DOI: 10.26346/1120-2726-225

A quantitative approach to clause type and syntactic change in two Indo-European corpora

Luigi Talamo,[a] Annemarie Verkerk,[b] Iker Salaberri[c]

[a] Department of Language Science and Technology, Universität des Saarlandes, Germany <luigi.talamo@uni-saarland.de>
[b] Department of Language Science and Technology, Universität des Saarlandes, Germany <annemarie.verkerk@uni-saarland.de>
[c] Department of Linguistics and Basque Studies, University of the Basque Country, Spain <iker.salaberri@ehu.eus>

The aim of this paper is to empirically test the claim that subordinate clauses tend to preserve conservative features in language change. To this end, the diachronic behavior of two well-understood and frequently adduced features of grammar, namely null subject pronouns and order of subject, object and verb, is analyzed for main and adverbial clauses in a balanced corpus of 45 Indo-European languages. This study combines qualitative and quantitative analysis by drawing on individual descriptive grammars and parallel corpora respectively. Additionally, diachronic change is modeled using phylogenetic comparative methods. The data suggest that adverbial clauses can in some cases develop asymmetries with respect to their independent counterparts, either through innovation or through preservation of conservative features, possibly due to a communicative need to distinguish clause types by means of grammar. However, the general tendency is for adverbial clauses to change much in the same way as main clauses. This finding contradicts previous claims and calls for a reassessment of studies on the diachronic nature of distinct clause types.

KEYWORDS: subordination, word order, null subject pronouns, Indo-European, corpus linguistics.

1. Introduction

Claims concerning the diachronic behavior of different clause types have often been substantiated by means of data from Indo-European languages (Stockwell & Minkova 1991: 399-400, Luraghi & Pinelli 2015, Ledgeway 2021, Hock 2021: 508-509, Jing *et al.* 2023 among others). However, and despite this sizeable body of work, most studies have either performed quantitative analysis by focusing on one single language (Jucker 1990, Vance 1997) or have looked into a deliberately limited number of changes and linguistic features in a reduced number of languages from a qualitative point of view (Bybee *et al.* 1994: 230-231, Bybee 2002). Consequently, except for Jing *et al.*'s (2023) recent study,

 (received June 2023)

there is as yet no research on this topic drawing on large-scale comparison of languages and empirically testable data.

In view of the state of affairs just described, this study aims to test the claim that subordinate clauses are conservative by analyzing the rate of null subjects and order of subject, object and verb in main and adverbial (temporal, causal, concessive, etc.) clauses in a balanced sample of 45 Indo-European languages; see Section 3 for more details on corpus and methodology. Adverbial clauses have been selected due to the fact that – unlike other kinds of clauses such as complements – they are available in the annotation scheme we peruse for all the languages under study, which enables cross-linguistic comparison. In turn, the motivation behind looking at null subject anaphora and order of subject, object and verb is that these two features are well-understood in comparison to other parts of grammar (cf. Dryer 2005a-b). Moreover, these two variables have already been used to account for the diachronic behavior of different clause types, i.e. Stockwell & Minkova (1991: 372-381) and Hock (2021: 508-509) for order of subject, object and verb, and Vance (1997: 294-321) for null subjects.

For our study we make a questionable but necessary assumption, namely what exactly constitutes 'conservative' grammar in Indo-European languages. The issue of the basic word order of the proto-language has been much debated, with proponents particularly of SOV (Lehmann 1974) and SVO (Friedrich 1975). The problem here resides in the fact that the earliest attestations from the various branches of the family show different patterns, with predominant verb-final order in Anatolian, Indo-Iranian, Italic and Tocharian, verb-initial in Celtic and an unclear data situation wavering between verb-medial and verb-final in Albanian, Armenian, Balto-Slavic, Germanic and Greek (Comrie 1998: 89). The question is further complicated by the fact that the first texts written in languages belonging to the different branches are in some cases separated by millennia. Furthermore, not all subgroups have developed in the same direction: while Balto-Slavic, Germanic and Italic have moved from more verb-final to more verb-medial, the opposite tendency can be observed in Armenian and Indo-Iranian.

Despite the former, it is now quite widely acknowledged that Proto-Indo-European must have had flexible word order with (perhaps) a preference for verb-final order (Ringe 2006: 64, Viti 2014: 81, Krisch 2017: 115). Therefore, for this study verb-final order has been considered a conservative trait, non-verb-final order an innovative one. Much the same is true for null subjects: even though the data are not very clear in this respect, older Indo-European languages seem to have allowed for null subjects and, perhaps to a lesser extent, null objects (Gamkrelidze &

Ivanov 1995: 317-318, Krisch 2009: 211). Accordingly, we have taken a high rate of zero pronouns to be a conservative trait, a low rate to be an innovation.

This is not to imply that the languages under analysis have not changed in more than one way in the course of their histories, including in the direction towards what we consider conservative traits. Welsh, for example, is believed to have had a high frequency of verb-initial order during the Old Welsh period, then to have developed verb-medial order in the Middle Welsh period, and to have changed back to verb-initial order in the Modern Welsh period (Willis 2010: 146). For the purpose of simplicity, we have focused only on the earliest diachronically traceable (Proto-Indo-European) and most recent stages of the languages in question.

This paper is structured as follows: Section 2 discusses the most relevant comparative concepts used throughout the paper, followed in Section 3 by an outline of the corpus and research methodology. Section 4 is dedicated to presenting the results and discussing the generalizations that can be made on the basis of the data. Finally, in Section 5 conclusions are drawn.

2. The domains of inquiry

This study is partly based on data from the Universal Dependencies (henceforth, UD) (Marneffe *et al*. 2021) treebank collection and the Corpus of Indo-European Prose (CIEP) (Talamo & Verkerk 2022),[1] and partly on data from descriptive grammars (cf. Section 3 for details). The results of our study thus inevitably reflect how basic comparative terms, particularly 'clause', 'subordination' and 'adverbial clause' are defined in the data sources. A 'clause' is defined in UD as minimally consisting of a predicate plus its core dependent arguments, in addition to optional non-core modifiers. This conception is in line with Haspelmath (2021: 41), who defines a 'clause' as "a combination of a predicate (full verb or nonverbal predicate) and its arguments plus modifiers", as well as with the definitions – whether implicit or explicit – provided in the descriptive grammars.

In turn, according to UD one can speak of subordination whenever a dependent of the main (independent) predicate is formally realized as a clause. This dependency-based approach to subordination is likewise common in most grammars we have consulted. Yousef (2018: 165, 293), for example, argues that in Persian adverbial clauses with non-finite

verb forms like *āmadani* 'while coming' and *gerefte budand* 'while holding' are dependent on the main clause predicate, as in (1a-b).[2]

(1) a. *Āmadan-i,* *Ali* *rā* *tuye* *rāh* *didam*
come.PTCP-while_ Ali OBJ on way see.1SG.PST
'While coming, I saw Ali on the way'
(Yousef 2018: 165)

b. *Dar* *hālike* *dast-e* *ham-digar* *rā* *gereft-e* *bud-and*
at while hand-of one-another OBJ take-PTCP be-3PL
vārede *khāne* *shod-and*
enter house AUX-3PL
'They entered the house while holding each other's hands'
(Yousef 2018: 293)

Even though 'dependency' has been traditionally regarded as a distinctive feature of subordinate clauses, in practice it does not work well as a criterion for subordination. Dependency is usually defined in terms of the presence of subordinating conjunctions like 'that' and 'because', reduced morphosyntactic properties such as absence of TAM and indexing markers, and the impossibility to occur in isolation (Cristofaro 2003: 15). According to these criteria, however, *āmadani* 'while coming' in (1a) should not be considered a subordinate clause because there is no subordinating conjunction accompanying it. Neither should *dar hālike daste hamdigar rā gerefte budand* 'while holding each other's hands' in (1b) be regarded as subordinate, since the predicate *gerefte budand* 'they were holding' is indexed for the subject's person and number. These facts also illustrate the issue that the criteria for subordination in UD and in descriptive grammars do not always coincide, which partly explains the mismatch between the results of our quantitative and qualitative analysis (cf. Section 4 below).

The third relevant notion under discussion here is 'adverbial clause'. A subordinate clause is characterized as adverbial if the two events expressed by the main and the subordinate clause have an asymmetric figure-ground relation. Using a metaphor taken from Gestalt psychology, Croft (2001: 328-335) argues that the subordinate clause expresses the ground event, against which the event expressed by the main clause stands out as a figure. The asymmetric figure-ground relation can be of different semantic types, describing temporal, clausal, purpose (final), means and additive relations (Croft 2022: 481-483). As mentioned above, this study relies on the UD framework for the annotation and the retrievement of adverbial clauses from corpora; along with the above-mentioned semantic types, the relevant dependency tag for

adverbial clauses additionally includes concessive, equative and conditional clauses, which are often treated separately in the literature.

