization, (inter)subjectivity and (inter)subjectification. Specific questions that we want to address include:

- which meanings do pragmatic markers express in the right periphery and which meanings do they convey in the left periphery?
- what, if any, are the functional differences between pragmatic markers in the left and the right periphery?
- what, if any, are the functional (dis)similarities between left/right periphery pragmatic markers in different languages or varieties?
- which source constructions end up in the left periphery, in the right periphery or in both and are there any cross-linguistic tendencies?
- does any diachronic directionality exist for pragmatic markers that can occur in both clause-initial and clause-final position?

References

Aijmer, Karin (2002), *English Discourse Particles: Evidence from a Corpus*, Amsterdam: John Benjamins.

Aijmer, Karin and Anne-Marie Simon-Vandenbergen (eds), (2006), *Pragmatic Markers in Contrast*, Oxford: Elsevier.

Beeching, Kate (2016), *Alors*/*Donc*/*Then* at the Right Periphery: Seeking Confirmation of an Inference, *Journal of Historical Pragmatics* 17(2), 208–230.

Beeching, Kate and Ulrich Detges (eds), (2014), *Discourse Functions at the Left and Right Periphery: Cross-linguistic Investigations of Language Use and Language Change*, Leiden: Brill.

Blakemore, Diane (1987), *Semantic Constraints on Relevance*, Oxford: Blackwell.

Blakemore, Diane (2002), *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*, Cambridge: Cambridge University Press.

Bolly, Catherine and Liesbet Degand (2013), Have you Seen what I Mean? From Verbal Constructions to Discourse Markers, *Journal of Historical Pragmatics* 14(2), 210–235.

Brinton, Laurel (1996), *Pragmatic Markers in English: Grammaticalization and Discourse Function*, Berlin: Mouton de Gruyter.

Celle, Agnès and Ruth Huart (eds), (2007), *Connectives as Discourse Landmarks*, Amsterdam: John Benjamins.

Chor, Winnie, Foong Ha Yap and Tak-Sum Wong (2016), Chinese Interrogative Particles as Talk Coordinators at the Right Periphery. *Journal of Historical Linguistics* 17(2), 178–207.

Degand, Liesbeth (2011), Connectieven in de Rechtperiferie: Een Contrastieve Analyse van *Dus* en *Donc* in Gesproken Taal, *Nederlandse Taalkunde* 16(3), 333–348.

Degand, Liesbeth, Bert Cornillie and Paola Pietrandrea (eds), (2013), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam: John Benjamins.

Degand, Liesbeth and Jacqueline Evers-Vermeul (2015), Grammaticalization or Pragmaticalization of Discourse Markers? More than a Terminological Issue, *Journal of Historical Pragmatics* 16(1), 59–85.

Fagard, Benjamin (2010), *É Vida, Olha...*: Imperatives as Discourse Markers and Grammaticalization Paths in Romance: A Diachronic Corpus Study, *Languages in Contrast* 10(2), 245–267.

Ghesquière, Lobke, Lieselotte Brems and Freek Van de Velde (2012), Intersubjectivity and Intersubjectification: Typology and Operationalization, *English Text Construction* 5(1): 128–152.

Fedriani, Chiara and Andrea Sansò (eds), (2017), *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*, Amsterdam: John Benjamins.

Fischer, Kerstin (ed), (2006), *Approaches to Discourse Particles*, Oxford: Elsevier.

Fried, Mirjam and Jan-Ola Östman (2005), Construction Grammar and Spoken Interaction: A Case of Pragmatic Particles, *Journal of Pragmatics* 37(11), 1752–1778.

Heine, Bernd (2013), On Discourse Markers: Grammaticalization, Pragmaticalization, or Something Else? *Linguistics* 51(6), 1205–1247.

Lewis, Diana (2011), A Discourse-Constructional Approach to the Emergence of Discourse Markers in English, *Linguistics* 49(2), 415–443.

Onodera, Noriko (2007), Interplay of (Inter)Subjectivity and Social Norm, *Journal of Historical Pragmatics* 8(2), 239–267.

Prévost, Sophie (2011), *A Propos* from Verbal Complement to Discourse Marker: A Case of Grammaticalization? *Linguistics* 49(2), 391–413.

Rhee, Seongha (2016), LP and RP in the Development of Discourse Markers from "What" in Korean, *Journal of Historical Pragmatics* 17(2), 255–281.

Schermer, Ina (2007), Wat Doet dat Woord daar, zeg? Over het Tussenwerpsel *Zeg*, *Voortgang* 25, 373–385.

Schiffrin, Deborah (1987), *Discourse Markers*, Cambridge: Cambridge University Press.

Siepmann, Dirk (2005), *Discourse Markers across Languages: A Contrastive Study of Second level Discourse Markers in Native and Non-native Text with Implications for General and Pedagogic Lexicography*, London: Routledge.

Traugott, Elizabeth C. (2010), (Inter)Subjectivity and (Inter)Subjectification: A Reassessment, in Kristin Davidse, Lieven Vandelanotte and Hubert Cuykens (eds), (2010), *Subjectification, Intersubjectification and Grammaticalization*, Berlin: Mouton de Gruyter, 29–74.

Traugott, Elizabeth C. (2012), Intersubjectification and Clause Periphery, *English Text Construction* 5(1), 7–28.

Traugott, Elizabeth C. (2016), On the Rise of Types of Clause-final Pragmatic Markers in English, *Journal of Historical Pragmatics* 17(1), 26–54.

Traugott, Elizabeth C. and Richard B. Dasher (2002), *Regularity in Semantic Change*, Cambridge: Cambridge University Press.

Van der Wouden, Ton and Ad Foolen (2015), Dutch Particles in the Right Periphery, in Sylvie Hancil, Alexander Haselow and Margje Post (eds), (2015), *Final Particles*, Berlin: De Gruyter, 221–247.

Van Olmen, Daniel (2012), The Imperative of Intentional Visual Perception as a Pragmatic Marker: A Contrastive Study of Dutch, English and Romance, in Pieter Lauwers, Gudrun Vanderbauwhede and Stijn Verleyen (eds), (2012), *Pragmatic Markers and Pragmaticalization*, Amsterdam: John Benjamins, 95–116.

Van Olmen, Daniel (2013), The Imperative of Say as a Pragmatic Marker in English and Dutch, *Journal of Germanic Linguistics* 23(3), 247–287.

Waltereit, Richard (2002), Imperatives, Interruption in Conversation and the Rise of Discourse Markers: A Study of Italian *Guarda*, *Linguistics* 40(5), 987–1010.

ꕥ

Discourse segmentation, boundaries and discourse marker use: A corpus-based study of spoken French

Liesbeth Degand & Ludivine Crible
(University of Louvain)

Key words: discourse markers, discourse segmentation, spoken language, prosodic boundary, clause periphery

The study of Discourse Marker (DM) use goes hand in hand with the influence their syntagmatic position may have on their function. Specifically, subjective and intersubjective meanings of DMs tend to be respectively associated with the left and right periphery (e.g. Beeching & Detges 2014), although other studies have nuanced these patterns (Haselow 2012, Traugott 2012). An aspect that has not received enough attention in this debate is the importance of defining concretely the unit of analysis, of which we study the peripheries (see, e.g. Degand 2014, Salameh et al. 2018). In other words, are we looking at the peripheries of a (syntactic) clause, of an intonation unit, of a (pragmatic) illocution, of a turn, or else? This leads us to formulating our research question as follows: To what extent does the unit of linguistic analysis influence the type and function of DMs on their periphery? Thus, it is necessary to operationally define this unit of analysis before comparing the functions of DMs at the left and right periphery.

In this presentation, we will focus on the distribution of all DMs (N = 1780) extracted from the LOCAS-F corpus, a multi-genre corpus of spoken French segmented into several types of units (Degand, Martin, Simon 2014): (i) clauses, defined as a verbal or averbal head and its dependencies; (ii) intonation units, defined as stretches of discourse marked by a prosodic boundary and an intonation contour; (iii) "basic discourse units" or BDUs, characterized by a coinciding syntactic and prosodic boundary. Table 1 presents the topological distribution of the DMs according to the units of analysis at stake.

	Initial	Medial	Final	Isolated
BDU	697 (39%)	833 (47%)	163 (9%)	87 (5%)
Clause	1345 (83%)	88 (5%)	179 (11%)	/
Intonation Unit	715 (40%)	797 (45%)	181 (10%)	87 (5%)

Table 1: Position of DMs in Locas-F

Strikingly, the topological distribution is strongly influenced by the type of unit of segmentation. Moreover, a preliminary functional analysis in terms of ideational, rhetorical, sequential and interpersonal domains of use (Crible 2018, Crible and Degand in press) indicates that there is a stronger association between function and position at the clausal level than at the other levels. This seems to indicate that clausal peripheries are linked more strongly to specific discourse functions. For instance, initial position strongly favors ideational and sequential uses, while final clause position attracts interpersonal DM use. Such topological regularities do not occur in BDUs.

The presentation will develop in more details the form and function of DMs in peripheral position, and how these vary according to the unit of analysis.

References

Beeching, Kate, and Ulrich Detges, eds. (2014), *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Studies in Pragmatics, Volume 12. Leiden ; Boston: Brill.

Crible, Ludivine (2018), *Discourse Markers and (Dis)Fluency. Forms and Functions across Languages and Registers*. John Benjamins. https://doi.org/10.1075/pbns.286.

Crible, Ludivine and Liesbeth Degand (in press), Reliability vs. Granularity in Discourse Annotation: What Is the Trade-Off? *Corpus Linguistics and Linguistic Theory*. https://doi.org/10.1515/cllt-2016-0046.

Degand, Liesbeth (2014), 'So Very Fast Very Fast Then' Discourse Markers at Left and Right Periphery in Spoken French." In *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*, edited by Kate Beeching and Ulrich Detges, 151–78. Studies in Pragmatics, Volume 12. Leiden ; Boston: Brill.

Degand, Liesbeth, Laurence Martin, and Anne Catherine Simon (2014), "LOCAS-F : Un Corpus Oral Multigenres Annoté." In *CMLF 2014 – 4ème Congrès Mondial de Linguistique Française*, 2613–26. Berlin. https://doi.org/10.1051/shsconf/20140801211.

Haselow, Alexander. (2012). "Subjectivity, Intersubjectivity and the Negotiation of Common Ground in Spoken Discourse: Final Particles in English." *Language & Communication* 32 (3): 182–204. https://doi.org/10.1016/j.langcom.2012.04.008.

Salameh Jiménez, Shima, Estélles, María & Pons Bordería, Salvador (2018). Beyond the notion of periphery : An account of polyfunctional discourse markers within the Val.Es.Co. model of discourse segmentation. In: Kate Beeching, Chiara Gezzin & Piera Molinelli (eds.). *Positioning the Self and Others. Linguistic Perspectives.* [Pragmatics and Beyond New Series 292]. Amsterdam: John Benjamins, 105-125.

Traugott, Elizabeth Closs (2012). Intersubjectification and clause periphery, In Lieselotte Brems, Lobke Ghesquière, and Freek Van de Velde, eds., Intersections of Intersubjectivity, special issue of *English Text Construction* 5: 7-28.

Pragmatic markers in English from LP to RP: The role of discourse information structuring

Diana M. Lewis
(University of Aix Marseille and Laboratoire Parole et Langage)

Keywords: pragmatic markers, information structure, periphery, grammaticalization, English

Recent work on pragmatic markers (PMs) in English and other European languages has paid attention to the significance of the position of the marker at the left and right peripheries (LP and RP). Much of this work has focused on the semantics of the markers, debating whether the two positions show functional specialization, such as interpersonal vs. textual meanings (e.g. Beeching and Detges 2014; Hancil et al 2015; Haselow 2012). This corpus-based study addresses PMs at LP and RP in English, and argues that these positions may result in PMs acquiring primarily discourse information-structuring roles.

A number of English PMs can occur both before and after their host (e.g. *actually, after all*, *anyway*, *at least*, *in fact*, *of course*, *then*, etc.). The differences include: (i) the RP variant tends to emerge later than the LP one; (ii) for several of these forms the RP marker host is often a fragment or sub-clausal consituent, while at LP a full clause is normally found; (iii) in the RP usage the host unit often has a tight discourse bond with the previous unit that is not necessarily the case in the LP usage. But the positions are not seen to correlate with clear differences in meaning (Traugott 2012). This paper argues that there is, however, evidence of a functional split along discourse information structure lines, whereby at LP the marker acquires a presentational function while at RP its host tends to be informationally subordinate to (often a comment on) a previous idea, accompanying either a subordinate structure or a hypotactic discourse structure. The relevant unit for position is taken to be the unit over which the PM has scope, which may be a sub-clausal unit.

A case study of PM *of course* (i.e. excluding the polarity use) is presented, tracing its recent historical development from the VP through epistemic adverb and resultative connective to its present-day usage in hedging, concession and discourse-management at both LP and RP. The historical analysis is based on corpora of informal English (letters, drama, diaries, journals, reported speech) of the C17th-C20th. It is argued that the recent evolution of *of course* towards discourse information structuring functions (Aijmer 2013) involves positional specialization along with the fading of its epistemic meaning, and can be seen as a further grammaticalization of the PM. A comparison is drawn with *obviously*, which has a similar distribution to *of course* and which has trebled in frequency over thirty years, leap-frogging *of course*, but which is arguably less grammaticalized and may be replacing epistemic *of course* in some contexts. The comparison is based on data from British English conversation of the 1980s (BNC 1994) and the 2010s (BNC 2014). More generally, it is suggested that the rise in RP for PMs and epistemic sentence adverbs may be a symptom of a wider reorganization of sentence-adverbial syntax in English.

References

Aijmer, Karin (2013), Analyzing modal adverbs as modal particles and discourse markers, in Liesbeth Degand, Bert Cornillie and Paola Pietrandrea (eds) (2013), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam: John Benjamins, 89-106.

Beeching, Kate and Ulrich Detges (2014), Introduction, in Kate Beeching and Ulrich Detges (eds) (2014), *Discourse Functions at the Left and Right Periphery*, Leiden: Brill, 1-23.

Hancil, Sylvie, Margje Post and Alexander Haselow (2015), Introduction: Final particles from a typological perspective, in Sylvie Hancil, Alexander Haselow and M. Post (eds) (2015), *Final Particles*, Berlin: Walter de Gruyter, 3-35.

Haselow, Alexander (2012), Discourse organization and the rise of final *then* in the history of English, in Irén Hegedüs and Alexandra Fodor (eds) (2012), *English Historical Linguistics 2010*, Amsterdam: John Benjamins, 153-175.

Traugott, Elizabeth C. (2012). Intersubjectification and clause periphery, *English Text Construction* 5(1), 7-28.

Functional asymmetry and left to right movement: Speaking of peripheries

Yinchun Bai
(Albert-Ludwigs-Universität Freiburg and Universiteit Antwerpen)

Keywords: pragmatic marker, periphery, functional asymmetry, (inter)subjectification, grammaticalization

It has been noted in numerous works that the (meta-)linguistic functions of elements appearing at the LP and RP slots feature asymmetry (e.g. Andersen 2015, Beeching and Detges 2014, Ferarra 1997, Haselow 2011, Higashiizumi and Onodera 2013, and Traugott 2015). This paper confirms this observation with results from a corpus-based study of the "*speaking of X* (SPOX)" pragmatic marker, e.g. (1)-(3).

(1) *" I'll bet you a can of corn against a bite of canned pie that he's from New York, " Jack Bosworth observed. " Speaking of pie, " Frank cut in, " there's a little restaurant on Beekman street where they serve hot pies at noon for a dime.*[…] (1911, FIC, BoyScoutsInPhilippines)
(2) *And in the finale Tuesday, during the finale, round one – speaking of songs -- will be the contestant favorite.* (2009, SPOK, CNN Newsroom)
(3) *It cuts spending in the places we don't want it. It doesn't cut spending in the entitlement programs, speaking of other powerful interest groups.* (2013, SPOK, PBS)

Using the spoken data from the Corpus of Contemporary American English (COCA) from 1990 to 2015, the study shows that SPOX features functional asymmetry when used at different peripheral positions. The left periphery (LP) has been observed with turn-taking and attention-getting, linking, grounding, and topicalization functions, which feature a closer interactional bond to the previous discourse, while at the medial periphery (MP) and the right periphery (RP), SPOX is more often used for proposition specification, which has a stronger connection to the introduced utterance instead. Furthermore, the pragmatic content at the MP and the RP positions also feature stronger intersubjectivity than those at the LP position.

Besides the functional asymmetry of SPOX at different peripheral positions, this paper also reveals that there is a slow movement from the LP to the MP and the RP during the recent decades. It thus addresses the issue of left to right movement raised by Beeching and Detges (2014: 1): i.e. when once syntactically or propositionally embedded lexical elements become pragmatic markers, they tend to start at LP, serving textual or subjective functions, and then gravitate towards the RP, where they play a more intersubjective or modalising role. This trend of development conforms to the hypothesis arising from Traugott (1982) and Traugott and Dasher (2002) that when semantic change takes place in the process of grammaticalization, it is more likely to move along the clines of "less personal -> more personal" and "propositional –> textual -> expressive/interpersonal" than the reverse. The findings in this paper thus add to the body of work on subjectification and intersubjectification and the mechanism of semantic change in general.

References

Andersen, Gisle (2015), Discourse topic markers from a cross-linguistic and variational perspective, Paper presented at the 14th conference of the International Pragmatics Association (IPrA 14), Antwerp, 26-31 July.

Beeching, Kate and Ulrich Detges (2014), *Discourse Functions at the Left and Right Periphery:*

Crosslinguistic Investigations of Language Use and Language Change, Leiden/Boston: Brill.
Ferrara, Kathleen W. (1997), Form and function of the discourse marker *anyway*: implications for discourse analysis, *Linguistics* 35, 343-378.
Haselow, Alexander (2011), Discourse marker and modal particle: the functions of utterance-final *then* in spoken English, *Journal of Pragmatics* 43, 3603-3623.
Higashiizumi, Yuko and Noriko O. Onodera (2013). Cross-Linguistic approach to form-function-periphery (LP and RP) mapping: With a special focus on 'exchange structure' and 'action structure', Paper presented at the 13th conference of the International Pragmatics Association (IPrA 13), New Delhi, 8-13 September.
Taugott, Elizabeth Closs (1982), From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization, in Winfred P. Lehmann and Yakov Malkiel (eds), *Perspectives on Historical Linguistics*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 245-271.
Traugott, Elizabeth Closs (2015), Investigating "periphery" from a functionalist perspective, *Linguistics Vanguard*, 1(1), 119-130.
Traugott, Elizabeth Closs and Richard B. Dasher (2002), *Regularity in Semantic Change*, Cambridge: Cambridge Univercity Press.

Another 'look!' (to the left and to the right): The Latvian particle *lūk* in parliamentary discourse

Nicole Nau
(Adam Mickiewicz University in Poznan)

Keywords: discourse particle; presentative particle; perception verbs; stance; Latvian

The Latvian particle *lūk* is derived from the imperative of the verb *lūkot* 'try; look'. It is used in formal and informal varieties of spoken and written discourse as a presentative particle and marker of stance. In this paper I will analyze its use and functions as attested in the corpus SAEIMA, compiled from transcripts of sittings of the Latvian parliament. Prosodic features will be examined with selected examples from the recordings of the sittings. In SAEIMA, *lūk* is much more frequent than in the balanced corpus LVK2018 (298.2 vs. 59.92 occurrences per million). It is used for presenting, evaluating and stressing facts within an argumentation. A main goal of this paper is to establish in which way different functions are bound to different positions of the particle and whether *lūk* is developing a "right-margin" use.

Lūk is typically found within a longer turn of one speaker, rarely starting or closing a turn. As a presentative particle (cf. Petit 2010; Porhiel 2012) it may point ahead to the following stretch of speech (like French *voici*), or backwards to what has been stated before (like French *voilà*). In the first case its position is either at the left margin of a clause or non-clausal unit (1), or parenthetically within a clause (2). When pointing backwards, the particle most often forms an intonation unit of its own (3), but may also occur at the beginning of a clause.

(1) ***Lūk***, *piemērs:*
PTC example.NOM.SG '**Here is** an example:'

(2) *Es* *gribētu,* ***lūk,*** *ko.*
1SG.NOM want.IRR PTC what.ACC '**Here is** what I would like.'

(3) *Ja jūs maldāties, man nav jāpiekrīt jums.* ***Lūk****!*
'If you are wrong, I don't have to agree with you. **LŪK**!' ('That's how it is!')

Another function of the particle is within represented speech, mostly with a negative stance towards the presented statement. In this function its position is at the beginning of a clause (4) or as parenthesis.

(4) *Tāpat šajā rakstā viņš teica, ka,* ***lūk****, Latvijas krievi ir tie labākie krievi*
'He also says in this paper that **LŪK** Latvia's Russians are the better Russians'

Lūk seems to be always speaker-oriented. Where it occurs, speakers are not seeking agreement nor invite a response. Maybe the lack of such intersubjective meanings prevents *lūk* to fuse with the previous utterance and become an element of the right margin, or acquire a turn-yielding function.
This paper will contribute to the cross-linguistic study of functions of pragmatic markers in different positions (Traugott 2012; Beeching & Detges, eds. 2014) and to that of particles originating in forms of perception verbs (Fagard 2010; Aijmer & Elgemark 2013). To my knowledge, the Latvian data have not been described before (but see Chojnicka 2012 for other markers of stance in Latvian parliament discourse).

References

Aijmer, Karin & Anna Elgemark (2013), The pragmatic markers *look* and *listen* in a cross-linguistic perspective, in N.-L. Johannesson, Gunnel Melchers, and Beyza Björkman (eds.), (2013), *Of butterflies and birds, of dialects and genres. Essays in honour of Philip Shaw*, Stockholm: University of Stockholm, 333-348.

Beeching, Kate & Ulrich Detges (eds.) (2014), *Discourse functions at the left and right periphery: Cross-linguistic investigations of language use and language change*. Leiden: Brill.

Chojnicka, Joanna (2012), *Linguistic markers of stance in Latvian parliamentary debates*. Saarbrücken: Akademieverlag.

Fagard, Benjamin (2010), É vida, olha...: Imperatives as discourse markers and grammaticalization paths in Romance: A diachronic corpus study, *Languages in Contrast* 10(2): 245-267.

LVK2018 = Balanced corpus of contemporary Latvian, 10 million words. Available at www.korpuss.lv.

Petit, Daniel (2010), On presentative particles in the Baltic languages, in N. Nau, and N. Ostrowski (eds.), (2010), *Particles and connectives in Baltic*, Vilnius: Vilnius University, 151-170.

Porhiel, Sylvie (2012), The presentative *voici/voilà* – Towards a pragmatic definition. *Journal of Pragmatics* 44 (2012) 435–452.

SAEIMA = Corpus of short hand transcripts of Latvian Parliament sittings, 21 million words. Available at www.korpuss.lv.

Traugott, Elizabeth Closs (2012), Intersubjectification and clause periphery. *English Text Construction* 5(1): 7-28.

A cross-linguistic study of pragmatic markers of visual perception at the clause peripheries

Vittorio Tantucci & Daniel Van Olmen
(Lancaster University)

Keywords: Chinese, Italian, (inter)subjectivity, West Germanic, '(you) see/look'

The imperative of intentional visual perception has been examined as a pragmatic marker in some detail for various European languages, especially in relation to 'listen!' (e.g. Waltereit 2002, Fagard 2010 and Aijmer and Elgemark 2013). English in (1) and Italian in (2) can serve as examples.

(1) **Look**, it was obviously suicide – you said yourself she was unhealthily obsessed with the damn thing. (Aijmer and Elgemark 2013: 340)
(2) *No **guarda** allora **guarda** se se io potessi te lo te lo regalerei volentieri io.* (Fagard 2010: 246)
'No look now look if I could I would gladly give it to you.'

Less attention has been paid to 'look!' versus other second person forms of visual perception verbs like 'see', which conveys an unintentional experience. The research has also not explored in detail how such inherently addressee-oriented pragmatic markers may contribute to the debate about clause peripheries and (inter)subjectivity (e.g. Degand 2011, Beeching and Detges 2014 and Traugott 2016) and particularly the tendency for pragmatic markers to express more subjective meanings at the left periphery and more intersubjective ones at the right periphery.

The present paper seeks to fill these gaps by contrasting the use of 'look!', 'see!' and 'you see' in comparable corpora of speech in English, Italian and Dutch and in Afrikaans and Chinese. The data come from the International Corpus of English – Great Britain (Survey of English Usage 2006), the Perugia Corpus (Spina 2014), the Corpus Spoken Dutch (Nederlandse Taalunie 2004), the Corpus Spoken Afrikaans (North-West University 2018) and the Peking University Corpus (Yu et al. 2002). All pragmatic marker cases of the verbs 'look' and 'see' are analyzed in terms of form, position, frequency and function.

Our initial results reveal some interesting facts about the languages under investigation. Afrikaans *kyk* 'look!', for instance, is found to display the same functions as its better-studied Dutch and English equivalents (e.g. taking/maintaining turns, introducing reported speech) but to be more frequent. The data also show that Chinese possesses no clause-peripheral uses of 见 *jiàn* 'see' as a pragmatic marker while left-peripheral 你看 *nǐkàn* 'you look' hinges on soliciting agreement with a distinctive anaphoric reference. We can still conclude, however, that, cross-linguistically, 'look!' exhibits a stronger tendency to occur in the left periphery whereas '(you) see' can be attested in the left as well as the right periphery. The latter's flexibility is argued to be at least partially due to its use – related to the semantic extension of 'see' to 'understand' (e.g. Sweetser 1993) – as a way to check comprehension at the end of a clause. Perhaps most importantly, our initial findings also challenge the supposedly typical connections between the two peripheries and (inter)subjectivity. More specifically, the pragmatic markers under discussion generally fulfill addressee-involving functions like attention-getting in left-peripheral position. Moreover, left-peripheral forms are often employed to pre-emptively tackle the addressee's reaction to potential face-threats.

References

Aijmer, Karin and Anna Elgemark (2013), The Pragmatic Markers *Look* and *Listen* in a Cross-

linguistic Perspective, in Nils-Lennart Johannesson, Gunnel Melchers and Beyza Björkman (eds), *Of Butterflies and Birds, of Dialects and Genres: Essays in Honour of Philip Shaw*, Stockholm: Stockholm University, 333–348.

Beeching, Kate and Ulrich Detges (eds) (2014), *Discourse Functions at the Left and Right Periphery: Cross-linguistic Investigations of Language Use and Language Change*, Leiden: Brill.

Degand, Liebeth (2011), Connectieven in de Rechtperiferie: Een Contrastieve Analyse van *Dus* en *Donc* in Gesproken Taal, *Nederlandse Taalkunde* 16(3), 333–348.

Fagard, Benjamin (2010), *É Vida, Olha...*: Imperatives as Discourse Markers and Grammaticalization Paths in Romance: A Diachronic Corpus Study, *Languages in Contrast* 10(2), 245–267.

Nederlandse Taalunie (2004), *Corpus Gesproken Nederlands*, release 1.0, The Hague: Nederlandse Taalunie.

North-West University (2018), *Korpus Gesproke Afrikaans 2003-2005*. Vanderbijlpark: North-West University.

Spina Stefania (2014), Il Perugia Corpus, una Risorsa di Riferimento per l'Italiano: Composizione, Annotazione e Valutazione, in Roberto Basili, Alessandro Lenci and Bernardo Magnini (eds), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014*, Pisa: Pisa University Press, 354–359.

Survey of English Usage (2006), *International Corpus of English: The British Component*, release 2, London: Survey of English Usage.

Sweetser, Eve. 1993. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*, Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs (2016), On the Rise of Types of Clause-final Pragmatic Markers in English, *Journal of Historical Pragmatics* 17(1), 26–54.

Waltereit, Richard (2002), Imperatives, Interruption in Conversation and the Rise of Discourse Markers: A Study of Italian *Guarda*, *Linguistics* 40(5), 987–1010.

Yu, Shiwen, Huiming Duan, Xuefeng Zhu and Bin Sun (2002), The Basic Overview of the Development of the Corpus of Modern Chinese of Peking University. *Journal of Computational Chinese* 16(5), 49–64.

A cross-linguistic look at the right periphery: Utterance-final adverbials in English, Spanish and Lithuanian

Anna Ruskan & Marta Carretero
(Vilnius University & Complutense University of Madrid)

Keywords: left periphery, right periphery, adverbials, discourse markers, cross-linguistic

Functional diversification of the right periphery has been revealed in a number of studies on final particles (Haselow 2012, 2013, 2015; Hancil et al. 2015; Sato 2017), sentence adverbials (Ureña Gómez-Moreno 2015), comment clauses (Aijmer 1997) and discourse markers (Izutsu, Izutsu 2013; Degand 2014) in English, German, Dutch, French, Norwegian, Finnish and other languages. The formal and functional features of various elements in the right periphery have also been considered from a cross-linguistic perspective: for example, English *then* is contrasted with its Norwegian counterparts *da* and *altså* in Fretheim (2015) and with its French counterparts *donc* and *alors* in

Beeching (2016). The present study aims to compare the formal and functional features of utterance-final adverbials in English, Spanish and Lithuanian:

(1) Mr. Russell-Brown: *No, I lean forward all the time. I'm sorry!*
Ms. Leila J. Gosselin: *I apologize for assuming* ***then***. (Strathy)

(2) Bárbara: *¿Algún problema en el despacho?*
Juan: *Los de siempre.*
Bárbara: *¿Todo bien,* ***entonces****?*
Juan: *Sí, bien.* (CREA)
Bárbara: 'Any problem in the office?'
Juan: 'The usual.'
Bárbara: 'Everything all right, **then**?'
Juan: 'Yes, all right.'

(3) A: *Rašiau, rašiau viską.*
B: *Ai, nu tai gerai* ***tada***.
A: 'I was writing down everything.' (CCLL)
B: 'Oh, it is fine **then**.'

The study will compare the frequency and functional profile of adverbials of different semantic categories (temporal, e.g. *then*; evidential, e.g. *obviously*, *apparently*; actuality and reality, e.g. *actually*, *really*) occurring in utterance final position across the three languages, and thus shed light on cross-linguistic and language-specific functions of the right periphery. The study will approach the contribution of final adverbials to the utterances that they modify on the illocutionary, propositional, subjective and interpersonal levels (Haselow 2012: 189), and will also compare this contribution to that of the same markers in the left periphery. The data will be drawn from the spoken and fiction sub-corpora of the Corpus of Canadian English (Strathy), the Corpus de Referencia del Español Actual (CREA), the Corpus del Español del Siglo XXI (Corpes XXI) and the Corpus of the Contemporary Lithuanian Language (CCLL).

Functional distribution of the adverbials in clause peripheries across the three languages supports the view that elements in the right periphery express "a greater degree of response-relevant intersubjectivity" and in the left periphery they foreground "the speaker's subjective stance" (Sato 2017: 95). In the right periphery the adverbials under study tend to elicit the hearer's response and activate common knowledge, whereas in the left periphery they may develop authorial argumentation. Although the cross-linguistic evidence points to a number of distinct functional features of the adverbials analysed in clause peripheries, instances of their functional overlap have also been observed, which may confirm the absence of "a one-to-one correspondence between form and function" (Fischer, Heide 2018: 531). The adverbials in utterance final position across the three languages have been found to differ in frequency, types and range of (inter)subjective meanings as well as in the degree of semantic bleaching of their core meanings (temporal, evidential, reality and actuality).

References

Aijmer, Karin (1997), *I think* – an English modal particle, in T. Swan & O. J. Westvik (eds), (1997), *Modality in the Germanic Languages: Historical and Comparative Perspectives*, Berlin: Mouton de Gruyter, 1-47.

Beeching, Kate (2016), *Alors*/*donc*/*then* at the right periphery: Seeking confirmation of an inference, *Journal of Historical Pragmatics* 17 (2), 208-230.

Degand, Liesbeth (2014), 'So very fast then' discourse markers at left and right periphery in spoken French, in K. Beeching & U. Detges (eds), (2014), *Discourse Functions at the Left and Right Periphery: Cross-linguistic Investigations of Language Use and Language Change*, Brill: Leiden, 151-187.

Fischer, Kerstin and Maike Heiden (2018), Inferential processes in English and the question whether English has modal particles, *Open Linguistics*, 4(1), 509-536.

Fretheim, Thorstein (2015), A relevance-theoretic perspective on the Norwegian utterance-final particles *da* and *altså* compared to their English counterpart *then*, in S. Hancil, A. Haselow & M. Post (eds), (2015), *Final Particles*, Berlin/Boston: Walter de Gruyter. 249-284.

Hancil, Sylvie, Margje Post and Alexander Haselow (2015), Introduction: Final particles from a typological perspective. in S. Hancil, A. Haselow & M. Post (eds), (2015), *Final Particles*, Berlin/Boston: Walter de Gruyter. 3-36.

Haselow, Alexander (2012), Subjectivity, intersubjectivity and the negotiation of common ground in spoken discourse: Final particles in English, *Language and Communication* 32, 182-204.

Haselow, Alexander (2013), Arguing for a wide conception of grammar: The case of final particles in spoken discourse, *Folia Linguistica* 47, 375-424.

Haselow, Alexander (2015), Left vs. right periphery in grammaticalization: The case of *anyway*, in A. D.M. Smith, G. Trousdale & R. Waltereit (eds), (2015), *New Directions in Grammaticalization Research*, Amsterdam/Philadelphia: Benjamins. 157-186.

Izutsu, Katsunobu and Mitsuko Narita Izutsu (2013), From discourse markers to modal/final particles: What the position reveals about the continuum, in L. Degand, B. Cornillie & P. Pietrandrea, (eds), (2013), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam/Philadelphia: Benjamins. 217-235.

Sato, Shie (2017), On establishing *I think* as a final particle in interactions: Some comparisons with sentence-final particles in Japanese, *Journal of Pragmatics* 110, 83-93.

Ureña Gómez-Moreno, Pedro (2015), Sentence-final adverbials: Recurrent types and usage. in S. Hancil, A. Haselow & M. Post (eds), (2015), *Final Particles*, Berlin/Boston: Walter de Gruyter. 39-54.

Mapping the left periphery of the Dutch clause through discourse particles

Ton van der Wouden & Ad Foolen
(Meertens Instituut & Radboud Universiteit)

Keywords: Dutch, discourse particle, left periphery

Dutch is a SOV language with V2 in the main clause. The first sentence position, right before the inflected verb, can be filled flexibly by the subject or any one of the other sentence constituents ('topicalisation'). In addition, there is one extra position in front of the first position (Haeseryn et al. 1997). This left peripheral position ("aanloop") can remain empty or be filled by a constituent which is anaphorically referred to in the main clause ('left-dislocation'), cf. (1):

(1) *Z'n broer, die heeft altijd al graag een glas bier gedronken* (Haeseryn et al. 1229)
His brother, that has always already please a glass beer drunk
'His brother, he has always liked to drink a glass of beer'

Our focus will be on the linear position of discourse particles in relation to the two 'referential' positions distinguished above. At least 3 positions can be distinguished: before a left-dislocated constituent, following it, and between first sentence position and the verb (cf. 2-4, discourse particles are underlined):

(2) *Kortom, [die man], [die] mag ik niet*
(3) *[Die man], kortom, [die] mag ik niet*
in short, I don't like that man
(4) *[Op dat moment] echter had ik geen tijd*
At that moment however had I no time

If the first sentence position is not filled with constituents belonging to the propositional content of the sentence, then the position is free for being taken by a discourse particle, cf. (5):

(5) *Eigenlijk heb ik geen tijd*
In fact have I no time
In fact, I don't have time

If the sentence does not contain a left-dislocated constituent, a discourse particle in the left periphery seems to take that position, analogous to what we assume for the first sentence position, cf. (6):

(6) *Kortom, alles is goed gegaan* (Dutch spoken Corpus)
In-short, everything is good went
'In short, everything went well'

These observations result in (minimally) 5 positional possibilities for discourse particles at the beginning of Dutch main clauses:

- Preceding the "aanloop" (as in 2)
- In the "aanloop" (6)
- Between "aanloop" and "first position" (3)
- In "first position" (5)
- Between "first position" and the inflected verb (4)

In the paper, we will focus on the following questions:

- Which particles can go into which positions?
- Which combinations of occupied positions are possible?
- If a particle can occupy different positions, does this lead to a different meaning or a different contribution to the discourse?
- Can we attribute specific functions to each of the linear positions?

By exploring these questions for the left periphery, we complement our earlier findings for the positional possibilities of discourse particles in the right periphery as sketched in Van der Wouden & Foolen (2015).

References

Haeseryn, Walter e.a. (eds.). 1997. *Algemene Nederlandse Spraakkunst.* 2nd ed., Groningen and Deurne, Martinus Nijhoff and Wolters Plantijn.

Van der Wouden, Ton & Ad Foolen. 2015. Dutch particles in the right periphery. In: Sylvie Hancil, Alexander Haselow, and Margje Post (Eds.), *Final particles.* Berlin: De Gruyter Mouton, 221-247.

ø

Comparing Norwegian *da* with English *then*: Right-dislocation as a grammatical trigger for the development of pragmatic particles in tag position

Kaja Borthen & Elena Karagjosova
Norwegian University of Science and Technology, Freie Universität Berlin

Keywords: pragmatic particles; pragmatic markers; right-dislocation

We propose a semantic-pragmatic analysis of the Norwegian tag *da* and compare it to the English tag *then*. The two particles are historically related, have functions in common and both forms have additional uses as temporal/propositional adverbs. Still, the tag *da* has functions not found for *then*, and the question is why. We argue that the differences in function are due to 1) the existence of middle field pragmatic particles in Norwegan not found in English, and 2) a more general right-dislocation construction in Norwegian.

Fretheim (2015) proposes that *da* and *then* have the same semantics, as one might suspect from (1). But this leaves unexplained why the two forms sometimes differ, as in (2).

(1) Så du dem ikke? sa Ugla temmelig forbauset. Kom og se på dem nå, **da**.
Saw you them not? said Owl quite stunned. Come and look on them now, DA
'Didn't you see them? said Owl quite stunned. Come and look at them now, THEN.'

(2) Ingen kjeder seg i London – hvis de ikke går hardnakket inn for det, **da.**
Nobody gets.bored REFL in London – if they not go hard.necked in for it, DA
'Nobody gets bored in London – unless they try really hard, THAT IS/ OF COURSE.'

In (1) *da* points to a contextual motivation for the request, and expresses an opposition between the requested situation and the present one. Similar functions are found for the English tag *then* (Haselow 2011). In (2), *da* strengthens the epistemic status of the utterance, it makes it clear that the speaker is making a correction, and it compensates for the preceding break/pause. This function is not found for *then.* This is no exceptional example; in a corpus-based translation study that we performed, the tag *da* was translated into *then* in only 14% of the cases.

We believe data such as (1) and (2) rely on two facts: Unlike English, Norwegian has middle field pragmatic particles (Fretheim 1991). Moreover, Norwegian allows for duplication and right-dislocation of not only NPs, but also referential adverbs and middle field particles (Askedal 1987; Borthen 2018). Since the form *da* corresponds to both a referential adverb and an epistemic middle-field modal particle (Berthelin 2018), a right-dislocated *da* can have two distinct resumptive elements:

(3) **Da** drar vi, **da**.
THEN leave we, THEN

'Then we are leaving'
(4) Du har **da** råd, **da**.
You have DA afford, DA
'You surely have enough money'

Correspondingly, we claim that the tag particle *da* has two historical origins: right-dislocation of a topicalized referential *da* as in (3) or right-dislocation of a middle-field modal particle *da* as in (4). This gives rise to the two tag functions illustrated in (1) and (2), of which only the first is reminiscent of English *then* (Haselow 2011). Our analysis relies on Givón (1983) and his notion of marked constructions. Being a marked construction, right-dislocation is expected to signal and compensate for discontinuity and contrast, which is what we find for the tag *da*.

References

Askedal, J.O. 1987. On the morphosyntactic properties and pragmatic functions of correlative right dislocation (right copying) in modern colloquial Norwegian. In Lilius, P. and M. Saari (eds.): *The Nordic languages and modern linguistics 6.* Helsinki: Helsinki University Press, 93-110.

Berthelin, S.R. 2018. Midtstilt *da* – en semantisk-pragmatisk redegjørelse og en sammenlikning med etterstil *da. Norsk Lingvistisk Tidsskrift* 36 (2), 353-401.

Borthen, K. 2018. Pronominal høyredislokering i norsk, det er et interessant fenomen, det. *Norsk Lingvistisk Tidsskrift* 36 (2), 403-450.

Fretheim, T. 2015. A relevance-theoretic perspective on the Norwegian utterance-final particles *da* and *altså* compared to their English counterpart *then*. In Hancil, S., A. Haselow & M. Post (eds.) *Final Particles*, 249–283.

Fretheim, T. 1991. Formal and functional differences between S-internal and S-external modal particles in Norwegian. *Multilingua* 10 (1/2), 175–200.

Givón, T. (ed.). 1983. *Topic Continuity in Discourse. A Quantitative Cross-language Study.* Amsterdam: John Benjamins.

Haselow, A. 2011. Discourse marker and modal particle: The functions of utterance-final *then* in spoken English. *Journal of Pragmatics* 43 (14), 3603-3623.

Hedged performatives and (inter)subjectivity

Ilse Depraetere & Gunther Kaltenböck
(Université de Lille & University of Graz)

Keywords: hedging, performative, intersubjectivity, left/right periphery, politeness

Hedged performatives (HPs) are combinations of a (semi-)modal verb and a performative verb, as illustrated in (1) and (2), and were originally discussed by Fraser (1976). Despite their frequent occurrence, especially in spoken discourse, hedged performatives have received surprisingly little attention so far (e.g. Schneider 2010).

(1) ***I have to confess**, I don't watch Downton Abbey.* (COCA)
(2) *Well you've you've named a whole bunch of really funny comics **i must say*** (Fisher)

(3) MICHELE-KELEMEN# *Well, definitely. I mean,* ***I have to say****, you know, foreign countries have always sort of hedged their bets on this* (COCA)

Based on data from the Corpus of Contemporary American English and the Fisher Corpus the present study investigates a set of high-frequency HPs with high mutual information scores in spoken English, viz. *I have to/must admit* (205/112 instances), *I have to/must confess* (76/81 instances), *I have to/must say* (200/200 instances). The aim is to explore their use in discourse, specifically in interactive contexts.

In doing so, the focus will be on their positional preferences (initial, medial, final) and their discourse functions, particularly with a view to subjective and intersubjective uses. It will be argued that a comprehensive account of their functions requires two different views of (inter)subjectivity: in interactional terms (turn-taking vs. turn-yielding) and in terms of politeness (face-preserving/boosting/damaging).

With regard to position, HPs clearly prefer left-periphery, where they have important turn-taking function. Position is shown to be significantly affected by the choice of *must* vs. *have to* but not by the choice of predicate (*say, admit, confess*). In right-peripheral position their function adopts a more intersubjective quality.

As for their discourse functions, HPs are shown not to be limited to hedging only (as suggested by Fraser). Instead, it is possible to identify three different functions: shield (hedge), emphasis (booster), and discourse maker uses. The shield and emphasis functions, illustrated in (1) and (2) respectively, are the result of a number of interacting co(n)textual parameters, notably speaker/hearer orientation and positive/negative host clause, and have different effects in terms of speaker and hearer face. The discourse marker uses are particularly prominent with *I have to say*, as illustrated in (3). In this use HPs frequently collocate with other discourse markers and function as discourse structuring devices for the purpose of stalling, turn-taking, floor-keeping etc.

References

Fraser, Bruce (1976), Hedged performatives, in P. Cole (ed.), (1976), *Syntax and Semantics 3. Speech Acts,* New York, NY: Academic Press, 187-210.

Schneider, Stefan (2010), Parenthetical hedged performatives, in G. Kaltenböck, W. Mihatsch, and S. Schneider (eds.), *New Approaches to Hedging,* Bingley: Emerald, 267-287.

