
Daltonian atoms, Steiner's curve and Voltaic sparks – eponymous terms in a diachronic corpus of English scientific writing

Katrin Menzel

Department of Language Science and Technology, Saarland University

k.menzel@mx.uni-saarland.de

This poster has a focus on eponymous academic and scientific terms in the first 200 years of the Royal Society Corpus (RSC, ca. 9,800 English scientific journal articles from the Royal Society of London, 1665-1869, cf. Kermes et al. 2016). It is annotated at different linguistic levels and provides a number of query and visualization options. Various types of metadata are encoded for each text, e.g. text topics / academic disciplines. This dataset contains a variety of eponymous terms named after English, foreign and classical scholars and inventors. The poster presents the results of a corpus study on eponymous terms with common structural features such as multiword terms with similar part of speech patterns (e.g. adjective + noun constructions such as *Newtonian telescope*) and terms with shared morphological elements, e.g. those that contain possessive markers (e.g. *Steiner's curve*) or identical derivational affixes (e.g. *Bezoutic*, *Hippocratic*). Queries have been developed to automatically retrieve these terms from the corpus and the results were manually filtered afterwards.

There are, for instance, around 3,000 eponymous adjective + noun constructions derived from ca. 160 different names of scholars. Some are used as titles for institutions or academic events, positions and honours (e.g. *Plumian Professor*, *Jacksonian prize*) while most refer to scientific concepts and discoveries (e.g. *Daltonian atoms*, *Voltaic sparks*). The terms show specific distribution patterns within and across documents. It can be observed how such terms have developed when English became established as a language of science and scholarship and what role they played throughout the following centuries. The analysis of these terms also contributes to reconstructing cultural aspects and language contacts in various scientific fields and time periods. Additionally, the results can be used to complement English lexicographical resources for specialized languages (cf. also Menzel 2018) and they contribute to a growing understanding of diachronic and cross-linguistic aspects of term formation processes.

References: • Kermes, Hannah, Stefania Degaetano-Ortlieb, Ashraf Khamis, Jörg Knappen & Elke Teich (2016). The Royal Society Corpus: From Uncharted Data to Corpus. *Proceedings of LREC 2016*. Portoroz, Slovenia, p. 1928-1931. • Menzel, Katrin (2018). Using diachronic corpora of scientific journal articles for complementing English corpus-based dictionaries and lexicographical resources for specialized languages. *Proceedings of Euralex2018*, 17-21 July 2018 in Ljubljana, Slovenia, p. 363-372. • The RSC is available for free download and online query via the CLARIN-D center at Saarland University under the persistent identifier: <http://hdl.handle.net/11858/00-246C-0000-0023-8D1C-0>.