Once we have defined adverbial clauses for their function, we can discuss how they are formally realized in the world's languages, i.e. their morpho-syntactic encoding. Descriptive and traditional grammars usually rely on the finiteness *vs* non-finiteness notion, by means of which verbal forms are categorized as either finite or non-finite. Even though this distinction can be useful at the language-specific level (Shagal *et al.* 2022: 529-530), no cross-linguistically valid set of criteria has so far been laid out for (non-)finiteness (Cristofaro 2007: 92, Stassen 2009: 260; for proposals see Langacker 1991: 417-438, Givón 2001: 327-387, and Halliday 2014: 23-24, 101). While our study focuses on Indo-European, its implications are meant to apply to general linguistic theory. Accordingly, in the following we work out and operationalize a cross-linguistically valid criterion.

Following previous typological work on complex sentences (Koptjevskaja-Tamm 1993, Cristofaro 2003, Stassen 2009), we analyze morpho-syntactic strategies in the encoding of adverbial clauses by using the notion of balancing *vs* deranking (Stassen 1985, 2009, Croft 2022: 476). According to Stassen (2009: 256), a balanced coding strategy is found when subordinate clauses use the same morpho-syntactic categories for their predicate forms as declarative main clauses, while in a deranked coding strategy the predicate form of one of the two clauses is limited to a set of morpho-syntactic categories that cannot be used in the predicate of a declarative main clause. Balanced strategies are used in both coordination and subordination, whereas deranking strategies are confined to subordinate clauses. Accordingly, the more verb forms confined to subordinate clauses a language has, the closer it is to being a 'pure' deranking language; the fewer, the closer it is to being an 'exclusive' balancing language.[3] Compare the following three complex sentences in Italian (2a-c):

(2) a. *Mi sforzo, ma non riesco a trovare una soluzione*
REFL try.1SG.PRS but NEG be_able.1SG.PRS to find.INF INDF solution
'I try, but I cannot find a solution.' (coordination)

b. *Per quanto mi sforzo, non riesco a trovare una soluzione*
for how_much REFL try.1SG.PRS NEG be_able.1SG.PRS to find.INF INDF solution
'No matter how I try, I cannot find a solution.' (subordination)

c. *Per quanto mi sforzi, non riesco a trovare una soluzione*
for how_much REFL try.1SG.PRS.SBJV NEG be_able.1SG.PRS to find.INF INDF solution
'No matter how I try, I cannot find a solution.' (subordination)

Sentences (2a) and (2b), in which the verbs *sforzarsi* 'to try' and *trovare* 'to find' are in the indicative, use a balanced strategy. The difference between these two complex sentences is in the subordinate conjunction *per quanto* 'no matter how', which however is not a mark of deranking. In turn, sentence (2c) shows a deranked strategy, whereby the conjunction *per quanto* is combined with the subjunctive form of the verb *sforzarsi* and the verb of the declarative main clause is in the indicative. In Italian, the subjunctive mood is rarely used in declarative main clauses.

Deranking can involve any category of the predicate form, including verbal categories such as subject/object indexation and TAM; ultimately, deranking may cause the loss of some of these verbal categories (Stassen 2009: 257). Indeed, this is often reported for non-finite forms, as well as for some nominal categories such as case, gender and number; likewise often mentioned are forms differently addressed as verbal nouns, participles, nominalized forms, etc. (Malchukov 2006). Along with special forms ("dependent moods", Stassen 2009: 261) like the Italian subjunctive and the loss of verbal categories, a third criterion for deranking is represented by special markers on the predicate form, such as the *-ing* suffix in English. More examples of balancing and deranking follow in Section 4.1.

According to Stassen (1985), the identity (same-subject) *vs* diversity (different-subject) of the subject in the matrix and adverbial clauses may condition in some languages the choice between balanced and deranked strategies; under the different-subject condition, some languages are denied the use of deranked strategies (absolute deranking), while still allowing for deranking under the same-subject condition (Stassen 2009: 264-265). For our study we have not, however, looked at the identity or the diversity of the subject in matrix and adverbial clauses, due to limitations of space and time and because this topic is rarely if ever mentioned in descriptive grammars, which form the basis for qualitative analysis (cf. Section 3). Moreover, in order to consider the same-subject *vs* different-subject distinction quantitatively, a corpus annotated for anaphoric reference would be necessary, which currently does not exist for most Indo-European languages. Therefore, the impossibility to address this point should be regarded as a limitation of our study.

3. Corpus analysis and methodology

This study investigates syntactic variability in main and adverbial clauses in a balanced sample of 45 Indo-European languages; see Appendices 1 and 2 for a comprehensive list. As stated in Section 2, the

data on these languages have been drawn from three sources: descriptive grammars, the CIEP and the UD treebank collection. We analyzed five measures or variables, namely, (1) balanced *vs* deranked verb forms, (2) null subjects, (3) order of S and V, (4) order of O and V, (5) order of S, O and V,[4] and modeled change in verb form, null subject rate and word order variability using phylogenetic comparative methods. Additionally, we assessed differences concerning the two latter variables in adverbial clauses in terms of placement within the sentence (see more details below).

3.1 Data sources

Our methodology combines qualitative analysis with quantitative analysis. As for the qualitative analysis, we have consulted descriptive grammars for 45 Indo-European languages; the results and full list of consulted grammars can be found in Appendices 1 and 2. For the quantitative analyses we used the CIEP corpus and the UD treebanks. The CIEP is a parallel corpus consisting of translations of up to 18 literary texts in 30 languages, with a size varying between ~2M tokens (~120,000 sentences: 18 texts) and ~226,000 tokens (~20,000 sentences: 3 texts); the corpus is automatically annotated for lexical information, syntactic dependencies and morpho-syntactic features using the Stanza parser, which has pre-trained UD models for each of the CIEP languages.

We are aware of a number of drawbacks in using parallel texts for comparative and historical linguistic analysis. This approach implies a bias towards planned (conscious) language use, religious and legalese registers, large and well-described languages, standardized varieties and translated (rather than genuine) language use (Wälchli 2007: 99, 132). In order to cover other linguistic registers and to expand our sample beyond the languages featured in CIEP, we have considered UD treebanks, which are collections of annotated sentences that vary widely in register and size. The selection of UD treebanks and languages was based mostly on their size; we chose languages that had sizable treebanks. In case a language had more than one treebank, we usually selected the biggest treebank. In a few cases smaller treebanks were chosen; for instance, for Icelandic we chose the Modern treebank (80,000 sentences) over the IcePaHC (985,000 sentences), because the latter is a historical corpus of Icelandic. In total, we analyzed UD treebanks for all 45 languages, however, we excluded Dutch and Old French from further consideration below, as their treebanks do not feature information on the mood category. In addition, we excluded Bhojpuri as its treebank turned out to be too small (357 sentences). Hence, the analyses conducted on UD treebanks are on 42 languages; on CIEP, on 30 languages.

For the phylogenetic analyses on UD treebanks, Nynorsk Norwegian, Old Russian, Upper Sorbian, and Old French are disregarded, as they do not feature in Bouckaert *et al.*'s (2012) phylogenetic tree set.

3.2. Annotation

UD is a framework for annotating linguistic structures at different levels: these include syntax (UD Relations), morphosyntax (UD Features) and the lexicon (universal part-of-speech tags (UPOS) and lemmatization). UD Relations is a dependency grammar that establishes head-dependent relations between two words and defines these relations using a cross-linguistically valid set of syntactic annotations, such as nominal modification (nmod), as in Spanish *La casa* [*de papel*] 'the paper house', or subject (nsubj) and object (obj), as in Dutch [*Robin*] (nsubj) *kijkt* [*een film*] (obj) 'Robin watches a film'. Annotation begins at the root node (root), which is identified as the verbal predicate of the main clause. In order to be cross-linguistically consistent, the root can also be a property word (adjective) or an object (a noun), which corresponds to non-verbal predication. We recognize an instance of adverbial clause whenever we find an adverbial clause relation between two tokens, where 'advcl' stands for the syntactic relation describing an adverbial clause in UD or, more precisely, for the (non-)verbal predicate of an adverbial clause.

A final word on annotation concerns mood values. The mood values mentioned in descriptive grammars of individual languages do not always coincide with those given by UD for the same languages. For example, Kalnača & Lokmane (2021) state that Latvian has five mood categories, namely indicative (Ind), imperative (Imp), conditional (Cnd), oblique (Obl) and debitive (Deb). However, UD lists the first three in its Latvian treebank,[5] where Obl is labeled 'quotative' (Qot) and Deb is called 'necessitative' (Nec). Whenever there is a mismatch we stick to the UD tags, i.e. Ind, Imp, Cnd, Qot and Nec.

3.3. Data extraction and analysis

The quantitative analysis goes as follows. For each adverbial clause we extracted the following information from the syntactic annotation: (i) subject and object arguments of the head and the adverbial clause, and (ii) (non-)verbal predicate of the head and the adverbial clause, including additional elements such as copula and auxiliary; these two points (i-ii) suffice in order to extract information on null subjects and the order of constituents. Here a caveat is in order: while it is true that UD has no empty nodes, we considered null subjects instances of clauses whose predicate does not have a nominal subject (UD: nsubj) among its dependents. Consider the

Italian sentence displayed in Figure 1; the adverbial clause has an overt subject and its main clause has a null subject and an overt object.