Pragmatic markers in progress: What text messages can tell us about LP/RP differences in German

Heike Wiese & Oliver Bunk
(Humboldt University of Berlin)

Keywords: pragmatic markers, sentence peripheries, texting, multilingualism, German

Phenomena of language variation can highlight ongoing tendencies of change, and this is also true for left-peripheral (LP) and right-peripheral (RP) associations of pragmatic markers: if we find different pathways for different markers "in progress", this might indicate functional preferences for different loci. A promising empirical domain for this is informal language use among adolescents, which combines the dynamics of vernaculars with that of adolescence (Tagliamonte 2016). Within this domain, texting in particular has a special potential to reveal patterns of variation and change, given its

prevalence as a means of communication among young people in the absence of normative, standard-language restrictions. And since multilingual communities are often further advanced and can thus put a spotlight on patterns of ongoing change, it is particularly interesting to compare productions of multilingual and monolingual speakers (Wiese 2013).

We discuss findings from an ongoing study on such data, based on German WhatsApp messages. The following, from a speaker with Russian as a heritage language, gives an illustration:

> "Dikka Brat du weißt nicht grad was passiert ist ja… einfach *Fatso*
> *bro you know not just what happened is yes simply*
> eine Frau ist über Rot gefahren und ein auto hat sie erwischt [EMOTICON]
> *a woman is over red driven and a car has her caught*
>
> …einfach so traurig ja…Autofahrer hilft ihr zwar aber er könnte
> *simply so sad yes car.driver helps her* MODAL.PARTICLE *but he could*
>
> auch bremsen [EMOTICON]. Andererseits ist die Frau dumm.
> *also break on.the.other.hand is the woman dumb*
> 'Man bro you won't believe just what happened well … simply a woman crossed a red light and a car got her …simply so sad yeah…car driver is helping her all right, but he could also break. On the other hand, the woman is stupid.'

As illustrated here, our data contains a wealth of noncanonical phenomena (preserved in the idiomatic translation), including those pointing to developing pragmatic markers, e.g., recurring usages of *ja* and of *einfach*, and of the three-dot sign as a possible register-specific device. The data is part of a larger, register-differentiated corpus, which allows us to employ further comparisons targeting possible effects of (in-)formality, modality, and age.

First results point to a number of interesting patterns of relocalisation, in particular:

- A tendency towards the left periphery for pragmatic markers expressing a subjective stance, e.g., *einfach* 'just' moving from the middle field to the left periphery when marking a surprising incident with high emotional involvement, or *kp* (for *keinen Plan* 'no idea') developing from a superordinate clause to a left-peripheral marker of an overwhelming, upsetting event.
- A possible path from the right periphery to an inter-sentential position for intersubjective markers, such as *ja* 'yes' as a hedging device, and the three-dot sign '…' as a textual, discourse structuring marker.

In our study, we investigate whether this points to general patterns for pragmatic markers, and discuss the implications for their functional associations with different peripheral domains.

References

Tagliamonte, Sali (2016), *Teen Talk - The Language of Adolescents*. Cambridge: Cambridge University Press.

Wiese, Heike (2013), What can new urban dialects tell us about internal language dynamics? The power of language diversity. *Linguistische Berichte*, Special Issue 19, 208-245

["Dialektologie in neuem Gewand. Zu Mikro-/Varietätenlinguistik, Sprachenvergleich und Universalgrammatik", eds. Werner Abraham and Elisabeth Leiss].

Functions of *anyway* in the left and right periphery: a cross-varietal perspective

Karin Aijmer

In the last few decades, discourse markers have been studied with regard to their position in the left and right periphery. The hypothesis is that the placement of a discourse marker in initial or final position is motivated by different discourse needs and interactional functions (Beeching & Detges 2014). Research so far has suggested that absolute claims about the relation between position and function cannot be made and that we need to study discourse markers in many language and varieties to test the hypothesis. Aijmer (2016) showed that *anyway* was used differently in ICE-GB and in several Asian varieties. The aim is to extend this study to describe the functions of the discourse marker *anyway* in British English, American English and Irish English in order to test if *anyway* has the same discourse functions in the left or the right periphery. The analysis will be based on the ICE-corpora for British and Irish English and the Santa Barbara Corpus of Spoken American English. The American and Irish varieties were chosen because *anyway* was shown to have different frequencies in the left and right periphery.

Anyway was least frequent in American English (116 examples) and the examples occurred mainly in the left periphery (73%). The SBCSAE used *anyway* for a rich range of tasks oriented to changing, shifting or resuming topic and organising the different parts of a story. Moreover, *anyway* in the left periphery could be extensively elaborated (*anyway so, anyway OK but*). On the other hand, in the right periphery *anyway* had only a single function (dismissal) and it was never combined with other discourse markers.

ICE-IRE had the highest frequency of *anyway* (259 examples) overwhelmingly found in final position (84%). In the left periphery *anyway* was unusual and never found on its own but always in combination with a connective (*but, so, and*) indicating its textual function. In addition to the retrospective function to dismiss something as unimportant *anyway* was used in final position when the speaker was uncertain about how to continue or how to end providing a possibility for the hearer to take the turn. In the latter function, the final *anyway* was often elaborated (*anyway but, anyway so, anyway like, anyway though*). These results can be compared with the ICE-GB (142 examples) where *anyway* was most frequent in the right periphery (65%). Unlike Irish English *anyway* could occur alone without a connective in the left periphery.

The preliminary results indicate that the position of *anyway* in the left or right periphery has to do with interactional factors and that position can be viewed as a shared resource in the interaction to help the processing of the message. However varieties are not similar in the ways and extent to which they choose to exploit the left and the right periphery.

References

Aijmer, K. 2016. Pragmatic markers as constructions. The case of *anyway*. In Kaltenböck, G., Keizer, E. and A. Lohmann (eds), *Outside the clause*. Amsterdam/Philadelphia: John Benjamins, 29-58.

Beeching, K. and U. Detges (eds). 2014. *Discourse functions at the left and right periphery: Crosslinguistic investigations of language change*. Brill: Leiden.

Functional development of the Lithuanian focus particles *net* 'even' and *tik* 'only' and clause periphery

Erika Jasionytė-Mikučionienė
(University of Vilnius)

Keywords: focus particle, discourse marker, clause periphery, multifunctionality, (inter)subjectification

The paper deals with the Lithuanian focus particles *net* 'even' and *tik* 'only' and their functions at clause peripheries. Cross-linguistically, focus particles have been studied in the languages of Europe (Gast and van der Auwera 2011): English and French (Beeching 2017), German (König 1991; Diewald 2013) and other languages. In Lithuanian, the studies on focus particles are limited to some observations on their etymology (cf. Nau and Ostrowski 2010) and lack the application of efficient qualitative and quantitative methods of analysis (cf. also Valančė 2017). Besides, little is known about the emergence of the particles under consideration as discourse markers and the correlation between their discourse functions and position at left or right periphery. Drawing on both corpus data of Contemporary Lithuanian and the earliest Old Lithuanian writings (from the 16-17th centuries), the paper attempts at investigating functional distribution of the particles *net* and *tik* at left and right periphery and the correlation between their syntactic position and discourse functions.

In Old Lithuanian texts, *net* and *tik* typically appear clause-medially: *net* functions as a temporal conjunction or particle. It can also be associated with the semantic domain of contrast. The use of *net* as a scalar additive particle in the earliest Lithuanian texts is relatively rare. *Tik*, in its turn, more frequently functions as a focus particle and marks emphasis. In contemporary Lithuanian, *net* and *tik* are positionally mobile: they subsequently shifted to the clause-initial and -final positions and, as a result, gained new uses. Beside their use as focus particles, *net* and *tik* have developed discourse-structuring and interpersonal functions that relate to the expression of speaker's attitude (1-2). *Tik*, in addition, can intensify the force of a directive or an assertive speech act (3), cf.:

(1) - *Žinai, ką šiąnakt sapnavau? Jau tokį baisų košmarą, kad* ***net****...* (CCLL-Fic)
'Guess, what I dreamt last night? Such a terrible nightmare that…'
(2) *Alia, vaikeli, sako, mačiau glamonėjas, glamonėjas,* ***net i[r]*** *bučiavosi!* (CCLL-Sp)
'But, my little, he said, I saw them hugging one another, they even kissed each other!'
(3) ***Tik*** *nesiskubink, - vėl išgirsta Mortos kuždesį.* (CCLL-Fic)
'Don't be in a hurry, - he can hear Morta's whisper again.'

The preliminary results of the analysis reflect the hypothesis provided by Beeching, Degand, Detges, Traugott and Waltereit (2009) who claim that the left periphery typically hosts discourse-coherence markers while the the right periphery is associated with interpersonal (or addresse-oriented) functions. When occuring at left periphery, the Lithuanian particles *net* and *tik* tend to realise discourse-structuring functions, while at right periphery they highlight speaker-hearer interaction. For example, the speaker may intend to elaborate on the preceding utterance or to clarify it by adding some new, more specific information. From a diachronic perspective, *net* and *tik* may be associated with the processes of subjectification and intersubjectification (Traugott 2010).

References

Beeching, Kate (2017), Just a suggestion. *just*/*e* in French and English, in Ch. Fedriani and A. Sansò (eds.) (2017), *Pragmatic markers, Discourse Markers and Modal Particles*. *New Perspectives*, Amsterdam / Philadelphia: John Benjamins, 459–480.

Diewald, Gabriele (2013), *Same same but different* – Modal particles, discourse markers and the art (and purpose) of categorization, in L. Degand, B. Cornillie and P. Pietrandrea (eds.), (2013), *Discourse Markers and Modal Particles. Categorization and description*, Amsterdam / Philadelphia: John Benjamins, 19–46.

Gast, Volker and Johan van der Auwera (2011), Scalar additive operators in the languages of Europe, *Language* 87(1), 2–54.

König, Ekkehard (1991), *The Meaning of Focus Particles. A Comparative Perspective*, London / New York: Routledge.

Nau, Nicole and Norbert Ostrowski (2010), Background and perspectives for the study of particles and connectives in Baltic languages, in N. Nau and N. Ostrowski (eds.), Particles and Connectives in Baltic, Vilnius: Vilnius University, "Academia Salensis".

Valančė, Danguolė (2017), *Pabrėžiamųjų dalelyčių funkcijos ir vartosena dabartinėje rašytinėje lietuvių kalboje*, PhD thesis, Kaunas: Vytautas Magnus University.

WORKSHOP 17

Schedule: Fri 9.00-16.55 (Room 13)

Predicative possession in a cross-linguistic perspective

Iliyana Krapova & Gréte Dalmi
(Ca' Foscari University, Venice & University of Technology, Budapest)

Keywords: predicative possession, BE-possessives, HAVE-possessives, Definiteness Effect, Genitive of Negation, possessive marker

Description of the topic

Aim and scope

This thematic workshop aims to bring together linguists working on predicative possession primarily in Slavic and Finno-Ugric. The predominant way of expressing predicative possession in Slavic languages is to use HAVE-possessives, while in Finno-Ugric languages BE-possessives are more prevalent. Though typologically remote, there is a fair amount of interaction between these two language families, especially in the areas of direct language contact (see MacAnnalen 2011, Mazzitelli 2017). A good example of such interaction is Balto-Slavic, where both HAVE-possessives and BE-possessives are used, with certain semantic restrictions on these two forms. The workshop will address the following issues:

1. Definiteness Restriction in existential and possessive sentences
2. Genitive of Negation in negated existential and possessive sentences
3. Lexical-semantic factors determining the choice between BE-possessives and HAVE-possessives
4. Agreement in possessive sentences
5. The case of the possessor and the possessee in BE-possessives
6. The argument structure of BE-possessives
7. Types of possessive sentences in a cross-linguistic perspective

Current approaches

Predicative possession is realized in various ways cross-linguistically. Stassen (2009) establishes four major types of predicative possession: locational BE-possessives, BE WITH-possessives, topic possessives and HAVE-possessives. While in Slavic languages HAVE-possessives are predominantly used, Uralic languages employ locational BE-possessives with the possessor taking a more prominent VP-internal position and bearing oblique case. In the Slavic language family, West Slavic predominantly uses HAVE-possessives while East Slavic has locational BE-possessives. Old Church Slavic displays locational BE-possessives, which is replaced by HAVE-possessives in present-day Bulgarian. In Balto-Slavic the two types co-occur with some semantic restrictions on the possessee. These areal and lexical considerations give rise to the question whether the choice between HAVE-possessives and locational BE-possessives is the syntactic reflex of parametric variation or whether the two forms are derivationally related.

The derivational approach relates BE-possessives and HAVE-possessives to copular sentences. This view goes back to Benveniste's (1966) claim that possessive sentences are nothing but inverted

copular sentences. This is also reflected in Freeze (1992), Kayne (1993), Belvin & Den Dikken (1997), and Myler (2016). Although these approaches differ in the structure of the derivational input, they share the basic concept that BE selects a small clause complement in all these sentence types. This makes existential and possessive BE-sentences look exactly like copular BE-sentences.

Freeze's (1992) locative hypothesis seeks to derive all existential, locative and possessive constructions from the same underlying structure. Kayne (1993), building on Szabolcsi (1983), proposes that BE selects a possessive DP, which is the source of the possessive relation in his theory.

Other approaches seek to relate existentials to possessives. For example, Belvin & Den Dikken (1997) claim that an analysis of existential THERE constructions can be extended to possessive HAVE constructions, arguing that the distinction between existential statements and possessive HAVE statements is made by the ability to assign Case, which they assume is due to preposition incorporation into Agr.

Myler (2016), building on Jung (2011), argues that, "copulas exist to "sentencify" fundamentally non-sentential syntactic units. The copula is usually realized as BE. To achieve this, he introduces the VoiceP functional projection, surmounting the lexical layer of copular BE and he employs the mechanism of delayed gratification to introduce the possessor in this functional projection. The differences between existential/possessive BE-sentences vs. copular BE-sentences in his theory are not attributed to the BE-predicate itself, even in languages with morphologically distinct BE predicates in this domain; rather, to the different functional layers and the syntactic mechanisms employed in them. The same meaningless BE verb is involved in each case, and the existential/possessive content comes from elsewhere.

Under the lexicalist approach, BE-possessives and HAVE-possessives are structurally distinct constructions. Both sentence types have two participants, the possessor and the possessee (see Paducheva 2000, Blaszczak 2007, Partee & Borschev 2008). BE-possessives share a whole range of syntactic and semantic properties with BE-existentials. The Definiteness Restriction in affirmative sentences, GEN NEG in Slavic negated sentences, binding relations, verb agreement and case marking on the possessor and the possessee are the most important questions to be raised in relation to predicative possession.

Paducheva (2000) proposes that in addition to copular BE, Russian has a distinct existential BE. This type of BE-sentence has several properties in common with possessive BE-sentences. The author argues that Russian existential BE appears in existential and possessive BE-sentences alike, and it invariably takes two arguments: a location and a theme in the former, and a possessor and a theme in the latter. She offers semantic, syntactic and prosodic tests, including Definiteness Restriction, GEN NEG, sentence stress, and word order, in support of this.

Błaszczak (2007) distinguishes between three different meanings of BE (different BEs) each of which has its own selectional properties/argument structure:

(i) existential BE: $[_{vP}\ PP_{LOC}\ [_{v'}\ v\ [_{VP}\ V\ NP_{THEME}]]]$

(ii) locative BE ('agentive' reading): $[_{vP}\ NP_{AGENT}\ [_{v'}\ v\ [_{VP}\ V\ PP_{LOC}]]]$

(iii) locative BE ('simple position' meaning): $[_{vP}\ v\ [_{VP}\ NP_{THEME}\ [_{V'}\ V\ PP_{LOC}]]]$

The postulation of three different types of BE, Błaszczak (2007) argues, can help us explain the puzzling phenomenon of GEN NEG.

Partee & Borschev (2008) assume that existential BE is also present in possessive BE-sentences. They derive the difference between locative copular sentences and locative existential sentences from their different perspectival center. While in locative copular sentences the subject is interpreted as

definite/specific, existential and possessive sentences impose the Definiteness Restriction, hence the theme argument must be indefinite/non-specific. Genitive of Negation is found in negated existential and possessive BE-sentences but never in copular BE-sentences.

The multiplicity of theories attempting to capture these puzzling facts indicates that predicative possession is a complex phenomenon. The comparative approach taken in this workshop makes it possible to test the various theories against some, so far under-estimated, empirical data.

References

Belvin, R & Dikken den, M. 1997. There, happens to, *be*, *have*. *Lingua* 101: 151–183.

Benveniste, E. 1966. *Être* et *avoir* dans leurs fonctions lingistiques. In: Benveniste, E., *Problèmesde linguistique générale*. Vol. 1. 187-207. Paris: Gallimard.

Błaszczak, J, 2007. *Phase Syntax. The Polish Genitive of Negation*. Habilitation treatise. University of Potsdam.

Den Dikken, M. 1997. The syntax of possession sentences and the verb *have*. *Lingua* 101: 129– 150.

Freeze, J. 1992. Existentials and other locatives. *Language* 68(3): 553–595.

Jung, H. 2011. *The syntax of BE-possessives. Parametric variation and surface diversities.* Amsterdam: Benjamins.

Kayne, R, 1993. Towards a modular theory of auxiliary selection. *Studia Linguistica* 47: 3-31.

McAnnalen, J. 2011. The history of predicative possession in Slavic: Internal development vs. language contact. PhD dissertation, UCL, Berkeley.

Mazzitelli, L. 2017. Predicative possession in the languages of the Circum-Baltic area. *Folia Linguistica* 51(1): 1–60.

Myler, N. 2016. *Building and Interpreting Possession Sentences*. Cambridge, MA: MIT Press.

Paducheva, E. 2000. Definiteness Effect: the case of Russian. In: K. von Heusinger & U. Egli (eds): *Reference and Anaphoric Relations.* 133-146. Dordrecht: Kluwer.

Partee, B. & Borschev, V. 2008. Existential sentences, BE and GEN NEG in Russian. In: Comorowski, I., & von Heusinger, K. (eds), *Existence: Semantics and Syntax* 147–191. New York: Springer.

Stassen, L. 2009. *Predicative possession.* Oxford: Oxford University Press.

Szabolcsi, A. 1983. The possessor that ran away from home. *The Linguistic Review* 3: 89-102

Locative, existential and possessive constructions: Major issues and approach. An overview

Joanna Błaszczak
(University of Wrocław)

Keywords: locative, existential, possessive, structure, analysis

1. Cross-linguistically, there is evidence for a close affinity between locative, existential and possessive sentences (Clark 1970, 1978). In many languages the only difference between locative and existential sentences seems to be a different arrangement of the locative and the nominal phrases: while in locative sentences the nominal phrase precedes the locative phrase, in existential sentences the opposite is the case, i.e., it is the locative PP that precedes the NP. In addition, in many languages

also possessive sentences closely resemble existential/locative sentences in that the possessor is realized by means of a prepositional (locative) phrase.

2. It is often assumed (that existential, locative and possessive constructions are all derived from the same underlying structure (Hoekstra and Mulder 1990, Freeze 1992, den Dikken 1997, 2006, Moro 1997, Witkoś 2000, and Harves 2002). The different types of constructions arise as the result of moving either the NP_{THEME} (locative sentences) or the PP_{LOC} (existential/possessive sentences) into a sentence-initial position (mostly understood as [Spec,IP], thus is the "surface subject" position).

However, in many languages possessive sentences have a nominal possessor instead of a prepositional one, and HAVE as predicate (Slavic and Finno-Ugric, the two families tackled by the workshop, more or less reflect this HAVE/BE split). In order to account for these differences, it is usually assumed that BE and HAVE are not lexical verbs, but rather spell-outs of (various) functional heads in syntax (den Dikken 2006). More precisely, HAVE is a result of syntactic incorporation of a(n abstract) prepositional locative head into BE, giving rise to an NP possessor as in English (see Myler 2016 for another version of this view going back to Benveniste 1966).

Such unified analysis might seem appealing and attractive at first glance, but on a closer inspection it turns out to be "too simple". An alternative would be to assume that locative, existential/possessive sentences differ in terms of the Perspective Structure (Partee and Borschev 2007). Another option would be to take an existential BE not to be a copula but a lexical verb and assume that different constructions might involve different verbs BE with different argument structures (Błaszczak 2007, 2018). The possessor argument can also be introduced in different ways into the structure (Myler 2016).

3. The goal of my introductory paper is to give an overview of the major issues and their possible accounts proposed in the literature, listed below.

(i) Can the existing analyses account for all predicative possession structures cross-linguistically?

(ii) If locatives, existentials and possessives have the same underlying argument structure, how can their meaning differences be accounted for?

(iii) Is existential BE comparable to copular elements in other predicative structures? Is HAVE just a spellout of P (or some other functional element) incorporated into BE?

(iv) Does the locative PP occupy the same syntactic position as the theme NP in these and how can oblique case/prepositional case and agreement in these constructions be accounted for?

(v) Why do we observe a definiteness restriction in some languages/some of the constructions and not in others?

References

Belvin, Robert, and den Dikken, Marcel (1997), *There*, happens, *to*, *be*, *have*, *Lingua* 101, 151–183.

Benveniste, Émile (1966), *Problèmes de linguistique générale*, Paris: Gallimard.

Bjorkman, Bronwyn, and Cowper, Elizabeth (2015), Where *there* is, and why. Paper presented at the Congrès de l'ACL de 2015 / 2015 CLA meeting, May 30 – June 1, 2015.

Błaszczak, Joanna (2007), *Phase Syntax: The Polish Genitive of Negation*, Habilitation thesis, University of Potsdam.

Błaszczak, Joanna (2018), Clause structure, case and agreement in Polish existential, possessive and locative sentences: A phase-based account, *Poznań Studies in Contemporary Linguistics* 54(4), 637–696.

Boneh, Nora, and Sichel, Ivy (2010), Deconstructing possession, *Natural Language & Linguistic Theory* 28, 1–40.

Borschev, Vladimir, and Partee, Barbara (2002), The Russian genitive of negation in existential sentences: The role of theme-rheme structure reconsidered, in E. Hajičova, P. Sgall, J. Hana,

and T. Hoskovec (eds), *Travaux de Cercle Linguistique de Prague (nouvelle série)*, Amsterdam: John Benjamins, 185–250.

Clark, Eve V. (1970), Locationals: A study of 'existential,' 'locative', and 'possessive' sentences, *Stanford University Working Papers in Language Universals* 3, 1–36.

Clark, Eve V. (1978), Locationals: Existential, locative and possessive constructions, in J. H. Greenberg (ed), *Universals of Human Language*, vol. 4: *Syntax*, Stanford, CA: Stanford University Press, 85–126.

den Dikken, Marcel (1997), The syntax of possession and the verb 'have', *Lingua* 101 (3/4), 129–150.

den Diken, Marcel (2006), *Relators and Linkers. The Syntax of Predication, Predicate Inversion, and Copulas*, Cambridge, MA: The MIT Press.

Freeze, Ray (1992), Existentials and other locatives, *Language* 68, 553–595.

Hartmann, Jutta M., and Milićević, Nataša (2008), The syntax of existential sentences in Serbian, in A. Antonenko, J. Bailyn, and Ch. Bethin (eds), *Formal Approaches to Slavic Linguistics 16: The Stony Brook 2006 Meeting*, Ann Arbor: Michigan Slavic Publications, 168–184.

Harves, Stephanie (2002), *Unaccusative Syntax in Russian*, Ph.D. Dissertation, Princeton University.

Heine, Bernd (1997), *Possession: Cognitive Sources, Forces, and Grammaticalization*, Cambridge: Cambridge University Press.

Hoekstra, Teun (1994), HAVE as BE plus or minus, in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (eds), *Paths Towards Universal Grammar; Studies in Honor of Richard S. Kayne*, Washington DC: Georgetown University Press, 199–215.

Jung, Hakyung (2011), *The Syntax of the Be-Possessive*, Amsterdam: John Benjamins.

Kayne, Richard (2018), Notes on expletive *there*, Manuscript, New York.

Moro, Andrea (1997), *The Raising of Predicates*, Cambridge: Cambridge University Press.

Muromatsu, Keiko (1997), Two types of existentials: Evidence from Japanese, *Lingua* 101, 245–269.

Myler, Neil Joseph (2016), *Building and Interpreting Possession Sentences*, Cambridge, MA MIT Press.

Paducheva, Elena (2000), Definiteness Effect: The case of Russian, in K. von Heusinger, and U. Egli (eds), *Reference and Anaphoric Relations*, Dordrecht: Kluwer, 133–146.

Partee, Barbara, and Borschev, Vladimir (2007), Existential sentences, BE, and the genitive of negation in Russian, in I. Comorovski, and K. von Heusinger (eds), *Existence: Semantics and Syntax*, Dordrecht: Springer, 147–190.

Stassen, Leon (2009), *Predicative Possession*, Cambridge: Cambridge University Press.

Witkoś, Jacek (2000), Nominative-to-genitive shift and the negative copula *nie ma*: Implications for checking theory, *Journal of Slavic Linguistics* 8(1/2), 295–327.

Genitive of Negation (GoN) in Polish possessive and locative existential sentences: A testing tool for Case Overwriting, Case Projections and derivational phases

Jacek Witkoś
(AMU, Poznań)

Keywords: syntax, minimalism, case, genitive, derivational phase

The GoN obligatorily replaces an ACC-marked nominal object in the context of clausal negation, both within the same clause (1) and long-distance, across an infinitive boundary (2):

(1) Maria czyta gazetę/ nie czyta *gazetę/gazety.
Maria$_{NOM}$ reads newspaper$_{ACC}$ /NEG reads newspaper$_{*ACC/GEN}$
'Maria is reading/not reading a newspaper.'

The GoN has served as a testing tool for each framework of generative grammar (from Babby 1980 to Kagan 2012), I follow suit and put three proposals to the GoN test: Pesetsky's (2013) case theory based on overwriting, Caha's (2009, 2010a-b, 2013) case theory relying on nano-syntax and the concept of the derivational phase (Chomsky 2001). It is shown that the two case theories cope with engineering details of the GoN but the nano-syntactic theory easily predicts that only accusative (but not oblique) is replaced with genitive, while Pesetsky's theory cannot independently constrain case overwriting, unless it resorts to a liberal application of Spell-Out. Finally, the GoN is treated as a diagnostic for the notion of the derivational phase and its extensions. As the relation between NEG and the NP object in (2) is Agree it should not be hampered by any phase boundaries (typically CP/vP), obeying the Phase Impenetrability Condition (PIC, Chomsky 2001). Ex. (2-3) combine GoN and movement, such that an intermediate position is necessary for reconstruction, in the spirit of Fox (1999) and Lebeaux (2009):

(2) Jan$_1$ nie kazał Marii$_2$ [PRO$_2$ pokazywać mu$_3$ [listów Tomka$_{*3/4}$ do siebie$_{1/2/4}$]].
Jan$_{NOM}$ NEG told Maria$_{DAT}$ show$_{INF}$ him$_{DAT}$ letters$_{GEN}$ Tom$_{GEN}$ to self
'Jan did not tell Maria to show him Tom's letters to him/her/himself.'

(3) [[ilu listów Tomka$_{3/4}$ do siebie$_{1/2/4}$] Jan$_1$ nie kazał Marii$_2$ [PRO$_2$ [$_{vP}$ t' pokazywać
how many letters$_{GEN}$ Tom$_{GEN}$ to self Jan$_{NOM}$ NEG told Maria$_{DAT}$ show$_{INF}$
mu$_3$ t wczoraj]]]
him$_{DAT}$ yesterday
'How many of Tom's letters to him/her/himself did Jan not tell Maria to show him yesterday?'

Ex. (2) shows long-distance GoN and three interpretations for the reflexive, where *siebie* 'self' can be bound by the main clause subject Jan$_1$, PRO$_2$, controlled by Maria$_{2,DAT}$, and the agentive Tomek$_{*3/4,GEN}$. Importantly, Tomek$_{*3/4,GEN}$ must be obviative w.r.t mu$_{3,DAT}$, as the indirect object c-commands the direct one and Principle C holds. In (3) wh-movement fronts the direct object to the main clause and an additional interpretation appears, where Tomek$_{3/4.GEN}$ can be coindexed with the indirect object of the embedded predicate. Assuming that one copy is interpreted at LF, this copy must occupy a position where the indirect object mu$_{3,DAT}$ does not c-command Tomek$_{3/4,GEN}$ and at the same time PRO$_2$ c-commands siebie$_{1/2/3}$ to provide for interpretation marked by index 2. This position is t' at the edge of the embedded vP in (3). The classic phase theory says that the wh-phrase needs to move cyclically through phase edges, vP is a phase and t' is its edge. In contrast, the Agree relation between NEG and the embedded NP cannot tolerate a phase in between.

References:

Babby, L. 1980. *Existential Sentences and Negation in Russian*. Ann Arbor, MI: Karoma Publishers.
Caha, Pavel. 2009. *Nanosyntax of case*. Ph.D. dissertation, University of Tromso.
Caha, Pavel. 2010a. The parameters of case marking and Spell-Out driven movement. http://ling.auf.net/lingbuzz/001026 (accessed 15 Feb 2011).
Caha, Pavel. 2010b. The German locative-directional alternation. A Peeling account. *Journal of Comparative German Linguistics* 13:179-223.

Caha, P. 2013. Czech numerals and no bundling. http://ling.auf.net/lingbuzz/001723 (accessed 10 Feb 2013).

Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, edited by Kenstowicz, M. 1-52. Cambridge, MA: The MIT Press.

Fox, D. 1999. Reconstruction, Binding Theory and the Interpretation of Chains. *Linguistic Inquiry* 30: 157-196.

Kagan, O. 2012. *Semantics of genitive objects in Russian: A study of Genitive of Negation and intensional genitive case*. Dordrecht/New York: Springer.

Lebeaux, D. 2009. *Where Does Binding Theory Apply?* Cambridge, MA: The MIT Press.

Pesetsky, D. 2013. *Russian case morphology and the syntactic categories*. Cambridge: The MIT Press.

Extraction of possessive NP-complements and the internal structure of the nominal domain in Polish

Piotr Cegłowski
(Adam Mickiewicz University in Poznań)

Keywords: extractability, NP-complements, phases, left periphery, antilocality

The phenomenon of *extractability*, i.e. the possibility of displacing of the specific NP-internal constituents to positions outside the NP, including all its intricacies, has for long constituted one of the major points of reference in the discussion of the internal structure of nominal phrases. Their behaviour with respect to the specific types of extraction (including, but not limited to the extraction of possessive (genitive) NP-complements, see (1a-b)) has stimulated extensive cross-linguistic research and yielded several insightful comparative accounts of the data from languages that allow it (e.g. English) and those that apparently don't (e.g., Serbo-Croatian (Bošković 2005), Russian (Pereltsvaig 2007), Polish (Rappaport 2001; Migdalski 2001, 2003; Rutkowski 2007). Extractability, among other phenomena, is taken to be one of the defining characteristics of the so-called 'DP-languages (Bošković 2008).

The aim of this talk is to present a possible scenario for the extraction of possessive (genitive) possessive NP-complements in Polish (2c) that sets a more elaborate architectural setting than the radically parsimonious, multiple adjunction – based account of Bošković (2014, 2015). In particular, it is assumed that the nominal structure (NP) in Polish is insulated with a number of functional projections (cf. Migdalski 2001; Rutkowski 2007) that exert influence on the extraction process and, to a large extent, determine its grammaticality. The reasoning is primarily based on the observation that Polish does not consistently meet the criteria of an 'NP-language' in the sense of Bošković (2008). This claim is supported with the results of an extractability survey (extraction of possessive (genitive) NP-complements, extraction across numerals, etc.) carried out on the native speakers of Polish. A closer look at the relevant data, coupled with the facts regarding *fronting* (2a-b), gives substance to the intuition that the sort of operation is to a large extent discourse-driven (hence the presence of TopP and FocP in the 'nominal spine', cf. Ishane & Putskás 2001; Aboh 2004; Cetnarowska 2013). Given the size and the number of the postulated projections (3), it is also argued that the account of Antilocality in the spirit of Grohmann (2000/2003) appears better suited to handle the displacement-related facts than the radically 'parsimonious' labelling model postulated by Bošković (2014, 2015).

(1) a. Which city$_i$ did you see a picture of t_i?

b. ?/*[Ovog grada]$_i$ sam pronašla [$_{NP}$ sliku t_i]
This city.GEN am found picture
'Of this city I found a/the picture' (Serbo-Croatian, Bošković 2014b: 48)

c. [Którego miasta]$_i$ znalazłeś [zdjęcie t_i]?
which city (you) found picture
'Which city did you find a picture of?' (Polish)

(2) a. ten Hanki kolejny narzeczony (to PIJAK)
this Hanka.GEN next fiancé COP drunkard.FOC
'as for this next fiancé of Hanka, he's a drunkard' (Polish)

b. [[Tego miasta]$_i$ zdjęcie t_i] ukazało się wczoraj w gazetach
this.GEN city.GEN picture appeared REFL yesterday in newspaper
'A picture of this city appeared in the newspapers yesterday' (Polish)

(3) [$_{DP}$ D [$_{TopP}$ Top [$_{FocP}$ Foc [$_{AgrP}$ DemP [$_{AgrP}$ Agr [$_{QP}$ NumP [$_{QP}$ F_Q [$_{NP}$ (wh/AP) [$_{NP}$ N]]]]]]]]

References:

Aboh, Enoch (2004) Topic and Focus wihtin D. *Linguistics in the Netherlands* 21, 1–12. Amsterdam/New York: John Benjamins.

Bošković, Željko (2005), On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59 (1–45).

Bošković, Željko (2008), What will you have, DP or NP? *Proceedings of the North East Linguistic Society* 37 (101–114).

Bošković, Željko (2009), More on the no-DP analysis of article-less languages. *Studia Linguistica* 63 (187–203).

Bošković, Željko (2014), More on the Edge of the Edge. In: Chapman, C., O. Kit, and I. Kučerová (eds.). (2014), *Formal Approaches to Slavic Linguistics: The McMaster Meeting 2013 (FASL 22)*. Michigan, MI: Michigan Slavic Publications, 44–66.

Bošković, Željko (2015), On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. Manuscript: http://ling.auf.net/lingbuzz/002452, Date of access: 27 Nov. 2015.

Cetnarowska, Bożena (2013), On topic movement within Polish noun phrases. [Talk presented at 29th Poznań Linguistic Meeting, Poznań, Poland, 29 Aug – 1 Sept. 2013].

Grohmann, Kleanthes (2000/2003), *Prolific domains: on the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Ishane, Tabea & Genovefa Puskás (2001), Specific is not definite. *Generative Grammar in Geneva* 2, 39–64.

Migdalski, Krzysztof (2001), A Determiner Phrase approach to the Structure of Polish Nominals. In: Przepiórkowski, A. & P. Bański (eds.), (2001), *Generative Linguistics in Poland: Syntax and morphosyntax*. Warsaw: IPIPAN. 135-148

Pereltsvaig, Asya (2007), On the Universality of DP: A View from Russian. *Studia Linguistica* 61(1), 59–94.

Rappaport, Gilbert (2001), Extraction from Nominal Phrases in Polish and the Theory of Determiners. In: Willim, E. and P. Bański (eds.), (2001), *Formal Approaches to Polish Syntax*. Special edition of *Journal of Slavic Linguistics* 8(3).

Rutkowski, Paweł (2007), *Hipoteza frazy przedimkowej jako narzędzie opisu składniowego polskich grup imiennych*. Doctoral dissertation. Warsaw: Warsaw University.

Copular BE, existential BE, and transitive HAVE: The case of predicative possession in Belarusian

Egor Tsedryk
(Saint Mary's University)

Keywords: <HAVE possession, BE possession, existential, copular, Belarusian>

Belarusian is known as a language "in a state of transition" between BE and HAVE types (Isačenko 1974), or "a mixed language" in the Circum-Baltic area (Mazzitelli 2017). Previous work on predicative possession in Belarusian (Mazzitelli 2015) has mostly focused on the competition between Location and Action schemas (Heine 1997) (cf. locational and HAVE possessives in Stassen 2009). As Isačenko (1974) points out, the BE/HAVE alternation is complicated by the fact that BE sentences can be copular or existential. This split in relation to HAVE has remained understudied.

In this talk, I aim at answering the following two questions: (a) How are copular BE, existential BE and HAVE distributed across possessive relations? (b) What does this distribution imply for a compositional analysis of possessives?

I rely on purely formal diagnostics: the copular and the existential BE in Belarusian have syncretic forms in the past and the future, but they are formally distinct in the present tense (null form vs. *ësc'*, respectively). Thus, if *ësc'* cannot be overtly used in a present tense BE clause, the latter has a copular BE. For example, (1a) has a copular BE, whereas (2b) has an existential one. As we see, HAVE can alternate with both. Table 1 summarizes the overall distribution.

(1) a. Hanna mae pryhožyja voču.
Hanna.NOM has [beautiful eyes].ACC
'Hanna has beautiful eyes.'

b. U Hanny (*ësc') pryhožyja voču.
at Hanna.GEN be.EXIST [beautiful eyes].NOM
'= (1a)'

(2) a. Hanna mae voču (jana sama ŭsë ŭbačyc').
Hanna.NOM has eyes.ACC (she self everything will.see)
'Hanna has eyes (she'll see everything herself).'

b. U Hanny ësc' voču (jana sama ŭsë ŭbačyc').
at Hanna.GEN be.EXIST eyes.NOM (she self everything will.see)
'= (2a)'

I argue that the complementary distribution between the copular and the existential BE is a surface manifestation of two set-theoretic options: (i) inclusion of a whole set and (ii) intersection of two sets, respectively. In syntax, the first option results from a functional application directly involving the locational PP and an NP. The second option involves an existential head that relates these two phrases, based on the "existence is relative" principle (Borschev & Partee 1998 et seq.). As for HAVE, I link its relative flexibility to transitivity. Assuming that the structural accusative case is assigned under agreement, I take the valued agreement features as a variable of type *e* in the logical form. Whenever this variable is existentially closed (a default scenario), HAVE has an existential reading. Otherwise, HAVE does not lexically encode an existential operator, nor does HAVE result from incorporation of P into the existential BE (cf. Myler 2016; contra Freeze 1992 and Kayne 1993).

Table 1: BE/HAVE across possessive relations

Relations	Copular BE	Existential BE	HAVE
Kinship		✓	✓
Body parts as descriptions (cf. (1))	✓		✓
Abstract possession		✓	✓
Diseases	✓		
Psychological conditions	✓		
Permanent possession (ownership)		✓	✓
Temporary possession	✓		?

References

Borschev, Vladimir, and Barbara H. Partee (1998), Formal and lexical semantics and the genitive in negated existential sentences in Russian, in: Ž. Bosković, S. Franks, and W. Snyder (eds.), (1998), *Formal Approaches to Slavic Linguistics: The Connecticut meeting 1997*, Ann Arbor, Michigan: Slavic Publications, 75–96.

Freeze, Ray (1992), Existentials and other locatives, *Language* 68, 553–595.

Heine, Bernd (1997), *Possession: Cognitive sources, forces and grammaticalization*, Cambridge: Cambridge University Press.

Isačenko, Alexandr (1974), On *have* and *be* languages: A typological sketch, in: M. Flier (ed.), (1974), *Slavic forum: Essays in linguistics and literature*, The Hague: Mouton, 43–77.

Kayne, Richard (1993), Towards a modular theory of auxiliary selection, *Studia Linguistica* 47, 3–31.

Mazzitelli, Lidia Federica (2015), *The expression of predicative possession: A comparative study of Belarusian and Lithuanian*, Studia Typologica, Volume 18, Berlin: Mouton De Gruyter.

Mazzitelli, Lidia Federica (2017), Predicative possession in the languages of the Circum-Baltic area, *Folia Linguistica* 51 (1), 1–60.

Myler, Neil (2016), *Building and interpreting possessive sentences*, Cambridge, MA: MIT Press.

Stassen, Leon (2009), *Predicative possession*, New York: Oxford University Press.

The historical development of predicative possession in Bulgarian

Iliyana Krapova
(University of Venice, Italy)

Keywords: predicative possession, Bulgarian, comitative, dative, alienable/inalienable

In this talk, I will trace the development of predicative possession constructions of contemporary Bulgarian, a language with a split possession system:

a) X *has* Y (default possessive); see (1)
b) Y is at X (locational possessive); see (2)
c) X is with Y (comitative possessive); see (3)

The co-occurrence of these three distinct patterns, and their partial functional overlap, represents a rather unusual typological situation when compared to the rest of Slavic. I will argue that two featural dimensions can be held responsible for the observed variation: (i) the length of time during which the possessed item has been *located* in proximity to the possessor (temporary vs. permanent possession),

and (ii) the extent of *control* that the possessor has over the possessed item (alienable vs. inalienable possession) (Bugenhagen 1986, Stassen 2009).

I will argue that these features were instrumental in the following historical processes of reanalysis of the Old Church Slavonic/Old Bulgarian possessive system: a) older *be* + dative possessives (the GOAL schema, cf. Heine 1997, Grković-Major 2011) got reanalyzed as *have*-possessives at the time when the dative emerged as a general DP-internal case marking mechanism; b) older locatives (the LOCATIVE schema, McAnallen 2011) were reanalyzed as temporary animate locations at the time when definiteness was already a well-established grammatical category since they do not show the definiteness effects of *have*-possessives. I will also argue that the use of the *with*-possessive (the COMPANION schema, cf. Heine 1997, Stassen 2009), which basically reanalyzes temporary vs. permanent co-location (Heine and Kuteva 2006) in terms of a syntactic distinction between alienable and inalienable possession, is an innovation in Bulgarian (within the Slavic family) with a probable Balkan origin. For the areal distribution of *with*-possessives, the Bulgarian data point that this split is characteristic of the extreme areas of the Indo-European boundaries (the Balkan languages, Portuguese, Icelandic, and some of the languages of the Circum-Baltic area, cf. Stolz 2001, Stolz et al. 2008) where temporary vs. permanent distinctions turn out to be more important than the alienable-inalienable ones.

I will show that a), b) and c) above cannot be reduced to a single underlying structure (cf. Freeze 1992), neither synchronically nor diachronically: a) *have*-structures are derived from a DP-internal small clause (Kayne 1993, Benveniste 1973) headed by a preposition P/D with the possessor c-commanding the possessee (4); b) temporary possession structures are derived from a locative underlying structure like (5); c) comitative possessives are derived from a layered PP structure with a richer functional content (cf. Levinson 2011) (6).

Examples:

(1) Ivan ima sini oči/brat/kniga/kâšta.
Ivan has blue eyes/brother/book/house
'John has blue eyes/a brother/a book/a house'

(2) Ključât/knigata/*kâštata ti e u Ivan.
key.DET/book.DET/*house.DET your is at Ivan
'Ivan has your key/book/*house'

(3) Deteto e s edin roditel/sâs dalga kosa/xarakter/očila/grip
child.DET is with one parent/with long hair/character/glasses/flu
'The child has got one parent/long hair/a strong character/glasses/the flu'

(4) BE D/P [DP [$_{AGRP}$ Possessor AGR Possessum]

(5) BE [$_{SC}$ DP-possessum PP-possessor] (cf. Hornstein et al. 1995, Boneh and Sichel 2010).

(6) Possessor [*p*P [$_{PP}$ [P Possessum] (cf. Levinson 2011)

References

Benveniste, E. (1966), *Problèmes de linguistique générale*. Paris: Gallimard

Bugenhagen, R. D. (1986), Possession in Mangap-Mbula: Its syntax and semantics. *Oceanic Linguistics* 25, 124–66.

Boneh, N. and I. Sichel (2010), Deconstructing Possession, http://linguistics.huji.ac.il/IATL/23/Boneh-Sichel.pdf

Freeze, R. (1992), Existentials and other locatives. *Language*. 68, 553 – 595.

Grković-Major, J. (2011) The Development of Predicative Possession in Slavic Languages, in Nomachi Motoki (ed.). *The Grammar of Possessivity in South Slavic Languages: Synchronic and Diachronic Perspectives*, Slavic Eurasian Studies, No 24, Sapporo, 35-54.

