

Introduction: Paronyms revisited empirically

This volume outlines theoretical and practical approaches to the linguistically largely neglected phenomenon of paronymy, which causes uncertainties among native speakers and foreign language learners alike. This chapter on “preliminaries” is a general introduction to “Paronymie im deutschen Sprachgebrauch” (Paronyms in German language use). It provides a guide to the papers in this special issue and explains their objectives.

Chapter 1 is a general discussion of easily confused words, so-called paronyms, and it explains how to deal with lexical confusion from a usage-based perspective. Paronyms are a specific type of confusables covering a heterogeneous group of two or three lexical items, all of which share similarities in spelling, sound and/or sometimes, meaning (e. g. *konzeptuell/konzeptionell*, *conceptual/conceptional*; *Methode/Methodik/Methodologie*, *method/methodology*). The causes of communicative uncertainties vary as individual items exhibit different semantic-pragmatic traits in context. Hence, it is important to investigate paronyms in their lexical-semantic environments, i. e. in authentic communication, where their meaning is exposed by their co-occurring elements. Usage-based investigations reveal the contextual circumstances, influences and impacts of confusion. At the same time, lexical rivalry and semantic adjustments can be identified. Therefore, it has been argued that paronymy is a flexible relation and a potential trigger of semantic change. To study the semantic behaviour of confusable terms empirically, a new project was launched, its aim is not only to provide a theoretical framework but also to present its findings in a dynamic, multifunctional e-dictionary. Overall, this introductory chapter depicts some of the challenges we encountered in terms of semantic theory, practical corpus methods and the editorial documentation of paronyms.

In the second chapter, we briefly review the project’s lexicological research results. As paronymy has been of no interest in corpus-semantic or cognitive fields, this project follows empirical approaches, employing contemporary corpus-based data to uncover contextual patterns of German paronyms in use. As such, it is the first attempt to undertake a large corpus-guided investigation of commonly confused pairs. An essential part of our research comprises its theoretical grounding and the embedding of the phenomenon in a larger semantic framework. This also includes the development of a new classification model that takes conceptual and cognitive facets, communication functions, context-bound aspects and discourse features into account instead of allocating confusing pairs to formally distinct classes based on morphological features, as was traditionally the case (cf. Lăzărescu 1999). Such new approaches involve looking at commonly confused terms descriptively (with the potential of being the impetus for semantic change) instead of prescriptively (as a form of lexical-semantic “derailment”). Paronymy only exists under specific conditions where categories of linguistic and non-linguistic knowledge and experience interfere. Therefore, it is the principles of conceptualisation, ontological adaptation, priming and metalinguistic reference that require special attention. The detection of both lexical and extra-linguistic knowledge and their specific way of interacting in the mental lexicon play a vital role in deriving a theory and definition of this complex phenomenon.

Chapter 3 highlights the impact of corpus-linguistic approaches on the exploration of language, descriptive views and dynamic paronym behaviour. Confusables can develop dynamically under the influence of semantic vagueness and can be subject to essential conceptual adjustments. Cognitive models are therefore in a position to offer more categories that account for situations of lexical confusion. Moreover, cognitive models consider lexical relations as conceptual relations. Our belief is that this can be supported by corpus-guided findings relating to lexical (syntagmatic) structures and collocational profiles, thus offering the necessary metalinguistic frame. Usage-based research reveals the discrepancies inherent in traditional prescriptive studies which prefer to use the normative classes of a linguistic system (see Storjohann/Schnörch 2017).

An effective analysis of paronyms in use requires the detection of patterns, lexical preferences and communicative constellations. For this purpose, we employed innovative explorative and contrastive corpus procedures to facilitate a comparison of meaning and use of pairs of items. In chapter 4, the compilation of the corpus material and the choice of corpus methods are delineated. In a first step, we set up a large, publically accessible corpus comprising contemporary texts, exemplifying German in written public discourse and containing 2.38 billion tokens. The project compiled the first systematic and corpus-driven lemma list of commonly confused words in German by searching for items following certain patterns and string sequences, e. g. *-haft/-lich* (*jugendhaft/jugendlich* = *boyish/youthful*), *-isch/-onisch* (*elektrisch/elektronisch* = *electric/electronic*). Using a triple-blind procedure, these were then edited manually according to specific parameters (Schnörch 2015). Subsequently, every pair underwent a corpus investigation involving an analysis of their collocational behaviour (Belica 1995) and a comparison with the profiles of their corresponding partners, a procedure referred to as *Contrasting-Near-Synonyms* (Belica 2006). Both methods guarantee the detection of semantic nuances and degrees of similarity. The outcomes of these investigations are documented in a new dictionary following an unprecedented concept and new lexicographic practices.