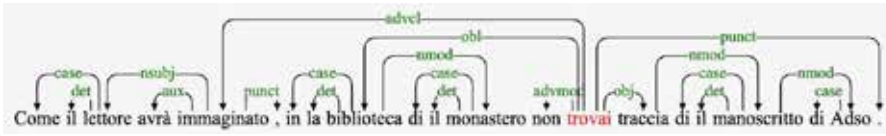

Figure 1. UD parsing of a sentence from the original text of Umberto Eco's *Il nome della Rosa*: 'As the reader must have guessed, in the monastery library I found no trace of Adso's manuscript' (English translation by William Weaver).

In order to determine whether the verb form is balanced or deranked, we need to look at another level, UD Features, specifically, the 'Mood' variable. With the aid of grammars we checked for each language of the sample which moods are available for main declarative clauses. Then, for each adverbial clause we checked whether its (non)verbal predicate and additional elements (copula, auxiliaries) are marked for a main declarative clause mood (balanced) or another mood (deranked). For instance, the adverbial clause of the sentence displayed in Figure 1 features an indicative mood, i.e. *avrà*.3SG.FUT *immaginato*.PTCP.PST 's/he will have guessed', thus using a balanced strategy. If a language has a higher proportion of non-declarative moods in adverbial clauses, while main clauses are mostly marked for a declarative mood (typically the indicative), then we consider the language to have deranked verb forms, i.e. to be a deranking language. Inversely, if the proportion of non-declarative moods in adverbial clauses is the same or lower than in main clauses, then we consider the language to have balanced verb forms, i.e. to be a balancing language.

Moreover, we also looked at the differences between the values for balancing and deranking; specifically, the larger the difference in value between main and adverbial clauses, the more likely it is that a language has deranked verb forms. For example, in Italian the value in CIEP for main clauses is 0.07 (proportion of predicates marked for non-declarative moods) *vs* 0.52 for adverbial clauses. This means that the difference in usage of non-declarative mood is 0.45, i.e. quite considerable, which suggests that Italian is a deranking language. In turn, in the same corpus the values for Russian are 0.32 (proportion of predicates marked for non-declarative moods) for main and 0.24 for adverbial. This indicates that Russian is a balancing language.[6] The script we designed and used for our study specifies that data collection stops at the second level of embedding.

Our choice of sources is intertwined with how the data have been examined: the results of qualitative analysis of descriptive grammar data

are contrasted with the results of quantitative data extracted from the CIEP and the UD treebanks, with the aim of gaining insights that would otherwise go unnoticed in a purely qualitative or purely quantitative study. Among other things, quantitative analysis allows us to observe syntactic variation directly, for example with respect to languages with variable uses of null subjects such as Belarusian and Ukrainian. This last point is related with the conception, advocated among others by Biagetti *et al.* (2023) and Levshina *et al.* (2023), that word order – and, by extension, other syntactic features such as rate of null pronouns – should not be regarded as a categorical but as a continuous variable, since variability in word order is the rule rather than the exception in the world's languages. It follows from this that labels such as 'verb-medial' and 'verb-final' should not be taken at face value. Instead, they should be gradable, i.e. languages should be considered 'more' or 'less' verb-medial or verb-final.

There are several advantages to adopting a gradient approach to typological variables. First of all, it allows us to express word order values numerically, which is more precise and informative than using labels such as 'SVO'. Second, this method is compatible with usage-based conceptions of language, according to which linguistic forms, meanings and the ways in which they are matched reflect various kinds of associations, the strength of which may vary (Diessel 2019). Third, a gradient approach can be matched with new data sources in the form of large corpora and with quantitative analyses based on statistical and computational methods. Fourth, this approach allows researchers to refrain from classifying word order patterns by means of problematic, non-clear-cut categorical notions like 'flexible', 'rigid' and 'dominant'.

In accordance with this view, and drawing on the proportion of verb-medial clauses in each language, we modeled word order by means of a general gradient measure ranging from 1 to 3, where 1 is equal to a language with exclusive verb-initial order, 3 is a verb-final-only language, and 2 is a language where all clauses follow verb-medial order. In turn, concerning the balanced *vs* deranked verb form and null subject variables we adopted gradient measures ranging from 0 to 1. In the former case 0 corresponds to languages which always use the declarative mood(s) for the (main or adverbial) clause, whereas 1 corresponds to languages always marking the (main or adverbial) clause with non-declarative mood(s). In the latter case the 0-to-1 measure reflects a natural cline in the proportion of null subjects in any given language.

3.4 Phylogenetic analysis

In order to make statistical inferences about our sample of related Indo-European languages, we use phylogenetic comparative methods. First, we test for phylogenetic signal using Pagel's (1999) lambda (λ) method as implemented in *phytools* (Revell 2012) in *R* (R Core Team 2022). Then, we move on to analyze correlations between adverbial and main clauses, and across datasets. In the past, for example in Verkerk (2014), PGLS (Phylogenetic Generalised Least Squares, Symonds & Blomberg 2014) has been used to model proportions. However, while proportions are continuous, they are capped by 0 and 1, where other types of continuous measures are not (temperature, weight). The data gathering procedure behind proportions comes down to measuring how many *A*s (for example, clauses) out of a sample of *n* (observed clauses in a corpus) equal *X* (for example, having a zero subject). This is very different from a true continuous measure such as temperature. Using PGLS or other 'simpler' regression models on proportion data would generate estimates that fall out of the [0,1] range and is hence a poor fit to proportion data (see also Douma & Weedon 2019).

Therefore, we adopt the models described by Winter & Bürkner (2021), which specifically deal with count data in linguistics. We assess correlations of the same measure between the two corpora and correlations of similar measures across main and adverbial clauses using logistic and beta regression in the *R* package *brms* (Bürkner 2017). All analyses include a phylogenetic term to deal with the genealogical autocorrelation present in the samples, based upon phylogenetic trees from Bouckaert *et al.* (2012). First, in logistic regression, we aim to avoid using proportions in the response term of the regression formula, but at the same time we need to allow for subcorpora to have different sizes – even in a full parallel corpus, there will be differences in the total number of clauses. We do this by modeling the count measure, for example the number of adverbial clauses with a zero subject, in terms of the total number of relevant clauses, using the *trials()* function. However, we cannot do this for the predictor measures. Hence, we decided on the following solution: each correlation we test is tested twice, so that each relevant measure can be modeled in the best way, namely as a count response variable. Predictor values are still proportions. As an example, here is how the correlation is tested between the number of null subject clauses in adverbial clauses with the number of null subject in main clauses:

brm(num_null_adv | trials(total_adv_cl) ~ prop_null_main, family = binomial)

brm(num_null_main | trials(total_main_cl) ~ prop_null_adv, family = binomial)

Second, for the verb score, we had to find a different solution, as this is not a proportion but rather a summary variable, capped at one and three, hence suffering from the same problem as our other measures. Verb scores were transformed to be on a scale from zero to one, and then beta regression in *brms* (Bürkner 2017) was used. In parallel to the logistic regression analyses above, we tested each correlation in both directions, once transformed and once untransformed. As an example, here is how the correlation is tested between word order in adverbial clauses with word order in main clauses:

```
brm(adv_verb_score_01 ~ main_verb_score, family = Beta())

brm(main_verb_score_01 ~ adv_verb_score, family = Beta())
```

When we report on the correlations between measures below, 'statistical significance', or rather, the relevance or importance of the predictor for the model, is assessed by seeing if its estimate excludes zero. Because we test each correlation twice, we should probably apply some correction, but since we are merely interested in the existence of a correlation and not its strength, we do not apply such a correction or other formal testing. All results are provided in Appendix 7, so the reader may review the output summaries directly.

4. Results and discussion

In this section we present the results of our study of the three variables laid out in Section 3: balancing *vs* deranking, word order and rate of null subjects. For each variable we first discuss the results of qualitative analysis, then the results of quantitative analysis, and after that we address the mismatches we come across, including those between qualitative (descriptive grammar) *vs* quantitative (CIEP + UD) analysis and those between CIEP-based and UD-based quantitative analysis. We conclude by presenting the results of phylogenetic analysis.

In general, the various analyses coincide in their results, and it can be claimed that the descriptions of grammars fit well with the corpus data. However, there are also exceptions. For instance, whereas descriptive grammars rarely mention whether or not a specific verb form is confined to subordinate clauses, the CIEP and UD-based data clearly show that the sample languages tend to be mostly deranking, next to a minority of balancing languages. Moreover, the differences between CIEP and UD data on balancing *vs* deranking are striking for a few languages, whereby

the CIEP data show a stronger tendency towards deranking. Nevertheless, the quantitative results on all three variables are correlated, given that the *brms* analyzes sampling of the 30 languages that overlap between the CIEP and UD corpora. Concerning languages with flexible word order, there are some divergent assessments in the grammars and in the corpora with respect to the basic word order pattern. Something similar occurs with languages that allow variable rates of null subjects.