Heine, B. (1997), *Possession: cognitive sources, forces, and grammaticalization*. Cambridge-New York: Cambridge University Press.
Heine, B. and T. Kuteva (2006), The changing languages of Europe, Oxford: Oxford University Press.
Hornstein, N.S., Rosen, S. & Uriagereka, J. (1995), Integrals. http://www.ling.umd.edu/Courses/Ling819/Papers/Integrals.pdf;
Kayne, R. S. (1993), Toward a modular theory of auxiliary selection. *Studia Linguistica* 47, 3–31. Levinson, L. (2011), Possessive *with* in Germanic: have and the Role of P. *Syntax*, 14, 4. 355−393;
McAnallen, J. (2011), Predicative possession in Old Church Slavic Bible Translations, in E. Welo (ed.) *Indo-European syntax and pragmatics: contrastive approaches*, Oslo Studies in Language 3(3), 155–172.
Stassen, L. (2009), *Predicative possession*. Oxford-New York: Oxford University Press,
Stolz, Thomas. (2001), To be with X is to have X: Comitatives, instrumentals, locative, and predicative possession. *Linguistics*. 39, 321–50;
Stolz, Thomas, Sonja Kettler, Cornelia Stroh, Aina Urdze (2008), *Split Possession*. Amsterdam: John Benjamins.

Existential Predicative possession: a subtype of non-canonical constructions in Western Iranian languages

Masoud Mohammadirad
(Université Sorbonne Nouvelle- Paris 3)

Keywords: predicative possession, be-possessives, have-possessive, Iranian languages, language change

The predicative possession constructions (PPCs) of Western Iranian languages (WILs) exhibit diverse properties in terms of marking strategies since Old Iranian, ranging from fronted oblique subjects and exhibiting canonical ergative features (1)-(2), to the loss of agreement with the possessed NP (3); and finally, to the levelling of the PPCs with other normalized transitive constructions (4).

dârayava[h]auš pausâ aniyaičiy ahantâ
Darius.GEN.M.SG son.NOM.M.PL other.NOM.M.PLexist.3PL.IPF.MID
'Darius had other sons.' (lit. 'To Darius were there other sons.' (Old Persian)

hakim-ek-î sê kur he-bo-n
prince-INDF-OBL three son exist.COP.PST-3PL
'A prince had three sons (three sons existed to him) (Badini Kurdish, Haig 2008:258)

dû minâl=î he-ye
two child=3SG:NS exist-COP.PRS.3SG
'He has two children.' (Central Kurdish, Mohammadirad, in prep)

(to) ye ketâb dâr-i
2SG a book have.PRS-2SG
'You have a book.' (modern Persian)

Evidence from Western Iranian languages suggests that wherever the existential stem, *ha-*, *he* has been preserved across modern languages, deviant marking of the possessor argument is attested in PPCs, whereas where languages have adopted the regular base *dâr-* 'to have', the relevant indexing pattern of the possessor argument is normalized to that of regular transitive verbs. In a way then, the deviant marking of the possessor argument is construction-based. In terms of diachrony, the data suggest that the older existential possessive (1)-(3) has been replaced by have-possessive (4) in some modern languages (see Figure 1). However, in two languages be-possessive and have-possessive can co-occur but the choice between them is restricted by the nature of possession as being 'inherent' and 'permanent' (as instances of inalienable possession) vs. temporary:

Bijar Southern Kurdish

a.	bizn-e	îšî	šîr=im	ni-ye
	goat-DEF	say.PRS.3SG	milk=1SG	NEG-exist.PRS
	'The goat says: "I don't have milk".'			
b.	îme	kewš	n-eyr-îmân	
	1PL	shoes	NEG-have.PRS.1PL	
	'We don't have shoes.'			

Sivandi

a.	ye	sultan-i	bi	ye	kor-i=š	bi
	a	king-INDF	be.PST	a	son-INDF=3SG	exist.PST
	'(Once) there was a king who had a son.'					
b.	ye	darediye	xub-i	dar-e		
	a	property	nice-INDF	have.PRS-3SG		
	'She has nice property.'					

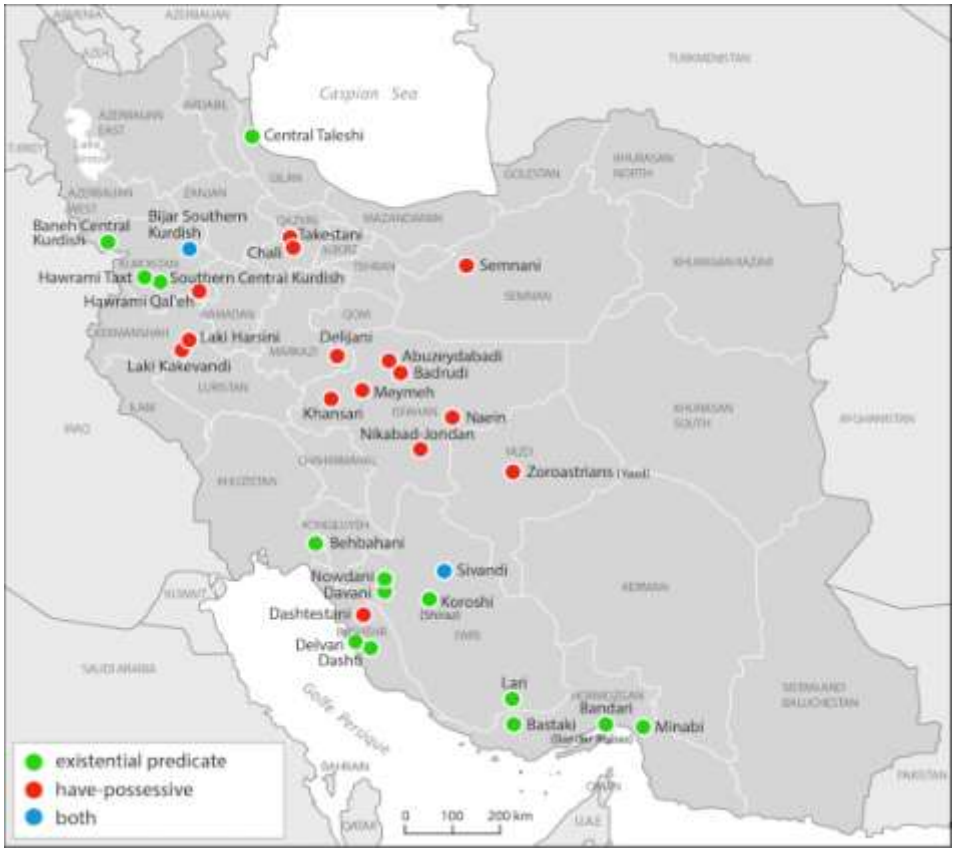


Figure 3: predicative possession constructions across WILs

As can be seen, WILs have can be said to have two primary types of PPCs: oblique possessives (resulting from existential predicates) and regular have-possessives. In addition, a further type occurs in which both these subtypes co-occur. This pattern of distribution goes contrary to the view that regards Iranian languages as having have-possessives (Stassen 2009, see also Stassen 2013).
Along with constructions of potentiality, necessity, and liking & non-controlled events, PPCs form a subset of non-canonical constructions They are placed in the middle of the continuum of non-canonical subjecthood across WILs and their presence in the grammar of WILs implies the existence of the constructions to the right in the below hierarchy.

Potentiality *and/or* Existential predicative possession < Necessity & wanting < Liking and non-controlled events

Acknowledgement
"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665850."

References

Haig, Geoffrey (2008), *Alignment change in Iranian languages: a construction grammar approach*, Berlin: De Gruyter Mouton.

Mohammadirad, Masoud, in prep, *Pronominal clitics in western Iranian languages: Description, mapping, and typological implications*. Paris: Sorbonne Nouvelle _ Paris 3. (Doctoral dissertation.)

Stassen, Leon (2009) *Predicative possession*, New York: OUP

Stassen, Leon (2013) Predicative Possession, in: M. Dryer, and M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/117, Accessed on 2019-01-03.)

Predicative possession as a clause type in Finnish

Maria Vilkuna
(University of Helsinki)

Keywords: locational possession, clause type, construction, subject

In the typological classification of predicative possession patterns defined by Heine (1997) and Stassen (2009), Finnish – like in Finnic languages in general – represents the Locational encoding strategy. The verb is 'be', which also functions as the copula in copular and locational sentences and as the tense auxiliary. The possessor is marked by an oblique case, adessive, which also has straightforwardly locative uses, often corresponding to 'at' or 'on'. The Finnish possessive clause shares most of its central properties with the existential clause. Both have the oblique (locative/possessor) element in the subject position, while the putative subject follows the verb. Unlike in Stassen's basic Locational Possession (2009: 49–50), however, the subjecthood of the possessee is highly controversial (e.g. Huumo & Helasvuo 2015). First, the possessee does not trigger agreement on the verb (1, 2); as testified by older varieties of Finnish and the closely related Estonian, this is a relatively new feature in Finnish. Also, the case-marking of the possessee could suggest object rather than subject properties (3) and therefore a degree of transitivization of the pattern (Stassen 2009: 230–

239). However, there is no sign of the possessor phrase taking on subject marking or triggering agreement, and accusative marking of the possessee as in (3) is restricted to the least typical possessees, personal pronouns.

(1) Anna-lla on uusi mekko.
Anna-ADE be.3SG new dress
'Anna has a new dress.'

(2) Anna-lla on uude-t vaattee-t.
Anna-ADE be.3SG new-PL cloth-PL
'Anna has new clothes.'

(3) Anna-lla on uus-i-a vaatte-i-ta / sinu-t.
Anna-ADE be.3SG new-PL-PAR cloth-PL-PAR you-ACC
'Anna has (some) new clothes / you.'

This paper adopts a constructional approach, i.e. a surface-oriented, declarative view of grammar as a network of constructions, conventionalised pairings of form and meaning or function (Goldberg 1995, 2006; Fried 2015; Hoffmann 2013). I argue that the relevant kind of construction, called *clause type* in Finnish grammar (Hakulinen and Karlsson 1979, Hakulinen et al 2004), is an argument structure construction in the sense defined by Goldberg (1995, 2006) but with a more inclusive scope. Clause types in the present sense are schematic argument structure constructions based on the status of the most prominent argument, or lack of it, in the construction. While so-called general clause types instantiated by e.g. transitive clauses refer to the subject argument and thereby to the subject agreement construction, more restricted types like the Possessive do not.

I will show how the constructional approach allows us to represent the connections between the Possessive clause and other clause types without having to resolve problematic grammatical relations. Still, although canonical instances of the possessive and existential and, on the other hand, these vs. the general clause types are distinctive enough, their borders are somewhat flexible. This will be illustrated by corpus examples.

Abbreviations

3SG = 3rd person singular, PL = plural, ACC = accusative case (on personal pronouns), ADE = adessive case, PAR = partitive case.

References

Goldberg, A (1995), *Constructions: A Construction Grammar Approach to Argument Structure*, Chicago: University of Chicago Press.

Goldberg, A. (2006), *Constructions at Work: the nature of generalization in Language*, Oxford: Oxford University Press.

Hakulinen, A. & Karlsson, F. (1979), *Nykysuomen lauseoppia* [Syntax of Modern Finnish], Helsinki: Suomalaisen Kirjallisuuden Seura.

Fried, M. (2015), Construction grammar. In: Alexiadou, A. & Kiss, T. (eds), *Syntax. Theory and Analysis. An International Handbook*, Vol. 2, 974–1003, Berlin/Munich/Boston: De Gruyter Mouton.

Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T. R. & Alho, I. (2004), *Iso suomen kielioppi* [Comprehensive Finnish Grammar], Helsinki: Suomalaisen Kirjallisuuden Seura. (Online: http://scripta.kotus.fi/visk/)

Heine, B. (1997), *Possession: Cognitive Sources, Forces, and Grammaticalization*, Cambridge: Cambridge University Press.

Hoffmann, T. (2013), Abstract Phrasal and Clausal Constructions, in Hoffmann, T. & Trousdale, G. (eds), *The Oxford Handbook of Construction Grammar*, Oxford: Oxford University Press, 307–328.

Huumo, T. & Helasvuo, M-L. (2015), On the subject of subject in Finnish, in Helasvuo, M-L. & Huumo, T. (eds), *Subjects in Constructions – Canonical and Non-Canonical*, Amsterdam / Philadelphia: John Benjamins, 15–41.

Stassen, L. (2009), *Predicative Possession*, Oxford: Oxford University Press.

ø

The argument structure of Hungarian BE-possessives

Gréte Dalmi
(Independent Researcher, Budapest)

Keywords: BE-possessives, dative possessor, nominative possessee, dyadic unaccusatives, *psych*-predicates

Current approaches to BE-possessives

Cross-linguistically, predicative possession often manifests itself *via* possessive HAVE or *via* existential BE (Stassen 2009). Hungarian, as most Finno-Ugric languages, uses BE-possessives, i.e. a BE-predicate plus an oblique possessor and a nominative possessee.

The mainstream derivational approach takes existential and possessive sentences akin to copular BE-sentences, adding some additional functional projections to the former two sentence types (Kayne 1993; Szabolcsi 1983, 1994; Belvin & den Dikken 1997; Den Dikken 1997; Myler 2016). The preverbal position of the possessor argument is attributed to extraction from the possessive DP.

The lexicalist approach, accepted here, views possessive predicates as dyadic unaccusatives with a possessor and a possessee argument (Paducheva 2000; Partee & Borschev 2008; Blaszczak 2007). BE-possessives under this approach resemble the *piacere*-class of *psych*-predicates, which have two VP-internal arguments (Belletti & Rizzi 1988; Grimshaw 1990; Harley 2002).

The syntactic and semantic properties of Hungarian BE-possessives

Hungarian BE-possessives resemble BE-existentials and differ from copualar BE-sentences wrt *(i)* Definiteness Restriction; *(ii)* clause negation. Furthermore, BE-possessives in standard Hungarian differ from possessive DPs in not showing anti-agreement (see Den Dikken 1999), *(iii)*.

(i) *Definiteness Restriction*

(1) Péter / A diák / Minden diák ∅ okos.
Peter / the student / every student COP clever
'Peter/ The student / Every student is clever.' BE_{COP}

(2) *A szobá-ban VAN Péter / a diák / minden diák.
the room-INESS BE_{EXIST} Peter / the student /every student
'There is Peter / the student / every student in the room.' BE_{EXIST}

(3) *Nekem VAN Péter-em / a diák-om / minden diák-om.
I-DAT BE_{EXIST} Peter-POSS1SG / the student-POSS1SG / every student-POSS1SG
'I have Peter / my student / every student.' BE_{POSS}

(ii) Clause negation

(4) Péter NEM ∅ okos / tanár. BE_{COP}
Peter NEG COP clever / teacher
'Peter is not clever / a teacher.'

(5) Az asztal-on NINCS kalap. BE_{EXIST}
the table-ALL NEG.BE_{EXIST} hat
'There is no hat on the table.'

(6) Péter-nek NINCS kalap-ja. BE_{POSS}
Peter-DAT NEG.BE_{EXIST} hat-POSS3SG
'Peter has no hat.'

(iii) Anti-agreement

(7) a fiúk-nak a kalap-ja /*-juk.
the boys-DAT the hat-POSS3SG /*POSS3PL
'the boys' hat'

(8) A fiúk-nak VAN kalap-juk /*-ja.
the boys-DAT BE_{EXIST} hat-POSS3PL /*POSS3SG
'The boys have a hat.'

(9) Nekik VAN kalap-juk /*-ja.
they.DAT BE_{EXIST} hat-POSS3PL /*POSS3SG
'They have a hat.'

The advantages of the proposed analysis

(i) In most current analyses, the possessor is extracted from the possessive DP, leaving a trace behind. In the present account, the oblique possessor starts out as a distinct argument of the dyadic unaccusative existential BE-predicate, hence no extraction is required.

(ii) The movement analysis cannot explain why an overt anaphoric possessor can co-occur with an overt oblique possessor in (10).

(10) Péter-nek$_i$ MEG-van [$_{DP}$ az $\ddot{o}_i$ saját problémá-ja].
Peter-DAT PFX-BE_{EXIST} the he.NOM own problem-POSS3SG
'Peter has got a problem of his own.'

Under the proposed analysis, such cases follow from Binding Principle A independently.

References

Belletti, Adriana & Rizzi, Luigi (1988), *Psych*-verbs and Theta-theory, *Natural Language and Linguistic Theory* 6: 291-352.

Belvin, R & Dikken den, Marcel (1997), There, happens to, *be*, *have*, *Lingua* 101: 151–183.

Błaszczak, Joanna (2007), *Phase Syntax. The Polish Genitive of Negation*, Habilitation treatise. University of Potsdam.

Den Dikken, Marcel (1997), The syntax of possession sentences and the verb *have*, *Lingua* 101: 129–150.

Den Dikken, Marcel (1999), On the structural representation of possession and agreement. The case anti-agreement in Hungarian possessed nominal phrases, In: Kenesei, I. (ed.), *Crossing Boundaries: Theoretical Advances in Central and Eastern European Languages*, 137-178, Amsterdam: John Benjamins.

Grimshaw, Jane (1990), *Argument Structure*, Cambridge, MA: MIT Press.

Harley, Heidi (2002), Possession and the double object construction, In: Pica, P. & Rooryck J. (eds), *The Yearbook of Linguistic Variation,* Vol 2: 21-69, Amsterdam: John Benjamins.

Kayne, Richard (1993), Towards a modular theory of auxiliary selection, *Studia Linguistica* 47: 3-31.
Myler, Neil (2016), *Building and Interpreting Possession Sentences*, Cambridge, MA: MIT Press.
Paducheva, Elena (2000), Definiteness Effect: the case of Russian, In: K. von Heusinger & U. Egli (eds): *Reference and Anaphoric Relations*, 133-146, Dordrecht: Kluwer.
Partee, Barbara & Borschev, Vladimir (2008), Existential sentences, BE and GEN NEG in Russian, In: Comomorwski, I., & von Heusinger, K. (eds), *Existence: Semantics and Syntax*, 147–191, New York: Springer.
Stassen, Leon (2009), *Predicative possession*, Oxford: Oxford University Press.
Szabolcsi, Anna (1983), The possessor that ran away from home, *The Linguistic Review* 3: 89-102.
Szabolcsi, Anna (1994), The noun phrase, In: É.Kiss, K. & Kiefer, F. (eds), *The Syntactic Structure of Hungarian*, 179-274, New York: Academic Press.

Predicative possession in Permic

Nikolett F. Gulyás
(Eötvös Loránd University)

Keywords: possession, be-possession, possessive suffix, Komi-Permyak, Udmurt

The present paper examines predicative possession in two Permic (Finno-Ugric, Uralic) languages, Komi-Permyak and Udmurt from a typological point of view. Although the domain of adnominal possession has received much attention in previous works (Bartens 2000, Winkler 2001, Ponomareva 2002), research on predicative possession has been of marginal interest. My aim is to provide a general overview on predicative possession using new data provided by native language experts, and to answer the following questions: i) what are the typical coding features of predicative possession in these languages, ii) under which syntactic or semantic conditions can the clauses deviate from the general pattern?

In both languages, predicative possessive clauses belong to the locational subtype according to Stassen's (2009) typology, since both nominal and pronominal possessors are marked by the genitive case. Additionally, the existential verb *em* 'be' in Komi-Permyak (1) and *vań* 'be' in Udmurt (2) is always present in the clause.

(1)	Nasta-lön	em	ńebög.
	Nastya-GEN	be.PRS.SG	book
	'Nastya has a book.'		
(2)	Peťa-len	apaj-ez	vań.
	Petya-GEN	sister-3SG	be.PRS
	'Petya has a sister.'		

Despite the close genealogic and areal relationship (Bereczki 1998) between the two languages, syntactic and semantic microvariation still can be observed. Possessees typically do not take a possessive suffix in Komi-Permyak, while possessive suffixes are generally overt in Udmurt. In Komi-Permyak, the presence of the possessive suffixes seems to depend on syntactic parameters, whereas semantic features do not affect coding properties. For instance, if the possessee is modified by a demonstrative, the possessive suffix cannot be omitted.

(3)	Nasta-lön	em	eta	perna-ys/*perna?
	Nastya- GEN	be. PRS.SG	this	cross-3SG/cross

‘Nastya has this cross.’

On the contrary, some Udmurt nouns denoting body parts, kinship and abstract terms take a possessive suffix containing the vowel *-y* as opposed to the canonic marker including the vowel *-e* (cf. Edygarova 2010). But this alternation can also be interpreted as morphonological variation instead of using a semantic criterion.

The present study was supported by the National Research, Development and Innovation Office under grant number NKFI K 125282 and NKFI 125206.

References

Bartens, Raija (2000), *Permilaisten kielten rakenne ja kehitys.* Mémoires de la Société Finno-Ougrienne 238. Helsinki: Suomalais-Ugrilainen Seura.

Bereczki, Gábor (1998) [1983], A Volga–Káma-vidék nyelveinek areális kapcsolatai, in Rédei, K. (ed.), *Ünnepi könyv Bereczki Gábor 70. születésnapja tiszteletére*, Budapest: ELTE Finnugor Tanszék, 179-205

Edygarova, Svetlana (2010), *Kategorija posessivnosti v udmurtskom jazyke*. Dissertationes Philologiae Uralicae Universitatis Tartuensis. Tartu: University of Tartu, Doctoral dissertation.

Ponomareva, Larisa (2002), *Fonetika i morfologija mysovskogo-lupinskogo dialekta Komi-permjatskogo jazyka*. Izhevsk: Udmurt State University, Doctoral dissertation.

Stassen, Leon (2009), *Predicative possession*. Oxford: Oxford University Press.

Winkler, Eberhard (2001), *Udmurt*. Munich: Lincom Europa.

WORKSHOP 18

Schedule: Thu 9.00-15.55 (Room 15)

Prehistoric languages beyond the limits of comparative linguistics

Antonio Benítez-Burraco & Ljiljana Progovac
(University of Seville, Spain; and Wayne State University, USA & visiting scholar, MIT, USA)

Keywords Prehistoric languages, comparative method, cultural evolution, adaptive value of languages, language phylogenies

Description of the topic and research questions

The comparative method has enabled linguists to reconstruct extinct languages from the past, but it has limitations and shortcomings that seemingly preclude going back too much in time and which plausibly boil down to its methodological assumptions (heavy reliance on cognates) and the very nature of language change (which tends to lack directionality, and which interacts with the effects of contact and borrowing). Beyond the limits of historical linguistics, what has been said about the languages spoken in our remote past remains controversial, if not plainly speculative. It is clear that there is a need not only of solid, informed hypotheses about remote prehistoric languages, but also of formulating them in such a way that they can be subjected to empirical testing. In truth, this is similar to what can be said of the study of language evolution as a whole. As recently stated in Fitch (2017, 1), despite some remarkable advances, "the field [of language evolution] remains plagued by an unfortunate but persistent belief that scientific work on the topic of language evolution is inherently and irredeemably speculative. This prejudice is unjustified: many other branches of science (e.g. geology or cosmology) study complex historical processes, buried in the past, but this does not stop researchers from developing, debating, and testing hypotheses and ultimately reaching scientific agreement on the basis of converging evidence…". Accordingly, the ultimate aim of this workshop is providing interested researchers an opportunity to debate around the nature of prehistoric languages, and particularly, to develop guidelines as to what counts as a solid, informed, and testable hypothesis regarding prehistoric languages, as opposed to an opinion.

As it is the case with other aspects of the study of language evolution, the consideration of evidence from other fields is expected to refine our view of how languages might have been in the past. Hence, there is also the need to engage in multidisciplinary approaches to prehistoric languages, taking into account, but also moving beyond hypotheses based on typological and philological data. This is in line with current research suggesting that the nature of languages and how they change depends not only on the internal dynamics of linguistic systems, but also on how our brain is and how are the physical and the cultural environments in which we live (Benítez-Burraco and Moran 2018). Accordingly, a second objective of the workshop is to find connections and cross-fertilization potential among approaches looking at different aspects not only of language, but also of human biological and cultural evolution that can improve the reconstruction/characterization of prehistoric languages. Several attempts along these lines deserve to be mentioned, including, but not limited to: using genetic data to support less clear or controversial, deep phylogenetic groupings (e.g. Colonna et al. 2010); exploring correlations between genetic and linguistic variation (e.g. Dediu and Ladd 2007); using language disorders as a proxy of previous stages in the evolution of languages (e.g. Code 2011); pushing back the internal reconstruction of languages relying on grammaticalization theory (e.g. Heine and Kuteva

2007) or on syntactic theory (e.g. Progovac 2015); exploring the links between typological variation and phylogenetic relationships (Dediu and Cysow 2013); linking sociological factors (e.g. nature of social networks, number of speakers) to language structure and complexity (e.g. Wray and Grace 2007); or linking physical factors (e.g. climate, orography, vegetal coverture) to language structure and complexity (e.g. Everett et al. 2016).

Overall, ongoing research in diverse areas suggests that, even if it may not be possible to reconstruct languages from distant periods with certainty, it is at least possible to improve our understanding of the basic structural features of prehistoric languages from remote epochs, and in particular, to advance testable hypotheses about how languages were in the past. Specific research questions explored by the participants of this workshop include (but are not limited to):

- The problems of the comparative method in linguistics and possible ways of circumventing them
- The reliability of deep phylogenetic trees under the light of new typological/philological data
- Linguistic theories that can shed light onto the nature of remote prehistoric languages
- The evolution of the human body and its effect on the structure of prehistoric languages
- The evolution of the human cognition and its effect on the structure of prehistoric languages
- Environmental changes in Prehistory and their effects on language structure
- Social and cultural changes in Prehistory and their effects on language structure
- The links between language change and cultural transitions in Prehistory

References

Benítez-Burraco, A, Moran S (2018) Editorial: the adaptive value of languages: non-linguistic causes of language diversity. Front Psychol. 9: 1827.

Code C (2011) Nonfluent aphasia and the evolution of proto-language. J Neuroling 24: 136-144.

Dediu D, Cysouw M (2013) Some structural aspects of language are more stable than others: a comparison of seven methods. PLoS One 8(1): e55009.

Dediu D, Ladd DR (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. Proc Natl Acad Sci U S A 104: 10944-10949.

Everett C, Blasí DE, Roberts SG (2016) Language evolution and climate: The case of desiccation and tone. J Lang Evol 1(1): 33-46.

Fitch WT (2017) Preface to the Special Issue on the Biology and Evolution of Language. Psychon Bull Rev 24: 1-2.

Heine B, Kuteva T (2007) The Genesis of Grammar. A Reconstruction. Oxford: Oxford University Press.

Progovac, L (2015) Evolutionary Syntax. Oxford University Press.

Wray A, Grace GW (2007) The consequences of talking to strangers: evolutionary corollaries of socio-cultural influences on linguistic form. Lingua 117: 543-578.

From physical aggression to verbal behavior: Language evolution and self-domestication feedback loop

Ljiljana Progovac & Antonio Benítez-Burraco
(Wayne State University, Detroit & University of Seville, Seville)

Keywords: language evolution; self-domestication; reactive aggression; verbal aggression; language disorders

The proposals correlating language evolution with self-domestication traits in humans (e.g. Benítez-Burraco & Kempe 2018, Thomas & Kirby 2018) typically assume a unidirectional causal effect for this correlation, namely that self-domestication (perhaps together with some (other) evolutionary developments affecting the brain) contributed to the emergence of human language readiness and of complex languages via a cultural mechanism. Here we consider the complementary possibility that the emergence of the simplest forms of language/grammar accelerated any processes of self-domestication already underway, in turn fueling the transition to a more complex language. This track gives more agency to language and is in line with claims that language learning has a feedback effect on our cognitive architecture (Deacon 2003, Clarke & Heyes 2017). Originally Darwin (e.g. 1874: 634) attributed an active, causing role to language, stating that "a great stride in the development of intellect will have followed, as soon as … language came into use … The largeness of the brain in man … may be attributed in chief part to the early use of some simple form of language" (see also Pinker & Bloom 1990, Deacon 2003). We thus ask: what might have been those simple forms of language that triggered this avalanche of cognitive changes?

Humans exhibit features of domestication compared to our primate relatives and extinct hominins, particularly, reduced aggression. According to e.g. Benítez-Burraco & Kempe 2018, the less aggressive behavior associated with self-domestication might have facilitated intergroup contacts and enhanced learning and teaching. In truth, our species has not eliminated aggression totally, but, to a large extent, replaced (reactive) physical aggression with verbal aggression. Verbal rituals have persisted throughout recorded history, and continue even today (Locke 2009, Locke & Bogin 2006). Such duels with words, as opposed to fists, are thought to be an adaptive way to discharge aggressive dispositions (Marsh 1978), and to compete without risking physical harm (Locke 2008).

To a great extent, even today verbal confrontation makes use of simple forms of language, including crude compounds consisting of just one verb and one noun (*kill-joy*, *pick-pocket, scatter-brain*, *turn-coat, cry-baby*; Serbian *cepi-dlaka* 'split-hair,', *vrti-guz* 'spin-butt' (fidget), *ispi-čutura* 'drink-flask' (drunkard); *jebi-vetar* 'screw-wind' (charlatan). Significantly, these compounds, which afford a particularly creative strategy for coining names with derogatory reference, have been analyzed as approximations of the earliest and crudest stages of syntax created by proto-Merge (Progovac & Locke 2009 and Progovac 2015, 2016).

Our hypothesis is that looking at the (gradual) emergence of verbal means of aggression (approximated by this kind of compound) helps illuminate the initial steps of language evolution and self-domestication feedback loop. These verbal items would have afforded an adaptive (non-violent) way to compete for status and sex, first by derogating existing rivals and placing prospective rivals on notice; and second by demonstrating verbal skills and quick wittedness, both directly relevant for sexual selection (Progovac & Locke 2009). Gradually, this (reactive) verbal behavior would have driven the selection away from (reactive) physical aggression, and thus enabled/reinforced the effects of self-domestication on human cognition and behavior.

References

Benítez-Burraco, A. & V. Kempe. (2018). The emergence of modern languages: Has human self-domestication optimized language transmission? *Frontiers in Psychology* 17;9:551. doi: 10.3389/fpsyg.2018.00551.

Clarke, E. & C. Heyes. (2017). The swashbuckling anthropologist: Henrich on the secret of our success. *Biology and Philosophy* 32: 289–305. doi: 10.1007/s10539-016-9554-y.

Darwin, C. M. A. (1874). *The Descent of Man, and Selection in Relation to Sex*. New edition, revised and augmented. New York: Hurst and Company.

Deacon, T. W. (2003). Multilevel selection in a complex adaptive system: The problem of language origins. In W. H. Bruce & D. J. Depew (eds.), *Evolution and Learning: The Baldwin Effect Reconsidered*. A Bradford Book. Cambridge, MA: The MIT Press, 81-106.

Locke, J. L. (2008). Cost and complexity: Selection in language and speech. *Journal of Theoretical Biology 251:* 640-652.

Locke, J. L. (2009). Evolutionary developmental linguistics: Naturalization of the faculty of language. *Language Sciences* 31: 33-59.

Locke, J. L. & B. Bogin (2006). Language and life history: A new perspective on the evolution and development of linguistic communication. *Behavioral and Brain Sciences* 29: 259–325.

Marsh, P. (1978). *Aggro: The illusion of violence*. London: Dent.

Pinker, S. & P. Bloom (1990). Natural language and natural Selection. *Behavioral and Brain Sciences* 13: 707–784.

Progovac, L. (2015). *Evolutionary Syntax*. Oxford Studies in the Evolution of Language. Oxford: Oxford University Press.

Progovac, L. (2016). A Gradualist scenario for language evolution: Precise linguistic reconstruction of early human (and Neandertal) grammars. *Frontiers in Psychology* 7:1714. doi: 10.3389/fpsyg.2016.01714.

Progovac, L. & J. L. Locke. (2009). The urge to merge: Ritual insult and the evolution of syntax. *Biolinguistics* 3.2-3: 337-354.

Thomas, J. & S. Kirby. (2018). Self-domestication and the evolution of language. *Biology and Philosophy* 33(1):9. doi: 10.1007/s10539-018-9612-8.

Typology reflects phylogeny

David Gil
(Max Planck Institute for the Science of Human History, Jena)

Keywords: compositional semantics, thematic roles, polity complexity, typology, phylogeny

Cross-linguistic variation may, in some cases, provide a model for the evolution of language. For a given feature, if one of its values is simpler than its more complex alternative, then the simpler value is likely to represent an earlier stage in the evolution of language.

Within the domain of compositional semantics, complexity is reflected in the extent to which the meaning of a juxtaposition of expressions is delimited by various grammatical rules making reference to morphosyntactic strategies such as agreement, flagging, linear order and others. Languages of *high articulation* have an abundance of such rules, substantially limiting the range of possible interpretations, while languages of *low articulation* have fewer such rules, and the range of possible interpretations is broader.

This paper presents the results of an ongoing cross-linguistic experiment measuring the complexity of compositional semantics in the domain of thematic role assignment In the experiment, subjects are asked to judge the truth conditions of sentences in their languages. Stimuli consist of written sentences, each accompanied by two pictures; speakers are asked which picture is correctly described by the sentence. The experiment consists of 32 stimuli measuring the relevance of thematic role

assignment to compositional semantics. For each language, at least 30 subjects are examined, all of lower socio-economic status, in order to control, for effects due to lifestyle and education. The experiment is ongoing; so far it has been conducted on 68 languages.

The main finding is that lower articulation correlates with lower polity complexity, in accordance with the following scale (in which each type is exemplified with one of the languages in the sample):

(1) *Polity Complexity*

high	1	National Languge: World	English
↑	2	National Languge: Primary	Japanese
	3	National Language: Secondary	Tagalog
	4	National Language: Regional Variety	Riau Indonesian
	5	Regional Language: 4 Tiers (Large States)	Javanese
	6	Regional Language: 3 Tiers (States)	Fongbe
	7	Regional Language: 2 Tiers (Larger Chiefdoms)	Tobelo
↓	8	Regional Language: 1 Tier (Petty Chiefdoms)	Yali
low	9	Regional Language: 0 Tiers (Autonomous Bands)	Tikuna

The primary bifurcation in the scale in (1) is between the official national languages of states on the one hand, and non-national languages typically associated with particular regions and/or ethnicities within a larger state on the other. National languages are further classified as world languages used for communication on an inter-continental scale; primary national languages of a state; secondary national languages that typically share their role with another language associated with a historical European colonizer; and colloquial varieties of national languages associated with a particular geographical region within the state. Regional languages are further classified in terms of "jurisdictional hierarchy beyond local community" as defined in the D-Place database providing for five levels of complexity of the polity associated with the language in question, in accordance with the number of tiers associated with the jurisdictional hierarchy.

Since we know independently that earlier sociopolitical structures were generally simpler than their contemporary counterparts, we may conclude that low articulation represents a prior stage in the evolution of compositional semantics.

ø

A test of different methods of homeland detection using synthetic data

Taraka Rama & Søren Wichmann
(Univ. of Oslo; & Univ. of Leiden and Kazan Federal Univ. and Beijing Language Univ.)

Keywords: linguistic homeland, simulations, Bayesian phylogeography, phylogenetics, computational historical linguistics

Application of quantitative methods to infer geographical homelands of language families is a recent enterprise. One method that has received much attention is the Bayesian phylogeography by Lemey et al. (2009). This takes cognate classes as input. Another method is that of Wichmann et al. (2010b), which implements the idea that the homeland probably corresponds to the area of highest linguistic diversity. Here the input is linguistic distances (a modified edit distance, cf. Wichmann et al. 2010a). It

is not clear how to characterize the strengths and weaknesses of these and other methods, which is what we intend to study through simulations.

Software for simulating the evolution of basic vocabulary down a (simulated) phylogenetic tree as well as concomitant semi-random movements of people in real geography is already available (Wichmann, 2017, Holman and Wichmann, 2017). The language migration model operates with movements constrained to locations that are populated today, using this as a proxy for locations that were probably inhabitable or passable in the past as well. This framework allows us to test whether a given method can pick up a preset homeland using information about the locations of current languages, the structure of a family tree, and lexical data. Tests were carried out on a range of methods, but not yet including Bayesian ones. Their performance is indicated here by the order in which they are listed, with the best performing method at the top.

- The homeland is identified following Wichmann et al. (2010b).
- The language with the greatest total linguistic distance to all other languages sits in the homeland.
- The homeland is assigned to the location of a random language.
- The homeland is assigned to the location of the language whose linguistic distances to the other languages correlate best with the geographical distances to other languages.
- The homeland is in the geographical center of the polygon represented by the location of the present languages.

While the method of Wichmann et al. (2010b) performs best among the ones studied, it is vulnerable in some situations, for instance when initial moves take languages in a specific direction away from the homeland rather than in various directions. Figure 1a shows a situation where the inferred homeland (large red square) is almost identical with the real homeland (green square), and Figure 1b shows a poor result (in both pictures red dots are the locations of current languages and white dots places visited by some lineage in the past). In Figure 1b, most movements were to the west, presumably because of a greater carrying capacity of that area as opposed to the mountainous eastern area. Among other, more systematic factors that weaken performance the speed of language migration is most important. In this paper we plan to include two Bayesian phylogeography methods—Landis et al. (2013) and Bouckaert et al. (2012) —in the battery of methods to be tested through synthetic data.

(a) Correct prediction (b) Incorrect prediction

Figure 1: Homeland predictions

References

Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J Drummond, Russell D Gray, Marc A. Suchard, and Quentin D. Atkinson (2012), Mapping the origins and expansion of the Indo-European language family. *Science*, 337(6097), 957-960.

Holman, Eric W. and Søren Wichmann (2017), New evidence from linguistic phylogenetics identifies limits to punctuational change, *Systematic Biology*, 66(4), 604-610.

Landis, Michael J., Nicholas J. Matzke, Brian R. Moore, and John P. Huelsenbeck (2013), Bayesian analysis of biogeography when the number of areas is large, *Systematic Biology*, 62(6), 789-804.

Lemey, Philippe, Andrew Rambaut, Alexei J. Drummond, and Marc A. Suchard (2009), Bayesian phylogeography finds its roots, *PLoS Computational Biology*, 5(9):e1000520.

Wichmann. Søren (2017), Modeling language family expansions, *Diachronica*, 34(1), 79-101.

Shared Mechanisms in the Emergence of Constructions in Language Evolution and Interaction

Michael Pleyer
(Universität Koblenz-Landau)

Keywords: language evolution, usage-based approaches, constructionist approaches, prehistoric languages, interaction

In this talk I argue that interactional and usage-based approaches (e.g, Goldberg 2006; Du Bois 2014; Hopper 2015; Brône & Zima 2014) can shed light on the nature of prehistoric languages. What these approaches all have in common is that they stress the importance of "the emergence of grammatical patterning in on-line production, dialogically, and as a cooperative achievement" (Fried & Östman 2005: 1756). If we want to elucidate the possible structure of prehistoric languages, we need to take into account the online, in-vivo interactional emergence of constructions and their subsequent conventionalization (cf. Pleyer 2017).

More specifically, interactionist and usage-based approaches can help to elucidate the cognitive and social mechanisms involved in the interactive emergence of structure. These mechanisms can not only help to explain how linguistic structures emerge in interaction, but also how these structures emerged and evolved in the course of human evolution. In addition, the can also provide insights into the effect of human cognition and interaction on the structure of early prehistoric languages. In interaction, cognitive processes such as priming, conceptual pacts, intersubjective alignment, local schematization and routinization, microentrenchment and resonance activation can lead to the emergence of a temporary local network of constructions (Brône & Zima 2014). In the case of frequently recurring routines and practices, temporary constructions that are frequently microentrenched also acquire increasing degrees of entrenchment and conventionalization (cf. Schmid 2016).

Over time, this leads to these constructions emerging more easily and more frequently in interactions. This in turn subsequently leads to them becoming increasingly entrenched and turning into structured, conventionalized constructions proper through processes of grammaticalization and constructionalization (cf. Heine & Kuteva 2007; Traugott & Trousdale 2013). In this way, temporary, emergent communicative routines turned into an inventory of entrenched community-wide communicative routines – protolanguage –, which then evolved into a fully grammaticalized and conventionalized structured inventory of constructions shared by a community – language. This process therefore also led to the emergence of the first prehistoric languages.

Crucially, these early, temporary, emergent protoconstructional templates should not be seen as form-meaning pairings in the same way as human constructions. Instead, they are more like pairings of a form with a number of meaning potentials in which communicative routines still have a high-degree of semantic ambiguity and openness. In order to function as communicative units within an interaction, they are still heavily reliant on context. This semantic openness, or underspecification, therefore would also have characterized prehistoric languages (cf. Scott-Phillips 2015; Gil 2017). This view is consistent with analyses of the gesture systems of non-human primates, whose meaning and

function is heavily context- dependent and underspecified (Liebal et al. 2014; Moore 2014). However, early prehistoric languages would have been much more semantically and expressively powerful, as they were based on much more powerful capacities for metacognition, advanced sociocognitive capacities and perspective-taking (Scott-Phillips 2015, Tomasello 2019).

References

Brône, G., & Zima, E. (2014). Towards a dialogic construction grammar: Ad hoc routines and resonance activation. *Cognitive Linguistics, 25*(3), 457-495.

Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25(3), 359-410.

Fried, M., & Östman, J.-O. (2005). Construction Grammar and spoken interaction: the case of pragmatic particles. *Journal of Pragmatics*, 37(11), 1752-1778.

Gil, D. (2017). Isolating-Monocategorial-Associational Language. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (2nd ed., pp. 471-510).

Amsterdam: Elsevier.

Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*.

Oxford ; New York: Oxford University Press.

Heine, B., & Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction*.Oxford: Oxford University Press.

Hopper, P. J. (2015). An Emergentist Approach to Grammar. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 314-327). Malden, MA: Wiley Blackwell.

Hurford, J. R. (2012). *The Origins of Grammar: Language in the Light of Evolution II*. Oxford: Oxford University Press.

Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014). *Primate Communication: A Multimodal Approach*. Cambridge: Cambridge University Press.

Moore, R. (2014). Ape Gestures: Interpreting Chimpanzee and Bonobo Minds. *Current Biology*, 24(14), R645-R647. doi:http://dx.doi.org/10.1016/j.cub.2014.05.072

Pleyer, M. (2017).

Scott-Phillips, T. C. (2015). *Speaking our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special*. Houndmills: Palgrave Macmillan.

Schmid, H.-J. (Ed.) (2016). *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*. Berlin: De Gruyter.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*.

Oxford: Oxford University Press.

Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard: Harvard University Press.

Can the anatomy of the vocal tract help us infer anything about the past linguistic diversity?

Dan Dediu

(Laboratoire Dynamique Du Langage UMR 5596, Université Lumière Lyon 2, Lyon, France)

Keywords: vocal tract anatomy, interindividual variation, linguistic diversity

Present-day linguistic diversity is the result of complex interactions among many forces, constraints and affordances across time (Bentz, Dediu, Verkerk, & Jäger, 2018; Dediu, Janssen, & Moisik, 2017; Everett, Blasi, & Roberts, 2016; Lupyan & Dale, 2016). However, understanding the linguistic past is notoriously difficult, especially when we move beyond a "horizon" a few millennia ago (Renfrew, McMahon, & Trask, 2000), but new approaches and data might (despite issues and controversies) help shed new light on the deep past (Dediu & Levinson, 2012; Jäger, 2015; Pagel, Atkinson, Calude, & Meade, 2013).

Here I suggest that biases (affordances and constraints) rooted in the *anatomy of the vocal tract* (VT), might add useful information. Due to our evolutionary past and demographic processes (Jobling, Hollox, Hurles, Kivisild, & Tyler-Smith, 2013), there is *widespread inter-individual and inter-group variation* (multivariate and distributed mostly continuously) in most aspects of our anatomy, including the VT (Dediu et al., 2017). This variation produces weak biases that are amplified by the repeated use and transmission of languages, with some examples being the influence of the alveolar ridge on click production (Moisik & Dediu, 2017), of the anterior hard palate on the North American English /r/ (Dediu & Moisik, 2019), and of food consistency on labiodentals (Blasi et al., 2019).

Certain aspects of VT anatomy (such as the jaw, dentition, cervical vertebrae and hard palate) are relatively well *conserved in the archeological record*, and their past distribution in certain populations can be reconstructed from currently available data (Baetsen, 2016; Bosman, Moisik, Dediu, & Waters-Rist, 2017). Moreover, there are *genetic influences* on VT anatomy (even if a lot remains to be learned about the precise genes and mechanisms involved) and massive amounts of ancient DNA are becoming available across the world spanning from a few decades to hundreds of thousands of years ago (Reich, 2018), allowing probabilistic inferences from the genetic makeup of past individual to their VT anatomy.

