

Prof. Dr. Martin Hilpert
(Université de Neuchâtel)

Does *fridge* mean the same thing as *refrigerator*? Distributional semantics and the meaning of English clippings

In this talk, I will discuss how corpus data and methods of distributional semantics can be used to study English clippings such as *dorm* (< *dormitory*), *memo* (< *memorandum*), or *quake* (< *earth-quake*). I investigate whether systematic meaning differences between clippings and their source words can be detected. The analysis is based on a sample of 50 English clippings and their source word counterparts. Pairs such as *cardio-cardiovascular*, *chemo-chemotherapy*, and *intro-introduction* are analyzed in terms of their collocational behavior. Each of the clippings is represented by a concordance of 100 examples in context that were gathered from the Corpus of Contemporary American English (Davies 2008). I compare clippings and their source words both at the aggregate level, and in terms of comparisons between individual clippings and their source words.

The aggregate comparisons reveal general distributional asymmetries that suggest a difference relating to involved vs. informational text production (Biber 1988). Clippings have a relatively greater tendency to appear in texts with contextual elements such as first or second person pronouns, demonstratives, or contractions. Clippings thus appear to be preferred in contexts in which there is substantial common ground between speaker and hearer, which aligns with the notion that clippings signal familiarity with the ideas that are conveyed (Wierzbicka 1984; Katamba 2005). For the individual comparisons between clippings and their source words, I draw on the distributional semantic method of token-based semantic vector spaces (Hilpert and Correia Saavedra 2020). The method allows us to pinpoint aspects of meaning that are specifically associated with a clipping, rather than its source word, and vice versa, while also revealing how their respective meanings overlap. The findings show that clippings such as *chemo* and *cardio* are semantically distinct from their source words, but I also document cases such as *fridge*, in which the collocational profile of the clipping is indistinguishable from that of the source word.

I interpret these findings against the theoretical background of Construction Grammar and specifically the Principle of No Synonymy. Following Levshina and Lorenz (2022), I argue that speakers' choices between clippings and their source words are motivated by meaning rather than by efficiency. For cases in which clippings overlap semantically with their source words, it can be shown that there are facets of meaning that are preferentially or even exclusively expressed by one of the two alternatives. I take these results as an indication that the Principle of No Synonymy holds up surprisingly well with regard to English clippings.