In order to carry out phylogenetic analysis we first tested the various measures under study for phylogenetic signal, which is a standard step before moving on to more complex analyses. The phylogenetic signal indicates the tendency of related languages to be more similar to each other than randomly selected languages from the same family tree. Therefore, this is a useful measure in order to make sure that the results are not conditioned by genetic relatedness. The signal turned out to be strong – i.e. over 0.6 in a 0-to-1 scale – in most of the variables, which means that these variables seem to change in accordance with the branches of the phylogenetic tree (see Section 3) in the sample languages.[7]

4.1. Balancing vs deranking

According to the grammars we consulted, out of 45 languages of our sample only Persian does not have a deranking strategy; the qualitative analysis is partially confirmed by the quantitative study. In the CIEP sample, the majority of languages (23/30) have deranked verb forms, whereas far fewer languages (7/30) present balanced forms, cf. Figure 2. The bars represent the proportion of clauses marked for deranked moods, blue for main clauses, and red for adverbial clauses.

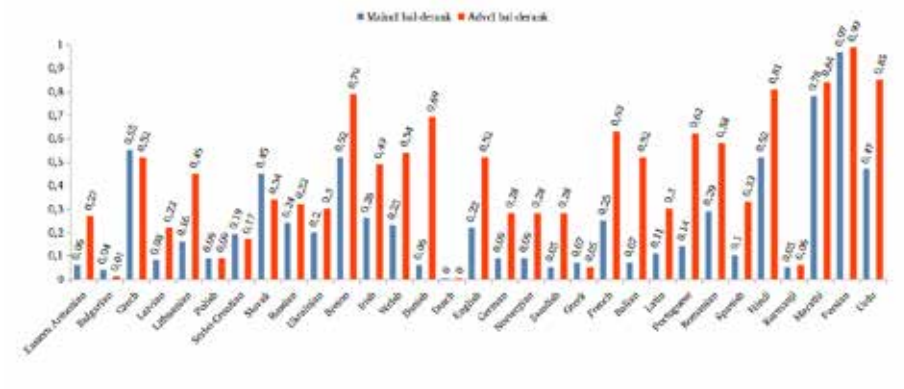

Figure 2. Balancing *vs* deranking in the sample languages (based on CIEP).[8]

Some of the balancing languages, including Belarusian, Bulgarian, Greek, Kurmanji, Polish, Scottish Gaelic and the above-mentioned Persian, show very similar values for main and adverbial clauses, which means that in these languages there are no (or hardly any) verb forms confined to subordinate clauses. By contrast, in languages like French (0.23 for main *vs* 0.68 for adverbial), Old Church Slavonic (0.05 (main) *vs* 0.65 (adverbial)) and Welsh (0.16 (main) *vs* 0.64 (adverbial)) the differences are considerable, which suggests that some verb forms are confined to subordinate clauses. This is illustrated by examples (3a-b) from French, whereby *eusse goûté* 'had tasted' in (3b) is a deranked form that cannot be used in main clauses.

(3) a. *Goûtez nos glaces, nous les préparons toutes avec des fruits frais*
taste.2PL.IMP our.PL ice_cream.PL 1PL 3PL prepare.1PL all with DEF fruit.PL fresh
'Try (lit. taste) our ice cream (types), all (of which) we prepare using fresh fruits' (Delatour *et al.* 2004: 70)

b. *La vue de la petite madeleine ne m'=avait rien rappelé*
DEF sight of DEF small madeleine NEG 1SG=AUX nothing remind.PTCP
avant que je n'=y eusse goûté
before that 1SG NEG=it AUX taste.PTCP
'The sight of the little madeleine had not reminded me of anything before I had tasted it' (Delatour *et al.* 2004: 139)

However, it should be pointed out that the UD and CIEP data on balancing *vs* deranking do not always coincide. For Danish adverbial clauses, for example, the CIEP yields a value of 0.69, i.e. clearly deranked in comparison with main clauses (0.06), whereas the difference is much smaller – and, therefore, seemingly much more balanced – concerning the values provided by UD (0.07 *vs* 0.31). Similar differences between CIEP and UD data occur to a lesser extent in Latin (CIEP = 0.11 (main) *vs* 0.3 (adverbial), UD = 0.33 (main) *vs* 0.2 (adverbial)) and Welsh (CIEP = 0.23 (main) *vs* 0.54 (adverbial), UD = 0.09 (main) *vs* 0.74 (adverbial)). These are at any rate exceptions, since in all other cases the divergence between both corpora, if any, lies at around 0.1.

4.2. Word order

According to the qualitative data that we have extracted from descriptive grammars, only a minority (13/45) of the sampled languages present verb-final order in adverbial clauses (recall that this is what has been reconstructed for the proto-language). In contrast, in the majority of sample languages (30/45) adverbial clauses have innovated verb-initial and verb-medial orders.

In two languages (Gothic and Scottish Gaelic) there is no clear picture, since adverbial clauses are sometimes verb-final and sometimes non-verb-final.[9] In addition, adverbial clauses can be conservative with respect to their independent counterparts. For example, in Scottish Gaelic declarative main clauses are usually verb-initial (4a), whereas adverbial clauses containing verbal nouns and usually labeled as 'non-finite' – deranked for present purposes – are verb-final (4b).

(4) a. *Chuir* *Iain* *leabhar* *gu* *Anna*
send.PST Iain book to Anna
'Iain sent a book to Anna'
(MacAulay 1992: 167)

b. *Sheas* *Anna* *an* *dèidh* *do* *Iain* *leabhar* *a* *cheannach*
stand_up.PST Anna to after to Iain book to buy
'Anna stood up after Iain had bought a book'
(MacAulay 1992: 172)

In turn, balanced adverbial clauses are verb-initial just like their independent counterparts (5a-b).

(5) a. *Is* *fhada* *bho* *nach* *fhaca* *mi* *thu*
be.PRS long since NEG see.PST 1SG 2SG
'It is (a) long (time) since I saw you (lit. since I haven't seen you)'
(Gillies 2010: 295)

b. *Dh'=fhalbh* *sinn* *mun* *do* *dh'=èirich* *a'=ghrian*
on=depart.PST 1PL before for on=rise.PST DEF=sun
'We departed before the sun rose'
(Gillies 2010: 295)

In Dutch and German main declarative clauses tend to be verb-medial, whereas all kinds of subordinate clause, including adverbial clauses, tend to be verb-final, as illustrated here by Dutch (6a-b).

(6) a. *Zij* *zag* *overal* *vreemde* *ding-en*
3SG.F see.PST everywhere strange thing-PL
'She saw strange things everywhere'
(Donaldson 2008: 141)

b. *Zij* *ging* *vroeg* *naar* *bed,* *omdat* *ze* *die* *dag* *een* *lange* *wandeling* *had*
3SG.F go.PST early to bed because 3SG.F DEF day INDF long walk have.3SG
gemaakt
make.PST.PTCP
'She went to bed early because that day she had gone for a long walk'
(Donaldson 2008: 142)

Despite the former, there are no cases in the sample where adverbial clauses have innovated a different order of S, O and V compared to main clauses. In the vast majority of sample languages (39/45) main

and adverbial clauses have either remained the same or jointly innovated. The results of qualitative analysis thus indicate that, even though adverbial clauses can occasionally preserve conservative features, in general they behave much in the same way as main clauses.[10]

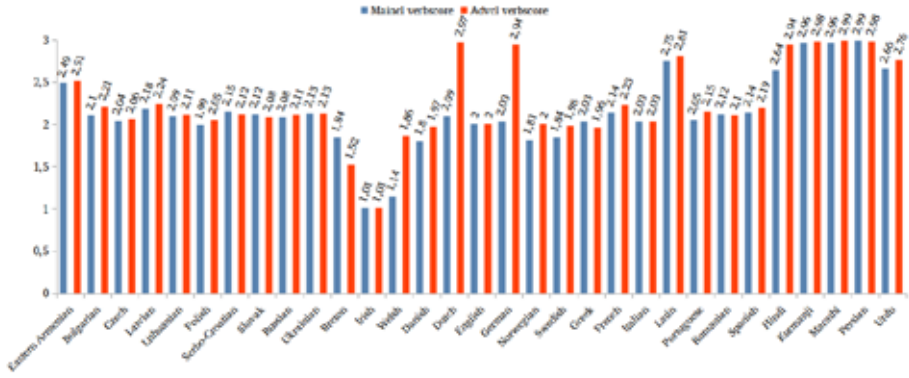

Figure 3. Verb score in the sample languages (based on CIEP).[11]

Data from the qualitative analysis are generally confirmed by corpus data, as can be seen in Figure 3 for the CIEP sample. However, as a caveat to the results of our analysis concerning the order of S, O and V in main *vs* adverbial clauses it should be pointed out that the qualitative and quantitative data do not always coincide. Specifically, concerning Ancient Greek there is a contrast between the relevant publications, which speak of verb-final order of constituents, and the UD data, which rather signal non-verb-final order. The inverse is true for Eastern Armenian and, to a certain extent, for Old French. All of these involve languages with rather flexible order of constituents, which may depend on author, register, and diachronic phase.

4.3. Null subjects

The qualitative data on null subjects indicate that only in a minority (13/45) of the sample languages they are no longer allowed in adverbial clauses. By contrast, in most of the languages in the sample (28/45) adverbial clauses permit null subjects.