Knowing the VT variation in past populations, the biases these variations generate on language, coupled with the application of quantitative methods for understanding the linguistic past (Bouckaert et al., 2012; Dediu, 2011), should allow us to make (inherently probabilistic) inferences about the properties of past languages and patterns of linguistic diversity. Maybe we can say something about the pre-Bantu-expansion distribution of click languages? The emergence of the typologically intriguing properties of the Australian languages? What vowels Neanderthal languages might have predominantly used? When did labiodentals emerge and spread? Given how little we know about the linguistic past, we should not dismiss out-of-hand any potential source of information, vocal tract anatomy included.

Acknowledgements: The work presented here was Funded by the Netherlands Organisation for Scientific Research (NWO) VIDI grant 276-70-022 (2012-2017), an European Institutes for Advanced Study (EURIAS) Fellowship (2017-2018), and an IDEXLyon Fellowship from the Université de Lyon, France (2018-2021).

References

Baetsen, W. A. (2016). Geometric morphometric analysis of the "skeletal vocal tract": A first step in involving osteoarchaeology in the search for a potential "genetic bias" for language, using two Dutch historical skeletal populations (MSc Thesis). University of Leiden, Leiden, The Netherlands.

Bentz, C., Dediu, D., Verkerk, A., & Jäger, G. (2018). The evolution of language families is shaped by the environment beyond neutral drift. Nature Human Behaviour, 2(11), 816. https://doi.org/10.1038/s41562-018-0457-6

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. Science, 363(6432), eaav3218. https://doi.org/10.1126/science.aav3218

Bosman, A. M., Moisik, S. R., Dediu, D., & Waters-Rist, A. (2017). Talking heads: Morphological variation in the human mandible over the last 500 years in the Netherlands. HOMO - Journal of Comparative Human Biology, 68(5), 329–342. https://doi.org/10.1016/j.jchb.2017.08.002

Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., … Atkinson, Q. D. (2012). Mapping the Origins and Expansion of the Indo-European Language Family. Science, 337(6097), 957–960. https://doi.org/10.1126/science.1219669

Dediu, D. (2011). A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. Proc R Soc B, 278, 474–479. https://doi.org/10.1098/rspb.2010.1595

Dediu, D., Janssen, R., & Moisik, S. R. (2017). Language is not isolated from its wider environment: vocal tract influences on the evolution of speech and language. Language and Communication, 54, 9–20. https://doi.org/doi:10.1016/j.langcom.2016.10.002

Dediu, D., & Levinson, S. C. (2012). Abstract Profiles of Structural Stability Point to Universal Tendencies, Family-Specific Factors, and Ancient Connections between Languages. PLoS ONE, 7(9), e45198. https://doi.org/10.1371/journal.pone.0045198

Dediu, D., & Moisik, S. R. (2019). Pushes and pulls from below: Anatomical variation, articulation and sound change. Glossa: A Journal of General Linguistics, 4(1), 7. https://doi.org/10.5334/gjgl.646

Everett, C., Blasi, D. E., & Roberts, S. G. (2016). Language evolution and climate: the case of desiccation and tone. Journal of Language Evolution, 1(1), 33–46. https://doi.org/10.1093/jole/lzv004

Jäger, G. (2015). Support for linguistic macrofamilies from weighted sequence alignment. Proceedings of the National Academy of Sciences, 112(41), 12752–12757. https://doi.org/10.1073/pnas.1500331112

Jobling, M. A., Hollox, E., Hurles, M., Kivisild, T., & Tyler-Smith, C. (2013). Human evolutionary genetics. New York: Garland Science.

Lupyan, G., & Dale, R. (2016). Why Are There Different Languages? The Role of Adaptation in Linguistic Diversity. Trends in Cognitive Sciences, 20(9), 649–660. https://doi.org/10.1016/j.tics.2016.07.005

Moisik, S. R., & Dediu, D. (2017). Anatomical biasing and clicks: Evidence from biomechanical modeling. Journal of Language Evolution, 2(1), 37–51. https://doi.org/doi:10.1093/jole/lzx004

Pagel, M., Atkinson, Q. D., Calude, A. S., & Meade, A. (2013). Ultraconserved words point to deep language ancestry across Eurasia. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.1218726110

Reich, D. (2018). Who We Are and How We Got Here: Ancient DNA and the new science of the human past. Oxford, UK: Oxford University Press.

Renfrew, C., McMahon, A., & Trask, L. (Eds.). (2000). Time depth in historical linguistics. McDonald Institute, UK.

Global phonetic tendencies point to patterns in speech that predate the African exodus

Caleb Everett

<pdf>

The grammaticalization approach to an etymological reconstruction: A case of Slavic epistemic lexis

Piotr Sobotka
(Nicolaus Copernicus University in Torun, Poland)

Keywords: etymology, Proto-Slavic, grammaticalization, internal reconstruction, comparative method

The comparative method appears to be a fundamental tool for reconstruction languages in the absence of written records. However, the method has limitations in various domains. Only really regular linguistic data can be formally and semantically usefully compared and reconstructed using the method; though, the complete regularity is rare (cf. Kruszewski 1995: 77, 88, Harrison 2003: 213). The limitations of the method are particularly troublesome for cognates that seem to have a different meaning. In exceptional cases, the comparative method may be replaced by the internal reconstruction of languages (cf. Kuryłowicz 1964) or an analysis considering the development trends of a specific group of words as a reconstruction relying on grammaticalization theory.

Etymological research tends to reconstruct both the form and the original meaning of the analyzed word and to investigate formal and semantic relationships that connect a word with another one which precedes it historically and from which it derives. The comparison is the principal methodology of etymology. Grammaticalization is a process considered in the description of the etymology of grammatical words (cf. Vykypěl 2013) or the explanation of possible changes of the word categories and in the interpretation of discourse contexts within which grammaticalization occurs (cf. Hopper 1996). However, an understanding of the mechanisms of grammaticalization could be helpful to draw hypotheses on the etymology of some words to which the comparative methods are considered to be limited, e.g. descendants the Proto-Slavic root **jьst-*, the derivatives of which mean ‘true’, ‘sure’, ‘capital’, ‘to exist’, ‘certainly’, ‘indeed’. There is no consensus between scholars on the original meaning of the root, cf. ‘to be’ (Meillet 1902–1905), ‘own’ (Stang 1949), ‘to see’ (Machek 1955), ‘clear’ (Trautmann 1923), etc.

Since the last cells in the historical development of these derivatives are grammaticalized words, an etymological hypothesis can be tested and verified through grammaticalization mechanisms. Epistemic particles (e.g. Czech ***zajisté*** ‘absolutely’) may be, *inter alia*, depronominal derivatives. The pronominal heritage of the Proto-Slavic root is evidenced by the isolated, however, reasoned Old Polish parallel examples:

(1a)	Proroka-li?	**tak**	ci	jest	powiadam
	prophet-NOM.SG.INTERROG.	yes/so	PATCL	be-PRES.3SG.	tell-PRES.1SG.
	wam (Mt 11:9, Murz.)				
	you-DAT.PL.				
(1b)	Prorokali?	**iście**	powiedam	wam (Mt 11:9, Czech.)	
	prophet-NOM.SG.INTERROG.	so	tell-PRES.1SG.	you-DAT.PL.	
	‘A prophet? Yes, I tell you’				

The root originally referred to the act of speaking: PIE **íd-tó-* ‘this precisely what one is talking about’. Some other examples (cf. PSlav. **pъvьn-* ‘trustworthy’, **pravě* ‘certainly’ → ‘nearly’, etc.) of unclear etymology from the comparative point of view can be explained by the use of grammaticalization mechanisms in the reconstruction.

The method allows to conclude the nature of analyzed words and their etyma as well as to reconstruct motivation paths fully from a proto-word to its reflexes if at least one of them is a grammaticalized unit. It also serves to determine the language development conditions that influence the selection of information that is subject to grammaticalization. Prehistoric languages seem to be rather discursive than grammaticalized, and the direction of the development proceeded from the ostensive-symbolic through communicational to grammatical functions.

The research project funded by the Ministry of Science and Higher Education of the Republic of Poland (National Programme for the Development of Humanities) between 2018 and 2023, no. 11H 17 0066 85.

References

Harrison, S. P. (2003), On the Limits of the Comparative Method, in B. D. Joseph, and R. D. Janda (eds), *The Handbook of Historical Linguistics* (Blackwell Handbooks in Linguistics), Oxford: Blackwell Publishing, 213–243.

Hopper, Paul (1996), Some Recent Trends in Grammaticalization, *Annual Review of Anthropology* 25, 217–236.

Kruszewski, Mikołaj (1995), *Writings in general linguistics* (Amsterdam Studies in the Theory and History of Linguistic Science Series I, Amsterdam Classics in Linguistics, 1800–1925, vol. 11). Amsterdam / Philadelphia: John Benjamins Publishing.

Kuryłowicz, Jerzy (1964), On the Methods of Internal Reconstruction, in H. G. Lunt (ed.), *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., August 27–31, 1962*, The Hague / London: Mouton, 9–36.

Machek, Václav (1955), Verbes slaves pour désigner les cinq sens, *Sbornik Prací Filosofické Fakulty Brnéské University (Studia Minora Facultatis Philosophicae Universitatis Brunensis)* A(3), 22–35.

Meillet, Antoine (1902–1905), *Études sur l'étymologie et le vocabulaire du vieux slave*, vol. 1–2, Paris: Librairie Émile Bouillon.

Stang, Christian S. (1949), L'adjectif slave *istъ*, *Norsk tidsskrift for sprogvidenskap* 15, 343–351.

Trautmann, Reinhold (1923), *Baltisch-slavisches Wörterbuch*, Göttingen: Vandenhoeck & Ruprecht.

Vykypěl, Bohumil (2013), Grammaticalization and etymology, in B. Vykypěl, and V. Boček (eds), *Etymology: An old discipline in new contexts* (Studia etymologica Brunensia, vol. 16), Praha: Nakladatelství Lidové noviny, 9–12.

Imagining a prehistoric world and its language: Creating the language of Beama for the film *Alpha*

Christine Schreyer
(University of British Columbia – Okanagan)

Keywords: prehistoric languages, constructed languages, endangered languages, proto-languages, Beama

In this paper I describe my work inventing the fictional prehistoric language of Beama for the Hollywood film Alpha (2018). Alpha is set in a prehistoric world, 20,000 years in the past in what is now Europe and as the movie unfolds we see one theory on the domestication of wolves into dogs. My

role in the film was language consultant and I was tasked with developing the language spoken by the Cro-Magnon characters. Using available data from Proto-Nostratic, Proto-Eurasiatic, and Proto-Dené-Caucasian, I blended academic research on linguistic typology with creative world-building, such as considering the imagined communities of speakers and their physical and social environments, to develop the language spoken in the film. Constructed languages, also known as conlangs, are often identified on their manner of creation, as well as their purpose. Languages that are developed without reference to other natural languages are known as *a priori* languages while those that utilize details from other languages are known as *a posteriori* languages. The language of Beama falls into the latter category as I drew on information from the phonology, morphology, and syntax of the proto-languages listed above. As well, the language of Beama falls into the category of artistic language, as its purpose was to be used in a fictional setting. There are many reasons a language may be developed for a film, including authenticity and character identity. In previous research I have argued that fandoms of created language communities may be models for communities of endangered languages due to the novel approaches they take in language learning (Schreyer 2011 and 2015). However, now that I have developed a prehistoric created language, I question what, if anything, invented prehistoric languages, particularly those situated in a well-developed imagined world such as the world of Alpha, might contribute to questions on the origins of natural prehistoric languages.

References

Schreyer, Christine (2011), Media, Information Technology and Language Planning: What can endangered language communities learn from created language communities?, *Current Issues in Language Planning* 12(3), 403-425.

Schreyer, Christine. (2015), The Digital Fandom of Na'vi Speakers, *Transformative Works and Cultures*,18, http://journal.transformativeworks.org/index.php/twc/article/view/610/512 http://dx.doi.org/10.3983/twc.2015.0610.

WORKSHOP 19

Schedule: Fri 9.00 – Sa 12.25 (Room 15)

Psycholinguistic aspects of complex words

Christina Gagné; Lívia Körtvélyessy; Thomas Spalding; & Pavol Štekauer
(University of Alberta; P.J.Safarik University; University of Alberta; & P.J.Safarik University)

Keywords: psycholinguistics, complex-word formation, complex-word-interpretation, complex-word processing, competition

The aim of the proposed workshop is to provide a state of the art in the field of psycholinguistic investigation of various aspects of complex-word formation and interpretation and to map the latest advancements in theoretical and empirical research in this field. The following brief overview of the directions and focus of previous research thus also serves as a scope of topics to be discussed at the workshop.

Complex words have been an object of intensive research within various theoretical frameworks since the 1960s, primarily in terms of their word-formation characteristics and their interpretation principles (e.g., Lees 1960, Zimmer 1971, Downing 1977, Levi 1978). This latter area attracted the interest of psycholinguists who came up with various models mainly (but not exclusively) accounting for the interpretation of novel context-free N+N compounds.

First, an approach putting stress on the role of the *head noun* (head concept) functioning as a *schema* with a certain number of *slots* that are filled by the modifier values. This model is called a *feature model* or *schema model* (e.g., Allen 1978, Cohen & Murphy 1984, Murphy 1988, 1990, Smith & Osherson 1984, Smith, Osherson, Rips & Keane 1988), and is based on the *semantic representations* of the constituent nouns and associated *encyclopaedic knowledge*. According to this approach, differences in the interpretability of context-free primary compounds are related to the *relative salience* of particular meaning aspects (slots in a schema-based model, attributes in the feature-based model). Thus, 'high-interpretable' compounds are based on more salient meaning aspects of their motivating constituent and 'low-interpretable' compounds are based on less salient meaning aspects (Coolen, van Jaarsveld & Schreuder 1991).

Second, the *relation model* (Gagné & Shoben 1997, Gagné 2001) emphasizes the central role of *thematic relations* between compound constituents and the language speaker's *linguistic knowledge* of the relative strength of the individual thematic relations bound to a particular *modifier concept*. This knowledge facilitates the interpretation of compounds by preferring the interpretation based on a thematic relation which is more readily available to the modifier concept.

Third, the *analogy-based model* (Derwing & Skousen, 1989, Skousen 1989, Ryder 1994, Wisniewski 1996, etc.) accounts for the interpretation of novel, context-free compounds primarily by lexicalized (i.e. established, institutionalized) compounds that serve as certain *interpretation patterns* or models).

Recent decades have witnessed an elaboration of these fundamental approaches and the introduction of new topics, such as the issue of *semantic transparency* (e.g., Libben 1998, Pollatsek and Hyönä 2005, Blais-Gonnerman 2012, El-Bialy, Gagné and Spalding 2013). In such an approach, it may be assumed that semantically transparent complex words are those whose constituents are used in

one of their fundamental meanings and the extent to which the meanings of the constituents pattern semantically with other members of the positional families of words (Libben 2014). Gagné and Shoben (1997) relate semantic transparency of a particular relation in conceptual combinations to the *frequency* of its occurrence with the *modifier concept*.

The definition of the notion of semantic transparency is accompanied with a number of related factors. Thus, it is assumed that the relative *modifier-head position* of the transparent-opaque constituents may also affect the interpretation of complex words (Libben et al. 2003, Marelli and Luzzatti 2012). Views also differ in assigning relative importance to complex word constituents for the process of interpretation. For example, while the CARIN model (Gagné and Shoben 1997, Gagné 2001, Spalding and Gagné 2008) attributes a crucial role to the modifier concept, its elaborated version, the RICE model (Spalding et al. 2010) assigns higher significance to the modifier only in suggesting potential relations, while both modifier and head are believed to be crucially involved in their evaluation. Tarasova (2013), too, assumes that both of them are important (even though with different roles) for the interpretation of compounds. Contrary to this, Körtvélyessy, Štekauer & Zimmermann (2015) maintain that semantic transparency cannot be restricted to semantic information provided by the complex word's constituents and that it depends on the presence/absence of morphological representation of the relational component in complex words.

Another new area of psycholinguistic research into complex words is the issue of *competition.* This direction of research can be set into the broader problem of competition in biological and social systems (cf. MacWhinney, Malchukov & Moravcsik, 2014). In psycholinguistic research into complex words it is manifested as competition among various strategies of forming new complex words (Körtvélyessy & Štekauer 2014), on one hand, and competition among various possibilities of interpretation of novel complex words, on the other. This direction of research is captured in various models such as the *Competition-Among-Relations-In-Nominals* (CARIN) theory of conceptual combination (Gagné and Shoben 1997, Gagné 2001, Spalding and Gagné 2008), its updated version in the form of the *Relational-Interpretation-Competitive-Evaluation* (RICE) model (Spalding, Gagné, Mullaly and Ji 2010), and the competition-based *meaning predictability* theory (Štekauer 2005). Another aspect of competition concerns possible competition among representations of the complex word and its constituents (e.g., Gagné, Spalding, Nisbet & Armstrong, 2018).

Much attention has also been paid to a multiplicity of problems related to the processing and representation of complex words (for a representative overview see Baayen 2014).

Furthermore, considerable attention has been paid to the role of *metaphor and metonymy* in complex-word formation and complex-word interpretation (e.g., Benczes 2006) and the inter-relation between word-formation and word-interpretation (Štekauer 2017).

This brief outline of psycholinguistic research into complex words thus establishes a framework for the basic scope of topics to be discussed within the proposed workshop:

- The role and the relative significance of the head, the modifier and the thematic relations in interpretation of complex words.
- Competition among various strategies of complex-word formation and various interpretation possibilities / meaning predictability of novel context-free complex-words, especially with regard to the struggle between the contradictory tendencies of economy of expression vs. semantic transparency.
- The role of metaphor and metonymy in complex-word formation and their influence upon complex-word interpretation.
- The interrelation between complex-word formation and complex-word interpretation.
- The concept of semantic transparency in relation to complex-word interpretation.

- The influence of psychological factors, such as creativity, upon complex-word formation and complex-word interpretation.
- The role of inference in complex-word interpretation.
- Processing and representation of complex words.
- Empirical and experimental methods of psycholinguistic research into complex-word formation and interpretation.

References

Allen, Margaret R. 1978. *Morphological Investigations.* Doctoral dissertation, University of Connecticut, Storrs, CT.

Baayen, Harald R. 2014. Experimental and psycholinguistic approaches to studying derivation. In: Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivation*. Oxford: Oxford University Press.

Benczes, Réka. 2006. *Creative Compounding in English: The Semantics of Metaphorical and Metonymical Noun-noun Combinations*. Amsterdam: John Benjamins.

Blais, Mary-Jane and Laura M. Gonnerman. 2012. The Role of Semantic Transparency in the Processing of Verb-particle Constructions by French-English Bilinguals. In: Naomi Miyake, David Peebles and Richard P. Cooper (eds.), *Building Bridges Across Cognitive Sciences Around the World*, 1338-1343. *CogSci 2012 Proceedings*.

Cohen, Benjamin and Gregory L. Murphy, 1984. Models of Concepts. *Cognitive Science* 8: 27-58.

Coolen, Riet, Henk van Jaarsveld and Robert Schreuder. 1993. Processing Novel Compounds: Evidence for Interactive Meaning Activation of Ambiguous Nouns. *Memory and Cognition* 21: 235-246.

Derwing, Bruce L. and Royal Skousen. 1989. Morphology in the mental lexicon: A new look at analogy. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology* 2, 55-71. Dordrecht: Foris.

Downing, Pamela. 1977. On the Creation and Use of English Compound Nouns. *Language* 4: 810-842.

El-Bialy, Rowan, Christina L. Gagné and Thomas L. Spalding. 2013. Processing of English compounds is sensitive to the constituents' semantic transparency. *Mental Lexicon*, Vol. 8/1: 75-95.

Gagné, Christina L. 2001. Relation and Lexical Priming During the Interpretation of Noun-Noun Combinations. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition* 1: 236-254.

Gagné, Christina L. and Edward J. Shoben, 1997. Influence of Thematic Relations on the Comprehension of Modifier-Noun Combinations. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition* 1: 71-87.

Gagné, Christina L., Thomas L. Spalding, Kelly A. Nisbet and Caitlin Armstrong. 2018. Pseudo-morphemic structure inhibits, but morphemic structure facilitates, processing of a repeated free morpheme. *Language, Cognition, and Neuroscience (online first)*.

Körtvélyessy, Lívia and Pavol Štekauer *2014.* Derivation in a social context. In: Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*. Oxford: Oxford University Press, 407-423.

Körtvélyessy, Lívia, Pavol Štekauer and Július Zimmermann. 2015. Word-formation strategies: semantic transparency vs. formal economy. In: Laurie Bauer, Lívia Körtvélyessy and Pavol Štekauer (eds.), 2015. *Semantics of Complex Words*. Dordrecht: Springer, 85-114.

Lees, Robert B. 1960. *The Grammar of English Nominalizations*. Bloomington, IN: Indiana University Press.

Levi, Judith N. 1974. On the alleged idiosyncrasy of non-predicate NP's. *Papers from the 10th Regional Meeting, Chicago Linguistic Society*, 402-415. Chicago: Chicago University Press.

Libben, Gary. 1998. Semantic Transparency in the Processing of Compounds: Consequence for Representation, Processing and Impairment."*Brain and Language* 61: 30-44.

Libben, Gary. 2014. The nature of compounds: a psychocentric perspective. *Cognitive Neuropsychology* 31/1-2: 8-25.

Libben, Gary, Martha Gibson, Yeo-Bom Yoon and Dominiek Sandra. 2003. Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language* 84/1: 50-64.

Marelli, Marco and Claudio Luzzatti. 2012. Frequency effects in the processing of Italian nominal compounds: modulation of headedness and semantic transparency. *Journal of memory and language* 66: 644-664.

MacWhinney, Brian, Andrej Malchukov and Edith Moravcsik (eds.). 2014. *Competing Motivations in Grammar and Usage.* Oxford: Oxford University Press.

Murphy, Gregory L. 1988. Comprehending Complex Concepts. *Cognitive Science* 12: 529-562.

Murphy, Gregory L. 1990. Noun phrase interpretation and conceptual combination. *Journal of Memory and Language* 29: 259-288.

Pollatsek, Alexander and Jukka Hyönä. 2005. The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes* 20: 261-290.

Ryder, Mary Ellen. 1994. *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*. Berkeley: University of California Press.

Skousen, Royal. 1989. *Analogical modelling of language*. Dordrecht: Kluwer.

Smith, Edward E. and Daniel N. Osherson. 1984. Conceptual combination with prototype concepts. *Cognitive Science* 8: 337-361.

Smith, Edward E., Daniel N. Osherson, Lance J. Rips and Margaret Keane. 1988. Combining prototypes: A selective modification model. *Cognitive Science* 12: 485-527.

Spalding, Thomas L. and Christina L. Gagné. 2008. CARIN theory reanalysis reanalyzed: a comment on Maguire, Devereaux, Costello, and Cater (2007). *Journal of Experimental Psychology, learning, Memory and Cognition* 34: 1573-1578.

Spalding, Thomas L., Christina L. Gagné, Alison C. Mullaly and Hongbo Ji. 2010. Relation-based interpretations of noun-noun phrases: a new theoretical approach. In: Susan Olsen (ed.), *New impulses in word-formation (Linguistische Berichte Sonderheft 17)*, 283-315. Hamburg: Buske.

Štekauer, Pavol. 2005. *Meaning Predictability in Word-Formation*. Amsterdam-Philadelphia: John Benjamins.

Štekauer, Pavol. 2017. Competition in natural languages. In: Juan Santana-Lario and Salvador Valera (eds.), *Competing Patterns in English Affixation*, 15-32. Bern: Peter Lang.

Tarasova, Elizaveta. 2013. *Some new insights into the semantics of English N+N compounds*. Unpublished PhD thesis. Victoria University of Wellington.

Wisniewski, Edward J. 1996. Construal and similarity in conceptual combination. *Journal of Memory and Language* 35: 424-453.

Zimmer, Karl E. 1971. Some General Observations about Nominal Compounds. *Working Papers on Language Universals.* Stanford University 5: 1-21.

On the influence of creativity upon the formation of complex words

Pavol Štekauer, Lívia Körtvélyessy & Pavol Kačmar
(P.J. Šafárik University, Košice)

Keywords: creativity, word-formation, experimental research, competition, economy vs. semantic transparency

Acknowledgement: The research presented in this paper has been implemented with support of the APVV-16-0035 grant.

Despite extensive morphological research into complex-word formation in recent decades language-users' creativity has not been studied as a word-formation factor yet. This paper addresses how language users of unequal creativity characteristics coin new complex words.

The paper reviews basic theoretical principles, describes the method of experimental research and discusses the results obtained from 309 university undergraduates whose mother tongue is Slovak) at two stages:

i) The informants undertook the *Torrance Test of Figurative Thinking* for assessment of four indicators of creativity: elaboration, fluency, flexibility and originality. Each indicator was evaluated and, by application of the median value, both age groups of informants were further divided into two subgroups for each indicator.
ii) The respondents (English proficiency at the level of B1-B2) were also tested for the formation of potential English complex words based on a word-formation experiment consisting of three sets of tasks aiming at the formalization of the category Agents:
    a) multiple choice from among applicable word-formation rules;
    b) naming based on a description of a person's unusual activity; and
    c) naming based on drawings of people in unusual situations.

The above ultimately targets the resolution of competition in natural languages (e.g. Aronoff 2013, Lindsay & Aronoff 2013, Mac Whinney et al. 2014, Štekauer 2017). In word-formation, it manifests itself as, *inter alia*, competition between various naming strategies reflecting the conflict between *economy of expression* and *semantic transparency*. We view semantic transparency as a degree to which the constituents of a new complex word enable a language user to predict its meaning, as in the following example:

i) *novel writer* is semantically transparent because it predicts the meaning as 'an Agent who performs an Action the Result of which is a novel', but its form is not economical;
ii) *novelist* is more economical but less transparent, because it may mean, for example, 'an Agent who performs an Action of writing/reading/stealing/selling/buying/collecting/…. books'
iii) *writer* is also more economical than the first option but less transparent. Here the Result of Action may be not only a book but also a letter, a poem, a newspaper article, etc;
iv) *write*$_{N}$ is the most economical and the least transparent option. This option offers dozens of possible interpretations.

The preference for any of these options reflects a language user's naming strategy, i.e. the preference for a more transparent or a more economical naming strategy. The results of our research thus

i) identify the similarities and differences between each two groups established by the *Torrance Test* in terms of the above-mentioned opposite tendencies, and compare the data for two age groups of respondents; and
ii) analyze the formation of new words at a more detailed level, i.e. in terms of the specific Agent-forming rules chosen by the individual groups of respondents, such as V+*er*, N+*ist*, N+*ian*, V+*ee*, N+*man*, V>N, synthetic compound, primary compound, exocentric compound, blend, etc.

References

Aronoff, Mark (2013), Competition and the Lexicon. To appear in *Proceedings of the Annual Meeting of La Società di Linguistica Italiana*. Prepublication paper.

Dressler, Wolfgang U. (2005), Word-formation in natural morphology, in P. Štekauer and R. Lieber (eds), *Handbook of word-formation*. Dordrecht: Springer, 267–284.

Libben, Gary (2014), *The nature of compounds: a psychocentric perspective*. Manuscript.

Lindsay, Mark and Mark Aronoff (2013), Natural Selection in Self-Organizing Morphological Systems, in F. Montermini, G. Boyé and J. Tseng (eds), *Morphology in Toulouse: Selected Proceedings of Décembrettes 7*, Munich: Lincom Europa, 133-153.

MacWhinney, Brian, Andrej Malchukov and Edith Moravcsik (eds) (2014), *Competing Motivations in Grammar and Usage*. Oxford: Oxford University Press.

Štekauer, Pavol (2017), Competition in natural languages, in J. Santana-Lario and S. Valera (eds), *Competing Patterns in English Affixation*. Bern: Peter Lang; 15-32.

Property inference from heads to compound words and modifier-noun phrases

Thomas L. Spalding & Christina L. Gagné
(University of Alberta & University of Alberta)

Keywords: modification effect, compound words, modifier-noun phrases, property verification, concepts

In this talk, we investigate how people infer properties of compound words and modifier-noun phrases from their knowledge of properties of the head. In general, one of the main purposes of concepts is to allow the inference of properties true of the concept in general to specific instances of that concept (see, e.g., Murphy, 2002). For example, knowing that birds in general fly, one can infer that a particular new bird has the capacity for flying, even though it is not flying at the moment. Similarly, one might reasonably infer that a new type or subcategory of bird would fly. On the other hand, people expect that different names for things reflect real underlying differences (Synonomy Avoidance, Carstairs-McCarthy, 2010; Principle of Contrast, Clark, 1993; Mutual Exclusivity Principle, Markman, 1989). Hence, one might imagine that a new type of bird is expected to be different from birds in general, and this expectation might impact the inferences that one would draw from the head noun to the compound word or phrase. We review a series of property inference studies in which we investigated the interplay between these conceptual inference and contrastive naming principles. These

studies use a variety of judgment tasks to investigate the extent to which people believe the same properties to be true of the head and of the compound word or modifier noun phrase. As one example judgment task, participants might be asked to judge the extent to which the statement "Candles are made of wax" is true or they might be asked to judge the extent to which the statement "Purple candles are made of wax" is true, such that the degree to which "made of wax" is believed to be true of candles can be compared with the degree to which it is believed to be true of purple candles. These studies show that people do infer properties from the head to the compound or phrase, but that they do so in ways consistent with their expectation of contrast between the head and the compound or phrase. In general, properties true of the head are less true of the compound or phrase, and properties false of the head are less false of the compound or phrase. In particular, property inference is affected by the extent to which the compound is established, by how semantically similar the compound is to the head, and by other variables that impact one's expectation of contrast.

References

Carstairs-McCarthy, A. (2010). *The evolution of morphology*. Oxford, UK: Oxford University Press.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge, UK: Cambridge University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press, Cambridge, MA.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Morphological superstates and the dynamic nature of morphological structure

Gary Libben

It is often assumed in psycholinguistic research on lexical processing that complex words in the mind are static structured representations. Yet, a great deal of on-line morphological processing research has yielded patterns of results in which the extent to which putative morphological constituents play a role in lexical processing depends on the interplay of a variety of stimulus and participant factors (Gagné, 2017).

In this presentation, it is claimed that this variability is in fact a core characteristic of morphological knowledge. Under this view, words in the minds of language users do not have univocal morphological structures and have many more of the properties of actions than the properties of representations. This perspective leads us away from a view of words in the mental lexicon as thing-like representations and more toward the kinds of explanations of mental activities originally outlined by William James (James, 1890). As actions, morphologically complex words may vary quite considerably both within and across individuals in the manner in which they are morphologically structured. In order for such variability to be possible, complex words must be described as though they were in a lexical superstate (Libben, 2017). Such a lexical superstate defines the kinds of morphological configurations that a word can have, based on patterns of language use. It is not, however, any one of those morphological configurations in particular.

In this report, the superstate perspective on morphological representation and processing is applied to an analysis of data from lexical comprehension and production in English, French, and German, using a combined progressive demasking and word typing paradigm. The consequences of this

perspective for the modeling of complex word knowledge and processing among bilinguals are also discussed.

References

Gagné, Christina. (2017), Psycholinguistic Approaches to Morphology, in M. Aronoff (ed), *Oxford Research Encyclopedia of Linguistics,* Oxford: Oxford University Press.

James, William, 1890, *The Principles of Psychology*, New York: Dover.

Libben, Gary (2017), The quantum metaphor and the organization of words in the mind, *Cultural Cognitive Science*, 1, 49-55.

Examining the scope of morpheme-based processing during complex word recognition: Which words have multiple morphemes, and who uses them?

Robert Fiorentino
(University of Kansas)

Keywords: morphological processing, compounding, inflection, priming, EEG

Alternative accounts make fundamentally different claims regarding to what extent the processing of complex words involves decomposition into morpheme-level representations, and combinatorial operations involving these decomposed representations. While there is increasing evidence that the processing of at least some complex words involves access to morpheme-level representations, questions remain regarding the conditions under which complex word recognition involves morpheme-level processing (see Amenta and Crepaldi, 2012 for a review). I will present findings from a set of experiments examining the scope of morpheme-based processing, focusing on test cases where debate remains. One issue that these experiments examine is to how and to what extent the semantic relationships between a complex word and its constituents constrain the morpheme-based processing of complex words. These experiments probe whether morpheme-based processing is limited to complex words with relatively transparent meaning relations between the complex words and its constituents, or extends to complex words with less transparent meaning relationships between the constituents and complex word. A second issue these experiments engage is the extent to which morpheme-based processing of words is modulated by lexicality, probing whether novel complex words undergo morpheme-based processing on par with known words. Novel complex words provide an interesting test case, as no previously-learned associations between the whole word and its constituents exist for novel complex words; research on novel complex words thus allows us to examine to what extent apparent effects of decomposition are likely to be due to the existence of pre-existing lexical associations, or are indeed due to across-the-board decomposition of words into potential morpheme constituents.

I will bring together previous findings and new evidence from psycholinguistic experiments, primarily using priming and lexical decision, together with electroencephalography (EEG) which suggests that the processing of both known English compounds (e.g., *teacup*) and novel compounds (e.g., *drugpack*) involves access to morpheme-level representations. The findings also establish that this access cannot be accounted for as resulting from purely orthographic or semantic relationships, or due to previously-established lexical associations between the whole compound and its constituents. I will also argue that morpheme-based processing extends to inflected words, demonstrating that

inflected verbs undergo morpheme-based processing during visual word recognition in native speakers, a finding we have also shown to extend to adult second language learners (Coughlin et al., 2019). Although the experiments mentioned above suggest a role for morphemes in the processing of complex words in the general case, I will also raise the question of to what extent the engagement of morpheme-based processing depends on properties of the individual. Although there has been increasing attention to the examination of individual differences in language processing (see, e.g., Boudewyn, 2015 for discussion), the nature and origins of individual variability in morphological processing remain open questions (e.g., Feldman et al., 2015). The findings from the investigation of morphological processing discussed in this presentation carry the potential to inform our understanding of the scope of morpheme-based processing during word recognition.

References

Amenta, Simona, & Crepaldi, Davide, (2012), Morphological processing as we know it: An analytical review of morphological effects in visual word identification, *Frontiers in Psychology* 3:232, doi: 10.3389/fpsyg.2012.00232

Boudewyn, Megan A., (2015), Individual differences in language processing: Electrophysiologial approaches, *Language and Linguistics Compass* 9(10), 406–419.

Coughlin, Caitlin E., Fiorentino, Robert, Royle, Phaedra, & Steinhauer, Karsten, (2019), Sensitivity to inflectional morphology in a non-native language: Evidence from ERPs, *Frontiers in Communication* 4:21, doi: 10.3389/fcomm.2019.00021

Feldman, Laurie B., Milin, Petar, Cho, Kit W., Moscoso del Prado Martín, Fermín, & O'Connor, Partrick A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience*, 9:111, doi: 10.3389/fnhum.2015.00111

The role of metaphor and metonymy in compound verbs in English and Bulgarian

Alexandra Bagasheva

Compound verbs (CVs) are a notoriously heterogeneous class word-formationally, arising from backformation, conversion, and compounding. There are contentions for minimal meaning generation conditions and detectable linguistic and cognitive procedures associated with particular word-formation patterns (Guevara and Scalise 2004 – compounding and back-formation; Farrell 2001 – conversion; Lieber 1981, 2004 – conversion and compounding; Nagano 2007 – conversion and back-formation; Sándor 2007 – conversion; Schmid 2011 – compounding, conversion and back-formation). In view of such claims, the first question addressed is whether there is any dependence between the role of metaphor and metonymy and the three derivation processes or whether the said cognitive mechanisms are operative across the board in the constitution of CVs.

Clark and Clark (1979) and Nagano (2007) analyze the meaning mechanisms involved in converted CVs, resulting in infelicitous lexical semantic interpretations: e.g. Nagano (2007: 65) classifies as Locatum the CVs *air-condition*, *face-lift*, *ill-treat*, *pressure-treat*, *tripletongue*, *turbocharge* and *valet-park*, while *tailor-make, jam-pack* and *hard-boil* appear as Goal and *prize-fight* as Sound-symbolism. Employing Cognitive Linguistics analytical procedures for the role of figurativity in CV composition and their lexical semantic interpretation, we demonstrate that the two mechanisms are indiscriminately operative in CV constitution (despite associations of metonymy exclusively with conversion (Sándor

2007; Schmidt 2011), e.g. *bear hug*, *cold shoulder*, *butt call*, etc.): *sun dry* and *rough dry*, *deep six*, *cherry pick*, etc. (all figuratively motivated, though derived via different processes). An explanation for this across-the-board utilization of metaphor and metonymy is sought in analogy as the basic "productive schema for the creation of genuine" (Lamberty and Schmid 2013) CVs. The role of these cognitive mechanisms in CVs, in this dictionary-cum-corpus study, is operationalized as 'degree of motivatability,' as defined by Leisi (1955) and used by Sanchez (2008) in terms of transparency. This is relativized against the domains to which CVs belong as naming units.

The second research question focuses on contrasts and similarities in the prominence of metaphor and metonymy in the CV lexicons of two languages: English and Bulgarian, in which CVs are consistently categorially and morphologically right-headed, but display wide variability in terms of semantic endo/exocentrcity (e.g. Engl. *red-shirt*, *deep-six*, *benchmark*, *cold finger*, etc.; Bulg. *glavoblaskam* "head trash" *worry too much or think things over*; *rakopolagam* "hand put" *ordain*, *umopomrachavam* "mind darken" *go crazy*, etc. Analyses of the data set (250 English and 75 Bulgarian CVs) reveal that metonymy and/or metaphor are operative in the formation of semantically exocentric compounds and this is correlated with the domains in which these are used as naming units. CVs are used in diverse semantic domains from purely physical activities such as Engl. *island-hop*, *chain-smoke*, *deep-fry* and Bulg. *vodosnabdyavam* (water provide) *supply with water* and *kravodaryavam* (blood donate) *donate blood*, etc. to abstract domains like emotional states and decision-making, as well as interpersonal relations, e.g. Engl. *hag-ride*, *cherry-pick*, *short-change*, *spoon-feed*, *browbeat*, etc., Bulg. *bogotvorya* "god create" *worship* and *slovobludstvam*, "word fornicate" *speak nonsense*. Metonymically and/or metaphorically motivated CVs in both languages abound in more abstract domains mapping the socio-cultural sphere, and are rarer in physical, directly actional domains.

References

Clark, Eve and Herbert Clark. 1979. When Nouns Surface as Verbs. In *Language* 55/4: 767–811.

CoCA: Davies, Mark. 2008-. The Corpus of Contemporary American English (COCA): 425 million words, 1990-present. http://www.americancorpus.org. Last visited: 28 December 2018

Farrell, Patrick. 2001. Functional Shift as category underspecification. In *English Language and Linguistics* 5: 109–130.

Guevara, Emiliano and Sergio Scalise. 2004. V-Compounding in Dutch and Italian. In *Cuadernos de Lingüística del Instituto Universitario Ortega y Gasset*, XI. Madrid: I.U. Ortega y Gasset.

Lamberty, Angela and Hans-Jörg Schmid. 2013. Verbal Compounding in English: A Challenge for Usage-Based Models of Word-Formation? In *Anglia* 131/4: 591–626.

Leisi, Ernst. 1955. *Das heutige Englisch: Wesenszüge und Probleme*. Heidelberg: Winter.

Lieber, Rochelle. 1981. Morphological Conversion within a Restrictive Theory of the Lexicon. In Mootgart, M., Hulst, H. and Hoekstra, T. (Eds.), *The Scope of Lexical Rules*. 161–200. Foris Publications.

Lieber, Rochelle. 2004. *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.

Nagano, Akiko. 2007. Marchand's analysis of back-formation revisited: back-formation as a type of conversion. In *Acta Linguistica Hungarica,* 54/1: 33–72.

Sanchez, Christina. 2008. *Consociation and Dissociation: An Empirical Study of Word-Family Integration in English and German*. Tübingen: Narr.

Sándor, Martsa. 2007. Construction of Meaning during Conversion. In *Proceedings HUSSE/8 Conference* held at the University of Szeged, January 25-28, 2007.

Schmid, Hans-Jörg. 2011. *English morphology and word-formation. An introduction*. Berlin: Erich Schmidt.

Macro-level competition in word-formation

Vincent Renner
(University of Lyon)

Keywords: word-formation, onomasiology, competition, English

The study of competition in word-formation has attracted wider interest in recent years, but most of the research is limited in scope to a micro-level of rivalry between individual patterns of affixation (see e.g. Bauer 2006, Lindsay & Aronoff 2013, Arndt-Lappe 2014, Santana-Lario & Valera-Hernández 2018). Perspectives have been broadened by pioneer work focusing on macro-level competition, i.e. rivalry between different processes of word-formation – affixation vs conversion (Bauer et al. 2010, Naya 2017), affixation vs conversion vs deaffixation (Nagano 2008), compounding vs blending (Renner 2006), compounding vs conversion (Tribout & Villoing 2014) –, but the field still lacks a substantial research foundation. In this paper, I would thus like to move beyond individual cases of competition and I aim to delineate a whole morphological ecosystem on the basis of an onomasiological-cum-ecological approach to word-formation, using English as an illustrative language.

I first draw the different correspondences between the main processes and functions of word-formation and show that there is no full-scale competition, but, on the contrary, a quasi-complementary distribution between the processes which prototypically encode the transcategorial function (affixation, conversion and deaffixation), the transconceptual function (affixation, compounding and blending) and the evaluative function (affixation, clipping and replication). In the second part of the paper, I then show that in each of the three demarcated competitive subsystems, the distributional tendencies can also be claimed to be of an essentially complementary nature. In the light of Mark Aronoff's recent work on affixal rivalry (2016, 2019), which is based on the evolutionary principle of competitive exclusion, I finally conclude that like affixes, individual word-formation processes can be seen as elements that tend to settle in distinct niches of the morphological ecosystem.

References

Arndt-Lappe, Sabine, 2014, Analogy in suffix rivalry: The case of English *-ity* and *-ness*, *English Language and Linguistics*, 18 (3).

Aronoff, Mark, 2016, Competition and the lexicon, in Annibale Elia, Claudio Iacobino & Miriam Voghera (eds), *Atti del XLVII congresso internazionale della società di linguistica italiana*, Rome, Bulzoni.

Aronoff, Mark, 2019, Competitors and alternants in linguistic morphology, in Franz Rainer, Francesco Gardani, Wolfgang Dressler & Hans Christian Luschützky (eds), *Competition in inflection and word-formation*, Cham, Springer.

Bauer, Laurie, 2006, Competition in English word formation, in Ans van Kemenade & Bettelou Los (eds), *The handbook of the history of English*, Malden, MA, Blackwell.

Bauer, Laurie, Salvador Valera & Ana Díaz-Negrillo, 2010, Affixation vs. conversion: The resolution of conflicting patterns, in Franz Rainer, Wolfgang Dressler, Dieter Kastovsky & Hans Christian Luschützky (eds), *Variation and change in morphology*, Amsterdam, Benjamins.

Lindsay, Mark & Mark Aronoff, 2013, Natural selection in self-organizing morphological systems, in Nabil Hathout, Fabio Montermini & Jesse Tseng (eds), *Morphology in Toulouse: Selected proceedings of Décembrettes 7*, Munich, Lincom Europa.

Nagano, Akiko, 2008, Categorial change of conversion and the process of relisting, *English Linguistics*, 25 (2).

Naya, Ryohei, 2017, Competition in word-formation: Deverbal nominalization by *-ment* vs. conversion, *JELS*, 34.

Renner, Vincent, 2006, *Les composés coordinatifs en anglais contemporain*, PhD dissertation, Université Lumière Lyon 2.

Santana-Lario, Juan & Salvador Valera-Hernández (eds), 2018, *Competing patterns in English affixation*, Bern, Peter Lang.

Tribout, Delphine & Florence Villoing, 2014, La composition VN et la conversion V>N en français : un nouveau cas de concurrence morphologique?, in Florence Villoing, Sophie David & Sarah Leroy (eds), *Foisonnements morphologiques*, Nanterre, Presses Universitaires de Paris Nanterre.