Chapter 5 looks more closely at users and their need for a reference guide to paronyms. As can be seen on learners' platforms and in internet forums, the interest in commonly confused words is large. Speakers give advice, recommend popular dictionaries or talk about their linguistic experience. They share their concern for linguistic "correctness" with great passion and controversy. An analysis of these forums helped the project to learn about potential users, their knowledge/intuitions, their consultation behaviour and, above all, their needs (cf. Storrer 2013). The results had a noticeable impact on the concept of the new dictionary, both with regard to content and presentation.

In chapter 6, we explain how the examination of forums has strongly influenced the details, structure and design of the new reference work. Generally, this section outlines the lexicographic concept behind the new online dictionary "*Paronyme – Dynamisch im Kontrast*" by accentuating two aspects. Firstly, we explain the cognitively driven ideas behind each meaning description. Secondly, we emphasise its dynamic e-lexicographic structure. The new dictionary aims to reconcile semantic theory and lexicographic practice (cf. Gries 2006) and is a step towards Cognitive Lexicography. The key elements of each entry are ontological references paired up with contextual patterns and examples. They link linguistic and encyclopaedic information, which users can study at different levels from different perspectives. In order to meet this demand, both text-technological and hypertextual possibilities needed to be exploited more fully, offering dynamic searches and individual

choices. For example, users can sort contextual instances by frequency, similarities or differences. Details can be hidden or expanded. Overviews and detailed views can be consulted, with the specific details of an entry being generated on demand. The user decides what to look for to what depth and what to focus on in which order. The dictionary breaks with established traditions and moves away from a static towards a dynamic and multi-functional resource.

The last chapter shines a light on future prospects and open research questions. These will go beyond the analysis of written and synchronic texts and beyond the investigation of public discourse. The focus of the years to come will be on spoken data, diachronic corpora, technical language and German for specific purposes. In the future, the reference guide will grow with new sections and more entries.

References

- Belica, Cyril (1995): *Statistische Kollokationsanalyse und -clustering. Korpuslinguistische Analysemethode*. Mannheim: Institut für Deutsche Sprache. <http://corpora.ids-mannheim.de> (accessed on: 13.12.2018).
- Belica, Cyril (2006): *Modellierung semantischer Nähe: Kontrastierung von nahen Synonymen. Korpusanalytische Methode*. Mannheim: Institut für Deutsche Sprache. <http://corpora.ids-mannheim.de/ccdb> (accessed on: 13.12.2018).
- Gries, Stefan Thomas (2006): *Corpus-based methods and cognitive semantics: The many senses of to run*. In: Gries, Stefan Thomas/Stefanowitsch, Anatol (eds.): *Corpora in cognitive linguistics corpus-based approaches to syntax and lexis*. Berlin/New York: De Gruyter, pp. 57–100.
- Lăzărescu, Ioan (1999): *Die Paronymie als lexikalisches Phänomen und die Paronomasie als Stilfigur im Deutschen*. Bukarest: Editura Anima.
- Schnörch, Ulrich (2015): *Wie viele Paronympaare gibt es eigentlich? Das Zusammenspiel aus korpuslinguistischen und redaktionellen Verfahren zur Ermittlung einer Paronymstichwortliste*. In: *Sprachreport* 31, 4, pp. 16–26.
- Storjohann, Petra/Schnörch, Ulrich (2017): *Sprachlicher Wandel der Gegenwart und seine Dokumentation in einem Wörterbuch*. In: Vachková, Marie/Šemelík, Martin/Kloudová, Vera (eds.): *Themenheft Lexikographie, Germanistica Pragensia, AUC Philologica* 4, pp. 133–172.
- Storrer, Angelika (2013): *Representing dictionaries in hypertextual form*. In: Gouws, Rufus et al. (eds.): *Dictionaries. An international encyclopedia of lexicography. Supplementary Volume: Recent developments with focus on electronic and computational lexicography*. Berlin/Boston: De Gruyter, pp. 1244–1253.