Four languages (Belarusian, Hindi, Russian, Ukrainian) present intermediate data, either because null subjects are only allowed to a certain extent, or because they are allowed in a subset of adverbial clauses.

Moreover, in a few cases adverbial clauses can be argued to be conservative with respect to their independent counterparts. Examples usually involve deranked constructions with obligatory null subjects as opposed to optional use of subject pronouns in balanced constructions in languages such as Hindi (7a-b).

(7) a. *Mɛdan* *mẽ* *k^{h}elte* *hue* *bəcce*
field.OBL in play.IPFV.PTCP PL.OBL child.M.PL
mere *kuttõ* *ko* *dek^{h}* *kər* *ruk* *gəye*
1SG.POSS.M.PL.OBL dog.M.PL.OBL OBJ see CP stop go.PFV.M.PL
'The children playing in the field stopped when they saw my dogs (lit. seeing my dogs)'
(Kachru 2006: 227)

b. *Vəh* *soc* *nəhĩ* *rəha,* *(vəh)* *so* *rəha* *hɛ*
3SG think NEG PROG.M.SG 3SG sleep PROG.M.SG PRS.SG
'He is not thinking, (he) is sleeping'
(Kachru 2006: 182)

In turn, there are no cases in the sample where at a given point in time adverbial clauses have innovated overt subject pronouns while main clauses have not, even though some languages of the sample like Old East Slavic have been argued to show a faster tendency for the loss of null subjects in subordinate clauses (cf. Luraghi & Pinelli 2015). In the vast majority of the sample languages (41/45) main and adverbial clauses have either not changed or have done so jointly. Therefore, the results of qualitative analysis suggest that in general terms there is no difference in the diachronic behavior of main and adverbial clauses.

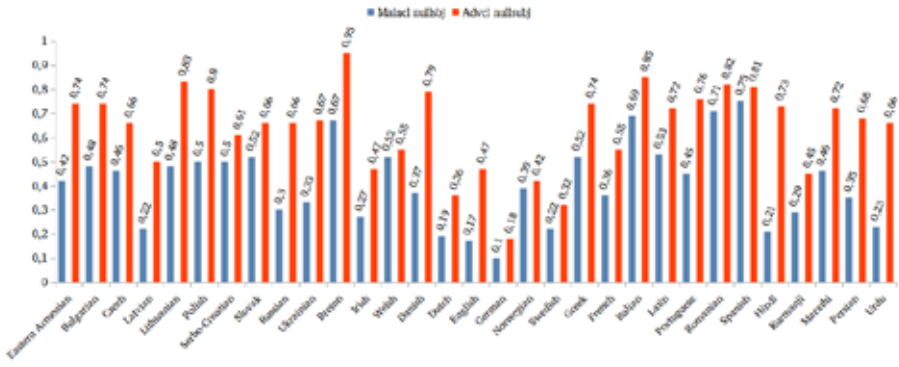

Figure 4. Null subjects in the sample languages (based on CIEP).[12]

Here, as well, a caveat is in order: the results for CIEP, which are visualized in Figure 4, and UD data do not match exactly in all cases. Concerning Danish adverbial clauses, the sources speak of a relatively low rate of null subjects, and this is supported by UD data (0.44), but not by the CIEP, which points toward a high rate (0.79). This may be due to the comparatively small number of texts included in the Danish treebanks, which might have led to an overestimation of texts with high null subject ratios. For all other sample languages there is no more than a 0.2 discrepancy (at most) between the CIEP and UD data.

As already pointed out in Section 2, we cannot account here for the relationship between null subjects and deranking; more specifically, we are not aware of cross-linguistic resources annotated for anaphoric reference, which would allow to test the same *vs* different subject condition as posited by Stassen (1985, 2009).

4.4 Phylogenetic analysis

Here we analyze the relationship between main and adverbial clauses regarding their rate of null subjects as well as their order of S, O, and V. As has become evident already from the previous discussion, these two variables are correlated: languages tend to have similar orders of S, O, and V across main and adverbial clauses and higher rates of null subjects in adverbial clauses than in main clauses. This holds up in all *brms* analyses across the CIEP and UD datasets.[13]

For the order of S, O, and V, beta regression of adverbial clause order and main clause order provides support for correlations for both CIEP and UD datasets. The coefficients are positive, showing that adverbial clauses are slightly more verb-final than main clauses. As is evident from Figure 5, this positive correlation is the effect of German, Dutch, and Welsh having divergent word orders in their adverbial clauses.

For the rate of null subjects, we obtain similar results. Logistic regression of the adverbial clause null subject rate and the main clause null subject rate provides relevant positive coefficients for both CIEP and UD datasets, showing that as languages allow for more null subjects in their main clauses, they generally allow for more null subjects in adverbial clauses accordingly (as shown in Figure 6).

In the remainder of this section we present figures illustrating change in these variables and further discussion. These figures were created using the function *contMap* in R (R Core Team 2022) package *phytools* (Revell 2012) and represent simple ancestral state reconstructions estimated using maximum likelihood. The ancestral state at the root of the tree represents a version of Proto-Indo-European excluding

Anatolian (because no relevant datasets are available). Nevertheless, we will refer to it using the common label PIE. The following discussion is intended as illustrative and tentative only, as we do not sample historical varieties to a large enough extent.

In Figure 6, PIE main clauses are reconstructed to allow for a slightly higher rate of null subjects than adverbial clauses, in line with the contemporary data (see Figure 4). Nevertheless, we observe in some languages and groups change in a similar direction, such as in Northern Germanic, whereas we observe in other languages and groups change such that main and adverbial clauses diverge. This seems to be the case in Hindi and Urdu, Lithuanian, and Western Germanic. We do not observe an across-the-board retention of conservative characteristics (in this case, a high rate of null subjects) in adverbial clauses as opposed to main clauses. As discussed above, languages that allow for more null subjects in their main clauses, generally allow for more null subjects in adverbial clauses as well. In turn, Figure 5 illustrates how similarly word order has changed in main and adverbial clauses. We can observe the well-known changes towards verb-initial word order in Celtic and verb-final word order in Indo-Iranian and Armenian, as well as a convergence to verb-medial word order elsewhere. The only exceptions are German and Dutch, Welsh, and, to some extent, Breton. The former are further discussed below.

Parallel Figures 7 and 8 display diachronic change for the sampled languages from the UD treebanks. While results are similar, it is worth noting that there are languages that change in a different direction from their sister languages, including Kurdish and Scottish Gaelic. There is also a prevalence of conservative characteristics in historical languages (Ancient Greek, Old Church Slavonic, Gothic), which has consequences for the role they play in the reconstruction of PIE, especially for the rate of null subjects. While the UD results converge with the results from CIEP, it is also clear that incorporating as many languages as possible, especially historical varieties, plays a big role in using phylogenetics for the study of language change. Future work should therefore focus on including attested historical varieties, given the availability of comparable annotation.

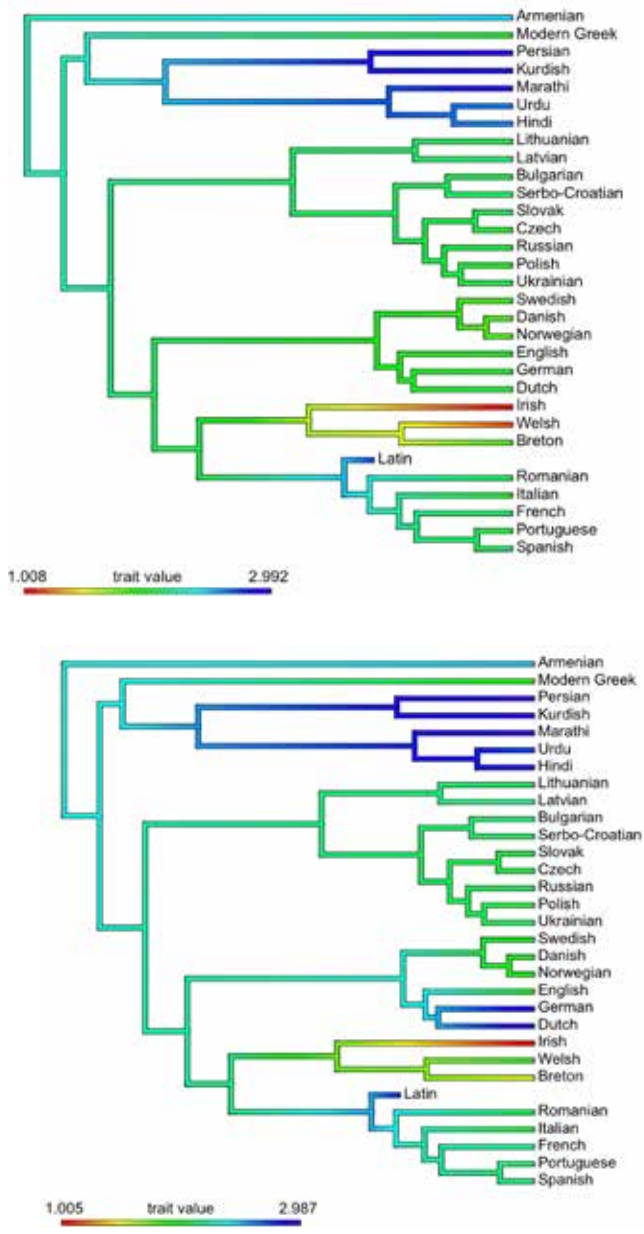

Figure 5. Illustrations of diachronic change in order of S, O, and V in main clauses (upper figure) and adverbial clauses (lower figure) (based on CIEP).