ꕥ

The typology of associative relations in (complex) word-formation

Pierre J.L. Arnaud & Steve Pepper
(Université Lumière, Lyon & University of Oslo)

Keywords: associative relations; semantic relations; metonymic relations; cognitive relations

In this talk we propose an integrated model of associative relations in word-formation, metonymy and lexical semantics, the PHAB model. We start by considering the unstated (or underspecified) relation ℜ that pertains between the major constituents of noun-noun compounds and other binominal lexemes (1).

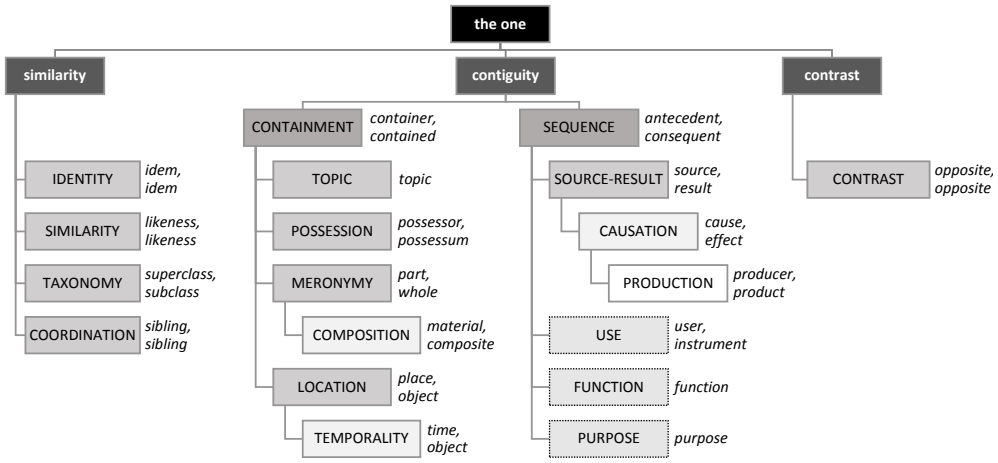
(1)	a. PRODUCTION	b. COMPOSITION	c. USE	d. RESEMBLANCE
	Seychelles Creole	Hebrew	Slovak	German
	mous dimyel	*mesil-at barzel*	*veter-ný mlyn*	*schlüssel-bein*
	fly honey	track-STC iron	wind-ADJZ mill	key-bone
	‘bee’	‘railway’	‘windmill’	‘collarbone’

The nature of ℜ has been the subject of considerable research, often with each new researcher reinventing the classificatory wheel (see Hacken 2016 for a recent summary). We focus on two classification schemes of the “reductionist” type (Søgaard 2005) which operate at different levels of granularity: Hatcher’s (1960) system of four logical relations and Bourque’s (2014) 25-way empirically-derived classification. Following Arnaud (2016), we show how these two systems can be mapped together into a two-tiered system (the “Hatcher-Bourque classification”). We argue that this resolves the dispute regarding the number of relations involved. That number depends on the requirements of the analysis, and the degree of granularity can range from one (as suggested by Bauer 1979) to unlimited (as opined by Jespersen 1942). Our resulting two-tiered system has been tested against a database of over 3,700 noun-noun compounds and their functional equivalents from 106 languages.

We then turn to metonymic relations, taking our cue from Janda (2011) and Peirsman & Geeraerts (2006). We show how most, if not all, such relations match the semantic relations found in binominals. For example, the PRODUCER FOR PRODUCT metonymy in *I’m reading <u>Shakespeare</u>* (Kövecses 2002) corresponds (at some level of abstraction) to the PRODUCTION relation between honey and fly in (1a).

All but one of the 23 metonymic patterns catalogued by Peirsman and Geeraerts can be accommodated by the Hatcher-Bourque classification, with only very minor tweaks (and the exception is exceptional in other ways, as we will show). Furthermore, even if the total number of relations should turn out to be unbounded, the vast majority, as Tratz & Hovy (2010) point out, fit within a relatively small set of categories.

Finally, we consider the seven-way system of "conceptual relations" developed by Andreas Blank and Peter Koch in various papers (e.g. Koch 2001; Blank 2003) and used to investigate polysemy, semantic shifts and other aspects of lexical semantics. According to Koch and Blank, all such relations ultimately go back to Aristotle's three associative relations of 'similarity', 'contiguity' and 'contrast'. We show how they, too, can be accommodated within a slightly extended version of the Hatcher-Bourque classification. The result is a single, integrated, multilevel system of associative relations that can be applied across subfields of linguistics as diverse as word-formation, metonymy and lexical semantics, the PHAB model. This model provides a common framework that can be tested and further elaborated by other researchers, who are thereby relieved of having to continually reinvent the wheel.



The PHAB model of associative relations

References

Arnaud, Pierre J.L. 2016. Categorizing the modification relations in French relational subordinative [NN]N compounds. In Pius ten Hacken (ed.), *The semantics of compounding*, 71–93. Cambridge: Cambridge University Press.

Bauer, Laurie. 1979. On the need for pragmatics in the study of nominal compounding. *Journal of Pragmatics* 3(1). 45–50.

Blank, Andreas. 2003. Words and concepts in time: Towards a diachronic cognitive onomasiology. In Regine Eckardt, Klaus von Heusinger & Christoph Schwarze (eds.), *Words in time: Diachronic semantics from different points of view*, 37–65. Berlin: de Gruyter.

Bourque, Yves Stephen. 2014. *Toward a typology of semantic transparency: The case of French compounds*. University of Toronto PhD dissertation.

Hacken, Pius ten. 2016. Introduction: compounds and their meaning. In Pius ten Hacken (ed.), *The semantics of compounding*, 1–11. Cambridge: Cambridge University Press.

Hatcher, Anna Granville. 1960. An introduction to the analysis of English noun compounds. *Word* 16(3). 356–373.

Janda, Laura A. 2011. Metonymy in word-formation. *Cognitive Linguistics* 22(2).

Jespersen, Otto. 1942. *A modern English grammar on historical principles. Part 6: Morphology*. London: George Allen.

Koch, Peter. 2001. Lexical typology from a cognitive and linguistic point of view. In Martin Haspelmath, Ekkehard König, Wolfgang Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: an international handbook*, 1142–1178. Berlin: de Gruyter.

Kövecses, Zoltan. 2002. *Metaphor: A practical introduction*. Oxford University Press.
Peirsman, Yves & Dirk Geeraerts. 2006. Metonymy as a prototypical category. *Cognitive Linguistics* 17(3). 269–316.
Søgaard, Anders. 2005. Compounding theories and linguistic diversity. In Zygmunt Frajzyngier, Adam Hodges & David S. Rood (eds.), *Linguistic diversity and language theories*, 319–337. Amsterdam: John Benjamins.
Tratz, Stephen & Eduard Hovy. 2010. A taxonomy, dataset, and classifier for automatic noun compound interpretation. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 678–687. Uppsala: Association for Computational Linguistics.

ꕥ

Are *tweople* different from *Twitter people*? Experimental study of semantic relations in compounds and blends

Elizaveta Tarasova and Natalia Beliaeva
(IPU New Zealand, Palmerston North; MCLaSS: Multicultural Learning & Support Services, Wellington)

Keywords: word-formation, N+N compounds, blends, semantic relations, native speaker perceptions

The present study analyses native speaker perceptions of the differences in the semantic structure of compounds and blends. Blends are no longer considered an unpredictable extragrammatical type of coining new words, but their stance in relation to other morphological categories such as compounds is still under discussion. The viewpoints on blending vary with some researchers considering it to be an instance of compounding (e.g. Kubozono, 1990), while others identifying blending as an interim word formation mechanism between compounding and shortening (e.g. López Rúa, 2004). Like compounds, blends are formed by joining together two (or more) words. The obvious difference is that in blends constituents are merged into one single lexeme, both graphically and prosodically, whereas compounds are comprised of discrete lexemes.

The question pursued in this study is whether the formal differences between compounds and blends are reflected on the semantic level. While the majority of registered blends are coordinative (Gries, 2012; Renner, 2006), e.g. *needcessity* < *need* + *necessity*, determinative blends such as *sarcastrophy* < *sarcasm* + *catastrophy* are widely attested and cannot be excluded from analysis. This means that blends may be comparable to compounds in terms of their semantics too, which is also evidenced in the semantic relations that are demonstrated to exist between the elements of blends (Bauer and Tarasova, 2013). In this study, the semantic characteristics of determinative blends and N+N subordinative compounds are compared by evaluating the differences in native speakers' understanding of the semantic relationships between constituents of the analyzed structures in order to answer the following questions.

- What are the differences between compounds and blends in terms of the semantic relations between their constituents?
- Do native speakers rely on semantic relations in accessing the meaning of compounds and blends, and if so, to what degree?
- What factors promote the speakers' use of semantic relations in the process of accessing the meaning of compounds and blends?

The data for the study is obtained from two web-based experiments, which use examples from the corpus of contemporary blends (Beliaeva, 2014) and the corpus of compounds (Tarasova, 2013), coded for semantic relations by means of *Bourquifier 2.0* software (Pepper, forthc.). The results demonstrate that:

- There are differences in the preference for the semantic relations realized in compounds and blends, with determinative blends being strongly focused on the causative relation (*snacksident* < *snack* + *accident* 'accident caused by a snack'), and compounds – on the purpose relation (animal clinic < 'a clinic for animals');
- Native speaker interpretations of meaning of compounds and blends show speakers' stronger reliance on the semantic relations in compounds;
- In blends, the presence of the semantic relation in the interpretation is determined by how typical a relation is. For compounds, the lack of familiarity of the speaker with the compound promotes the use of the relation in the interpretation.

Overall, the results of the study suggest that both compounds and blends are potentially compositional. Composition (as a cognitive operation) is employed when formal cues do not provide enough information on meaning.

References

Bauer, Laurie and Elizaveta Tarasova (2013), The meaning link in nominal compounds, *SKASE Journal of Theoretical Linguistics* 10(3), 1-18.

Beliaeva, Natalia (2014), A study of English blends: From structure to meaning and back again, *Word Structure*, 7(1), 29–54.

Gries, Stefan Th. (2012), Quantitative corpus data on blend formation: Psycho- and cognitive-linguistic perspectives, in V. Renner, F. Maniez, & P. J. L. Arnaud (eds), (2012), *Cross-disciplinary perspectives on lexical blending*, Berlin: De Gruyter Mouton, 145-167.

Kubozono, Haruo (1990), Phonological constraints on blending in English as a case for phonology-morphology interface, *Yearbook of Morphology*, 3, 1-20.

López Rúa, Paula (2004), The categorial continuum of English blends, *English Studies*, *85*(1), 63-76.

Pepper, Steven (forthc.), *The typology of binominal lexemes: Noun-noun compounds and their functional equivalents* (PhD thesis), University of Oslo.

Renner, Vincent (2006), *Les composés coordinatifs en anglais contemporain* (PhD thesis), Université Lumière-Lyon 2, Lyon.

Tarasova, Elizaveta (2013), *Some new insights into the semantics of English N+N* (PhD thesis), Victoria University of Wellington.

Orthography-Semantic Consistency and Semantic Transparency: Re-evaluating the semantic impact in masked morphological priming

Simona Amenta & Marco Marelli
(University of Trento and Ghent University; & University of Milano-Bicocca)

Keywords: morphological processing, form-meaning mapping, semantic transparency, word recognition, masked priming

The role of semantic transparency in masked morphological priming has been extensively studied, but results in the field are inconsistent (e.g, Rastle, Davis, and New, 2004; Feldman, O'Connor, and del Prado Martín, 2009 for a debate). Over and above the much debated semantic modulation of priming magnitude, a largely overlooked side effect is reported in most studies: a main effect of semantic transparency across priming conditions. That is, participants are faster at recognizing stems from transparent sets (e.g., widow from widower) in comparison to stems from opaque sets (e.g., whisk from whisker), regardless of the preceding primes (Marelli, Amenta, and Crepaldi, 2015).

We propose that this pattern of effects can be interpreted in terms of Orthography-Semantic Consistency (OSC), a measure of how well the meaning of a given word can be predicted from its form (Marelli et al., 2015; Marelli and Amenta, 2018). OSC is operationally defined as the degree of semantic relatedness between a word and its orthographic relatives, computed as the frequency-weighted average semantic similarity between the meaning of the target (e.g., whisk) and the meanings of all the words embedding it (e.g., whisker, whiskey, whisks). Moreover, we propose that OSC might have a key role in explaining the facilitation pattern observed in the majority of morphological priming studies (i.e., transparent > or ≥ opaque > orthographic pairs). In fact, since (1) the pool of orthographic relatives of a given target - on which the computation of OSC is based - includes words that could all in principle be used as primes, and (2) the contribution to OSC of each orthographic relative is determined by its frequency, a straightforward prediction is that priming magnitude should be modulated by an interaction between OSC and prime frequency.

The re-analysis of results from a number of masked priming studies (Rastle, Davis, Marslen-Wilson, and Tyler, 2000; Rastle, Davis, and New, 2004; Marslen-Wilson, Bozic, and Randall, 2008; Diependaele, Sandra, and Grainger, 2005; Andrews and Lo, 2013) indicates the central role of OSC in explaining the impact of semantic transparency in masked priming. Indeed, the "stem transparency" effect (the facilitation in processing stems of transparent words as opposed to stems of opaque words) is associated to an imbalance in the OSC of the target words included in the studies. Moreover, the magnitude of the observed priming effect is also better explained by the variability in the target OSC, rather than the traditional semantic-transparency classification. A further analysis confirms that the impact of the prime is crucially qualified by its frequency in the semantic cohort activated by the target orthography, as captured by OSC, as predicted by our hypothesis.

Taken together these results indicate that the traditional way of investigating semantic transparency in morphological priming offers a limited account of the semantic dynamics at early stages of morphological processing. Early semantic effects in morphological priming are in fact better characterized in terms of the semantic network activated by the prime-target pair by means of simple orthographic information.

References

Andrews, Sally, and Lo, Steson (2013), Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary, Journal of Memory and Language, 68(3), 279-296.

Diependaele, Kevin, Sandra, Dominiek, and Grainger, Johnathan (2005), Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition, Language and cognitive processes, 20(1-2), 75-114.

Feldman, Laurie Beth, O'Connor, Patrick A., and del Prado Martín, Fermín Moscoso (2009), Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition, Psychonomic bulletin & review, 16(4), 684-691.

Marelli, Marco, and Amenta, Simona (2018), A database of orthography-semantics consistency (OSC) estimates for 15,017 English words, Behavior research methods, 50(4), 1482-1495.

Marelli, Marco, Amenta, Simona, and Crepaldi, Davide (2015), Semantic transparency in free stems: The effect of Orthography-Semantics Consistency on word recognition, The Quarterly Journal of Experimental Psychology, 68(8), 1571-1583.

Marslen-Wilson, William D., Bozic, Mirjana, and Randall, Billi (2008), Early decomposition in visual word recognition: Dissociating morphology, form, and meaning, Language and Cognitive Processes, 23(3), 394-421.

Rastle, Kathleen, Davis, Matthew H., & New, Boris (2004), The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition, Psychonomic Bulletin & Review, 11(6), 1090-1098.

Rastle, Kathleen, Davis, Matthew H., Marslen-Wilson, William D., and Tyler, Lorraine K. (2000), Morphological and semantic effects in visual word recognition: A time-course study, Language and cognitive processes, 15(4-5), 507-537.

Interpretational diversity in novel compounds

Melanie J. Bell & Martin Schäfer
(Anglia Ruskin University & Universität Düsseldorf)

Keywords: compound interpretation, novel compound, semantic relation, polysemy, constituent family

Since Gagné and Shoben (1997), evidence has accumulated that humans use statistical semantic preferences associated with compound constituents in the interpretation of compound nouns. Most of this work has focused on the semantic relation between constituents, rather than interpretation of the constituents themselves. Yet due to the ubiquity of polysemy and homonymy, there are typically several concepts to choose from for any given constituent, and it is not clear at the outset which concept is to be used. This holds both for modifiers and heads, and for concepts that are unrelated (*plant* factory/organism) as well as those that are metaphorically linked (*gold* material/colour). Using attested novel compounds, we show that diversity of interpretation is much greater than suggested by studies that focus only on semantic relations and that the ambiguity of the constituents, especially the head, plays an important role in determining both difficulty and diversity of interpretation.

We generated a set of attested novel compounds by selecting 45 noun-noun compounds that occur only once in ukWaC, a 2 billion word corpus of British English. Twenty native speakers of British English provided free paraphrases of each compound presented in isolation, and rated the difficulty of doing so on a 6-point Likert scale. The paraphrases were subsequently coded and grouped according to whether or not they represented the same reading. We modelled the difficulty ratings using a mixed effects regression model with random intercepts for subjects and items and fixed effects for presentation order as well as various constituent-family-based predictors. We included head and modifier family sizes, number of different semantic relations in each family, and number of distinct constituent senses in each family. We also calculated the entropy of the probability distributions of the synsets and semantic relations for all modifier and head constituent families ($H = -\sum_{i=1}^{n} p_i \, log p_i$.

In the final model for difficulty rating, the only significant fixed effects are presentation order (there is a small tendency for participants to rate items as increasingly difficult) and synset entropy of

the head family: the more ambiguous the head noun, the more difficult the compound is to interpret. The average difficulty rating for each compound is also positively correlated with the number of different interpretations given ($r = 0.57$, $p < 0.001$). Furthermore, the diversity of interpretations is far greater than suggested by e.g. Middleton et al. (2011): in our data, only 34% of participants on average agree on the most frequent interpretation for any compound. In all cases, different interpretations involve different senses of at least one constituent, and in 85% of cases they involve different senses of both constituents. Current models of conceptual combination focus either on rather underspecified semantic relations or on 'slot-filling' in the semantics of the head; in either case, it is unclear how and at what stage disambiguation of the constituents takes place. As our results indicate, future incarnations of such models will need to address this challenge if they are to account for the interpretation not only of familiar but also of novel compounds.

References

Gagné, C. L. and E. J. Shoben (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(1), 71–87.

Middleton, E. L., K. A. Rawson, and E. J. Wisniewski (2011). How do we process novel conceptual combinations in context? *The Quarterly Journal of Experimental Psychology* 64(4), 807– 822.

NN compounding, semantic categories and low productivity

Jesús Fernández-Domínguez
(University of Granada)

Keywords: compounding semantics, corpus-based, marginal productivity, NN compounding, word-formation

NN compounds have whipped up wide interest across contemporary morphological theories due to a number of distinctive characteristics. One of them is the flexible meaning connection tying their constituents which, in most cases, has been studied by targeting the head of the compound (e.g. Allen 1978). While acknowledging the importance of heads, other studies have redefined the role of modifiers, thus disclosing their decisive effect on compound semantics. Among these, Gagné (2009) shows that the more compounds that share a given constituent, the more probable it is to have a variety of relations attached to it (see Libben 2006).

A second focus of attention for NN compounding is its virtually limitless syntax-like productivity. Regarding the two possible slots in a NN compound, Maguire et al. (2010) maintain that nouns are used with an identical degree of productivity both as modifiers and as heads, while for Baayen (2010) the members of lexicalized compounds are position-bound because of factors like constituent family size or lexical availability. This has been recently linked to the existence and nature of derivational paradigms and the place of compound words in them (Tarasova 2018).

Globally observed, hence, research on NN compounding has largely concentrated on two readily visible features: its most frequent semantic categories and its most productive morphological types. This paper resorts to a sample of NN formations from the BNC (Davies 2004–) in order to explore the underresearched connections between less common semantic categories in NN compounds, areas of low productivity and the role of left-hand constituents therein. This is done by assessing the

performance of various productivity models on the corpus sample and by correlating their results with the formal and semantic configuration of the NN units.

Acknowledgements: the present paper was funded by I+D Project FFI2017-89665-P (Spanish Ministry of Economy, Industry and Competitiveness).

References

Allen, Margaret R. (1978), *Morphological investigations*, Doctoral dissertation, University of Connecticut, Storrs, CT.

Baayen, R. Harald (2010), The Directed Compound Graph of English. An Exploration of Lexical Connectivity and its Processing Consequences, *Linguistische Berichte Sonderheft* 17, 383–402.

Davies, Mark (2004–), *BYU-BNC* (Based on the British National Corpus from Oxford University Press). Available online at https://corpus.byu.edu/bnc.

Gagné, Christina. L. (2009), Psycholinguistic perspectives, in R. Lieber, and P. Štekauer (eds.), (2009), *The Oxford handbook of compounding*, Oxford: Oxford University Press, 255-271.

Libben, Gary (2006), Why study compound processing? An overview of the issues, in G. Libben, and G. Jarema (eds.), (2006) *The representation and processing of compound words*, Cambridge: Cambridge University Press, 3-22.

Maguire, Phil, Edward J. Wisniewski, and Gert Storms (2010), A Corpus Study of Semantic Patterns in Compounding, *Corpus Linguistics and Linguistic Theory* 6(1), 49–73.

Tarasova, Elizaveta (2018), N+N Compound Paradigms: Productivity of the Form = Productivity of Meaning? Paper presented at the *Word-Formation Theories III and Typology and Universals in Word-Formation IV Conference*, Pavol Jozef Šafárik University (Košice), 27–30 June 2018.

WORKSHOP 20

Schedule: Fri 9.00-16.55 (Room 4)

(Semi-)automatic retrieval of data from historical corpora: Chances and challenges

Marianne Hundt, Melanie Röthlisberger, Gerold Schneider & Eva Zehentner
(University of Zurich)

Keywords: historical corpus linguistics, computational linguistics, (semi-)automatic retrieval, annotation, precision and recall

This workshop aims to bring together researchers working at the interface between historical linguistics and computational linguistics, zooming in on the particular issues that arise when retrieving data automatically or semi-automatically from historical databases.

Developments in historical corpus linguistics have taken a similar route as in corpus-based research on present-day languages: from the creation of small reference corpora to increasingly larger databases and from text-only to richly annotated resources. However, historical data have always posed particular challenges for the development of corpus resources, their annotation, and their analysis. Corpus representativeness and balancedness, for instance, has been impaired by the limited availability of texts, particularly for the very early stages of written attestation. This is often generally referred to as the 'bad data' problem of historical linguistics (cf. e.g. Labov 1994: 11). Additionally, the highly variable orthography typical of earlier texts has meant that the tools developed for more uniform data cannot be applied in a straightforward manner to historical corpora. In the case of smaller corpora, this has resulted in grammatical annotation through manual annotation or post-editing. For the increasingly larger resources, however, manual annotation is tedious, and researchers have developed tools for pre-processing like spelling normalisation (Baron & Rayson 2008) and lemmatisation (Burns 2013) to enable automatic tagging and parsing. Matters are complicated further by the fact that a range of different annotated resources exist (*Penn Treebank, Penn Parsed Corpora, Universal Dependency Treebanks*) and different parsing tools (e.g. Schneider 2012) have been applied to historical corpora, which are likely to require different retrieval strategies, which in turn make comparisons across corpora difficult. While the list of syntactic parsers is large (e.g. Schneider 2008 for English, Sennrich et al. 2009 for German, van Noord 2006 for Dutch, Alberti et al. 2017 for Universal Dependency parsing), few have been used on, or adapted to historical texts.

The goal of this workshop is to discuss the challenges that (semi-)automatic retrieval of data from historical corpora pose for the study of grammatical change, specifically in Germanic languages. In particular, we invite contributions addressing questions such as (but not limited to) the following:

- Mapping of different annotation schemes:
    - Many corpora use different annotation schemes: for example, the Penn Parsed corpora are annotated in a different way than the Penn Treebank (Marcus et al. 1993), or dependency-parsed corpora like ARCHER. If this is the case, up to which point, and for which phenomena can the data be compared?
    - How is change over time dealt with or reflected in annotation? For example, the annotation of certain items differs between the Penn-Parsed corpora (covering different periods),

indicating e.g. grammaticalisation processes. Is this problematic, and what can we do to make sure it does not affect retrieval?

- Automatic parsers use different annotations schemes, particularly across different languages and/or periods. What can be mapped easily, and where do we need additional manual decisions in the mapping? Can probabilistic mappers help us, or do they just extend the issue?
- One way to avoid language-specific annotation may be to employ tools like the Universal Dependency label (in progress).1 Using such highly underspecified, coarse sets of tags (and e.g. dependency labels) may indeed increase comparability, but may also lead to a loss of granularity. Is this worth the risk, and are the labels really directly mappable between different languages? Does historical data add further problems?

- Evaluation of bottom-up approaches to data retrieval for language change:
  - Data-driven approaches to language change (e.g. Hilpert and Gries 2016) can detect new patterns, increasing recall. But can we also detect changes in rare constructions, and how much insight does this really add?
  - How much benefit is there in using fine-grained annotations for exploratory research? Could we get the same (or more representative) results with automatically parsed data/ POS tags only?
  - Query tools like Stanford Tregex (Levy and Andrew 2006) allow us to come up with very elaborate, detailed queries to extract relevant data. This can be useful, but may also be problematic: Where can we draw the line between too generic and too specific queries?
  - What is the role of ambiguity in language change? Is there a correlation between ambiguity of a language model (e.g. low tagger and parser confidence) or human annotators (low inter-annotator agreement) and change?
- Issues of precision and recall in historical corpora:
  - Precision and recall is lower for most phenomena in historical corpora when using automatically annotated data. How much does this affect results? For example, an increase of a specific lexical item may be due to low recall caused by unidentified spelling variants in earlier texts.
  - Precision is often less affected than recall. Up to which point are we ready to infer conclusions from annotation which is largely correct but only annotates prototypical cases?
  - While automatic annotation typically involves more errors, it has the benefit of allowing us to deal with almost unlimited amounts of data. Up to which point can sheer size compensate and even overcompensate the errors?
  - Significance testing assumes no errors or homogenously spread noise in the data. What can we do when this assumption is clearly violated?

Ultimately, this workshop seeks to provide a platform for researchers working within these subject areas to exchange ideas and to jointly address the challenges (and chances) we are faced with.

References

Alberti, Chris, Daniel Andor, Ivan Bogatyy, Michael Collins, Dan Gillick, Lingpeng Kong, Terry Koo, Ji Ma, Mark Omernick, Slav Petrov, Chayut Thanapirom, Zora Tung, and David Weiss. 2017. "Syntaxnet models for the CoNLL 2017 shared task". *arXiv:1703.04929*, 2017.3

Baron, Alistair and Paul Rayson. 2008. "VARD 2: A tool for dealing with spelling variation in historical corpora". *Proceedings of the Postgraduate Conference in Corpus Linguistics*.

Burns, Philip. 2013. "MorphAdorner v2: A Java Library for the Morphological Adornment of English Language Texts." Evanston, IL. Northwestern University. (http://morphadorner.northwestern.edu/morphadorner/documentation/citation).

Hilpert, Martin and Stefan Th. Gries 2016. Quantitative approaches to diachronic corpus linguistics. In Merja Kytö and Paivi Pahta (Eds.), *The Cambridge Handbook of English Historical Linguistics*. Cambridge: Cambridge University Press, 36-53.

Kroch, Anthony, and Ann Taylor. 2000. The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, release 4 (http://www.ling.upenn.edu/ppche-release-2016/PPCME2-RELEASE-4).

Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME). Department of Linguistics, University of Pennsylvania. CD-ROM, first edition, release 3 (http://www.ling.upenn.edu/ppche-release-2016/PPCEME-RELEASE-3).

Labov, William. 1994. *Principle of change, internal factors*. Oxford: Blackwell.

Levy, Roger and Galen Andrew. 2006. "Tregex and Tsurgeon: tools for querying and manipulating tree data structures". *5th International Conference on Language Resources and Evaluation* (LREC 2006). (https://nlp.stanford.edu/pubs/levy_andrew_lrec2006.pdf).

Marcus, Mitch, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. "Building a Large Annotated Corpus of English: the Penn Treebank". *Computational Linguistics*, 19, 313-330.

Schneider, Gerold. 2008. *Hybrid Long-Distance Functional Dependency Parsing*. Doctoral Thesis. University of Zurich.

Schneider, Gerold. 2012. "Adapting a parser to Historical English". In Jukka Tyrkkö, Matti Kilpiö, Terttu Nevalainen & Matti Rissanen (Eds.) *Studies in Variation, Contacts and Change in English, Volume 10: Outposts of Historical Corpus Linguistics: From the Helsinki Corpus to a Proliferation of Resources*. Helsinki: VARIENG.

Sennrich, Rico, Gerold Schneider, Martin Volk, and Martin Warin. 2009. "A New Hybrid Dependency Parser for German". *Von der Form zur Bedeutung: Texte automatisch verarbeiten / From Form to Meaning: Processing Texts Automatically. Proceedings of the Biennial GSCL Conference 2009*, 115-124.

van Noord, Gertjan. 2006. "At last parsing is now operational". In Piet Mertens, Cedrick Fairon, Anne Dister & Patrick Watrin (Eds.) *TALN06. Verbum Ex Machina. Actes de la 13e conference sur le traitement automatique des langues naturelles*, 20-42.

Contact-induced change in the diachrony of English and semi-automatic retrieval of data from historical corpora of translated *vs* non-translated texts

Nikolaos Lavidas
(National and Kapodistrian University of Athens)

Keywords: Written language contact; History of English; Translations from Latin and French; Borrowing of word-formation patterns; Machine learning approach

The aim of this study is to investigate language contact-induced change in English through semi-automatic retrieval of data from historical corpora. We examine cases of borrowing of word formation patterns from Latin and French into English. We compare various translations of Boethius' *"De Consolatione Philosophiae"* from different periods of the history of English (see also Author 2018). We also examine various translations of the *New Testament* into English (see Taylor 2008 on types of transfer from biblical and non-biblical translations).

Translation typically crops up within the historical linguistic context in several cases (see, among others, Lehiste 1979, Blake 1992, Koller 1998). For instance, Fischer's (2013) explanation of the contrast between the status of Latin and French in the history of contact with English (with fewer loanwords from Latin than from French) refers to the written type of contact with Latin: "On the external side, it can be related to the fact that communication in terms of Latin was far more indirect (mostly via translators of texts)."

Based on a machine learning approach (with document classification; Mayfield & Rosé 2010, 2013), we locate which features are most characteristic of the class of texts that are translations from Latin (the features with the highest weights in a logistic regression are the most characteristic for the class), and in which degree earlier translations differ from later translations and from non-translated texts. The aim of a document classification approach is that the characteristic features of each class of documents (e.g., translated *vs* non-translated, early translations *vs* later re-translations, and early non-translated texts *vs* later non-translated texts) should be automatically learned from the manually classified documents in a corpus-driven way. For example, in Table 1, we compare the status of suffixes *-ity* and *-ness* in translations of Boethius' text that were prepared in different periods. Romaine (1985: 452) claims that "the suffix *-ity* […] makes its appearance in 14th-15th century loanwords from French, and later, in loanwords from Latin." (See also Gardner 2014).

We propose that the word formation patterns in this case reflect grammar competition, according to which transition stages of the diachrony of a language demonstrate the coexistence of more than one grammar and a competition between these grammars (Kroch 1989, 2001). Grammar competition is represented, for instance, with competing forms that appear in later translations (e.g., *fragileness/fragility*) and with semantically related sets (e.g., *madness/insaneness/insanity*) (Romaine 1985).

Table 1. Presence of *-ity* vs-*ness* formations in Chaucer's (14th century; probably based on a French translation) *vs* Queen Elizabeth's (16th century) translation of Boethius' "De Consolatione Philosophiae" (sample): Feature Weight of nouns in *-ity/-ness*

Chaucer's translation		*Queen Elizabeth's translation*	
Feature	Feature Weight	Feature	Feature Weight
blisfulnesse	0.8639	vniuersalitie	0.4255
wikkidnesse	0.5515	dignitie	0.4131
wrecchidnesse	0.4111	felicitie	0.3994
schrewednesse	0.3350	nobilitie	0.3832
unselynesse	0.3193	blessedness	0.2744
symplicite	0.2484	seueritie	0.1866
welefulnesse	0.2223	happynes	0.1847
adversitie	0.1928	darknes	0.1828
swetnesse	0.1848	stabilitie	0.1747
lightnesse	0.1773	affinitie	0.1395
dyvinite	0.1645		
necessite	0.1296		

References

Blake, Norman Francis (1992), Translation and the history of English, in M. Rissanen, O. Ihalainen, T. Nevalainen and I. Taavitsainen (eds). *History of Englishes: New Methods and Interpretations in Historical Linguistics*, Berlin: Mouton de Gruyter, 3-24.

Fischer, Olga (2013), The role of contact in English syntactic change in the Old and Middle English periods, in D. Schreier and M. Hundt (eds), *English as a Contact Language*, Cambridge: Cambridge University Press, 18-40.

Gardner, Anne-Christine (2014), *Derivation in Middle English: Regional and Text Type Variation*. Helsinki: Société Néophilologique.

Koller, Werner (1998), Übersetzungen ins Deutsche und ihre Bedeutung für die deutsche Sprachgeschichte, in W. Besch, A. Betten, O. Reichmann and S. Sonderegger (eds), *Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*, Berlin: de Gruyter, 210-229.

Kroch, Anthony (1989), Reflexes of grammar in patterns of language change, *Language Variation and Change* 1, 199-244.

Kroch, Anthony (2001), Syntactic change, in M. Baltin and C. Collins (eds), *The Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell, 699-729.

Lehiste, Ilse (1979), Translation from Russian as a source of syntactic change in Contemporary Estonian, in P. R. Clyne, W. F Hanks and C. L. Hofbauer (eds), *The Elements: A Parasession on Linguistic Units and Levels*, Chicago: Chicago Linguistic Society, 413-419.

Mayfield, Elijah and Carolyn Penstein Rosé (2010), Using feature construction to avoid large feature spaces in text classification. *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*, 1299-1306.

Mayfield, Elijah and Carolyn Penstein Rosé (2013), LightSIDE: Open source machine learning for text, in M. D. Shermis and J. Burstein (eds), *Handbook of Automated Essay Assessment: Current Applications and New Directions*, New York: Routledge, 124-135.

Romaine, Suzanne (1985), Variability in word formation patterns and productivity in the history of English, in J. Fisiak (ed.), *Papers from the 6th International Conference on Historical Linguistics*. Amsterdam: John Benjamins, 451-465.

Taylor, Ann (2008), Contact effects of translation: Distinguishing two kinds of influence in Old English, *Language Variation and Change* 20(2), 341-365.

Mapping of lemmatisation annotation to multiple Middle English corpora

Michael Percillier
(University of Mannheim)

Keywords: Middle English, historical corpus linguistics, (semi-)automatic retrieval, annotation, lemmatisation

The present paper describes the mapping of additional annotation originally undertaken for the *Penn-Helsinki Parsed Corpus of Middle English* (PPCME2, Kroch & Taylor 2000) to other Middle English (ME) corpora. Specifically, the lemmatisation of lexical verbs in the PPCME2 (see Percillier 2016, 2018), which also includes the determination of the verbs' origin as either French-based or native, is applied to texts from other ME corpora.

The motivation for applying this annotation layer to further corpora lies in the unbalanced coverage of the PPCME2, in which the period M2 (1250-1350) is underrepresented in comparison to the other

periods. Expanding the coverage of M2 by including texts from other ME corpora allows for a more comprehensive analysis of the transition from M1 (1150-1250) to M2, which is potentially pivotal as suggested by the *to* prepositional object construction overtaking the double object construction (Zehentner 2017:10, Elter 2018) as well as the sudden influx of French-based verbs (Percillier 2016:212).

Furthermore, the unbalanced coverage of the PPCME2 also applies to the diatopic dimension. M1 contains texts from the Midlands as well as Southern texts, whereas M2 is limited to Southern texts. Only in the later periods M3 (1350-1420) and M4 (1420-1500) do Northern texts appear. These gaps in coverage entail that any investigation of diachronic developments from M1-M4 may unwittingly blur the distinction between diachrony and diatopy, meaning that any diachronic development observed in the corpus may in fact be partly due to diatopic variation.

The coverage gaps in the PPCME2 are due to a focus on prose texts, justified by the influence that rhyme and metre in verse may exert on syntax. Expanding the coverage of the PPCME2 necessitates the inclusion of verse texts. Rather than choosing between coverage gaps or the influence of rhyme and metre as the "lesser evil", a solution is to treat text type (prose or verse) as an extralinguistic variable that may or may not influence a given linguistic feature, akin to how text genre (law, fiction, sermon, etc.) is already treated in the corpus.

The additional ME corpora used include the *Corpus of Middle English Prose and Verse* (CMEPV, McSparran et al. 2006), the corpus of the *Linguistic Atlas of Early Middle English* (LAEME, Laing 2013), more specifically its recently parsed version (PLAEME, Truswell et al. 2018), the *Parsed Corpus of Middle English Poetry* (PCMEP, Zimmermann 2018), as well as texts from the *Leuven English Old to New* corpus (LEON, Petré 2013). The annotation of these corpora ranges from non-existent (CMEPV) and only partial POS-tagging (LEON) to POS and constituent annotation that follow the PPCME2 scheme (PLAEME and PCMEP). This heterogeneity presents a challenge not only for the insertion of lemma information, but also for the formulation of queries that are valid across corpora.

A case study examining the possible relation between the increase of French-based verbs and the rise of the *to* prepositional object construction will serve as a working example to illustrate and test the mapping process.

Acknowledgements

Funding from the *Deutsche Forschungsgemeinschaft* for the research project *Borrowing of Argument Structure in Contact Situations* (grant TR555/6-2) is gratefully acknowledged.

References

Elter, Wiebke Juliane (2018), The rise of the to-dative: A language-contact approach to a phenomenon of structural language change. Mannheim: University of Mannheim MA-Thesis.

Kroch, Anthony & Ann Taylor (2000), *The Penn-Helsinki Parsed Corpus of Middle English, Second Edition (PPCME2), release 3*. Philadelphia: University of Pennsylvania. http://www.ling.upenn.edu/hist-corpora/.

Laing, Margaret (2013), *A Linguistic Atlas of Early Middle English, 1150–1325, Version 3.2*. Edinburgh: The University of Edinburgh. http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html.

McSparran, Frances, Paul Schaffner, John Latta, Alan Pagliere, Christina Powell & Matt Stoeffler (2006), *Corpus of Middle English Prose and Verse*. Ann Arbor: University of Michigan. https://quod.lib.umich.edu/c/cme/.

Percillier, Michael (2016), Verb lemmatization and semantic verb classes in a Middle English corpus, in *Proceedings of the 13th Conference on Natural Language Processing (KONVENS 2016)*, 209–214. https://www.linguistics.rub.de/konvens16/pub/26_konvensproc.pdf.

Percillier, Michael (2018), A Toolkit for lemmatising, analysing, and visualising Middle English Data, in A. U. Frank, C. Ivanovic, F. Mambrini, M. Passarotti and C. Sporleder (eds), (2018), *Proceedings of the Second Workshop on Corpus-Based Research in the Humanities CRH-2*, Vienna: Gerastree Proceedings, 153–160. https://www.oeaw.ac.at/fileadmin/subsites/academiaecorpora/PDF/CRH2.pdf.

Petré, Peter (2013), *LEON: Leuven English Old to New, version 0.3*. Leuven: KU Leuven. https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1820596&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US.

Truswell, Robert, Rhona Alcorn, James Donaldson & Joel Wallenberg (2018), *A Parsed Linguistic Atlas of Early Middle English*. University of Edinburgh. doi:10.7488/ds/2310.

Zehentner, Eva (2017), Ditransitives in Middle English: On semantic specialisation and the rise of the dative alternation. *English Language and Linguistics* 22(1), 149–175. doi:10.1017/S1360674316000447.

Zimmermann, Richard (2018), *The Parsed Corpus of Middle English Poetry (PCMEP)*. http://pcmep.net/.

Pattern matching or holistic retrieval: finding bare clefts in unannotated corpus data

Lara Verheyen, Sara Budts, Peter Petre and William Standing

In corpus linguistics it is common practice to compile a dataset by querying a corpus with regular expressions and manually sifting out relevant hits. While effective for construction-specific lexical items, this method becomes unwieldy when retrieving schematic constructions in larger corpora (Forsberg et al. 2014; Dunn 2016). Part-of-speech tagging, syntactic parsing and spelling normalisation might provide a handle, but off-the-shelve tools typically perform below ceiling when applied to historical data (Schneider et al. 2016; Yang & Eisenstein 2016). Systems trained on modern corpora may force historical data into categories tailored for present-day data, which entails the risk that historically peculiar but relevant samples will be systematically missed out on.
One construction suffering from a poor precision/recall trade-off are clefts without relative markers:

(1) Look at her. It is she I love.

While the construction itself is infrequent, querying its substantial parts ("it BE") will yield a plethora of spurious hits. Automated parsing is no satisfactory solution either, as the construction is too rare to be treated uniformly by probabilistic taggers.

We present a custom search engine that retrieves bare clefts from a corpus of Early Modern English. Rather than relying on discrete pattern matching, our search engine deploys holistic similarity as its relevance metric to retrieve new hits. Using machine learning techniques (Hochreiter & Schmidhuber 1997; Sutskever et al. 2014; Vanni et al. 2018), the engine converts all corpus sentences into low-dimensional real-valued vectors that implicitly represent the sentences' semantic and syntactic features. This low dimensionality is crucial to holistic search - if the dimensionality of the vectors had equalled the vocabulary size, the vectors would have been mere bags-of-words. Dimensionality reduction, by contrast, restricts the expressivity of the vectors, which forces the engine

to compress the information in the sentences by exploiting the similarity between tokens and generalizing over them. As such, these vectors encode patterns that emerge from *generalizations over token sequences* rather than the literal token sequences themselves, which enables them to capture relations that are not literally expressed, such as the equivalence between the antecedent of the cleft ("her") and the topic of the cleft ("she").

The user provides the engine with a (low-precision) query that maximizes recall, along with a small training set with samples labelled as either relevant or noisy. The engine compares the vectors of the labelled hits to those of the unlabelled hits, bearing on the assumption that unlabelled hits with vectors similar to the relevant items will be relevant themselves, and vice versa. Top-ranked sentences, plus a random set (to optimize recall), are iteratively presented to the user, who labels them as a hit or miss, further enriching the training set and increasing the engine's precision (cf. Forsberg & al. 2016).

This similarity-based search not only provides a massive speed-up compared to traditional manual filtering, its fuzzy inclusion criteria also enable it to retrieve relevant hits with non-prototypical formal characteristics (e.g. intervening words, synonyms), or even semantic or pragmatic characteristics that are otherwise hard to retrieve formally, such as negative polarity.

References

Dunn, Jonathan. 2017. Computational learning of construction grammars. 2017. *Language and Cognition* 9(2). 254-292.

Forsberg, Markus, Richard Johansson, Linnéa Bäckström, Lars Borin, Benjamin Lyngfelt, Joel Olofsson and Julia Prentice. 2014. From construction candidates to constructicon entries. *Constructions and Frames* 6(1). 114-135

Hochreiter, Sepp and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9(8). 1735-1780.

Schneider, Gerold, Marianne Hundt & Rahel Oppliger. 2016. Part-of-speech in historical corpora: Tagger evaluation and ensemble systems on ARCHER. *Konvens 2016, Bochum, 19-21 September 2016.*

Sutskever, Ilya, Oriol Vinyals and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*. 3104-3112.

Vanni, Laurent, Mélanie Ducoffe, Damon Mayaffre, Frédéric Precioso, Dominique Longrée, Veeresh Elango, Nazly Santos, Juan Gonzalez, Luis Galdo & Carlos Aguilar. 2018. Text Deconvolution Saliency (TDS): a deep tool box for linguistic analysis. *56th Annual Meeting of the Association for Computational Linguistics*, July 2018, Melbourne.

Yang, Yi and Jacob Eisenstein. Part-of-speech tagging for historical English. *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL)*.

Comparing annotation schemes across time: The problem of syntactic mapping

Eva Zehentner, Marianne Hundt, Melanie Röthlisberger and Gerold Schneider

<pdf>

Annotation quality assessment and error correction in diachronic corpora: Combining pattern-based and machine learning approaches

Tom S Juzek, Stefan Fisher, Pauline Krielke, Stefania Degaetano-Ortlieb & Elke Teich
(Saarland University)

We report on experiences from processing a diachronic corpus of scientific texts, the Royal Society Corpus (RSC ; Kermes et al. 2016). The RSC consists of contributions to the Philosophical Transactions of the Royal Society of London, i.e. mainly scientific articles, but also letters of scientific correspondence, reviews and summaries, and reports of experiments. Among other things, the RSC has been used to investigate how information density changes as new discourse types were emerging.