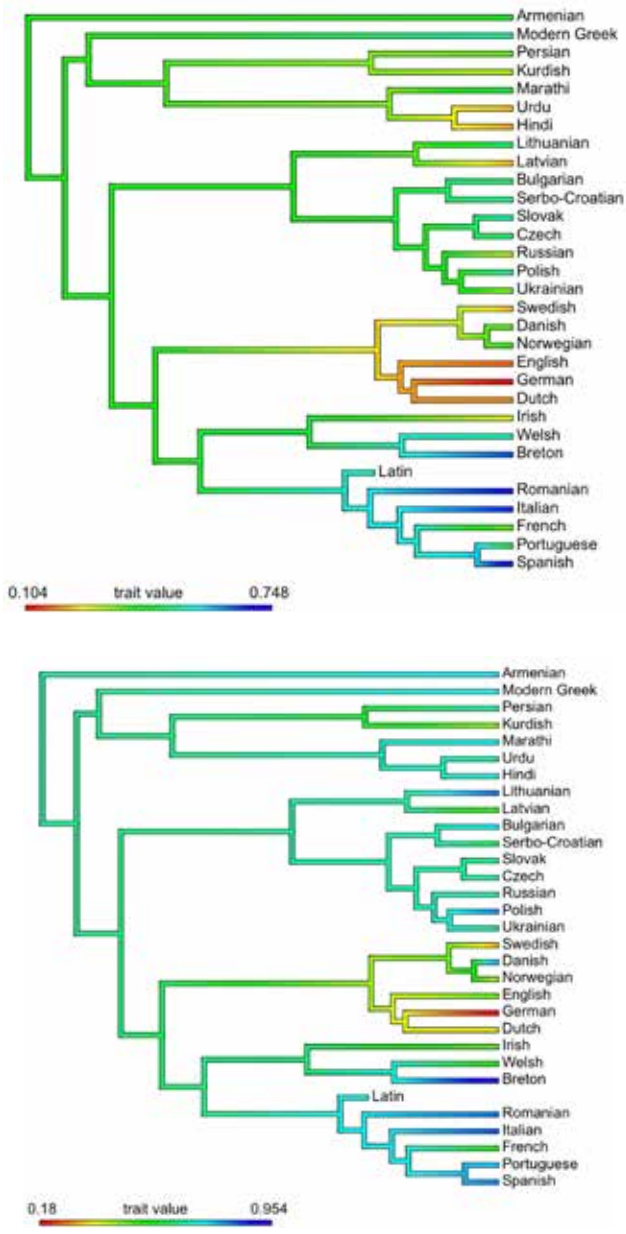

Figure 6. Illustrations of diachronic change in the rate of null subjects in main clauses (upper figure) and adverbial clauses (lower figure) (based on CIEP).

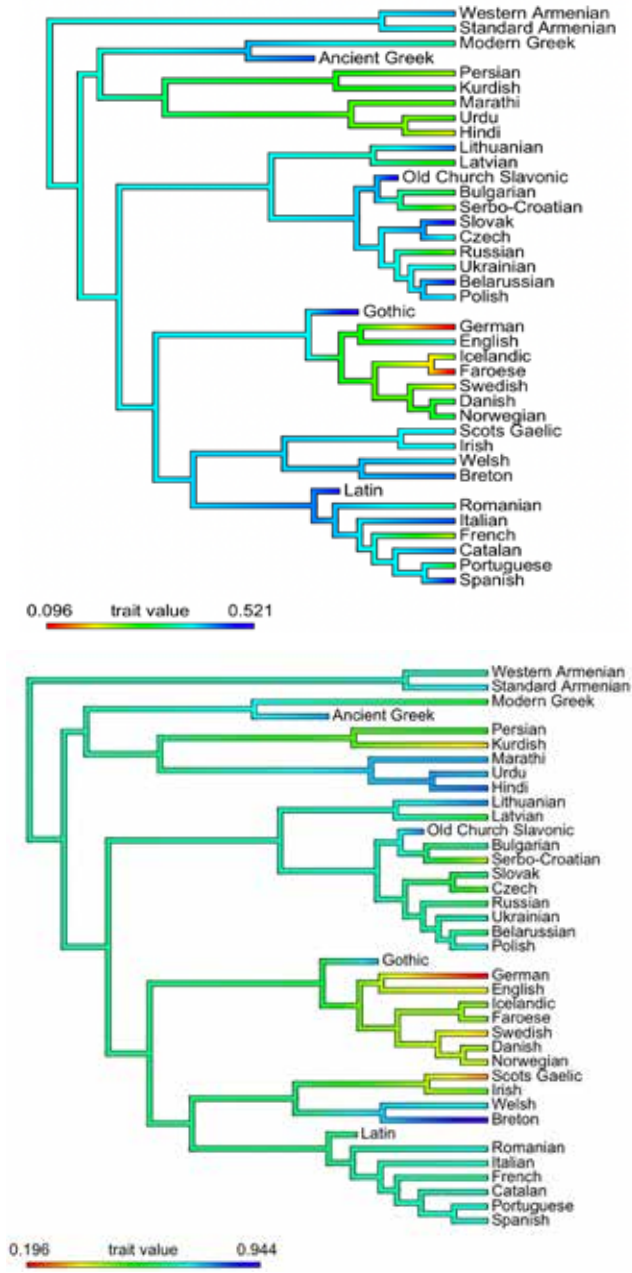

Figure 7. Illustrations of diachronic change in the rate of null subjects in main clauses (upper figure) and adverbial clauses (lower figure) (based on UD).

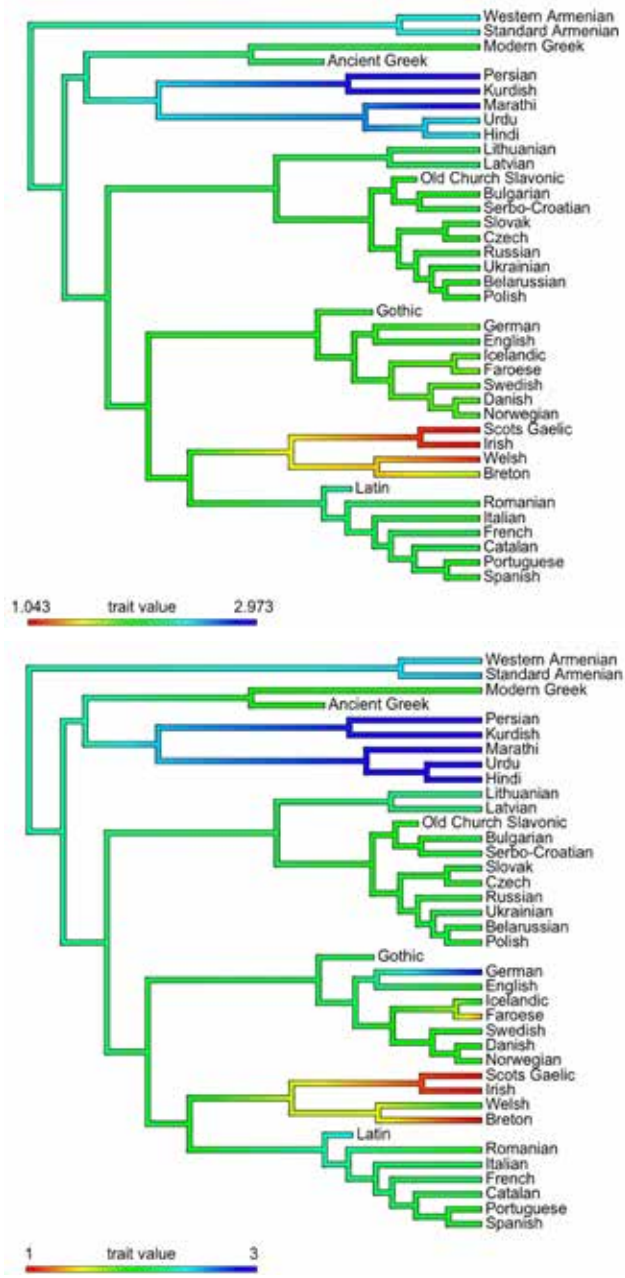

Figure 8. Illustrations of diachronic change in order of S, O, and V in main clauses (upper figure) and adverbial clauses (lower figure) (based on UD).

In short, then, phylogenetic analysis suggests that main and adverbial clauses change similarly in most of the sample languages. The few languages that display considerable differences between main and adverbial clauses are, particularly, Dutch, German and Welsh concerning verb score, and Hindi, Urdu, as well as, again, Dutch, German and Welsh concerning rate of null subjects.

With respect to word order it should be pointed out that verb-final order is reconstructed as the most frequent pattern in Proto-Germanic (Hopper 1975: 82-83), and also that verb-final order in subordinate (including adverbial) clauses in modern Dutch and German is considered an inherited retention (Hock 2021: 509). However, this diachronic connection is not straightforward: subordinate clauses in medieval Dutch and German used to allow for many more postverbal constituents than they do in the modern languages, and this changed only with the advent of normativizing efforts in the 16th century (Von Polenz 2009: 90). As mentioned in Section 1, the evolution of verb position in the history of Welsh is also anything but straightforward (Willis 2010: 146). Therefore, the slight diachronic conservatism of adverbial clauses suggested by our phylogenetic analysis must be relativized. There is also the fact that early Germanic languages are paratactic, i.e. there is barely any formal distinction between main and subordinate clauses, so generalization of verb-final order in subordinate clauses (as opposed to verb-second in main clauses) may have been a way to distinguish main from subordinate clauses by formal means.

In turn, concerning the rate of null subjects it should be mentioned that Hindi and Urdu make a clear distinction between correlative balanced adverbial clauses on the one hand, and participial deranked adverbial clauses on the other. The examples in Kachru (2006: 224-228) illustrate the fact that the former kind are often used when the subject in the main and the adverbial clause do not coincide, whereas the latter are often used in same-subject constructions. Thus the Hindi and Urdu data could simply be the result of a high rate of same-subject constructions, which as explained in Section 2 favor null subjects. Therefore, the higher rate of null subjects in adverbial clauses in some languages of our sample can and should also be relativized.

The results of our study concerning word order are in line with Jing *et al.* (2023), who adopt a phylogenetic analysis of rates of word order change in main and subordinate clauses across Indo-European languages and find no significant differences. There are, however, a number of discrepancies between Jing *et al.*'s (2023) and our study that hamper the comparability of results: while we adopt a gradient approach to word order, they compare relative orders of S, O and V from a probabilistic

perspective. Moreover, they consider various kinds of subordinate clause including adjectival and adverbial clauses, clausal subjects and complements, whereas our focus is on adverbial clauses. Therefore, the similarity of our findings may be coincidental and need not reflect a general lack of differences between the rates of change in main and subordinate clauses across Indo-European.

5. Conclusions

This study has presented the results of quantitative and qualitative analysis of the rate of null subjects and order of subject, object and verb in main and adverbial clauses in a balanced sample of Indo-European languages in three data sources (grammars and UD treebanks: 45 languages; CIEP: 30 languages). This contrasts with previous studies, which have either performed quantitative analysis of a single feature in a single language or looked qualitatively into a number of features in a few languages. The results reveal the existence of a few asymmetries between main and adverbial clauses, which may be explained by the fact that grammatical features like word order can be exploited by speakers for the functional-communicative role of distinguishing clause types. In general terms, however, both kinds of clause overwhelmingly change in the same manner, either in developing innovations or preserving retentions. Accordingly, there are no grounds to support the widespread claim that adverbial clauses tend to preserve conservative features in the face of language change.

Author contributions
LT: conceptualization, data collection, validation, methodology, writing Section 2 (original draft), writing (review and editing).
AV: conceptualization, phylogenetic analysis, validation, writing Sections 3.4 and 4.4 (original draft), writing (review and editing).
IS: conceptualization, data collection, validation, writing Sections 1, 3.1-3.3, 4.1-4.3 and 5 (original draft), writing (review and editing).
All authors contributed to the article and approve the submitted version.

Abbreviations
1, 3 = 1st, 3rd person; AUX = auxiliary verb; CP = conjunctive participle; DEF = definite; F = feminine; FUT = future; IMP = imperative; INDF = indefinite; INF = infinitive; IPFV = imperfective; M = masculine; NEG = negation; OBJ = object marker; OBL = oblique; PFV = perfective; PL = plural; POSS = possessive; PROG = progressive; PRS = present; PST = past; PTCP = participle; REFL = reflexive; SBJV = subjunctive; SG = singular; SUB = subordinator; UD = Universal Dependencies; UPOS = universal part-of-speech tags.

Acknowledgments
This paper has been supported by the research grant project TransGram PID2021-124769NB-I00, funded by the Spanish Ministry of Science and Innovation / Spanish Research Agency (AEI) / 10.13039/501100011033, the European Regional Development Fund (ERDF "A way of making Europe"), as well as by the research group *Diachronic Linguistics, Typology and the History of Basque* (DLTB) (IT1534-22), funded by the Government of the Basque Autonomous Community. In addition, this work has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through SFB 1102 (Project-ID 232722074). We are grateful to two anonymous reviewers for theoretical, methodological and formal suggestions, as well as to Bodo Winter for statistical advice and to Pauline Krielke for general comments on the paper. All remaining errors are our sole responsibility.

Notes

[1] The CIEP is a parallel corpus of 18 literary texts translated into numerous languages. For details see Talamo & Verkerk (2022).
[2] Glosses follow the *Leipzig glossing rules* (Bickel *et al.* 2015). A list of gloss meanings is provided at the end of this paper.
[3] While Stassen (1985, 2009) and Croft (2022), both qualitative studies, make a categorical balancing *vs* deranking distinction, we are in a position to quantify exactly how many balanced and deranked forms are attested in each sampled language. Therefore, in this study we conceive of this as a continuous feature. Accordingly, we consider that *sforzarsi* in (1a-c) could qualify for both strategies, as a reviewer points out.
[4] Only the results for variables (1), (2) and (5) are discussed below.
[5] <universaldependencies.org/lv/index.html>.
[6] Besides Russian, two more languages present higher balancing *vs* deranking rates for main clauses than for adverbial clauses: Czech (CIEP: 0.54 (main) *vs* 0.51 (adverbial), UD: 0.45 (main) *vs* 0.38 (adverbial)) and Slovak (CIEP: 0.45 (main) *vs* 0.34 (adverbial), UD: 0.41 (main) *vs* 0.35 (adverbial)). These three languages can be regarded as extreme examples of balancing.
[7] Lambda estimates for the various datasets: CIEP, main clauses – rate of null subjects 0.61; deranking/balancing 0.99; order of S, O and V 0.99; CIEP, adv. clauses – rate of null subjects 0.37; deranking/balancing 0.48; order of S, O and V 0.99; UD, main clauses – rate of null subjects 0.19; deranking/balancing 0.74; order of S, O and V 0.99; UD, adv clauses – rate of null subjects 0.86; deranking/balancing 0.51; order of S, O and V 0.94.
[8] For the UD results see Figure 3 in Appendix 5.
[9] This is a drawback of the verb score measure we introduced in this paper because it means that the verb score does not fully reflect the variation internal to these languages. Future studies should attempt to do this, perhaps by introducing verb scores not for whole languages, but for individual constructions.
[10] See Appendix 3 and Appendix 4 for plots of verb scores in the languages under study.
[11] For the UD results see Figure 5 in Appendix 5.
[12] For the UD results see Figure 7 in Appendix 5.
[13] There is only one exception: the credible interval for *model_comp_ciep_null_subj1*, where we correlate the frequency of null subjects in adverbial clauses with the proportion of null subjects in main clauses for CIEP. This credible interval is -0.05 ~ 3.13 and hence includes zero. The corresponding inverted model *model_comp_ciep_null_subj2* does show a credible interval that excludes zero.

Bibliographical References

Biagetti, Erica; Inglese, Guglielmo; Zanchi, Chiara & Luraghi, Silvia 2023. Reconstructing variation in Indo-European word order: A treebank-based quantitative study. *Language Dynamics and Change* 13,1. 1-34.

Bickel, Balthasar; Comrie, Bernard & Haspelmath, Martin 2015. Leipzig glossing rules. (Available online at <www.eva.mpg.de/lingua/resources/glossing-rules.php>, last access June 14th 2023.)

Bouckaert, Remco; Lemey, Philippe; Dunn, Michael; Greenhill, Simon J.; Alekseyenko, Alexander V.; Drummond, Alexei J.; Gray, Russell D.; Suchard, Marc A. & Atkinson, Quentin D. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337,6097. 957-60. <doi.org/10.1126/science.1219669>.

Bürkner, Paul-Christian 2017. Brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software* 80,1. 1-28. <doi.org/10.18637/jss.v080.i01>.

Bybee, Joan 2002. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In Bybee, Joan & Noonan, Michael (eds.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*. Amsterdam / Philadelphia: John Benjamins. 1-18.

Bybee, Joan; Perkins, Revere & Pagliuca, William 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago / London: University of Chicago Press.

Comrie, Bernard 1998. The Indo-European linguistic family: Genetic and typological perspectives. In Ramat, Anna G. & Ramat, Paolo (eds.), *The Indo-European languages*. London / New York: Routledge. 74-97.

Cristofaro, Sonia 2003. *Subordination*. Oxford: Oxford University Press.

Cristofaro, Sonia 2007. Deconstructing categories: Finiteness in a functional-typological perspective. In Nikolaeva, Irina (ed.), *Finiteness: Theoretical and empirical foundations*. Oxford: Oxford University Press. 91-114.

Croft, William 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press.