Our focus is on the late modern period (1665–1869) of the RSC, which consists of ca. 10 000 documents and 30 million tokens. On this basis, we are tracing the development of Scientific English, focusing on grammatical complexity. Specifically, we are interested in mapping common linguistic complexity measures, e.g. length in tokens, depth of embedding, and number of dependants, to estimates of information content. For this purpose, we need a high-quality annotation of the corpus at the grammatical level. To syntactically parse the RSC, we use the Stanford Parser with off-the-shelf settings (Klein and Manning 2003). The parser works best with normalised orthography, for which we use VARD (Baron and Rayson 2008), and with high-accuracy part-of-speech tagging, as per the parser's default POS tagger (Manning 2011). For contemporary English, such a set-up would usually result in parses with an accuracy rate of about 82% to 87% (cf. Rehbein and Ruppenhofer 2018).

In order to assess the performance of the parser on the diachronic data, we evaluated a sample of 100 parses. Sequences from 1665 are parsed with up to 45% accuracy, depending on the exact measurement criteria, and there is only a slight improvement to up to 53% for sequences from 1850. Apart from genuine parsing errors, we observe that the low accuracy rates are due to errors propagating bottom-up, such as OCR errors, spelling variants, wrong end-of-sentence detection, and wrong POS tagging.

For items sampled from 1665, about half of the errors are genuine parsing errors, the other half consists of propagating errors. For items sampled from 1850, the relative share of genuine parsing errors goes down to 38%. Presumably, the reason for this shift is that the English of 1850 more resembles contemporary English, on which the parser was trained. Spot-checks support this reasoning, as it can be fairly hard to manually parse sentences from 1665. Further error categories may be detected when we extend the sample.

Error correction is a separate step that needs its own procedures and tools. As a first step, we wish to implement advanced OCR correction (Klaus, Klakow, and Fankhauser 2019), which is based on the noisy channel model. Further, we are currently exploring majority vote approaches, using various taggers, and automatic, integrative approaches with human supervision to error detection and correction. In particular, we are exploring an approach inspired by Rehbein and Ruppenhofer (2017, 2018), who combine information-theoretic measures like entropy for error detection and semi-supervised machine learning, notably active learning, for error correction. This approach automatically detects genuine parsing errors with a good accuracy rate. The affected items are manually corrected and then used to retrain the parser. Our expectation is that this will reduce parsing errors considerably.

References

Baron, Alistair and Paul Rayson (2008). "VARD 2: A tool for dealing with spelling variation in historical corpora". In: Proceedings of the Postgraduate Conference in Corpus Linguistics. Birmingham, UK.

Kermes, Hannah et al. (2016). "The Royal Society Corpus: From Uncharted Data to Corpus". In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). Ed. by Nicoletta Calzolari (Conference Chair) et al. Portorož, Slovenia: European Language Resources Association (ELRA).

Klaus, Carsten, Dietrich Klakow, and Peter Fankhauser (2019). "OCR Post-Correction of the Royal Society Corpus based on the Noisy Channel Model". In: Deutsche Gesellschaft for Sprachwissenschaft (DGfS) Sektion Computer Linguistic - Poster-session. Bremen, Germany.

Klein, Dan and Christopher D. Manning (2003). "Accurate Unlexicalized Parsing". In: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1. ACL '03. Sapporo, Japan: Association for Computational Linguistics, pp. 423–430.

Manning, Christopher D. (2011). "Part-of-Speech Tagging from 97% to 100%: Is It Time for Some Linguistics?" In: Computational Linguistics and Intelligent Text Processing. Ed. by Alexander F. Gelbukh. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 171–189.

Rehbein, Ines and Josef Ruppenhofer (2017). "Detecting annotation noise in automatically labelled data". In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, Canada: Association for Computational Linguistics, pp. 1160–1170.

Rehbein, Ines and Josef Ruppenhofer (2018). "Sprucing up the trees – Error detection in treebanks". In: Proceedings of the 27th International Conference on Computational Linguistics. Santa Fe, New Mexico, USA: Association for Computational Linguistics, pp. 107–118.

When polysemy is what a construction is *(all) about*: Exploring the use of neural language models for semantic search and classification in (diachronic) corpora

Lauren Fonteyn

The growing size of historical corpora has led to more robust (statistically verifiable) results (Kytö 2011) and has enabled researchers to retrieve sufficient data to tackle even low-frequency phenomena (e.g. Neels 2017). However, for studies investigating infrequent functional-semantic subtypes of one particular construction in different stages of its development, large corpora are both a blessing and a curse, as larger data sets also involve larger batches of noise to wade through. To illustrate this problem, this paper focusses on the constructions exemplified in (1):

(1) I'm all **about** being friendly and open and flirty. (2007, COHA)
(2) East Lansing. They're very **into** big American cars there. (1994, COHA)

The structures in (1)-(2) illustrate (colloquial) English 'idioms' expressing a subject's state (e.g. involvement/knowledge/interest/fondness) regarding an object, which appear to have emerged roughly simultaneously in the 20th century (OED, 2018). The core element can often be characterized as a **preposition** (e.g. *into, about*); but in certain uses, they exhibit preposition-adjective underspecification (see 'very' in (2); cf. Denison 2013). Besides many unanswered questions regarding their emergence and development, these particular uses of *into* and *about* also pose challenges for traditional, structure-based methods of (semi-)automatic retrieval. This is because "more than any other word class, prepositions (…) are multiply polysemous" (Cuyckens&Radden 2002: xiii), and the vast majority of

tokens yielded by a structural search express other (mainly spatial/temporal) meanings. Thus, the retrieval of such infrequent functional-semantic subtypes of *into* and *about* would benefit substantially from a search method that detects **structural alongside semantic features**.

The method explored in this paper is based on a recent language representation model called BERT (Bidirectional Encoder Representations from Transformers; Devlin et al. 2018), which learns contextual relations between words in a text. Importantly, rather than generating embeddings of word **types** (cf. word2vec; Mikolov et al. 2013), BERT creates context-dependent vector representations from contextualized **tokens**. These token-vectors can be used to detect structural as well as semantic **similarities** between instances of *into* and *about* in the corpus (COHA) by comparing and ranking the cosine distances between an exemplary hit and the other hits (see Table 1):

Table 1 – similarity top-3 rank for two contextualized examples of *into* from COHA (2000-2009) calculated with BERT

HIT	> he was *into* his early eighties (***time stage***)	
rank	**cosine distance**	**similar example**
1	(0.23)	the project is also *into* its fifth season
2	(0.25)	quantum computing is barely *into* its proof-of-principle stage
3	(0.29)	I was *into* my second semester of college
	(…)	…
HIT	> Mia Kevin and I are so *into* each other, we didn't notice (***interested in/fond of***)	
rank	**cosine distance**	**similar example**
1	(0.14)	when I first start dating a guy, he's really *into* me
2	(0.21)	yes, George was *into* girls. girl-centipedes, that is.
3	(0.22)	unless she's really *into* asymmetrical green men
	(…)	…

With these similarity measures, we can address challenges purely structural searches cannot. This paper will demonstrate that the pos-tags in COHA (CLAWS-7, uncorrected) negatively affect recall compared to non-tagged *about* and *into*, and precision still being poor, relevant hits can only be found by (manually) filtering spatial, temporal, and other irrelevant uses (e.g. *about* occurs 724,224 times in COHA, with 0%-35% relevant hits/decade). However, the feature representations drawn from context-dependent Neural Language Models such as BERT offer a promising means of combining structural search with (semi-)automatic **functional-semantic classification** of the various senses of *about* and *into*.

References

"about, adv., prep.1, adj., and int." *OED Online*, Oxford University Press, December 2018, www.oed.com/view/Entry/527. Accessed 6 January 2019.

Cuyckens, Hubert & Günter Radden. 2002. Perspectives on Prepositions. Tübingen: Max Niemeyer Verlag.

Denison, David. 2013. Parts of speech: Solid citizens or slippery customers? *Journal of the British Academy*, 1, 151–185. DOI 10.5871/jba/001.151

Devlin, Jacob & Chang, Ming-Wei & Lee, Kenton & Toutanova, Kristina. 2018. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* https://arxiv.org/pdf/1810.04805.pdf

"into, prep. and adj." *OED Online*, Oxford University Press, December 2018, www.oed.com/view/Entry/98530. Accessed 6 January 2019.

Kytö, Merja. (2011). Corpora and historical linguistics. *Revista Brasileira de Linguística Aplicada*, 11(2), 417-457. https://dx.doi.org/10.1590/S1984-63982011000200007

Mikolov, Tomas, Sutskever, Ilja, Chen, Kai, Corrado, Greg & Jeff Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems* 26, 3111-3119.

Neels, Jakob. 2017. Low-frequency grammaticalisation and lifespan change: Insights from let alone and William Faulkner. *50th Annual Meeting of the Societas Linguistica Europaea, Zürich.*

Extraposition and Information Density in Early New High German Corpora

Sophia Voigtmann, Richard Gerbracht, Dietrich Klakow, Augustin Speyer
(Saarland University)

In Early New High German (ENHG) the nowadays common bracket structure of clauses is established. Still many apparent violations of this structure are found in the extraposition of phrases to the postfield (‘Nachfeld’) (Admoni 1990).

One of the factors responsible for that is Information Density (ID). ID (Shannon 1948) can be defined as the “amount of information per unit comprising the utterance” (Levy & Jaeger 2007, 1). It is measured via surprisal (Hale 2001). Frequent combinations have low surprisal-values whereas rare combinations have high surprisal-values.

We aim to prove a correlation between the extraposition of nominal, prepositional and adverbial phrases and ID for ENHG texts. The hypothesis is that extraposition occurs in the case of high surprisal-values and that it levels the distribution of information according to Levy’s and Jaeger’s “Uniform Information Density Hypothesis” (2007).

To build an information database for extraposition of phrases, we establish a web-based application that helps us annotating the texts imported from the DTA (Deutsches Textarchiv), because they are lemmatized, tokenized, normalized and POS-tagged. We generate a corpus of sermons from the 16th to the 18th century and scientific texts from the 17th.

With the annotated data of over 25000 sentences, we classify the occurrence of extraposition by using feature-based machine learning algorithms. To classify whether a word is inside an extraposition or not, we used Gradient Tree Boosting (Friedmann, 1999) with features based on the POS-tag of the word itself and the topology (or structure, see Wöllstein 2014) of the sentence, which we derive from the POS-tags of its words. This approach has the drawback of not considering the results for neighbouring words during the decision process.

Therefore, we shifted to using a sequence modelling algorithm in CRF (Conditional Random Fields, Lafferty et al. 2001). We base the features for the CRF-algorithm mainly on the context of the word, e.g. unigrams, bigrams and trigrams containing the POS-tag of the word and its direct neighbours, bigram for the predicted classes of the word and its predecessor. Adding the previously engineered features (which were based on topology) to the feature set does not yield a significant improvement to the results.

Since the occurrence of extraposition is lower than 5%, our training dataset suffers from class imbalance, which leads to high accuracy scores (>95%) for the classifier, while having poor precision (~57%) and recall (~32%) on the minority class, i.e. extrapositions (accuracy paradox). To combat the low recall values, we undersample the training dataset by deleting every n-th sentence that does not include at least one extraposition. Note that the smaller the n is, the higher the achieved recall will be.

The resulting classifier facilitates an active learning process (Cohn et. al 1994), in which we prioritize labelling sentences from the DTA, that the classifier is most uncertain about (uncertainty sampling) with the expectation that labelling those gives the largest improvement in precision and recall of the automatic classifier.

References

Admoni, Wladimir. 1990. Historische Syntax des Deutschen. Max Niemeyer Verlag. Tübingen.

Cohn, D., Atlas, L. and Ladner, R. (1994). Improving generalization with active learning. Machine Learning, 15(2), pp.201-221.

Friedman, Jerome. 1999. Greedy Function Approximation: A Gradient Boosting Machine. URL: https://statweb.stanford.edu/~jhf/ftp/trebst.pdf.

Hale, John. 2001. A probabilistic Early parser as a psycholinguistic model. Proceedings of the second meeting of the North American chapter of the Association for Computional Linguistics.

Lafferty, John, Andrew McCallum and Fernando C.N. Pereira. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. Proceedings of the 18th International Conference on Machine Learning 2001 (ICML 2001); 282-289.

Levy, Roger and T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. In Bernhard Schöllkopf, John Platt & Thomas Hoffman (ed.), Advances in Neural Information Processing Systems 19, 849-856. Cambridge Ma.

Shannon, C. E. 1948. A Mathematical Theory of Communication. Bell System

Technical Journal 27. 379-423; 623-656.

Wöllstein, Angelika. 2014. Topologisches Satzmodell. 2. überarb. Aufl., Heidelberg: Winter.

WORKSHOP 21

Schedule: Thu 14.00 – Fri 17.25 (Room 1)

Tense, aspect and mood categories across languages

Kilu von Prince, Ana Krajinović & Jozina Vander Klok

Keywords: tense; aspect; modality; mood; typology; formal approaches

There is a proliferation of terms in the study of tense, aspect and modality (TAM). But it is not always clear what the relation between those terms is.

In some cases, several terms appear to refer to phenomena that are functionally very similar. This might be the case with continuous and progressive aspect, or irrealis, potential and hypothetical mood. In other cases, one category might be a special case of another category, for example habitual sentences are sometimes described as a special case of generic statements (Krifka et al. 1995) or, more generally, of imperfective aspect (Comrie 1976). We would like to bring together perspectives from language documentation and language-specific description, typology, formal and functional approaches to semantics and pragmatics, as well as syntax and morphology to discuss relations between TAM terms cross-linguistically.

We also welcome contributions that specifically address discrepancies between linguistic subdisciplines, or to state it from another perspective, how different linguistic frameworks might constrain or create more freedom in their approach to accounting for TAM semantic properties.

For instance, typologists often come to different conclusions from formal semanticists when it comes to the classification of TAM categories. Typologists tend to assume that cross-linguistic differences in the distribution of particular TAM expressions are based on their lexical semantic definition. By contrast, in formal semantics some of these differences can be derived from various language-internal structures and processes, such as paradigmatic effects.
Examples for this include:

- English simple past has a discontinuous implicature in connection with stative verbs as in "Rose is in the hospital. She had trouble breathing" -- such an utterance implicates that Rose is now better able to breathe. Altshuler & Schwarzschild (2012) argue that this implicature is a result of the contrast of English simple past and simple present.

In some languages, such discontinuous interpretations are however not restricted to stative verbs. This has led Plungian & van der Auwera (2006) to diagnose a specific typological category of "discontinuous past". By contrast, Cable (2016) and von Prince (2017) have argued that in those languages, too, the discontinuous interpretation is a result of the paradigm in which the past markers find themselves, rather than a function of their lexical meaning.

- The category of "iamitive aspect" has been introduced by Olsson (2013) and Dahl & Wälchli (2016) similar to, but separate from, both perfect aspect and "already". One of its defining features is the change-of-state interpretation that iamitives show in connection with stative verbs, unlike Indo-European perfects. But more recent research argues that these interpretations may equally be expressed by perfect aspect (Krajinović 2018 for Nafsan) because of language-specific processes (also compare the analysis of "already" by vander Klok & Matthewson (2015) for Javanese). This poses a general question of whether the different functions of perfect attested cross-linguistically can be

explained by language-specific effects, and how different would these functions need to be to justify establishing a new typological gram.

- Languages that do not have obligatory tense marking are commonly categorized as tenseless. But Matthewson (2006) has argued that St'a´t'imcets has a zero tense morpheme, which restricts the temporal reference of a clause to non-future tense.
By contrast, Mucha (2012) argues for Hausa that it is genuinely tenseless and that temporal reference is determined by pragmatic defaults.

-This divide between sub-disciplines is also reflected in how modal meanings are categorized. Thus, the distinction between participant-internal and participant-external modalities, which is fundamental to typological studies of modality (e.g. Bybee et al. 1994, van der Auwera & Plungian 1998), is not reflected in the formal semantics discourse on modality (e.g. Portner 2009).

Independent from particular theoretical perspectives, there are many cases in which the relation between categories it not entirely clear, for example: How does iterativity relate to pluractionality? What is the relation between sequentiality, prospective aspect and perfect aspect, frequentatives and habituals, progressive and continuous aspect? What is the relation between "timitive", "avertive", "apprehensive" modality, also called "volitive of fear" (Vuillermet 2018)?
How many modal distinctions such as "deontic" and "circumstantial" do languages mark grammatically?

We invite contributions from a variety of backgrounds and perspectives, including language documentation and description, typology, semantics and pragmatics, syntax, and morphology. We particularly welcome submissions on understudied or underdocumented languages. Possible topics for submission include:
- Descriptive case studies of a specific TAM marker or paradigm;
- Typological studies of the distribution of certain categories;
- Formal and functional approaches to specific TAM categories;
- Formal and functional approaches to the grammaticalization/diachrony of specific TAM categories;
- Interdisciplinary studies that compare different approaches.

References

Altshuler, Daniel and Roger Schwarzschild. Moment of change, cessation implicatures and simultaneous readings. In Emmanuel Chemla, Vincent Homer, and Gr´egoire Winterstein, editors, Proceedings of Sinn und Bedeutung 17, Paris, 2012.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Dahl, Östen & Bernhard Wälchli. 2016. Perfects and iamitives: two gram types in one grammatical space. *Letras de Hoje* 51(3). 325–348. doi:http://dx.doi.org/10.15448/1984-7726.2016.3.25454.

van der Auwera Johan and Vladimir A. Plungian. Modality's semantic map. Linguistic typology, 2(1):79–124, 1998.

Cable, Seth. The implicatures of optional past tense in Tlingit and the implications for 'discontinuous past'. Natural Language & Linguistic Theory, 2016.

Comrie, Bernard. Aspect. Cambridge Textbooks in Linguistics. Cambridge University Press, 1976.

Krajinović, Ana. The semantics of perfect in Nafsan. Paper presented at The Semantics of African, Asian, and Austronesian Languages (Triple A) 5, 27–29 June 2018, Konstanz, 2018. https://semanticsofaaa2018.files.wordpress.com/2018/07/krajinovic_slides.pdf.

Krifka, Manfred, Francis J Pelletier, Gregory N Carlson, Gennaro Chierchia, Godehard Link, and Alice Ter Meulen. Introduction to genericity. In Gregory N Carlson and Francis J Pelletier, editors, The Generic Book, pages 1–124. University of Chicago Press, Chicago and London, 1995.

Matthewson, Lisa. Temporal semantics in a supposedly tenseless language. Linguistics and Philosophy, 29:673–713, 2006.

Mucha, Anne . Temporal reference in a genuinely tenseless language: the case of Hausa. Proceedings of SALT, 22:188–207, 2012.

Olsson, Bruno. Iamitives: Perfects in Southeast Asia and beyond. Stockholm University, 2013. MA thesis.

Plungian, Vladimir A. and Johan van der Auwera. Towards a typology of discontinuous past marking. Sprachtypologische Universalienforschung, 59:317–349, 2006.

Portner, Paul. Modality. Oxford University Press, 2009.

Vander Klok, Jozina and Lisa Matthewson. Distinguishing already from perfect aspect: A case study of javanese wis. Oceanic Linguistics, 54(1):172–205, 2015.

von Prince, Kilu. Paradigm-induced implicatures in TAM-expression: A view from the Daakaka distal. In Proceedings of Sinn und Bedeutung 21, Edinburgh, 2017.

Vuillermet, Marine. 2018. Grammatical 'fear' morphemes in ese ejja: Making the case for a morpho-semantic apprehensional domain. Studies in Language, 42(1), 2018.

Discontinuous past is real: evidence from Northwest Caucasian

Evgeniya Klyagina & Peter Arkadiev
(Higher School of Economics; Russian State University for the Humanities & Institute of Slavic Studies of the Russian Academy of Sciences; Russian State University for the Humanities)

Keywords: Northwest Caucasian, tense-aspect, discontinuous past, semantics

The work has been supported by Russian Foundation for Basic Research, grant #17-04-00444.

"Discontinuous past" is the term proposed in Plungian & van der Auwera (2006) for grams expressing that a past situation is no longer holding at speech-time. Recently, it has been argued by Cable (2017) that discontinuous past interpretation is only available in languages with optional past markers and is just an implicature arising from paradigmatic opposition between such markers and other tense grams (cf. also von Prince 2017). We challenge both these claims by the evidence from the Northwest Caucasian languages Abaza and Kabardian.

Polysynthetic NWC languages have elaborate systems of obligatory tense marking. In the past domain, perfective and imperfective tenses are distinguished. In addition to that, NWC languages distinguish between simple and retrospectivized tenses, the latter built from the former by adding one of the past tense markers (e.g. Kabardian general past *-a*, imperfective past *-t* vs. retrospectivized past *-a-t*). Retrospectivized past tenses in NWC fit squarely into the domain of discontinuous past marking. In contrast to the Tlingit past tense discussed by Cable (2017), they are mainly used in perfective contexts with telic verbs implying that the resultant state of the event was canceled.

In order to test whether the canceled result interpretation of the retrospectivized past forms in Abaza and Kabardian is an implicature or part of meaning, we have conducted an experiment asking

informants to evaluate constructed dialogues containing simple and retrospectivized past forms and subsequent utterances inconsistent with either the canceled result or the persistent result interpretations. Thus, in (1–4) from Kabardian, line A contains simple (1–2) or retrospectivized (3–4) past tense forms of the verb 'to lose', while line B gives an utterance inconsistent with either the canceled result (1,3) or the persistent result (2,4) interpretation. Our Kabardian informants judged the dialogues (2) and (3) to be logically inconsistent.

(1) simple past, persistent result: OK

A: *s-jə-pasportə-r* ***s-ʁeḳʷed-a-ŝ***
1SG.IO-POSS-passport-ABS 1SG.ERG-lose-PST-DCL
'I have lost my passport.'

B: [ok]*q̇ə-p-xʷe-s-ʁʷet-əw* *sə-p-deʔepəq̇ʷə-n?*
DIR-2SG.IO-BEN-1SG.ERG-find-ADV 1SG.ABS-2SG.IO-help-POT
'Can I help you find it?'

(2) simple past, canceled result: ill-formed

A: *s-jə-pasportə-r* ***s-ʁeḳʷed-a-ŝ***
1SG.IO-POSS-passport-ABS 1SG.ERG-lose-PST-DCL
'I have lost my passport.'

B: *#dene* *a-r* *q̇ə-z-de-b-ʁʷetə-ž-a-r?*
where DEM-ABS DIR-REL.IO-LOC-2SG.ERG-find-RE-PST-ABS
'Where did you find it?'

(3) retrospectivized past, persistent result: ill-formed

A: *s-jə-pasportə-r* ***s-ʁe-ḳʷed-a-t***
1SG.IO-POSS-passport-ABS 1SG.ERG-lose-PST-IPF
'I lost my passport.'

B: *#q̇ə-p-xʷe-s-ʁʷet-əw* *sə-p-deʔepəq̇ʷə-n?*
DIR-2SG.IO-BEN-1SG.ERG-find-ADV 1SG.ABS-2SG.IO-help-POT
'Can I help you find it?'

(4) retrospectivized past, canceled result: OK

A: *s-jə-pasportə-r* ***s-ʁeḳʷed-a-t***
1SG.IO-POSS-passport-ABS 1SG.ERG-lose-PST-IPF
'I lost my passport.'

B: [ok]*dene* *a-r* *q̇ə-z-de-b-ʁʷetə-ž-a-r?*
where DEM-ABS DIR-REL.IO-LOC-2SG.ERG-find-RE-PST-ABS
'Where did you find it?'

Despite some variation, our results show quite robustly that neither the persistent result interpretation of the simple past nor the canceled result interpretation of the retrospectivized past can generally be overridden in subsequent discourse. Hence, both interpretations should be considered as belonging to the meaning of the respective grammatical forms.

Abbreviations

ABS — absolutive; ADV — adverbial; BEN — benefactive; DCL — declarative; DEM — demonstrative; DIR — directional preverb; ERG — ergative; IO — indirect object; IPF — imperfect; LOC — locative preverb; POSS — possessive; POT — potential; PST — past; RE — refactive; REL — relativizer; SG — singular.

References

Cable, Seth (2017), The implicatures of optional past tense in Tlingit and the implications for 'discontinuous past', *Natural Language and Linguistics Theory* 35(3), 635–681.
Plungian, Vladimir, and Johan van der Auwera (2006), Towards a typology of discontinuous past marking, *Sprachtypologie und Universalienforschung* 59(4), 317–349.
von Prince, Kilu (2017), Paradigm-induced implicatures in TAM-expression: A view from the Daakaka distal, *Proceedings of Sinn und Bedeutung 21*.

ꕥ

(Im)perfectivity and phasal aspect in Äiwoo

Åshild Næss
(University of Oslo)

Keywords: Oceanic, aspect, perfective, perfect

The Oceanic language Äiwoo combines (near-)obligatory aspect-mood prefixes with optional phasal-aspect clitics in its verb complex. In this talk I will discuss the functions of these morphemes and their various possible combinations, and how they relate to established aspectual categories such as perfective, perfect, and aspectual particles like English 'already'. I will focus, firstly, on the aspect prefix *i-*, which seems to conform to established definitions of a perfective in some contexts (contrasting ongoing with intruding events as in (1), where imperfective *ki-* is used on the verb denoting the ongoing event and *i-* on the verb denoting the intruding event), but not in others. In particular, it is used with unchanging states (2), which may be compatible with Comrie's (1976) definition of perfectivity as «involv[ing] lack of explicit reference to the internal temporal constituency of a situation, rather than explicitly implying the lack of such internal temporal constituency», but not Timberlake's (2007): «[a] perfective presupposes a situation consisting of three phases: a prior situation in which there is no activity or no state holds, then a phase of change or transition, and an ensuing situation after which no more change is to be expected.» This suggests that perhaps the perfective/ imperfective distinction can be approached from two different angles from a crosslinguistic perspective: imperfectivity as the absence of perfectivity, i.e. the casting of a situation as bounded, vs perfectivity as the absence of imperfectivity in the sense of internal temporal structure. The latter would cover the functions of Äiwoo *i-*.

(1) *Lâ* *ku-wokâu=kâ* *mo* *temaale* *i-pu-eagâ-mä.*
DIST IPFV-bathe=DIST CONJ needlefish ASP-come-quiet-DIR:1
'While [the hermit crab] was bathing, the needlefish came quietly.'

(2) *Nuwosi* *nä* *manioki* *i-to.*
kind of.3MIN manioc ASP-exist
'There are [different] kinds of manioc.'

Secondly, I will address the semantics of the phasal-aspect clitic *=to*, and how it relates to categories such as perfect aspect and 'already'. I will suggest that *=to* shares with 'already' the assertion that a certain state of affairs holds during a time interval t, as suggested by Krifka (2000: 402); but not the presupposition also assumed by Krifka for 'already', namely that the state of affairs did not hold before t. The latter is a frequent interpretation of *=to* in combination with the imperfective prefix (3), but not a necessary one (4). Thus aspectual semantics in Äiwoo illustrates some ways in which aspect marking can involve variations on established crosslinguistic themes.

(3) *Ki-li-mei=**to**=wâ*
IPFV-3AUG-sleep=TO=DIST
'They fell asleep [i.e. are sleeping now but were not previously].'

(4) *Nugo nâbulä lâ memave=kâ,*
? upper.arm.3MIN DIST weak=DIST
*lâto ki-täve=**to**=wâ.*
thus IPFV-hang=TO=DIST
'His arms got weak [from hanging from the branch], so he kept hanging [because he wasn't able to get back up].'

Abbreviations
ASP aspect prefix (function to be discussed), AUG augmented number, CONJ conjunction, DIR directional, DIST distal deictic marker, IPFV imperfective, MIN minimal number

References
Comrie, Bernard. 1976. Aspect. Cambridge: Cambridge University Press.
Krifka, Manfred. 2000. Alternatives for aspectual particles: semantics of still and already. *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*, 401–412. Berkeley: Berkeley Linguistics Society.
Timberlake, Alan. 2007. Aspect, tense and mood. In Timothy Shopen (ed.), Language typology and syntactic description. Volume III: Grammatical categories and the lexicon, 280-333. Cambridge: Cambridge University Press.

ø

Diagnosing pluractionality: the view from Seri

Jérémy Pasquereau & Patricia Cabredo Hofherr
(Surrey Morphology Group, U. Surrey & UMR 7023 CNRS/U. Paris-8)

Keywords: Seri, verbal morphology, verbal paradigm, verbal plurality, imperfective aspect

We present a study of stem-alternations in Seri (isolate) based on first-hand data against the backdrop of the proliferation of terms in the field of event plurality.

The majority of Seri verbs have 4 stem forms. The stem alternations mark subject number and a second category involving multiple events (glossed MULT following Cabredo, Pasquereau & O'Meara 2019, henceforth CPO 2019). Our aim is threefold: (i) describe the semantic profile of Seri MULT-forms, (ii) examine the relationship between event plurality and related categories proposed in the literature, (iii) propose a list of questions/diagnostics for describing the semantic profile of pluractional forms cross-linguistically.

Seri MULT-forms have been analysed as imperfective aspect or as verbal number (e.g. Marlett 2016). (For a subset of verbs, MULT-forms have been argued to show interactions with object number; this does not apply to MULT-forms generally, see CPO 2019) To examine the exact contribution of Seri MULT-forms we therefore require operational diagnostics to distinguish event plurality marking from imperfective aspect. Applying the criteria in Cover & Tonhauser (2015) shows that MULT-forms do not mark imperfective aspect. Firstly, CPO 2019 show that MULT-forms do not have habitual or

continuous readings. Secondly, the MULT-form of a telic predicate like `go to X' can be used truthfully in a context where the eventuality-time is included in the topic-time (1). Finally, MULT-forms cannot be used truthfully in imperfective contexts (with topic-time included in the eventuality-time) if (sub-)events cannot be individuated in the context (2, context-1) .

(1) Context: Yesterday, I went to Puerto Libertad several times in the morning.
Moxima, Xpanohax co-nt-h-ay-a-**tim**.
yesterday Puerto_Libertad 3IO-AWAY-1SG-GO-RLS.YO-**MULT**
Yesterday, I went to Puerto Libertad (several times). **TRUE** [elicitation]

As shown in CPO 2019, like other markers of event plurality, the plurality of events expressed by Seri MULT-forms (i) cannot be counted by adverbial expressions of exact cardinality, and (ii) cannot multiply singular indefinites (see Laca 2006 for these diagnostics). Since Dressler 1968, Cusic 1981 we know that event-plurality markers are a heterogeneous class cross-linguistically. We examine MULT-forms in Seri in relation to other markers of event plurality. We show that MULT-forms in Seri are `distributive markers' in the sense of the Americanist linguistics tradition (Boas 1911:37, Mithun 1988:220 on Nass Tsimshian) – with distributive configurations providing individuation criteria for events in the plurality: e.g. (2, context 2) where contextual information has to make individuation of distinct subevents of running salient.

(2) Hoyacj quih coipanozx-**im** iti, hamac cánoj quih
1POS.brother DEF.FLX 3IO.3POS.OBL.NMLZ.run-**MULT** while fire SUBJ.NMLZ.buzz DEF.FLX
iicot cöyooctim.
3POS.between 3IO.RLYO.cut
While my brother was running, the light went out.
Context 1: Yesterday, my brother ran a race. While he did, the light went out. **FALSE**
Context 2: My brother was doing a treasure hunt. (=running between clues) While he did, the light went out. **TRUE** [elicitation]

Building on Laca 2006 we propose diagnostic contexts for the cross-linguistic description of the distributive configurations available for event-plurality markers wrt different pluralities (participants, locations, times) of different syntactic forms (implicit pluralities, plural definites, QPs, numeral+noun, coordinations) and roles (subject/object/adjunct).

References

Boas, Franz (1911), Introduction. *Handbook of American Indian Languages*, Vol. 1, 1-83.

Cabredo Hofherr, Pasquerau & O'Meara (2019), Event plurality in Seri. *Proceedings of Semantics of Under-Represented Languages in the Americas* 10: 1-16.

Cover, Rebecca T. & Judith Tonhauser (2015), Theories of meaning in the field: temporal and aspectual reference. In R. Bochnak & L. Matthewson (eds), *Methodologies in semantic fieldwork*, Oxford University Press, 306-350.

Cusic, David Dowell (1981), Verbal plurality and aspect. Doctoral dissertation, Stanford University.

Dressler, Wolfgang (1968), *Studien zur verbalen Pluralität*. Österreichische Akademie der Wissenschaften. Phil.-hist. Klasse. Sitzungsberichte. Bd. 259. Abh. 1. Wien: Bühlau in Kommission.

Laca, Brenda (2006), Indefinites, quantifiers, and pluractionals. What scope effects tell us about event pluralities, in ed. S. Vogeleer & L. Tasmowski (eds), *Non-definiteness and plurality*, Amsterdam/ Philadelphia: John Benjamins, 191–217.

Marlett, Stephen A. (2016), *Cmiique Iitom: the Seri language*. Unpublished grammar (2016 draft). http://arts-sciences.und.edu/summer-institute-oflinguistics/ faculty/marlett-steve/marlett-seri-grammar-latest-draft.pdf.

Mithun, Marianne (1988), Lexical categories and the evolution of number marking, in M. Hammond and M. Noonan (eds), *Theoretical morphology*, New York: Academic Press, 211–234.

ꕥ

Realistic vs. hypothetical modality in Jaminjung (Australia)

Patrick Caudal & Eva Schultze-Berndt
(CNRS and Université Paris-Diderot & University of Manchester)

Keywords: possibility modal; circumstantial modal; apprehensional; tense-modal interaction; negation

In this paper, we argue that Jaminjung, a Western Mirndi language spoken in Northern Australia, displays a clearcut formal distinction between a realistic (circumstantial) and a hypothetical (general/stereotypical possibility) modal. These modals are verbal affixes which between them carve up the entire semantic domain of non-epistemic modality and future time reference; epistemic modality has a distinct manifestation as a modal adverb. Like a number of other non-European languages (e.g. Davis et al. 2009; Deal 2011), Jaminjung does not lexicalize modal force: it only has possibility modals.

The realistic modal has an underspecified circumstantial conversational background in the sense of Kratzer (1981; 1991: 664), as refined by Thomas (2014: 439): accessible possible worlds are compatible with the speaker's perception of relevant facts in the actual world at the time of modal evaluation. In affirmative contexts, only this modal is compatible with additional deontic, teleological or bouletic conversational backgrounds. A typical (future-oriented) usage of the modal is illustrated in (1). This modal can also combine with past imperfective verb forms, with a past counterfactual reading.

(1) **na-w-ijga** warnanggal-bina, juwud=biya jarlag **na-w-iyaj**
2SG-CIRC-go doctor-ALL eye=then good 2SG-CIRC-be
(Discussing eye operation appointments:)
'you should go to the doctor, your eyes will/may become (lit. be) good then'

The hypothetical modal is incompatible with tense marking. In affirmative contexts, it is employed to describe natural laws but also stereotypical behaviours of natural kinds including humans. Example (2), unlike (1), is about a hypothetical encounter with a healer and its generic features. This modal also appears in all negative modal contexts except in past tense.

(2) mayi=biya warrng-warrng **yaniny-garrga**, (...)
person=then RDP-walk HYP:3SG>2SG-approach
jarlag **yaniny-gilinyma**
good HYP:3SG>2SG-make
(Discussing methods of traditional healers:)
'The person may walk up to you, (touch you, and) make you well'

Modals with the properties of the Jaminjung hypothetical modal have not received much attention in the linguistic literature, but may be widespread in Australian and other lesser described languages. Notions that come closest to capturing the semantics of this modal are stereotypical modality based on "the normal course of events" (Kratzer 1981: 45), DePraetere & Reed's (2011) "general possibility", Müller's (2012) "abstract X-possibility", and the aleatory uncertainty invoked in the literature on risk assessment (e.g. Paté-Cornell 1996).

The lack of interaction of the hypothetical modal with tense can be accounted for by the assumption that general possibilities are timeless (Müller 2012: 57–58). Considering realistic possibilities a subset of general possibilities (in other words, any possibility arising in the actual world has to be in line with natural laws and expected behaviours) can account for the sole use of the hypothetical modal under negation, since a negated hypothetical proposition entails the corresponding proposition involving the realistic modal.

Finally, we need to account for a further observation: the frequent use of the hypothetical modal in an apprehensional function to express possible but undesirable events, e.g. in warnings. We argue that this interpretation arises through pragmatic implicature: a possibility that is to be avoided is presented as a merely hypothetical, non-contingent possibility.

References

Davis, Henry, Lisa Matthewson, and Hotze Rullmann (2009), 'Out of control' marking as circumstantial modality in St'át'imcets, in L. Hogeweg, H. de Hoop, and A. Malchukov (eds), *Cross-linguistic Semantics of Tense, Aspect and Modality*, Amsterdam: John Benjamins, 205–244.

Deal, Amy Rose (2011), Modals without Scales, *Language* 87(3), 559-585.

Depraetere, Ilse, and Susan Reed (2011), Towards a more explicit taxonomy of root possibility in English. *English Language and Linguistics* 15(1), 1–29.

Kratzer, Angelika (1981), The notional category of modality, in H.-J. Eikmeyer, and H. Rieser (eds), *Words, Worlds, and Contexts*, Berlin: Mouton de Gruyter, 38-74.

Kratzer, Angelika (1991), Modality, in A. v. Stechow, and D. Wunderlich (eds), *Semantik / Semantics. An International Handbook of Contemporary Research*, Berlin: Mouton de Gruyter, 639-650.

Müller, Thomas (2012), Branching in the landscape of possibilities, *Synthese* 188, 41–65.

Paté-Cornell, M. Elizabeth (1996), Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety* 54(2-3), 95-111.

Thomas, Guillaume (2014), Circumstantial modality and the diversity condition. *Proceedings of Sinn und Bedeutung* 18, 433-450.

Tense, modality and aspect and the Krio verb cluster

Beatrice Nyampong, Anna Struck & Dorothee Beermann
(University of Trondheim, University of Tübingen & University of Trondheim)

Keywords: Tense-Mood-Aspect, Krio, verb clusters, preverbs

Krio is an English-based Creole spoken along the coast of West Africa (Finney 2008). Not unlike other creoles, it marks the past and the future tense while the present remains unmarked. TMA features are realised by verbal words which combined with the main predicate form a verb cluster expressing the language's Tense Mood Aspect system. We suggest that such a TMA system can be best

understood through an analysis of the relations between the verbs that form the language's verb cluster (Beermann 2019). At present the standard way to describe TMA features for Krio is by ascribing them to the main verbs, or by presenting verb clusters of the Krio on a par with the IE periphrastic tenses where a main verb combines with an auxiliary. Krio verb clusters behave differently. Take *dɔn de it*, it is by form a perfect continuous 'has been eating' which differs from its English counterpart in that none of the Krio verbs is inflected; dɔn is not finite, nor are the following verbs dependent forms. Take for example the unit *De it* which as we just saw may occur following *dɔn* in *dɔn de it.* Different from IE periphrastic constructions *de (it)* is not a dependent verb. Instead it can head main clauses expressing the continuous form of 'eat'. Notice that Krio main verbs can just be specified for aspect. This is different from English where main verbs need to be finite, that is, they need to be anchored in time.

TMA marked verbs, which in Krio are identical to verbs occurring in verb clusters, are not necessarily interpreted according to what their form suggests. A comprehensive representation of the Krio TMA system requires not only a morpho-syntactic analysis of the verb cluster but also a careful review of the conversational context in order to capture the interpretation of non-tensed verbs (e.g. Holm (2000), Velupillai (2003)). Let's look at the Krio perfect. It is formed with the help of *dɔn* which, as a tense marker, must occur in a preverbal position:

(1)	bifo	a	sɛt	mi	yai	ɛn	opin	am	a	**go**	**dɔn**	**dɔn**
	before	1SG	*close*	1SG	*eye*		*open*	3SG	1SG	**FUT**	**PRF**	***finish***

'Before I close my eyes and open them, I will have finished.'

The Krio perfect is not a 'present perfect', as it is sometimes called, it denotes an event that occurs prior to another event in the past, present or future. Considering (1), the expression *go dɔn dɔn* consists of the preverbs *go* and *dɔn*, and the main verb *dɔn*. The preverb *dɔn* locates the finishing event before the blinking event while *go* locates the event in the immediate future.

Anterior markers like *bin* and *dɔn*, that partly constitute preverbs of the Krio TMA system, can be found in varying versions in other creoles as well. On the basis of these observations, we will suggest that the Krio TMA system is modelled on the preverb clusters described for West African languages (Dakubu 2008, Dakubu et al. 2007).

References

Beermann, Dorothee (2019) Tense, Modality and Aspect in Krio – A corpus-based study. **J**ames Essegbey, Dalina Kallulli, Adams Bodomo (eds.): *The Grammar of Verbs and their Arguments*. Rüdiger Köppe Verlag.

Finney, Malcom, A. (2008) *The Origin, Historical Development, and Linguistic Properties of Krio.*

Holm, John. (2000). *Cambridge Textbooks in Linguistics: An Introduction to Pidgins and Creoles.* Cambridge, GB: Cambridge University Press.

Dakubu, May Esther Kropp, Hellan, Lars; Beermann, Dorothee. (2007) Verb Sequencing Constraints in Ga: Serial Verb Constructions and the Extended Verb Complex. *Proceedings of the International Conference on Head-Driven Phrase Structure Grammar*.

Dakubu, M.E.K. (2008). Ga verb features. In Ameka and Dakubu eds., *Aspect and Modalityin Kwa Languages.* John Benjamins Publishing Co. Pp. 91-134.

Velupillai, V. (2003). *Hawai'i Creole English: A Typological Analysis of the Tense-Mood-Aspect System.* Palgrave.

Gram types as clusters in grammatical space

Östen Dahl
(University of Stockholm)

Keywords: gram types, grammatical space, perfects, iamitives, parallel corpora

Earlier work in grammatical typology has argued for the postulation of "cross-linguistic gram types", under which language-specific grammatical items, or "grams" in the terminology proposed by Joan Bybee and her associates, could be subsumed. Such gram types need not be thought of as being specified in an innate universal grammar, but the assumption has tended to be that they are discrete entities, separate from each other. In a data-driven corpus-based approach, comparison between grams is not done in terms of postulated abstract meaning but rather by studying their distributions in parallel texts. Thus, language-specific grams can be seen as points in a multidimensional "grammatical space", where closeness between items indicates similarity in distribution, which is assumed to reflect similarity in meaning. Cross-linguistic gram types will now correspond to clusters of grams in such a grammatical space, at what is called "attractors" in dynamical system theory, that is, points towards which systems tend to evolve. Such clusters would reflect general tendencies in the evolution of grammatical systems but would not necessarily be neatly delineated from each other and could also be related hierarchically. A large and relative sparse cluster could contain smaller and denser subclusters, which might reflect areally restricted tendencies.

In order to study grammatical space, we need data about the distributions of grammatical items and ways to visualize clusters. The most convenient way of obtaining comparable data from different languages is via parallel corpora. The availability of parallel texts in minor and/or non-European languages is at present mainly restricted to translations of religious texts such as the New Testament, but in spite of the limitations, much valuable information can be extracted this way. The visualization of clusters can be done by dimension-reducing techniques such as Multidimensional Scaling and statistical cluster analysis.

Perfects or anteriors, as exemplified by the English Perfect and other similar grams, have been proposed as a cross-linguistic gram type with representative in many parts of the world. At the same time, the grams labelled as perfects show considerable variation. In Dahl & Wälchli (2016), a large-scale parallel corpus of Bible texts was used to look at items with distributions similar to perfects, including those recently labelled as "iamitives", which tend to represent grammaticalizations of words with meanings like 'already'. The study found no clear boundaries within a grammatical space including all these items, at the same time as there was a very large variation within the purported "iamitives", both with respect to frequency in the corpus (with grammaticalized items being up 50 times as frequent as their assumed sources) and with respect to their convergence with perfects of the traditional kind. Such a situation is natural if we assume that grammatical diversity is driven by a manifold of factors, both global and local. But it necessitates a more flexible view of "gram types". In particular, we need to be able to accommodate situations where grammaticalizations from different sources lead to partial convergence of grams with retainment of earlier uses.

References

Dahl, Östen & Bernhard Wälchli. 2016. Perfects and iamitives: two gram types in one grammatical space. *Letras de Hoje* 51(3). 325–348. doi:http://dx.doi.org/10.15448/1984-7726.2016.3.25454.

The future of *already* in Singapore English

Debra Ziegeler
(Université Sorbonne-Nouvelle Paris III)

Keywords: contact English, aspect, tense, grammaticalization, renovation.

The use of the adverb *already* in Singapore Colloquial English, or Singlish, has long been known as one of the most readily recognizable features defining the contact dialect, marking aspectual nuances such as anterior, completive, inchoative and inceptive functions, as noted by Bao (2005), for example. Bao (2015) discusses the functional parallels with the Mandarin Chinese perfective marker *le*, noting that all the four functions of *already* can be expressed by *le.* However, the syntactic ambiguities occurring in the use of *already* are not found in Mandarin; the completive, for example, occurs post-verbally, while the other three uses are marked clause-finally, for example:

(i) a. Xia4 le yu3
down ASP rain
'It rained' (completive)

b. Xia4 yu3 le
down rain ASP
'It started to rain/is about to rain' (inceptive/inchoative)
(Singlish: '(It) rain already.')

While the anterior use, and to some extent, the completive use, can both be considered as functionally equivalent to the perfect aspect in other languages (see, e.g. Bybee, Perkins & Pagliuca 1994), or the past tense, as shown above, the use of *already* as an inchoative or inceptive marker, expressing a change of state meaning in the former case, or a start of a new habitual state in the latter, are more familiar to the function of the adverb as a perfective marker, marking the starting boundaries of a new situation rather than the terminative boundaries of an old one. The co-existence of such functions indicates that the adverb is already grammaticalizing to express perfective as well as perfect functions. It is perhaps less often recognized, though, that the aspectual use of *already* co-exists with the variable marking for past tense in Singlish (as discussed by Ho & Platt 1993), and that both the aspectual adverb and the past tense may be seen to co-occur in the same construction.

This situation raises a number of questions, amongst them, whether the aspectual situation is as stable as it might seem at first, and whether the presence of the standard sub-variety of Singapore English may be exerting an influence over the present-day tense-aspect system of Singapore Colloquial English. The standard use of *already* is seen to co-occur in the same data, alongside the aspectual uses. Furthermore, it has already been noted, in an earlier study (Ziegeler & Lee 2019) that the present-day distribution of *already* is affected by lexical persistence from the lexifier, standard (Singapore) English. The frequency of *already* in its various functions is examined in a sub-section of 120 occurrences of spoken data from the ICE-Singapore Corpus, and 133 internet blog-postings of the *Flowerpod* Corpus, and the tendency for co-occurrence with the past tense is investigated across both corpora. It is hypothesized that, rather than grammaticalising onwards to become a past tense marker itself, as observed for creole Portuguese *ya* ('already') by Clancy Clements (2006), *already* is becoming increasingly restricted in its functional range in today's Singlish, and that its perfect and

completive functions may be at a stage of renovation by the use of the past tense in Standard Singapore English.

References

Bao, Zhiming (2005), The Aspectual System of Singapore English and the SystemicSubstratist Explanation, *Journal of Linguistics* 41, 237–267.

Bao, Zhiming (2015), *The making of vernacular Singapore English*, Cambridge: CambridgeUniversity Press.

Bybee, Joan, Revere D. Perkins & William Pagliuca (1994), *The evolution of grammar:Tense, aspect and modality in the languages of the world*, Chicago: University ofChicago Press.

Clancy Clements, J. (2006). The lexicalization-grammaticalization continuum, in J. ClancyClements, Thomas A Klingler, Deborah Piston-Hatlen & Kevin J. Rottet (eds.),(2006), *History, Society and Variation*, Amsterdam/Philadelphia: Benjamins, 77-101.

Ho, Mian Lian and John Platt (1993), *Dynamics of a contact continuum*. *Singaporean English*, Oxford: Clarendon.

Counterparts of perfect in Paraguayan Guaraní

Dmitry Gerasimov

<pdf>

Building aspectual futures: Evidence from Gitksan

Lisa Matthewson, Michael Schwan and Neda Todorović
(University of British Columbia)

Keywords: Progressive, inceptive, future, tense, Gitksan

Overview. Some languages have two grammaticized future constructions, one of which incorporates progressive aspect, e.g. English *will* vs. *be going to*. We expand the empirical coverage in this domain by examining future time reference in Gitksan (Tsimshianic). Gitksan has plain vs. progressive futures with similar semantics to the English futures; moreover, it extends the aspectual possibilities with an *inceptively*-inflected future, which conveys that the event is just about to happen. We examine the theoretical and typological consequences of the Gitksan system.

The issue. Copley (2009) observes a contrast between *will / be going to* in 'offer' contexts, as in (1). She proposes that the two differ in aspect: while both contain a universal modal, only the latter contains progressive aspect. The contrast is derived compositionally: when a progressive operator applies to the plain future, it supplies the modal even with worlds in which the hearer doesn't want something (e.g. oil-change), which clashes with her definition of an offer.

(1) [Advertising billboard on the highway]
We**'ll** / # We'**re going to** change your oil in Madera.

The contrast between plain and progressive futures recurs cross-linguistically, e.g. in Indonesian, Turkish (Copley 2009), and Blackfoot (Reis Silva 2009). Deeper cross-linguistic investigation is however required; in addition, the question arises whether other grammatical aspects combine with future morphemes, and what semantics results if they do.

Gitksan. Our data, gathered through fieldwork, show that Gitksan constructs progressive futures in a compositionally transparent way: the morpheme *yukw* (independently known to be a progressive, Schwan 2019) adds to the future morpheme *dim*. This cross-linguistically supports Copley's approach.

(2) **Dim** wis.
FUT rain
'It will rain.'

(3) **Yukw dim** wis.
PROG FUT rain
'It is going to rain.'

Moreover, Gitksan encodes a further distinction, involving *imminent* (progressive) futures. These contain the inceptive aspect marker *hlaa* (which in non-future contexts conveys that a change-of-state has recently happened).

(4) **Hlaa (yukw) dim** wis. ✓:if rain is imminent
INCEP (PROG) FUT rain #:if talking about tomorrow
'It is just about to rain.'

Analysis. We propose that *hlaa* applied to a proposition *p* asserts that a BECOME(*p*)-event just took place. *Hlaa* applied to a future proposition asserts that it has *just become the case* that the relevant event will occur in all accessible worlds. We hypothesize that this gives rise to a strong pragmatic inference: the reason the event has just become inevitable is due to evidence that it is on the brink of happening. If imminence is a pragmatic inference then inceptive futures should also be possible with *non*-imminent events. This is supported by (5), which does not convey that Tim will be sick immediately, but rather that his sickness has just become certain:

(5) **Hlaa dim** siipxw-s Tim. ✓:if I have infected Tim with measles
INCEP FUT sick-PROP.NOUN Tim
'Tim will get sick now.'

Consequences. The Gitksan data, and the Gitksan-English comparison, support a view whereby languages combine smaller semantic building blocks to create complex temporal/aspectual meanings. This research also highlights the importance of further cross-linguistic studies of aspect-future interactions.

References

Copley, Bridget 2009. *The Semantics of the Future.* New York: Routledge.
Reis Silva, Maria Amélia 2009. Two futures in Blackfoot. Ms., University of British Columbia.
Schwan, Michael 2019. *Yukw*-ing in Gitksan. Ms., University of British Columbia.

Are apples and pears comparable when they grow in random forests? The problem of identifying Finnish statistical correlates of the Polish verbal aspect

Edyta Jurkiewicz-Rohrbacher
(University of Regensburg & University of Helsinki)

Keywords: Polish verbal aspect, Finnish, predicate-argument structure, quantitative methods, parallel corpora.

The objective of my paper is language-specific category Polish verbal aspect (PVA) and the identification of its correlates in Finnish in indicative, affirmative clauses containing one finite verbal form.

PVA is treated as an obligatory verbal classifier based on a binary opposition Perfective (PFV) – Imperfective (IPFV) whose semantics becomes blurred due to broad set of functional categories realised by the members of the opposition, such as: progressive (IPFV), sequence (PFV) and simultaneity (IPFV) of events, general-factual (IPFV), concrete-factual (PFV), habitual (IPFV). In Finnish, no single verbal category covers a similar functional scope, but in each context, a different set of features allows expressing similar meanings as in Polish.

The present paper examines how the correlates of PVA in Finnish can be identified quantitatively using the data from a self-compiled bidirectional, parallel corpus consisting of 900 aligned clauses arising from modern texts. The data set is stratified in three sub samples according to text-type: literary-narrative, informative, to-be-spoken (clauses extracted from film subtitles, theatre plays, dialogues in literary texts).

The relation between PVA and the set of the most frequent, semantic and linguistic features of the Finnish clauses found in the data is examined by means of classifications trees and random forest models (Breiman 2001, Breiman et al. 1984). In other words, I try to predict the value of PVA basing on information related to the temporal-aspectual domain available in the Finnish clause aligned to the Polish one, and associate the language-specific features with the semantic concepts they represent.

The grammatical features in the model are tense and case marking of arguments, while the semantic features are the type of temporal quantification (individual situation localized to a definite or indefinite unit of time versus generic situations and patterns) and semantic roles of arguments.

The statistical analysis shows the appropriateness of two-level model of aspectuality for the used parallel data. The higher, outer level is associated with the temporal domain as its omission causes the higher rate of misclassification, in comparison to the lower, inner level related to the argument structure.

Within argument structure the most relevant predictor is the semantic role of subject which directly correlates with the well-known semantic distinction static-dynamic, but in the studied languages does not have any grammatical markers.

The next important distinction relates to the case marking of direct object (total-partitve case opposition), and oblique (essive-lative oppositions).

Thus, my study provides an empirical evidence for previous claims that Slavic style aspect and Finnish differential object marking at least partially represent the same semantic field (Dahl 1985, Tommola, 1986), but additionally I point two further, aspectually relevant grammatical categories in Finnish which are tense and rarely considered lative-essive case opposition (see Kangasmaa-Minn 1984).

Bibliography

Breiman, Leo (2001), Random forests. *Machine Learning* 45. 5 – 32.

Breiman, Leo, Joseph H. Friedman, Richard A. Olshen & Charles J. Stone (1984), *Classification and regression trees*, Belmont, CA: Wadsworth International Group.

Dahl, Östen (1985), *Tense and aspect systems*, Oxford: Blackwell.

Kangasmaa-Minn, Eeva (1984), Tense, aspect and Aktionsarten in Finno-Ugrian. In Casper De Groot & Hannu Tommola (eds.), *Aspect bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*, 77 – 96. Dordrecht: Foris Publications.

Tommola, Hannu (1986), *Aspektual'nost' v finskom i russkom jazykah*, Helsinki: Neuvostoliittoinstituutti.

The Finnic 'begin' verbs on the path towards future

Miina Norvik

The Uralic languages display different possibilities to express future time reference (FTR). One option is to use the main verb in the present tense and deduce future meaning from the broader context (e.g. in the Finnic languages); a language may contain a future copula or a future particle (e.g. Mansi, an Ob Ugric language); inflectional future can also be found (e.g. Nganasan, a Samoyed language) (Dahl 2000, Riese 2001, Wagner-Nagy 2018). A common characteristic of the Uralic languages is that they tend to contain at least one 'begin' verb (or suffix), which may be used as a future marking device conveying primarily FTR and not ingressivity (ex. 1) (see more in Majtinskaja 1973, Metslang 1996). To compare, crosslingustically, the most typical sources of FTR devices seem to be verbs of motion (see Bybee et al. 1994).

(1) Northern Mansi (OUBD)

xurəm	*xum*	*jot*	***te-ŋkʷe***	***pat-j-en***,	*naŋ,*	*law-i*
three	man	with	eat-INF	begin-PRS-2SG	2SG	say-PRS[3SG]

'You are going to eat with the three men', she says […]'

This paper concentrates on the 'begin' verbs in the Finnic languages but some parallels will be drawn with the more distant related cognate languages, too. The focus is on such 'begin' verbs that can be associated with the expression of FTR at least to some extent. Whereas there is research on the Estonian *hakkama* 'begin' verb, which is shown to have spread in the futurate uses (Metslang 2006), the other Finnic languages have received much less attention although their futurate uses may be described in overview articles (see Metslang 1996) or mentioned in grammar books (e.g. see Ariste 1948 for Votic *nõisa* 'rise; begin'). The objective is to compare the usage of the 'begin' verbs in the Finnic languages in order to analyse to what extent they have grammaticalised as future auxiliaries, what are their main differences/similarities in licensing the futurate interpretation or what could hold it back. Preliminary results seem to indicate that in most cases, ingressivity and futurity are intertwined, and the future interpretation tends to emerge from semantic rather than syntactic factors. Furthermore, in order to compare the futurate usage of 'begin' verbs in various Finnic languages, it turns out to be necessary to view them against the background of other possible FTR devices such as the usage of the present tense and future copulas.

(2) Estonian (Metslang 2006: 6)

”Novembri	*keskpaiga-ks*	*tea-me,*	*kus*	*Markko*	***sõit-ma***
november.GEN	mid_part-TRA	know-1PL	where	Markko	drive-INF
hakka-b,”	*lubas*	*mänedžer*		*Paul*	*Turner.*
begin-3SG	promise-PST.3SG	manager		Paul	Turner

‘In mid-November we will know, where Markko will drive,’ promised manager Paul Turner.

The linguistic data analysed for the purposes of the present study originate from various collections of text, text corpora, and fieldwork data of the Finnic languages. The approach taken in the study is functional-typological.

References

Ariste, Paul (1948), *Vadja keele grammatika* (Nõukogude soome-ugri teadused 9), Tartu: Teaduslik Kirjandus.

Bybee, Joan L., Revere Perkins, William Pagliuca (1994), *The Evolution of Grammar: tense, aspect, and modality in the languages of the world*, Chicago: The University of Chicago Press.

Dahl, Östen (2000), The grammar of future time reference in European languages, in Ö. Dahl (ed), *Tense and Aspect in the langauges of Europe*, Berlin / New York: Mouton de Gruyter, 309–328.

Majtinskaja 1973 = Майтинская, Клара Е. (1973), Будущее время в финно-угорских языках, *Советское финно-угроведение* 2, 81–90.

Metslang, Helle (2006) Predikaat ajastut kogemas, *Keel ja Kirjandus* 9, 714–727.

Metslang, Helle (1996), The development of the futures in the Finno-Ugric languages, in M. Erelt (ed), *Estonian: Typological Studies* I (Publications of the Department of Estonian of the University of Tartu 4), Tartu, 123–144.

OUBD = Ob-Ugric database. Available online at http://www.babel.gwi.uni-muenchen.de/ (15.01.2019).

Riese, Timothy (2001), *Vogul* (Languages of the World/Materials 158), Munich: Lincom.

Wagner-Nagy, Béata (2018), *A Grammar of Nganasan* (Grammars and Language Sketches of the World’s Languages: Indigenous Languages of Russia), Brill.

Less is more? Verbal diminutives, iterative aspect and event plurality

Eva van Lier, Jenny Audring & Sterre Leufkens
(University of Amsterdam, Leiden University & Utrecht University)

Keywords: Verbal diminutives, Aspect, Pluractionality, Morphosemantics, Typology

Verbal diminutives are verbs marked for a lower intensity of the denoted action. Examples are German *hüsteln* ‘cough lightly’ (< *husten* ‘cough’), Huave (isolate) *jujyuij* ‘shake gently’ (< *jajyaij* ‘shake’, Kim 2008: 322), and Lardil (Tangkic) *ŋaala* ‘to jab lightly’ (< *ŋac* ‘to spear’, Round 2015: 452). Such verb forms, while apparently not uncommon in the world’s languages, have received relatively little attention in the linguistic literature, certainly when compared to nominal diminutives.

Yet, verbal diminutives are typologically interesting because they live at the intersection of three larger and more widely investigated phenomena: evaluative morphology, aspect, and verbal number (pluractionality). As part of evaluative morphology, verbal diminutives can express a variety of pragmatic dimensions, including endearment, contempt, playfulness and pretence (Grandi & Körtvelyessy 2015). Our study focusses on the other two dimensions, i.e. the functions of verbal

diminutives in the domains of aspect and pluractionality, based on a variety sample (Bakker 2010) of 102 languages.

In this talk, we will explore the puzzling observation that verbal diminutive marking can serve contradictory semantic functions (cf. Jurafsky 1996; and see Moravcsik 1978, Regier 1994 and François 2004 for similar remarks regarding (nominal) reduplication): While the central meaning of diminution—be it in the nominal, the verbal or any other domain—is reduction (e.g. in size or intensity), diminutive markers can also be found to express notions of augmentation or increase. In the verbal domain, this includes iterative or frequentative aspect and event plurality. For example, in Lushootseed (Salishan) the diminutive form is marked by a specific type of CV-reduplication (*ǰə́səd* 'foot' < *ǰí-ǰəsəd* 'little foot'). When used with verbs, the same marker indicates either decreased duration of the event (*təd^{ʒ}il* 'lie in bed' < *tí-təd^{ʒ}il* 'lie down for a little while) or multiple occurrences of the event (*bə́č* 'fall down' < *bí- bə́č* 'drop in from time to time') (Urbanczyk 2006: 180). Another case in point is Central Alaskan Yup'ik (Eskimo-Aleut), which has a marker *-mciur(ar)* signifying event-internal pluractionality, that creates a diminutive meaning with particular verbs (e.g. *neremciurtuq* 'he is eating a little at a time' < *nere-* 'to eat') (Wood 1997: 129). Yet another example is Hungarian, where the verbal suffix *-gat/-get* creates a diminutive and iterative meaning at the same time (*lép* 'to take a step' > *lépeget* 'to take multiple small steps') (Szilvia Szita p.c.).

In order to shed light on this apparent paradox, we will provide a cross-linguistic survey of the formal and semantic characteristics of verbal diminutives and identify generalizations about the relation between the two. Specifically, we will see whether specific formal types of verbal diminutives are more likely to express augmentative functions. We will also consider the influence of the typological profile of the languages, as well as any parallels with a (synchronically or diachronically) shared diminutive-augmentative function in the nominal domain.

References

Bakker, Dik (2011), Language sampling. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 100-127.

François, Alexandre (2004), La réduplication en mwotlap: les paradoxes du fractionnement. In É. Zeitoun (ed.), *Les langues austronésiennes. Faits de langues* n°24, 177-194.

Grandi, Nicola & Lívia Körtvelyessy (eds.) (2015), *Edinburgh Handbook of Evaluative Morphology*. Edinburgh: Edinburg University Press.

Jurafsky, Daniel (1996), Universal tendencies in the semantics of the diminutive. *Language* 72(3), 533-578.

Kim, Yuni (2008), *Topics in the phonology and morphology of San Francisco del Mar Huav*e. PhD Thesis, University of California Berkeley.

Moravcsik, Edith (1978), Reduplicative constructions. In J. H. Greenberg (ed.), *Universals of Human Language, vol. 3: Word Structure*. Stanford, CA: Stanford University Press, 297-334.

Regier, Terrry (1994), A preliminary study of the semantics of reduplication. TR-94-019. International Computer Science Institute, Berkeley, CA.

Round, Erich (2015), Yukulta and its relatives Kayardild and Lardil. *Edinburgh Handbook of Evaluative Morphology*, ed. by Nicola Grandi and L. Kortvelyessy. Edinburgh: Edinburgh University Press, 448-452.

Urbanczyk, Suzanne (2006), Reduplicative form and the root-affix asymmetry. *Natural Language and Linguistic Theory* 24, 179-240.

Wood, Esther Jane (2007), *The semantic typology of pluractionality*. PhD Thesis, University of California Berkeley.

WORKSHOP 22

Schedule: Fri 9.00-17.25 (Room 8)

Usage-based approaches to language contacts

Anna Verschik & Antje Quick

Keywords: Usage-based, Contact Linguistics, bottom-up

The current workshop proposal is aimed at summarizing what has been achieved so far under the heading of a usage-based approach to contact linguistics (Backus 2015). Languages in contact almost always influence each other and the impacts are manifold such as structural and lexical borrowing, or code-switching. Many different disciplines have investigated these issues and each discipline has come up with their own explanations. For example, in structuralist approaches, different metalanguages are used for the description of contact phenomena in lexicon (code-switching, borrowing) and in morphosyntax (structural borrowing, convergence, grammatical interference, morphosyntactic restructuring), resulting in different models and different theoretical accounts all with the aim to find universal validity. However, empirical studies demonstrate that there is no strict separation between morphosyntax and lexicon on the cognitive level (insertion of a foreign lexical item can affect morphosyntax, contrarily to some influential models such as MLF). This suggests that structural accounts only provide useful descriptive generalizations: they might not tell us much concerning the cognitive mechanisms that are involved in contact phenomena. As a result, we get formal descriptions without explanations. Another reason why the structuralist approaches failed to find universal validity is that they do not take diversity of the bi- or multilingual individuals into account. However, variation is important as most of the time speakers in contact situations show little uniformity and this fact is seldom reflected in general descriptions.

On the other extreme, sociolinguistic studies of bilingualism have been focusing on communities, treating individual language users merely as representatives of these communities but not as something worth investigating in itself. It is true that in some instances the identification of the relevant communities themselves is the focus (e.g. Keim 2007) but this is not always a research goal; quite often it is language use, variation and change that a sociolinguist wants to investigate. On the community level one can attest already conventionalized innovations; however, it is instructive to understand how innovations start and spread. Innovations (including contact-induced ones) first appear in an individual's language use. What kind of innovations will appear, depends on experience and social interaction of a given speaker. An individual would opt for an innovation because it helps to achieve communicative goals in a given situation. Thus, linguistic and communicative competence emerges through the interaction of cognition and use. This is in accordance with the emergentist view on language acquisition and change: grammar is not something "out there" and innate but a process of identification of innovations through individual experience, and accumulation and internalization of these innovations. As everybody has unique linguistic experience, and a language user's experience with language is represented rather directly in competence, individual differences can be expected.

The role of the individual in contact-induced language change has been acknowledged (Weinreich 1953, Matras 2009, Blommaert and Backus 2011) but needs further elaboration. An innovation starts in one or many individuals and may spread gradually among a larger set of individuals (or not), until we recognize a change at the community level. Or the innovation does not gain currency and

eventually becomes obsolete. Investigating all this requires adding cognitive aspects to sociolinguistic research on contact-induced language change and sociolinguistic aspects to cognitive and linguistic research.

Therefore, in this workshop, we strive towards an elaboration of a holistic metalanguage that would unite the various phenomena at the descriptive level while also doing justice to what we know about cognitive processes. Methodologically, a usage-based approach presupposes bottom-up, data-driven explanations of contact phenomena. Usage-based approaches are thoroughly functionalist which means that human linguistic communication has a symbolic functional dimension. It is a cognitive-functional approach that sees speakers' grammars as fundamentally grounded in 'usage events'; emerging from language use and consequently shaped by individual usage, both historically and ontogenetically (Bybee, 2010; Croft, 2001; Langacker, 1987). Speaker's mental representations therefore, differ in their make-up and their competence is changing all the time: language competence is inherently based on language usage. Consequently, research should focus on language use. Constructions are the units in Usage-based approaches and are moving along a continuum of completely lexically fixed items (chunks), via partially schematic to fully schematic utterances. Exactly this feature blurs the lines of syntax versus lexicon, competence versus performance and functional versus lexical items. Constructions constantly differ in their degrees of entrenchment: the more frequently a construction is used, the higher its level of entrenchment and consequently the easier its activation.

Based on these assumptions, the workshop addresses the following topics:

- The cognitive question: what's behind the descriptive structural generalizations? What determines chunk status? What psycholinguistic evidence can back it up? What does it mean when speakers establish equivalence? Does that have to do with entrenchment 'jumping' to an equivalent structure in the other language?
- Community and individual: What is the granularity of community? Why is insertion common in certain forms in certain sociolinguistic settings? Why is there no clear preference for insertion vs. alternation in some settings? And do these concepts reflect cognitive reality?
- Comparative perspective: attractiveness across language contact situations (for instance, conjunctions, discourse particles etc), universal and common principles behind borrowability, contact between different versus contact between similar languages.
- Emergence, entrenchment and conventionalization (establishment of equivalence, "pivot matching", variation in adaptation), role of cognitive routinization; spread through a community.
- Methodological issues: what quantitative analyses are possible with relatively small corpora?

To summarize, in the following workshop we want to look at contact phenomena bringing together contact linguistics, cognitive linguistics and sociolinguistic research. We feel that bringing together these disciplines and applying a bottom-up, non-constraint approach would contribute to a better understanding of contact-induced language change mechanisms and outcomes.

References

Backus, Ad (2015) A usage-based approach to code-switching: a need to reconcile structure and function. In Gerald Stell and Kofi Yakpo (eds.), *Code-switching between structural and sociolinguistic perspectives*, 19–37. Berlin: Mouton de Gruyter.

Blommaert, Jan and Backus, Ad (2011) Repertoires revisited: 'Knowing language' in superdiversity. *Working Papers in Urban Language and Literacy* 67, available at: http://www.kcl.ac.uk/innovation/groups/ldc/publications/workingpapers/download.aspx

Keim, Inken (2007). The construction of a social style of communication and socio-cultural identity: Case study of a German Turkish youth group in Mannheim/Germany. In Peter Auer (ed.), *Social identity and communicative styles: An alternative approach to variability in language*, 155-186. Berlin & New York: Walter de Gruyter.
Matras, Yaron (2009) *Language Contact*. Cambridge, New York: Cambridge University Press.
Weinreich, Uriel (1953) Languages in Contact. The Hague: Mouton.

A usage-based approach to „language“ in code-switching

Philipp Wasserscheidt
(Humboldt-University of Berlin)

Keywords: Bilingualism, Construction Grammar, Usage-based, Language, Mental Lexicon

The use of named languages is still prevalent in current contact linguistics research. In this presentation, I argue that this cannot explain many contact situations and types of language knowledge. I propose a usage-based account of language instead.

The notion of “language” is notoriously difficult not only in bilingualism research, but in linguistics in general. There is no consensus as to what exactly constitutes the system of a language and how to draw boundaries between languages or varieties (Otheguy et al. 2015). Also, “using a language” can mean very different things, from knowing only single words to mastering several registers and sociolects. However, avoidance to define the term language and thus the very object of study is an obstacle to understanding language contact phenomena properly.

In this talk, I want to propose a usage-based approach to “language” which puts the single linguistic units in the focus. These units can be described as constructions that form networks and constitute the centre of linguistic knowledge. Hence “languages” are not supposed to be fixed systems, but flexible networks of constructions.

I propose a three-dimensional model of how these constructions can be clustered in order to be perceived as belonging to one language, variety or register. The dimensions are form, linguistic context and communicative situation. A language is thus a network of constructions which

- use the same phonology, syllable patterns, intonation patterns
- are used in the same linguistic context
- are used in the same situations

These dimensions are individual cognitive vectors that develop through (motor, combinatorial, and pragmatic) routinization but are rooted in the conventionalised linguistic behaviour of the communities of practice the individual speaker is part of.

In the presentation, I will discuss the possible combinations of the three dimension and test their predictiveness with a meta-study of code-switching from different contact scenarios, bilingual communities, speaker biographies and pairs of languages/varieties. The results will show, how different contact situations can inform us about the underlying conceptualisation of the linguistic networks involved.

References

Otheguy, Ricardo; García, Ofelia; Reid, Wallis (2015): Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. Applied Linguistics Review 6(3), 281-307.

A usage-driven analysis of naturalistic code-switching in three bilingual 2-3 year olds

Maria Frick, Dorota Gaskins, Elina Palola & Antje Endesfelder Quick
(University of Oulu; University of London; University of Oulu; & University of Leipzig)

Keywords: slot-and-frame; code-switching; children; cross-linguistic

Several usage-based studies have traced monolingual children's early language back to that used is their previous speech to establish which elements are being recycled (the so-called *frames*) and which combined with them in a novel way (the so-called *slot fillers*) and they have shown that such frame-and-slot patterns dominate spontaneous language use (Lieven, Pine & Baldwin 1997; Lieven, Salomo & Tomasello 2009). An example of a recyclable *frame* which could become frozen through repeated use is *This is;* an example of a slot (here marked as *X*) could be one placed utterance finally (as in *This is X*) and filled initially with nouns and subsequently with more complex phrases. We apply the trace-back method to data from three bilingual two-to-three year olds with English as one of their languages and Polish, German, or Finnish as the other to examine a) what these children's code-switching has in common and b) how it differs between the children in light of the typological distance of the languages used. Their bilingual constructions are derived from individual corpora of naturalistic interactions of each child video recorded during play time and traced back to both monolingual and bilingual constructions produced previously to establish whether they are completely frozen or partially schematic. First and foremost, our results show that each child's code-switching is dominated by partially schematic combinations with an open slot (64%, 63% and 65% respectively) but that frozen bilingual chunks also feature heavily in child speech (32%, 12% and 25%). Further to this, we propose a model of switching which helps us to distinguish between the qualitative aspects of bilingual use in these two types of combinations: we show that not all bilingual chunks continue to be used as such in longer partially-schematic constructions, possibly because separation of languages in the input promotes development of monolingual, not bilingual units in child speech. Whatever bilingual combinations remain frozen in the frames of longer units can be explained by phonological overlap of the children's two languages, including *Daj mi that X* 'give me that X', *Ich want X* 'I want X' and *Me tosi X* 'me very X' where *me* is phonologically close to the respective Finnish pronoun (mina ~ mä(ä)). The highest use of bilingual frames is observed in the acquisition of English-German and explained by the typological proximity of the two languages which is manifested in high numbers of cross-linguistic cognates.

References

Lieven, Elena, Julian Pine, and Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. Journal of Child Language 24, 187–220

Lieven, Elena, Dorothè Salomo, and Michael Tomasello. 2009. Two-year-old children's production of multiword utterances: a usage-based analysis. Cognitive Linguistics 20, 481–507.

Analyzing Estonian-English-Japanese code-copying in Facebook communications: A usage-based perspective

Geidi Kilp
(Tallinn University)

Keywords: code-copying, usage-based approach, contact linguistics, multilingual communication, Estonian-English-Japanese

This paper investigates Estonian-English-Japanese multilingual communication within the framework of usage-based approach to contact-induced language change (Backus 2015). Applying a usage-based perspective to contact data enables the analysis of morphosyntax and lexicon under the same terminological framework, as morphosyntax and lexicon are not separated (Verschik 2019). This can be the case with Japanese insertions that affect Estonian morphosyntax. Code-copying and code-alternation cases are analyzed, utilizing Johanson's code-copying framework (2002), as it is compatible with usage-based approach and treats all contact phenomena as various types of copying.

The combination of these languages is rare, with very few users, Estonians who speak English and Japanese. The data consists of synchronous Facebook messages (private chat, group chat and public posts), with about 10,000 tokens in the corpus. As the corpus is small and it is necessary to identify contact-induced impact, a qualitative analysis is applied for observing changes, their entrenchment and conventionalization. I concentrate on Japanese-Estonian and Japanese-English copying in cases where elements of all three languages are present in one conversation. In some cases it is difficult to establish a base language due to frequent alternations (primarily between Estonian and English). This phenomenon is common in computer-mediated communication, and presents a challenge because a large number of copies could also be considered alternations.

The paper analyzes what kind of items and combinations thereof are likely to be copied, which can be, if they are, semantically specific, have strong expressive connotation or are relevant for organizing discourse (pragmatic particles). Japanese has a complex system of politeness markers, and if Estonian-speakers have mastered them, they tend to copy the markers into Estonian and English. This may occur because their perception has changed and they feel a necessity to mark politeness in a Japanese way (pragmatic gap, Verschik 2010). In addition, mediated copying is common, where English copies in Japanese are copied back into English or Estonian, with a slightly different, often more specific meaning. Some cases are not as clear-cut, because some copies are ambiguous and could belong to both Estonian and English, for example, such as *pm* (Eng. *pretty much*, Est. *põhimõtteliselt*, 'pretty much'), and some global copies border on alternation (primarily due to frequent line changes and the brevity of utterances).

As for results, global copies are the most frequent, and selective and mixed copies are rare, as is the case with other research on the topic (such as Verschik, Kask 2019), indicating that the distribution of the types of copying are similar in trilingual communication. Alternations are vastly more frequent between Estonian-English and Estonian-Japanese, than English-Japanese. A bottom-up approach allows to identify the emergence of new copies, their entrenchment and conventionalization in the language use of the micro-community.

References

Backus, Ad (2015). A usage-based approach to code-switching: a need to reconcile structure and function. In Gerald Stell, & Kofi Yakpo (Eds.) *Code-switching between structural and sociolinguistic perspectives*, 19–37. Berlin: Mouton de Gruyter.

Johanson, Lars (2002). Contact-induced linguistic change in a code-copying framework. In Mari C. Jones, Edith Esch (Eds.) *Language change: The interplay of internal, external and extra-linguistic factors* (Contributions to the Sociology of Language, 86), 285–313. Berlin: Mouton de Gruyter.

Verschik, Anna (2010). Ethnolect debate: Evidence from Jewish Lithuanian. *International Journal of Multilingualism* 7 (4): 285–305.

Verschik, Anna; Kask, Helin (2019). Estonian–English Code Alternation in Fashion Blogs: Structure, Norms and Meaning. In Sanita Lazdiņa, Heiko F. Marten (Eds.) *Multilingualism in the Baltic States: Societal Discourses and Contact Phenomena.* Palgrave Macmillan UK, London.

Verschik, Anna (2019). English-Estonian code-copying in blogs: combining contact linguistic and cognitive approach. In Zenner, Eline; Backus, Ad; Winter-Froemel, Esme (Eds.) *Cognitive Contact Linguistics*, Mouton de Gruyter.

Bilingual language acquisition from a usage-based perspective: A corpus study on the code-mixing of a German-English bilingual child

Antje Endesfelder Quick, Stefan Hartmann Ad Backus, and Elena Lieven
(University of Leipzig, Bamberg University, Tilburg University, University of Manchester)

Keywords: Language acquisition, multilingualism, traceback, corpus linguistics, acquisition of constructions

Lexically specific patterns such as *What's this?* and patterns consisting of a lexically specific "frame" and an open slot such as *look, an* x*!* play an important role in language acquisition scenarios (Cameron-Faulkner, Lieven, and Tomasello 2003).

The ubiquity of conventionalized chunks and partially schematic patterns such as *I want x* supports the idea that children construct their early utterances out of concrete pieces they have heard and stored before (Lieven, Salomo, and Tomasello 2009). Recently, Quick et al. (2018) have shown that partially schematic patterns also play an important role in the code-mixing of a German-English bilingual child, e.g. *I want* x as in *I want die paint* 'I want the paint' suggesting that children's code-mixing is influenced by the child's recent linguistic experience.

Following up on this observation, the present study investigates whether lexically specific as well as partially schematic patterns in the code-mixing of a German-English bilingual child (n=1,024 code-mixed utterances out of 47,812 utterances in total) aged 2;3 to 3;11 can be traced back to patterns found in the input (n=61,077). In line with usage-based accounts of language acquisition (e.g. Tomasello 2003), this would suggest that bilingual children construct their code-mixed utterances on the basis of concrete linguistic material they have heard before.

In our operationalization of utterance-initial chunks as well as partially schematic patterns, we follow e.g. Dąbrowska & Lieven (2005) or Cameron-Faulkner et al. (2003): Repeated strings are considered patterns if they occur at least four times in the corpus. In a first step, only utterance-initial *n*-grams were retrieved automatically to identify recurring lexical patterns both in the child's utterances and in the input material. In line with the operationalization mentioned above, they were considered frame patterns if they occurred at least 4 times. (n-grams embedded in other n-grams were of course subtracted, i.e. if 7 utterances start with *I* and 4 of these sentences start with *I want*, then *I*

want qualifies as a frame pattern while *I* doesn't as it does not reach the threshold of 4 with 7–4 = 3 attestations.) In a second step, these patterns were checked manually.

The results suggest that a large proportion of the code-mixed data contain utterance-initial lexical chunks (82.2%), a large proportion of which (72.7 %) can be traced back to the parental input. A more in-depth study of a subset of the data using manual annotation, which allows for identifying partially schematic rather than just lexically specific chunks, reveals that code-mixing often involves the filling of an open slot in a partially schematic construction. In addition, much of the code-mixing turns out to be primed by the occurrence of the same forms in the immediately preceding discourse.

In sum, these findings lend further support to usage-based theories of language acquisition, and they can help understand code-mixing in early bilingual language acquisition from a construction-based perspective.

References

Cameron-Faulkner, Thea, Elena Lieven & Michael Tomasello (2003), A construction based analysis of child directed speech. *Cognitive Science* 27(6). 843–873.

Dąbrowska, Ewa & Elena Lieven (2005), Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16(3). 437–474.

Lieven, Elena, Dorothé Salomo & Michael Tomasello (2009), Two-year-old Children's Production of Multiword Utterances: A Usage-Based Analysis. *Cognitive Linguistics* 20. 481–507.

Quick, Antje Endesfelder, Elena Lieven, Malinda Carpenter & Michael Tomasello (2018), Identifying partially schematic units in the code-mixing of an English and German speaking child. *Linguistic Approaches to Bilingualism* 8(4). 477–501.

Stoll, Sabine, Kirsten Abbot-Smith & Elena Lieven (2009), Lexically Restricted Utterances in Russian, German, and English Child-Directed Speech. *Cognitive Science* 33(1). 75–103.

Tomasello, Michael (2003), *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Contact between different varieties of Basque: dialectal constructions in L2 Basque speakers' linguistic varieties

Eeva Sippola & Hanna Lantto

The sociolinguistic dynamics in the Basque Country have changed completely during the last forty years. While education was previously only available in Spanish, the new generations of Basques are educated in the minority language. For many of them, Basque is their second language acquired in the classroom environment. The linguistic variety used is these classrooms is the standard *Euskara Batua,* which has been perceived as artificial and too formal for informal interactions. Authenticity and intimacy, in turn, have been connected to the Basque dialects that the native Basque speakers learn in their primary socialization (Urla et al. 2016). This is why some L2 Basque speakers make their linguistic varieties more informal by incorporating dialectal constructions into their speech (Ortega et al. 2015).

This paper examines conventionalized patterns in the use of dialectal constructions in L2 Basque speakers' varieties, and addresses the language contact between the standard and the dialect by both quantitative and qualitative methods. The data is based on interviews with twenty-five L2 Basque speakers from the city of Bilbao, where 18,6% of the population speaks Basque, the great majority of

the Basque speakers have learned Basque as their second language (Basque Government 2016), and the local vernacular dialects have become virtually extinct. All these factors make the city an excellent language laboratory to examine the emergence of new contact varieties of Basque. The questions to be focused on are 1) the type of dialectal constructions that appear in the speech of the L2 speakers of Basque: do they mainly show phonological or morphological dialectal features?, 2) the provenance of the dialectal constructions: are they adopted from the surrounding Biscayan dialects or are they from several dialectal sources?, and 3) the structural salience of the constructions, ie. are the dialectal constructions similar in meaning and form with their standard equivalents or do the L2 speakers favour dissimilarity and structural salience as a means to mark their speaker authenticity? The study identifies the most common patterns in the use of dialectal constructions in the speech of the informants and discusses if they can be considered widespread enough to be a part of a new emerging Basque variety.

Basque Government (2016) Sociolinguistic Survey.
http://www.euskadi.eus/contenidos/informacion/argitalpenak/es_6092/adjuntos/Resumen_VI_Encuesta_Socioling%C3%BC%C3%ADstica_EAE_%202016_1.pdf (Accessed 7 January, 2019)

References

Ortega, Ane, Amorrortu, Estibaliz, Goirigolzarri, Jone, Urla, Jacqueline and Uranga, Belen (2015). Linguistic identity among new speakers of Basque. En O‟Rourke, Benadette, Pujolar, Joan and Ramallo, Fernando (Eds.) (2015) Special Issue: New Speakers of Minority Languages: The Challenging Opportunity. *International Journal of the Sociology of Language 231*.

Urla, Jacqueline, Amorrortu, Estibaliz, Ortega, Ane and Goirigolzarri, Jone (2016). *Authenticity and Linguistic Variety among New Speakers of Basque.* In LD&C Special Publication No. 9: Language Documentation and Conservation in Europe

Studying a lack of usage in a usage-based approach: The social meaning of English insertions in Belgian Dutch family dinner table interactions

Eline Zenner & Dorien Van De Mieroop
(KU Leuven)

Keywords: English; weak contact; dinner table conversations; usage-based; negative evidence

Belgian Dutch elementary school children acquire a substantial amount of English vocabulary before the start of EFL in school contexts (see De Wilde & Eyckmans 2017, and De Wilde, Brysbaerts & Eyckmans 2018). Crossing insights from usage-based approaches to language acquisition (e.g. Tomasello 2003) and language socialization research (Ochs & Schieffelin 2014), this contribution studies the role of caregiver speech in this incidental acquisition of English vocabulary. The following research questions are put forward: (RQ1) how frequently do we find English insertions in Belgian Dutch parents caregiver speech; (RQ2) how can we explain the answer to RQ1?

To address these questions, we scrutinize a corpus of self-recorded dinner table conversations that consist of over 25,000 utterances for eight families, complemented by sociolinguistic interviews with the primary caregivers of the families in the database. After mining the corpus for English insertions (which we identify through the grapheme phoneme-mapping rule described in Onysko 2007), we

present a quantitative exploration of the attested insertions, revealing how no more than 1% of the utterances per family contain English insertions. In order to determine whether this result can be interpreted as individual parents' attempts to socialize their children towards Dutch, and what this reveals about their language regards (Preston 2013), we scrutinize selected fragments by using multimodal discourse analysis.

After exploring the implications of our findings for the position of English in Flanders, we additionally discuss them against the theoretical background of child-directed speech and developmental sociolinguistics (Smith et al. 2013; De Vogelaer & Katerbow 2017), and against the methodological background of dealing with negative evidence in a usage-based approach (linking up to – amongst others – the notion of negative entrenchment; Stefanowitch 2008).

References

De Vogelaer, Gunther & Matthias Katerbow (2017), *Acquiring Sociolinguistic Variation*. Amsterdam: Benjamins.

De Wilde, Vanessa, Marc Brysbaert & June Eyckmans (2018), Incidental learning of English in children: How do word-related variables influence receptive vocabulary learning? Presented at *Psycholinguistics in Flanders*.

De Wilde, Vanessa & June Ecykmans (2017), Game on! Young Learners' Incidental Language Learning of English Prior to Instruction. *Studies in second language learning and teaching* 7(4): 673-694.

Ochs, Elinor, and Bambi B. Schieffelin (2014), The Theory of Language Socialization, in Alessandro Duranti, Elinor Ochs, Bambi B. Schieffelin (eds.), *The handbook of language socialization*, Chichester: Wiley-Blackwell, 1-21.

Onysko, Alexander (2007), *Anglicisms in German. Borrowing, lexical productivity and written codeswitching*, Berlin/New York: Walter de Gruyter.

Preston, Dennis (2013), The influence of regard on language variation and change, *Journal of Pragmatics* 52, 93–104.

Smith, Jennifer, Mercedes Durham, and Hazel Richards (2013), The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation, *Linguistics* 51(2), 285-324.

Stefanowitsch, Anatol (2008), Negative entrenchment: A usage-based approach to negative evidence, *Cognitive Linguistics* 19 (3), 513-531.

Tomasello, Michael (2003), *Constructing a language: A usage-based theory of language acquisition*, Cambridge (MA): Harvard University Press.

Loanword non-adaptation from Southern California to South India: The role of (de)entrenchment

Savithry Namboodiripad
(University of Michigan, Ann Arbor)

Keywords: language contact, usage-based, loanword phonology, language ideology, borrowing

Usage-based approaches assume that linguistic generalizations are emergent, bottom-up abstractions across exemplars (e.g., Bybee 2006, Goldberg 2006). As such, contact effects are not fundamentally different from other linguistic phenomena: the existence of subpatterns is expected, and small- and

large-scale contact effects should arise by similar processes, like entrenchment (Backus 2012). As a demonstration of the potential for similarities across contact situations when taking this approach, I compare non-adaptation of Spanish words by English-speakers in the United States to non-adaptation of English words by Malayalam-speakers in Kerala, India. Increased experience with the historical source languages leads to de-entrenchment and the introduction of novel forms in both cases, but differences in the contact situations lead to differences in social meaning of the non-adapted forms.