Delatour, Yvonne; Hennepin, Dominique; Léon-Dufour, Maylis & Teyssier, Brigitte 2004. *Nouvelle grammaire du français: Cours de civilisation française de la Sorbonne*. Paris: Éditions Hachette.

Diessel, Holger 2019. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.

Donaldson, Bruce 2008. *Dutch: A comprehensive grammar* (2nd ed.). London / New York: Routledge.

Douma, Jacob C. &. Weedon, James T. 2019. Analyzing continuous proportions in ecology and evolution: A practical introduction to Beta and Dirichlet regression. In Warton, David (ed.), *Methods in Ecology and Evolution* 10,9. 1412-30.

Dryer, Matthew S. 2005a. Order of subject, object and verb. In Haspelmath,

Martin; Dryer, Matthew S.; Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*. Oxford: Oxford University Press. 330-333.
Dryer, Matthew S. 2005b. Expression of pronominal subjects. In Haspelmath, Martin; Dryer, Matthew S.; Gil, David & Comrie, Bernard (eds.), *The world atlas of language structures*. Oxford: Oxford University Press. 410-413.
Friedrich, Paul 1975. *Proto-Indo-European syntax: The order of meaningful elements*. Washington: Journal of Indo-European Studies Monograph Series.
Gamkrelidze, Thomas V. & Ivanov, Vjačeslav V. 1995. *Indo-European and the Indo-Europeans: A reconstruction and historical analysis of a proto-language and a proto-culture*, Part I: *The text* (English translation by Nichols, Johanna). Berlin / New York: De Gruyter.
Gillies, William 2010. Scottish Gaelic. In Ball, Martin J. & Müller, Nicole (eds.), *The Celtic languages* (2nd ed.). London / New York: Routledge. 230-304.
Givón, Talmy 2001. *Syntax: An introduction*, Vol. 2. Amsterdam / Philadelphia: John Benjamins.
Halliday, Michael A. K. 2014. *Introduction to functional grammar* (4th ed.). London / New York: Routledge.
Haspelmath, Martin 2021. Towards standardization of morphosyntactic terminology for general linguistics. In Alfieri, Luca; Arcodia, Giorgio F. & Ramat, Paolo (eds.), *Linguistic categories, language description and linguistic typology*. Amsterdam / Philadelphia: John Benjamins. 35-58.
Hock, Hans H. 2021. *Principles of historical linguistics* (3rd ed.). Berlin / Boston: De Gruyter.
Hopper, Paul J. 1975. *The syntax of the simple sentence in Proto-Germanic*. The Hague / Paris: De Gruyter.
Jing, Yingqi; Widmer, Paul & Bickel, Balthasar 2023. Word order evolves at similar rates in main and subordinate clauses: Corpus-based evidence from Indo-European. *Diachronica* 40,4. 532-556.
Jucker, Andreas H. 1990. Word order changes in Early Middle English: Some evidence against the conservatism of subordinate clauses. *Studia Anglica Posnaniensia* 23. 31–42.
Kachru, Yamuna 2006. *Hindi*. Amsterdam / Philadelphia: John Benjamins.
Kalnača, Andra & Lokmane, Ilze 2021. *Latvian grammar*. Riga: University of Latvia Press.
Koptjevska-Tamm, Maria 1993. *Nominalizations*. London / New York: Routledge.
Krisch, Thomas 2009. On the 'syntax of silence' in Proto-Indo-European. In Hinterhölzl, Roland & Petrova, Svetlana (eds.), *New approaches to word order variation and change in Germanic*. Berlin: De Gruyter. 192-220.
Krisch, Thomas 2017. Proto-Indo-European syntax. In Kapović, Mate (ed.), *The Indo-European languages* (2nd ed.). London / New York: Routledge. 111-152.
Langacker, Ronald W. 1991. *Foundations of cognitive grammar*, Vol. 2: *Descriptive application*. Stanford: Stanford University Press.
Ledgeway, Adam 2021. V2 beyond borders: The *Histoire Ancienne jusqu'à César*. *Journal of Historical Syntax* 29,5. 1-65.
Lehmann, Winfred 1974. *Proto-Indo-European syntax*. Austin: University of Texas Press.
Levshina, Natalia; Namboodiripad, Savithry; Allassonnière-Tang, Marc; Kramer,

Matthew A.; Talamo, Luigi; Verkerk, Annemarie; Wilmoth, Sasha; Garrido Rodríguez, Gabriela; Gupton, Timothy; Kidd, Evan; Liu, Zoey; Naccarato, Chiara; Nordlinger, Rachel; Panova, Anastasia & Stoynova, Natalia 2023. Why we need a gradient approach to word order. *Linguistics* 61,4. 825-883

Luraghi, Silvia & Pinelli, Erica 2015. The loss of null subjects in Russian: What subordinate clauses can tell us. Paper presented at the *Slavic corpus linguistics: the historical dimension* conference, Arctic University of Norway, April 21-22.

MacAulay, Donald 1992. The Scottish Gaelic language. In MacAulay, Donald (ed.), *The Celtic languages*. Cambridge: Cambridge University Press. 137-248.

Malchukov, Andrej J. 2006. Constraining nominalization: Function/form competition. *Linguistics* 44,5. 973-1009.

Marneffe, Marie-Catherine; Manning, Christopher D.; Nivre, Joakim & Zeman, Daniel 2021. Universal dependencies. *Computational Linguistics* 47,2. 255-308.

Orme, David 2013. The caper package: Comparative analysis of phylogenetics and evolution in R. *R Package Volume* 5,2. 1-36.

Pagel, Mark 1999. Inferring the historical patterns of biological evolution. *Nature* 401. 877-884.

R Core Team 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <www.R-project.org>.

Revell, Liam J. 2012. Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3,2. 217-223. <doi.org/10.1111/j.2041-210X.2011.00169.x>.

Ringe, Don 2006. *A linguistic history of English*, Vol. 1: *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.

Shagal, Ksenia; Rudnev, Pavel & Volkova, Anna 2022. Multifunctionality and syncretism in non-finite forms: An introduction. *Folia Linguistica* 56,3. 529-557.

Stassen, Leon 1985. *Comparison and universal grammar*. Oxford: Basil Blackwell.

Stassen, Leon 2009. *Predicative possession*. Oxford: Oxford University Press.

Stockwell, Robert P. & Minkova, Donka 1991. Subordination and word order change in the history of English. In Kastovsky, Dieter (ed.), *Historical English syntax*. Berlin / New York: De Gruyter. 367-408.

Symonds, Matthew R. E. & Blomberg, Simon P. 2014. A primer on phylogenetic generalized least squares. In Garamszegi, László Z. (ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology*. Berlin / Heidelberg: Springer. 105-30.

Talamo, Luigi & Verkerk, Annemarie 2022. A new methodology for an old problem: A corpus-based typology of adnominal word order in European languages. *Italian Journal of Linguistics* 34,2. 171-226. <doi.org/10.26346/1120-2726-197>.

Vance, Barbara S. 1997. *Syntactic change in medieval French: Verb-second and null subjects*. Dordrecht: Springer.

Verkerk, Annemarie 2014. The evolutionary dynamics of motion event encoding. Doctoral dissertation, Radboud University Nijmegen.

Viti, Carlotta 2014. Reconstructing syntactic variation in Proto-Indo-European. *Indo-European Linguistics* 2. 73-111.
Von Polenz, Peter 2009. *Geschichte der deutschen Sprache* (10th ed.). Berlin / New York: De Gruyter.
Wälchli, Bernhard 2007. Advantages and disadvantages of using parallel texts in typological investigations. *Sprachtypologie und Universalienforschung* 60,2. 118-134.
Willis, David 2010. Old and Middle Welsh. In MacAulay, Donald (ed.), *The Celtic languages*. Cambridge: Cambridge University Press. 117-160.
Winter, Bodo & Bürkner, Paul-Christian 2021. Poisson regression for linguists: A tutorial introduction to modelling count data with brms. *Language and Linguistics Compass* 15,11.
Yousef, Saeed 2018. *Persian: A comprehensive grammar*. London / New York: Routledge.

Appendix

The appendix files for this paper can be accessed via several anonymous links to the Zenodo repository:

Appendix 1: A .docx file with a summary of the qualitative data in Appendix 2 (https://zenodo.org/records/10480308);
Appendix 2: An unprocessed .xlsx document with the results of quantitative and qualitative analysis on the 45-language sample (https://zenodo.org/records/10480311);
Appendix 3: A .pdf file that plots verb scores (1 to 3) in main and adverbial clauses in the sample languages covered by the CIEP (https://zenodo.org/records/10079582);
Appendix 4: A .pdf file that plots verb scores (1 to 3) in main and adverbial clauses in the sample languages covered by UD (https://zenodo.org/records/10084915);
Appendix 5: Bar plots for the null subject, verb score and balancing *vs* deranking variables in the languages covered by UD and the CIEP (https://zenodo.org/records/10604225);
Appendix 6: A compilation of all additional materials used for this study (https://zenodo.org/records/10107866).
Appendix 7: Data, code, and results of the phylogenetic analyses (https://zenodo.org/records/10476156).