English-Malayalam contact has only increased since the colonial period. Kala (1977) describes the integration of English into conversational Malayalam; currently, about one-third of students in Kerala attend English-immersion schools (Kerala Educational Census 2011). As experience with and use of English has increased, more English-like novel forms have been introduced:

(1) a. aːpiːsə (older)
b. oːfiːsə (newer)
'office'

(2) a. peːna (older)
b. penːə (newer)
'pen'

(3) a. kaːpi (older)
b. koːfi (newer)
'coffee'

(4) a. kuːpa (older)
b. kapːə (newer)
'cup'

The older forms were highly entrenched, and the only forms associated with those meanings. The newer pronunciations arose concurrent with increased experience with English at a societal level. This, along with English's hegemonic status, influenced the use of the more English-like forms.

Spanish toponyms in Southern California are highly entrenched, and, thus, their historical origins are often masked. However, "re-Latinization" (Estrada 2013), more positive attitudes towards Spanish, and an influx of migrants from other parts of the US, have increased the salience of the Spanish-ness of these words, resulting in de-adaptation:

(5) a. los fiːlus (older)
b. los feliːs (newer)
'Los Feliz'

(6) a. sæn pʰiːdɻo (older)
b. sæn pʰeːdɻo (newer)
'San Pedro'

The (b) forms are shibboleths associated with more recent residents, and they are also closer to the pronunciation in Spanish.

In each case, a form which was previously entrenched became less entrenched via speakers having more experience with and/or more positive attitudes towards the historical source language, resulting in a form closer in its phonological form to that language. However, different social meanings are associated with the de-adapted forms. For Spanish in US English, de-adapted forms can carry with them marked social meanings, indicating mockery, foreignness, or pretension, and this is mediated by the racial presentation of the speaker; speakers who appear to be LatinX are perceived differently if they use more Spanish-like pronunciations than others are (Hill 1993). For Malayalam, certain de-adaptations carry no social meaning at all, while others signify a lack of knowledge of Malayalam and a non-traditional outlook.

Taking the view that language boundaries should be disassociated from linguists' classifications (Höder 2012), I argue that these un-adapted words are less entrenched and represent novel subpatterns; ideology is a part of language use when it comes to motivating contact-induced change.

References

Backus, A. (2014). Towards a Usage-Based Account of Language Change: Implications of Contact Linguistics for Linguistic Theory. In *Questioning language contact* (pp. 91-118). Brill.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711-733.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press on Demand.

Hill, J. H. (1993). Hasta la vista, baby: Anglo Spanish in the American Southwest. *Critique of Anthropology*, *13*(2), 145-176.

Höder, S. (2012). Multilingual constructions. *A diasystematic approach to common structures*, 241-258.

Kala, C. V. (1977). Hybrid Conversational Malayalam the Role of English in the Kerala Social Situation. *Anthropological Linguistics*, *19*(6), 265-273.

Kerala Department of Education (2011). Kerala Department of Education Statistical Report. Technical report.

The dynamics of fusion: The case of vernacular Kildin Saami

Nikolay Hakimov and Michael Riessler

Any description of language mixture needs to distinguish between language mixing and language fusion. This is a particularly non-trivial task in situations of long-standing bilingualism. The major purpose of this paper is thus to propose and test a methodology for discerning language fusion from conventionalized mixing. Additionally, the paper examines the hypothesis that the fusion of unbound elements evolves from alternational mixing. These goals are addressed through an analysis of a vernacular variety of Kildin Saami, an East-Saamic (Uralic) language spoken on the Kola Peninsula in Northwest Russia, as a partially fused lect due to contact with Russian. In our corpus of Kildin Saami conversational data, Russian-origin elements regularly appear in the otherwise Kildin Saami discourse. In order to identify those Russian-origin elements that have fused with Kildin Saami grammar, we are developing and will showcase the following diagnostic criteria: (a) regularization of the donor language items' usage patterns in the mixed variety; (b) functional reduction, or functional extension, of the donor language element, and/or of its inherited native equivalent; (c) the introduction of new constructions involving the donor language grammatical elements by way of loan translation. We argue that these criteria may reliably be employed as indicators of fusion in future studies of contact varieties with little, or undocumented, linguistic histories. Finally, we argue that the numerous parallels existing between the distribution of Russian-origin items in vernacular Kildin Saami and alternational mixing provide support for the fusion hypothesis.

WORKSHOP 23

Schedule: We 11.00-17.55 (Room 13)

V2 effects in the world's languages

Maia Duguine & Michael Zimmermann
(CNRS-IKER UMR 5478 & University of Konstanz)

Keywords: syntax, left periphery, V2 effects, adjacency effects, information structure

Germanic languages, with the major exception of Modern English, are characterized by the regular occurrence of the finite verb in second position in main clauses, particularly declaratives, a property widely known as 'verb second' (V2). In the classical formal approach this configuration is assumed to result from the attraction of the finite verb to the highest clausal head (C) along with the appearance of an element in its outer edge (Thiersch 1978, den Besten 1983).

Old Romance languages, even though they display V>2 orders quite robustly, have also been characterized as obeying a similar syntactic constraint: the occurrence of the finite verb in *at least* second position (cf. Jouitteau 2009). While some researchers assume that Old Romance languages have a non-V2 system with V2 configurations following from stylistic, pragmatic or information-structural factors (e.g. Kaiser 1999, 2002, Sitaridou 2011, Elsig 2012, Varga 2017 for Old Romance and Haeberli 2005 for Old English), others propose to conceive of Germanic V2 as part of a wider set of V2 effects obtaining in various other languages (e.g. Old English (Pintzuk 1993), Old Romance (e.g. Benincà 2013, Salvi 2012, Wolfe 2015)). In recent generative work, the latter approach has been framed in the rich model of the sentential left periphery in the wake of Rizzi (1997).

Modern English and Modern Romance languages likewise display V2 effects in specific contexts, in particular main *wh*-interrogatives (cf. (1), (2) and (3)):

(1) Which battery type would you recommend? *English*

(2) Quel livre a-t-elle acheté ? *French*
which book has-she bought
'Which book did she buy?'

(3) ¿Qué libro ha comprado María? *Spanish*
which book has bought Mary
'Which book did Mary buy?'

These languages have been labelled 'residual' V2 (Rizzi 1996), the implications being that the V2 effects are the residue of an earlier general V2 grammar, and that, in these constructions, the finite verb moves to C to enter into a Spec-Head relation with the fronted *wh*-phrase. However, these languages also deviate to various degrees from V2 order in the contexts at issue, a state of affairs that challenges the classical V2 approach in terms of verb-movement to the left periphery (e.g. Suñer 1994, Barbosa 2001, Goodall 2004). A case in point is the order of the subject and the finite verb in Spoken French and Brazilian Portuguese, in which the former regularly precedes the latter (e.g. Ambar 2008, Kato 2012):

(4) Quel livre Marie a acheté ? *Spoken French*
which book Mary has bought

(5) Que livro a Maria comprou? *Brazilian Portuguese*
which book DET Mary bought
'Which book did Mary buy?'

Among the many other languages that exhibit similar V2 effects is Basque, an SOV language that is typologically different from, yet, for centuries, in close contact with Romance. In Basque, the *wh*-phrase as well as the focal-phrase occur immediately left-adjacent to the verbal cluster (e.g. Ortiz de Urbina 1989, 1995, 1999, Irurtzun 2007):

(6) Nork eman dio Mireni liburua? *Basque*
who.ERG give AUX Miren.DAT book.the.ABS
'Who gave the book to Miren?'

The adjacency effects of Basque have been associated with a general ban on verb-first (or tense-first) sentences (*V1) (Ortiz de Urbina 1989, 1994, Uriagereka 1999, Elordieta & Haddican 2017):

(7) a. *Dator Miren. b.Miren dator.
come.3SG Miren Miren come.3SG
'Miren comes.' 'Miren comes.'

These adjacency effects have occasionally led to the classification of Basque, along with numerous other SOV languages, as a *wh*-in-situ language (e.g. Richards 2010, Dryer 2012). Alternative accounts argue for the analysis put forth regarding 'residual V2-languages', involving the leftward movement of the *wh*-/focal phrase and the verb (Ortiz de Urbina 1989, 1995, 1999)), and are hereby able to concomitantly tackle the fact that, in the context of long-distance movement, the *wh*-/focal phrase appears left-adjacent to the matrix verb (Irurtzun 2007, 2008):

(8) Nork_i uste duzu [esan duela Jonek t_i erosi duela liburua]?
who.ERG think AUX say AUX.thatJohn.ERG buy AUX.that book.ABS
Lit. 'Who do you think that Jon said that bought the book?'

The notion of V2 thus possibly constitutes a lens through which to explore the question of *wh*-in-situ in SOV languages.

The aim of this workshop is to bring together scholars working within a variety of theoretical frameworks and adopting either a diachronic or modern synchronic perspective to shed light on V2 effects in languages that are not commonly considered to have a strict V2 grammar. In particular, the typology of V2 effects is to be explored cross-linguistically to determine whether there is a common syntactic basis as well as, more generally, to gain deeper insights into the structure of the left periphery of the clause. Questions that the workshop seeks to address from theoretical and experimental angles include the following:

- What languages, beyond the Germanic ones, are relevant to the V2 typology?
- How do 'at least V2' languages as well as '*V1 languages' fit into this typology?
- What is to be made of 'residual V2 languages', in which V2 effects are confined to particular syntactic contexts?
- Do V2 effects constitute a uniform phenomenon cross-linguistically or, at least, in particular language groups/families (e.g. Modern Romance)?
- How can the tendency, noted in the typological literature, for *wh*-phrases and foci to surface in a position adjacent to the verb in SOV languages be theoretically conceived of?

- Does language contact play a role in V2 effects being an areal phenomenon in e.g. Europe?
- What do changes in word order in the history of particular languages teach us on the nature of V2 as well as its typology?

References

Ambar, Manuela (2008), On some special adverbs, word order, and CP: Variation vs. micro-variation, *Canadian Journal of Linguistics/Revue canadienne de linguistique* 53, 143-179.

Barbosa, Pilar (2001), On inversion in wh-questions in Romance, in A. Hulk, and J.-Y. Pollock (eds), *Subject Inversion in Romance and the Theory of Universal Grammar*, Oxford: Oxford University Press, 20-60.

Benincà, Paola (2013), Caratteristiche del V2 romanzo. Lingue romanze antiche, ladino dolomitico e portoghese, in E. Bidese, and F. Cognola (eds), *Introduzione alla linguistica del mòcheno*, Torino: Rosenberg & Sellier, 65-83.

Besten, Hans den (1983), On the interaction of root transformations and lexical deletive rules, in W. Abraham (ed), *On the Formal Syntax of the Westgermania. Papers from the "3rd Groningen Grammar Talks". Groningen, January 1981*, Amsterdam / Philadelphia: John Benjamins, 47-131.

Dryer, Matthew S. (2012), On the position of interrogative phrases and the order of complementizer and clause, *UCLA Working Papers in Linguistics* 17, 72-79.

Elordieta, Arantzazu & William F. Haddican (2017), Truncation feeds intervention: Two clause type effects in Basque, *Natural Language and Linguistic Theory* 36(2), 403-443.

Elsig, Martin (2012), Subject-verb inversion in 13th century German and French. A comparative view, in:K. Braunmüller, and C. Gabriel (eds), *Multilingual Individuals and Multilingual Societies*, Amsterdam / Philadelphia: John Benjamins, 223-239.

Goodall, Grant (2004), On the syntax and processing in *wh*-questions in Spanish, in V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser (eds), *WCCFL 23: Proceeding from the 23rd West Coast Conference on Formal Linguistics*, Somerville: Cascadilla Press, 237-250.

Haeberli, Eric (2005), Clause type asymmetries in Old English and the syntax of verb movement, in M. Batllori, M.-L. Hernanz, C. Picallo, and F. Roca (eds), *Grammaticalization and Parametric Variation*, Oxford: Oxford University Press, 267-284.

Irurtzun, Aritz (2007), *The Grammar of focus at the interfaces*, Vitoria-Gasteiz: University of the Basque Country dissertation.

Irurtzun, Aritz (2008), A Derivational Approach to the Focus-Structure. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 53, 327-354.

Jouitteau, Mélanie (2009), A typology of V2 with regard to V1 and second position phenomena: An introduction to the V1/V2 volume, *Lingua* 120, 197-209.

Kaiser, Georg A. (1999), A ordem das palavras e a posição do verbo finito no português antigo, in P. Ferenc (ed), *Actas do Congresso Internacional por motivo dos vinte anos do português no ensino superior*, Budapest: Departamento de língua e literatura portuguesas da Faculdade de Letras da Universidade Eötvös Loránd de Budapeste, 248-261.

Kaiser, Georg A. (2002), *Verbstellung und Verbstellungswandel in den romanischen Sprachen*, Tübingen: Niemeyer.

Kato, Mary Aizawa (2012), Brazilian Portuguese and Caribbean Spanish. Similar changes in Romania Nova, in C. Galves, S. Cyrino, R. Lopes, F. Sandalo, and J. Avelar (eds), *Parameter Theory and Linguistic Change*, Oxford: Oxford University Press, 117-132.

Ortiz de Urbina, Jon (1989), *Parameters in the Grammar of Basque. A GB Approach to Basque Syntax*, Dordrecht: Foris.

Ortiz de Urbina, Jon (1994), Verb-initial patterns in Basque and Breton, *Lingua* 94, 125-153.

Ortiz de Urbina, Jon (1995), Residual verb second and verb first in Basque, in K. E. Kiss (ed), *Discourse Configurational Languages*, New York: Oxford University Press, 99-121.

Ortiz de Urbina, Jon (1999), Focus in Basque, in G. Rebuschi, and L. Tuller (eds), *The Grammar of Focus*. Amsterdam / Philadelphia: John Benjamins, 311-334.

Richards, Norvin (2010), *Uttering Trees*, Cambridge: MIT Press.

Salvi, Giampaolo (2012), On the nature of the V2 system of Medieval Romance, in: L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, and C. Poletto (eds), *Functional Heads. The Cartography of Syntactic Structures, Volume 7*, Oxford: Oxford University Press, 103-111.

Sitaridou, Ioanna (2011), Word order and information structure in Old Spanish, *Journal of Catalan Linguistics* 10, 159-184.

Suñer, Margarita (1994), V-movement and the licensing of argumental wh-phrases in Spanish, *Natural Language and Linguistic Theory* 12, 335-372.

Thiersch, Craig L. (1978), *Topics in German Syntax*, Ph.D. thesis. Cambridge: Massachusetts Institute of Technology.

Uriagereka, Juan (1999), Minimal restrictions on Basque movements, *Natural Language and Linguistic Theory* 17, 403-444.

Varga, Eva (2017), *Verbstellungsmuster im Altfranzösischen. Ein Beitrag zur historischen Syntaxforschung aus diskurstraditioneller Perspektive*, Berlin: De Gruyter.

Wolfe, Sam (2015), *Microvariation in Medieval Romance Syntax: A Comparative Approach*, Ph.D. thesis. Cambridge: University of Cambridge.

V2 in the history of Portuguese – a corpus-based study

Charlotte Galves

<pdf>

How the components of V2 came apart in late Old French

Richard Ingham

Structural treatments of V2 phenomena standardly posit two associated syntactic processes targeting the left periphery, finite V movement to a functional head and XP movement to its specifier. The moved XP blocks raising of another XP to the left periphery, thus creating a 'bottleneck' accounting for the V2 constraint (Haegeman 1996, Holmberg 2015). Cartographic approaches (Rizzi 1997, Benincà & Poletto 2004) offer the moved XP potential left-peripheral target positions with discoursal properties, and propose first-merge sites for certain adverb types, allowing V3 surface order to comport with structural V2. This paper considers how these three strands of the V2 scenario fare when V2 declined in medieval French. It is shown that they are not uniformly lost together. First merge of adjuncts in the left periphery, rather than movement there, became common in the later 13th c. (Vance 1997, Ingham 2008), at a time when initial tropicalised objects overwhelmingly retained V2. This development was followed by V->C loss after an inital Object, without a co-referential clitic, e.g.:

(1) Et tout ceu que dessus est devisé nos avom promis e otreié… *La Rochelle,* 1256
'And everything mentioned above we have promised and granted…'

Objects in the left periphery could also appear in the Clitic Left-dislocation structure (CLD), e.g.:

(2) Ces lxx verghes de tiere, Pieres devant dis les a en couvent a-aquiter a Gontier le Sauvage.
'These 70 yards of land, P. aforesaid has agreed to cede to G. le S.' *Cherq (Hainault),* 1269

Troberg (2004) argued that OSV as in (1) was a variant of CLD with a null object pronominal, and without Object fronting. However, OSV was found to occur with a focused object, excluding a CLD analysis with a null resumptive:

(3) Et riens il n-i a retenu. *Barbençon, Hainault* 1248
'And he kept nothing there'

Kroch (1989) and Vance (1997) saw V2 in later Old and Middle French as being in competition with CLD.

We show that 13th c. localised Old French charters dated between 1236 and 1271 display interesting dialectal variability regarding these constructions. Three regions were compared, offering over 200 main clause contexts. V2 occurred rarely in South-Western charters, whereas those from the Picard domain maintained V2 at around 70%. CLD is not found in charters from the South-West. The dissociation here between V2 loss/retention and use of CLD indicates, contrary to Kroch (1989) and Vance (1997) that V2 was lost independently of the establishment of LD as a topic construction (cf. also de Andrade 2018). Charters from the Champagne region show intermediate values, with less V2 than in the North, and use of both OSV and CLD. The sequence of V2 decline was thus that XP raising of adjuncts weakened first, then V-> C, then XP raising of Objects, in preference for CLD. These results are discussed in relation to a diachronic process favouring Merge rather than Move (cf. Roberts & Roussou 2003).

References

Benincà, Paola & Poletto, Cecilia. 2004. Topic, focus and V2: defining the CP sublayers. In Rizzi, L. (ed.) The Structure of CP and IP. *The Cartography of Syntactic Structures, vol. 2*. OUP, pp. 52–75.

De Andrade, Aroldo. 2018. Aboutness Topics in Old and Middle French: A corpus-based study on the fate of V2. *Canadian Journal of Linguistics,* 63 (2), 194-220.

Glessgen, Martin-Dietrich. 2005. Diskurstraditionen zwischen pragmatischen Regeln und sprachlichen Varietäten. In Angela Schrott & Harald Völker (eds.), *Historische Pragmatik und historische Varietätenlinguistik in den romanischen Sprachen,* Universitätsverlag Göttingen, 207-228.

Haegeman, Liliane. 1996. Verb second, the split CP and null subjects in early Dutch finite clauses. *GenGenP* 4(2), 135–175.

Holmberg, Anders. 2015. Verb-second. In Tibor Kiss & Artemis Alexiadou (eds.) *Syntax – theory and analysis: an in- ternational handbook*, 342–382. Berlin: de Gruyter.

Ingham, Richard. 2008. L'ordre syntaxique V3 aux débuts du moyen français. In B. Fagard, S. Prévost, B. Combettes & O. Bertrand (eds.) *Évolutions en français. Étude de linguistique diachronique*. Bern: Peter Lang, 153-170.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman, (ed.) *Elements of Grammar*. Dordrecht: Kluwer.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change. A minimalist approach to grammaticalization.* CUP.

Troberg, Michelle. 2004. Topic-comment resumptive pronouns in Modern French and Old and Middle French *Toronto Working Papers in Linguistics* 23(1), 133-156

Vance, Barbara. 1997. *Syntactic Change in Medieval French: Verb-Second and Null Subjects.* Kluwer, Dordrecht.

An Information Structure scenario for the loss of V2 in Medieval French

Pierre Larrivée
(Université de Caen Normandie)

Keywords: Information Structure, Verb-second word order, Preverbal complements, Productive syntax, Medieval French

The purpose of this talk is to revisit the trajectory of V2 in Medieval French in relation to its informational contribution. Medieval French V2 is generally understood as a configuration where as other V2 constructions, both the initial XP and the verb are in the left periphery of the sentence (Wolfe 2016 *i.a.*). The question that this understanding raises is the cause of the posited movement. It has been proposed that verb movement is motivated by the licensing requirements of the subject at a period where subject expression is optional (Roberts 1993 *i.a.*), and that the verb movement drives the XP movement. However, such proposals have to explain why the verb movement would persist beyond the period where the expressed subject has become the default option (around 1200, Zimmermann 2014, Balon and Larrivée 2016), and on what basis it forces XP movement. An alternative scenario is envisaged here by which it is the informational value of the XP that accounts for its movement and drives the movement of the verb (see also Samo 2018). This predicts that for the period when V2 is syntactically productive, the initial XP has a categorical informational value, the XPs susceptible to have an informational value (such as complements and particles) are predominant, and that on the whole there is no verb movement without a XP, effectively excluding V1. By contrast, when V2 starts being manifested by an increasing proportion of formulaic sequences, the categorical informational value should gradually be lost, and there should be a predominance of XPs that are not susceptible to an informational value (such as coordinators). These expectations are found to be supported by the chronology, rates of use and informational value of the three subtypes of XP (particles, complements and coordinators) found in a new corpus of legal prose material from one continental region for the period 1150-1450. The suggestion therefore is that perceptible informational motivations is necessary for V2 to be maintained. This accounts for the contrast between Breton and Welsh (Jouitteau 2016, Willis 1998), where particles triggering V2 are lost in the latter, but not in the former; as well as that between French and English on the one hand, and German on the other, where the former loses both V2-related particles and V2 (Wolfe 2018, van Kemenade 2015) whereas the latter maintains both.

References

Balon, Laurent & Pierre Larrivée. 2016. « L'Ancien français n'est déjà plus une langue à sujet nul. Nouveau témoignage des textes légaux ». *Journal of French Language Studies* 26,2, 221-237.

Jouitteau, M. (2016), 'Rethinking V2 typology from expletives, EPP inquiries on stylistic fronting,

excorporation and verb-doubling', Theresa Biberauer, Sam Wolfe & Rebecca Woods (éds.). *Rethinking V2*.
Samo, Giuseppe. 2018. Towards a criterial V2. PhD Thesis, Université de Genève.
Roberts, Ian. 1993. *Verbs and diachronic syntax*. Dordrecht: Kluwer.
van Kemenade, Ans. 2015. Modal particles in the history of English from a West-Germanic perspective. MS, Radboud University.
Willis, David. 1998. Syntactic Change in Welsh: A Study of the Loss of Verb-Second. Oxford: Oxford University Press.
Wolfe, Sam. 2018. Probing the Syntax of a Problematic Particle: Old French 'si' Revisited *Transactions of the Philological Society.*
Wolfe, Sam. 2016. *On the left periphery of V2 languages: evidence from romance fin and force V2 systems*. MS, University of Oxford.
Zimmermann, Michael. 2014. *Expletive and referential subject pronouns in French.* Berlin: De Gruyter

Optional V2 in Estonian

Anders Holmberg, Heete Sahkai & Anne Tamm

<pdf>

Why break the 'residual V2' in Basque: diachronic and synchronic evidence

Aritz Irurtzun

<pdf>

Emphatic restrictions on V2 constructions across languages and clause types

Andreas Trotzke
(Universität Konstanz & Universitat Autònoma de Barcelona)

Keywords: emphasis; left periphery; questions; syntax; V2 languages

From a minimalist perspective, V2 constructions in many V2 languages can be analyzed as CPs with C heads that only have an 'edge feature' (Chomsky 2008), meaning that there are no pragmatic restrictions on filling SpecCP (e.g., Fanselow 2016). According to this view, marked interpretations in declarative configurations arise from interactions at the prosody-pragmatics interface only and are not due to word order at all; filling SpecCP is thus analyzed as 'Formal Fronting' because the syntactic

movement does not have any pragmatic impact and merely serves to satisfy the V2 constraint in declaratives.

Crucially, object frontings (1a) can be analyzed as purely formal movements as well once we postulate that (i) there is a topic projection in the middle field in scrambling languages (1b) and (ii) movement to SpecCP obeys a strict locality constraint: only the highest phrase in TP moves to SpecCP. (1a) thus indicates that object fronting is in no way different to frontings of subjects or higher adverbs, which also have no pragmatic impact (see Frey 2004):

(1) a.[Den Max]$_i$ sollte t_i unsere Gruppe t_i unterstützen.
the.ACC Max should our group support
b.... dass ('that') den Max unsere Gruppe unterstützen sollte.

In this talk, I focus on non-assertive V2 cases where material appearing in SpecCP is obligatory associated with both focal stress and a special pragmatic interpretation, which I will describe as 'emphasis for intensity' (Trotzke 2017). By 'emphasis for intensity', I refer to a general linguistic strategy whereby the meaning of linguistic objects is made salient by means of boosting its value along a scalar dimension (e.g., likelihood, noteworthiness, etc.). By 'obligatory associated', I mean 'encoded in grammar' (i.e., word order obligatory yields the relevant pragmatic interpretation and prosodic stress).

I will outline a unified account of marked non-assertive V2 cases such as polar questions with declarative V2 word order and fronted foci ([2]; see Bianchi & Cruschina 2016 for Italian counterparts) and frontings of non-contrastable elements such as discourse particles in V2 questions in languages like Basque and German ([3]; Trotzke & Monforte 2019; Bayer & Trotzke 2015):

(2) Ans Meer fährst du morgen?
to-theseaside go you tomorrow

(3)a.Wie nur habe ich den Schlüssel verlieren können?
how PART have I the key lose could
b.Non ote utzi dut egunkaria?
where PART leave AUX newspaper

(4) [$_{\text{ForceP}}$ Force^0 [$_{\text{EmpP}}$ Emp^0 [$_{\text{FinP}}$ …[$_{\text{VP}}$ …]]]]

In contrast to cartographic approaches that postulate dedicated functional heads for different kind of 'special interrogatives' within or even beyond the Force domain of the clause (e.g., Obenauer 2006; Pan 2015), my structural claim will be that the respective operator must be located below Force (4) and that 'emphasis for intensity' is thus located in the regular focus position.

References

Bayer, Josef & Andreas Trotzke. 2015. The derivation and interpretation of left peripheral discourse particles. In Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds.), *Discourse-oriented Syntax*, 13-40. Amsterdam: John Benjamins.

Bianchi, Valentina & Silvio Cruschina. 2016. The derivation and interpretation of polar questions with a fronted focus. *Lingua* 170. 47-68.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, 133-166. Cambridge, MA: MIT Press.

Fanselow, Gisbert. 2016. Syntactic and prosodic reflexes of information structure in Germanic. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford Handbook of Information Structure*, 621-641. Oxford: Oxford University Press.

Frey, Werner. 2004. The grammar-pragmatics interface and the German prefield. *Sprache und Pragmatik* 52. 1–39.

Obenauer, Hans-Georg. 2006. Special interrogatives: Left periphery, *wh*-doubling, and (apparently) optional elements. In Jenny Doetjes & Paz González (eds.), *Romance Languages and Linguistic Theory 2004*, 247–273. Amsterdam: John Benjamins.

Pan, Victor Junnan. 2015. Mandarin peripheral construals at the syntax-discourse interface. *The Linguistic Review* 32. 819–868.

Trotzke, Andreas. 2017. *The Grammar of Emphasis: From Information Structure to the Expressive Dimension*. Berlin: Mouton de Gruyter.

Trotzke, Andreas & Sergio Monforte. 2019. Basque question particles: Implications for a syntax of discourse particles. *Linguistic Variation*. Online-First, doi: 10.1075/lv.18002.tro

ø

Supporting tensed verbs: *T1, V2 and embedded negation

Arantzazu Elordieta
(University of the Basque Country UPV/EHU)

Keywords: clause type verb second stress intervention truncation

Basque displays V2 effects in interrogative and focal configurations (1). However, it differs from Germanic V2 languages in that T in root clauses seems not to appear in strictly 2nd position but rather in ≥ 2. In fact, in neutral declarative clauses T usually appears in sentence final position (2). A further difference is that the ban on sentence-initial position (*T1) also applies to yes/no questions (a.o. Ortiz de Urbina 1994, 1995).

Another similarity with V2 languages is that *T1 also arises in certain embedded contexts. Nevertheless, it differs in that *T1 is not related to the absence/presence of pragmatic illocutionary force in the embedded clause (a.o. Heycock 2006; Julien 2009, Wiklund et al. 2009, Holmberg 2015), but rather, it seems to correlate with the specific complementizer morpheme selected (3).

We present an analysis of these effects, by linking it to another clause type-sensitive phenomenon, namely word order variation in negated embedded contexts (Neg-Aux-V/V-Neg-Aux), which also correlate with the C selected and with the presence of operators (4), (Elordieta and Haddican 2018). Firstly, assuming a morphophonological clitic status of T in Basque, we propose that *T1 is a consequence of a PF-condition on a high left-peripheral head, presumably Force, which requires its Spec to have phonetic content so that it provides prosodic support to T, which is stressless. Secondly, assuming the functional sequence TopP >ForceP >FocusP >P > FinP >TP, we propose an account of the *T1 asymmetries occurring in embedded (negated) contexts which combines both truncation (Haegeman 2006) and intervention (Haegeman & Ürögdi 2010, Haegeman 2012, 2014, et seq.). Specifically, whether the clause typing feature is merged as a separate Force head, or is merged instead on Fin (Rizzi 1997). In the latter case, Force is not projected and hence *T1 does not arise, and operators usually merged in ForceP will be merged lower in FinP, where they will intervene in Neg movement to P (Laka 1990), a polarity-related head with an EPP-feature which also attracts predicate fronting in affirmative clauses.

(1) a. Gaur dator Jon
today come-3sg. Jon
'Jon is coming TODAY'
b. Noiz dator Jon?
when come-3sg. Jon
'When is John coming?'

(2) Jon-ek Miren-i liburua eman dio (neutral contexts)
Jon-ERG Miren-DAT book.ABS give.PERF AUX
'Jon has given the book to Miren'

(3) a. Jon-ek *[datorr-ela] / [ba-datorr-ela] esa-n du
Jon-erg come.3.abs-C / ba-come.3.abs-C say- PERF AUX
'Jon has said that he is coming'
b. Jon-ek ahaz-tu du *[datorr-ela]/ [ba-datorr-ela]
Jon-erg forget- PERF AUX come.3sg.abs-C/ ba-come.3sg.abs-C
'Jon has forgotten that he is coming'
c. Ez daki-t [datorr-en /% ba-datorr-en] ala ez
neg know-1sg come.3sg.abs-C /ba-come.3.abs-C or not
'I don't know if she's coming or not' (embedded yes/no question)

(4) a. Jon-ek uste du [ez de-la etorri-ko/*etorri-ko ez de-la]
Jon-erg think aux neg aux-C come-fut/come-fut neg aux-C
'Jon thinks that (s)he will not come'
b. Ez dakit [Jon lehenago etorri-ko ez d-en/ez d-en etorriko
neg know.1sg Jon earlier come-fut neg aux-C/neg aux-C come-fut
'I don't know whether Jon will not arrive earlier'

References

Elordieta, Arantzazu and Bill Haddican (2018). Truncation feeds intervention: two clause type effects in Basque. *Natural Language and Linguistic Theory* 36: 2, 403-443.

Haegeman, Liliane (2006), Conditionals, factives and the left periphery. *Lingua* 116 (10): 1651–1669.

Haegeman, Liliane (2012), The syntax of MCP: Deriving the truncation account. In *Main clause phenomena: New horizons*, eds. L. Aelbrecht, L. Haegeman, and R. Nye, 113–134. Amsterdam: John Benjamins.

Haegeman, Liliane (2014), Locality and the distribution of main clause phenomena. In *Locality*, eds. Obladé Aboh, M-T. Guasti, and . Roberts, 186–222. Oxford: Oxford University Press.

Haegeman, Liliane, and Barbara Ürögdi (2010), Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36 (2–3): 111–152.

Heycock, Caroline (2006), Embedded root phenomena. In *The Blackwell companion to syntax*, eds. M. Everaert and H. Van Riemsdijk, 174–209. Oxford: Blackwell.

Holmberg, Anders (2015), Verb second. In *Syntax—theory and analysis: An international handbook*, eds.T. Kiss and A. Alexiadou, 342–383. Berlin: de Gruyter.

Julien, Marit (2009), Embedded V2 in Norwegian and Swedish. *Working Papers in Scandinavian Syntax* 80: 103–161.

Laka, Itziar (1990), Negation in syntax, on the nature of functional categories and projections. PhD diss,Massachusetts Institute of Technology.

Ortiz de Urbina, Jon (1994), Verb-initial patterns in Basque and Breton. *Lingua* 94 (2): 125–153.

Ortiz de Urbina, Jon (1995), Residual verb second and verb first in Basque. In *Discourse configurational languages*, ed. K. É Kiss, 99–121. New York: Oxford University Press.

Rizzi, Luigi (1997), The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson, and Þorbjörg Hróarsdóttir (2009), On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119 (12): 1914–1938.

Embedded V2 is syntactically integrated

Mélanie Jouitteau

In both Germanic and Celtic, V2 is a root phenomenon, mandatory in matrix sentences and exceptional in embedded domains. Embedded V2 orders are thus tentatively analyzed as paratactic strategies due to fresh grammaticalisations of complementizers from declarative verbs or coordination markers (Jouitteau 2005:199) or influence of complementizers from the SVO language in contact (Favereau 1997:§598). However, those diachronic correlations have no explanative power and offer no formal analysis of embedded V2. A recurrent intuition is that embedded V2 orders are lesser integrated into the syntactic structure. This is confirmed for peripheral adverbial clauses (Haegeman 2012), but their particular semantics could allow them to be attached higher in the structure, leaving other embedded V2 unexplained. Wurmbrand (2014) proposes that all Germanic embedded V2 clauses are derived by late Merge of two separated roots with signs of weak syntactic integration. I show that this hypothesis is not tenable for Breton, where embedded V2 clauses are unambiguously integrated into the syntactic structure.

Celtic languages are VSO, but in the middle ages the Brittonic branch in contact with Middle English or Middle French developed an extra V2 step. Welsh later switched back to VSO, but Modern Breton still shows consistent V2, with a wide array of last resort V2 strategies from expletive insertion to Stylistic Fronting in matrix sentences. Most embedded domains are C-VSO, which is not surprising because complementizer heads count as V2 satisfiers in matrix sentences (*linear V2*, Borsley & Kathol 2000). However, there are some embedded V2 exceptions (Rivero 1999:81-82). Much like in Germanic, embedded V2 appears in different adjunct clauses denoting cause, complements of verbs of saying and thinking, including forms of the complementizer ‘if’ (1), protasis of conditionals or relatives of temporal nouns.

I present a battery of tests including extraction out of an embedded V2 clause (2), which shows unambiguously that **syntactic embedding is realized syntactically**.

(1) a. *N' ouzon ket hag-eñ* ***en deus*** *lennet al levr.*
b. *N' ouzon ket ha* <u>*lennet*</u> ***en deus*** ____ *al levr.*
NEG know NEG if {expl/read} Fin has read the book
'I don't know if she has read the book.'

(2) ***Petra*** *an archerien e neus soñjet* [<u>*ar seurezed*</u> ***a houie*** __] *?*
what the policemen Fin has thought the sisters Fin knew
‘What did the policemen think that the sisters knew?’

Previous research established that Breton V2 is neither phonological nor syntactic (Rivero 1999). Jouitteau (2012, forthcoming) observes that last-resort V2 operations are prior to auxiliary *do*-support insertion and that verb-doubling shows morphological idiosyncrasy. She concludes that the last-resort operations for V2 take place in a post-syntactic morphological component before Spell-Out, at linearization.

Taking stock, I propose that linearization is clause bound and obeys morphological exponence at the clause level. Embedded domains linearize most of the time with their initial complementizer, leading to C-V orders in Breton and V-final in German. Some other times, speakers linearize without the complementizer, thus leading to obligatory V2 repairs. This optionality is the source of dialectal variations.

References

Borsley, Robert & Kathol, Andreas, 2000. 'Breton as a V2 language', *Linguistics* 38(4), 665-710.

Favereau, F. 1997. *Grammaire du breton contemporain*, Morlaix: Skol Vreizh.

Haegeman, Liliane. 2012. *Adverbial Clauses, Main Clause Phenomena, and Composition of the Left Periphery*, *The Cartography of Syntactic Structures* 8, Oxford Studies in Comparative Syntax.

Jouitteau, M. 2012. 'Verb doubling in Breton and Gungbe; obligatory exponence at the sentence level', *The Morphosyntax of Reiteration in Creole and Non-Creole Languages*, Aboh, Enoch O., Norval Smith et Anne Zribi-Hertz (éds.), John Benjamins, 135-174.

Jouitteau, M. forthcoming. 'Left Edge Filling Trigger; V2 as obligatory exponence at the sentence level and typological implications', Theresa Biberauer, Sam Wolfe & Rebecca Woods (éds.), *Rethinking V2*.

Jouitteau, M. 2005. *La syntaxe comparée du Breton*, PhD ms.

Rivero, M.L. 1999. 'Stylistic verb-movement in yes-no Questions in Bulgarian and Breton', *Crossing Boundaries, advances in the theory of Central and Eastern European languages*, Kenesei, Istvan (éd.), Amsterdam, John Benjamins, 67-90.

Wurmbrand, Susi. 2014. 'The Merge Condition: A syntactic approach to selection', Peter Kosta, Lilia Schürcks, Steven Franks, and Teodora Radev-Bork (eds.), *Minimalism and Beyond: Radicalizing the interfaces*, Amsterdam: John Benjamins, 139–177.

Embedded V2: Experimental evidence from Estonian relative clauses

Marju Kaps
(University of California, Los Angeles)

Estonian has been classified as a verb-second (V2) language (Holmberg, 2015), but shows considerable variability between V2 and verb-final order in embedded clauses, particularly in relative clauses (RCs) and embedded interrogatives (Lindström, 2005). Variability in embedded V2 has previously been studied experimentally in Scandinavian, where theoretical proposals (Simons, 2007) have linked the availability of embedded V2 to the at-issueness of the clause, whereby V2 order indicates that the content of the clause is asserted and addresses a salient discourse question. However, at-issueness effects have been difficult to experimentally demonstrate independently of effects contributed by the embedding predicate (Djärv et al., 2017). The present study reduces effects of factivity arising from the matrix clause by instead manipulating the at-issueness of RCs using RC-internal adverbs. Not-at-issue appositive RCs (Potts, 2005) were marked by using speaker-oriented adverbs such as "unfortunately" (de Vries, 2012) and compared to no-adverb controls, which were ambiguous between a restrictive and appositive interpretation. Crossing the presence of a speaker-oriented adverb (i.e. RC type) with RC word order (V2, verb-final) yielded four experimental conditions.

Sample item:

'The audience clapped for the singer who (unfortunately) performed for the last time.'

V2 conditions

Publik	*plaksutas*	*lauljale,*	*kes*	***esines***	*(kahjuks)*	*viimast*	*korda*
audience	clapped	to singer	who	**performed**	unfortunately	last	time

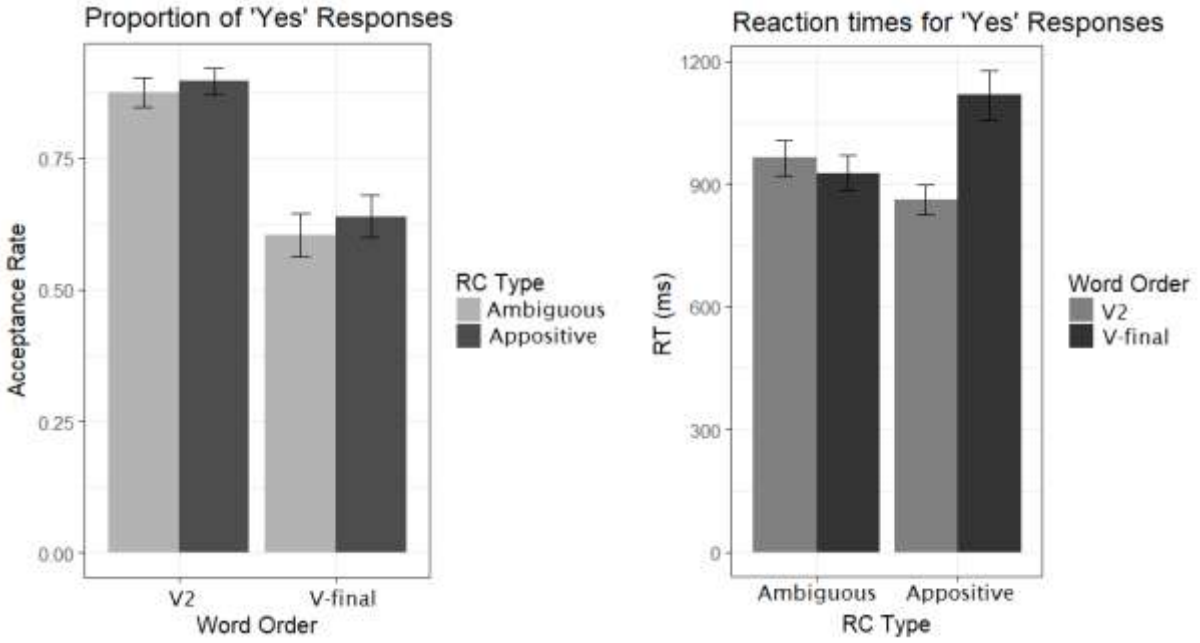
Verb-final conditions

Publik	*plaksutas*	*lauljale,*	*kes*	*(kahjuks)*	*viimast*	*korda*	***esines***
audience	clapped	to singer	who	unfortunately	last	time	**performed**

The experiment consisted of a speeded acceptability task (N=36), where sentences were presented visually, word-by-word. Participants' judgements of acceptability ("yes"/"no") and response times were recorded. I tested two contrasting hypotheses. The at-issueness account of embedded V2 would predict V2 to be *less acceptable* in appositive RCs (which contribute not-at-issue content), compared to controls. Alternatively, appositive RCs have been argued to be syntactically independent of their host clause (e.g. Haegeman, 2009), potentially licensing root phenomena (including V2) *more* than controls. Comparing V2 and verb-final appositives and control RCs, the experiment revealed a uniform V2 preference across clause types in terms of acceptance rates (p<.05 in GLMER models). "Yes" responses were more frequent to V2 RCs (88.5%) than verb-final RCs (62.1%). Despite the lack of an effect of clause type in acceptance rates, reaction times to "yes" responses showed an interaction between clause type and word order (p<.05 in LMER models), with verb-final order penalized in appositives (diff=257ms) but not in ambiguous controls (diff=37ms), indicating processing difficulty associated with verb-final appositive RCs (see Figures).

The experiment failed to support a semantic at-issueness account for embedded V2 (see Simons, 2007), instead pointing to syntactic factors such as root clause status as a potential explanation for V2 in RCs. In addition, V2 appears to present as the preferred order for Estonian RCs, raising interesting questions about the possibly pragmatic conditions under which it can be deviated from. Embedded V2 phenomena can be difficulty to study due to speaker judgements being subtle, and this study highlights experimental approaches, particularly those relying on reaction times and other measures sensitive to processing difficulty, as a promising direction for future work.

Figures:



References

Djärv, K., Rohde, H. & Heycock, C. (2017). Assertion and Factivity: Towards explaining restrictions on Embedded V2 in Scandinavian. In Laura Bailey & Michelle Sheehan (eds.) *Order and Structure in Syntax (Open Generative Syntax).* Berlin: Language Science Press.

Haegeman, L. (2009). Parenthetical adverbials: The radical orphanage approach. In *Dislocated elements in discourse*. Routledge.

Holmberg, A. (2015). Verb second. *Syntax–Theory and Analysis. An International Handbook of Contemporary Syntactic Research*, 342-383.

Lindström, L. (2005). *Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles*. Tartu.

Potts, C. (2005). The logic of conventional implicatures (No. 7). Oxford University Press.

Simons, M. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua, 117*(6), 1034-1056.

de Vries, M. (2012). On appositives and quasi-relatives. *Main Clause Phenomena: New Horizons, 190*, 177.