

# English clippings and their source words: A constructional perspective

Martin Hilpert, David Correia Saavedra & Jennifer Rains  
Université de Neuchâtel, martin.hilpert@unine.ch

**Keywords:** clipping, morphology, meaning, distributional semantics, constructions

This paper uses corpus data and methods of distributional semantics in order to study English clippings such as dorm (< dormitory), memo (< memorandum), or quake (< earthquake). We investigate whether systematic meaning differences between clippings and their source words can be detected. Alber and Lappe (2012: 314) observe that semantic questions have received relatively little attention in the study of clippings, and they remark that systematic studies of meaning in truncatory processes are virtually absent. The present paper tries to address that gap.

Our 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). We 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, we 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. Our findings show that clippings such as chemo and cardio are semantically distinct from their source words, but we also document cases such as fridge, in which the collocational profile of the clipping is indistinguishable from that of the source word.

We interpret these findings against the theoretical background of Construction Grammar and specifically the Principle of No Synonymy. Following Levshina and Lorenz (2022), we 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, we documented facets of meaning that are preferentially or even exclusively expressed by one of the two alternatives. We take these results as an indication that the Principle of No Synonymy holds up surprisingly well with regard to English clippings.

## References

- Alber, Birgit, and Sabine Arndt-Lappe. 2012. Templatic and subtractive truncation. In J. Trommer (ed.), *The Phonology and Morphology of Exponence – the State of the Art*, 289-325. Oxford: Oxford University Press.
- Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Davies, Mark. 2008. The Corpus of Contemporary American English (COCA). Available online at <https://www.english-corpora.org/coca/>.
- Hilpert, Martin and Correia Saavedra, David. 2020. Using token-based semantic vector spaces for corpus-linguistic analyses: From practical applications to tests of theoretical claims. *Corpus Linguistics and Linguistic Theory* 16/2, 393-424. <https://doi.org/10.1515/cllt-2017-0009>
- Katamba, Francis. 2005. *English Words*. 2nd edition. New York: Routledge.
- Levshina, Natalia and David Lorenz. 2022. Communicative efficiency and the Principle of No Synonymy: Predictability effects and the variation of *want to* and *wanna*. *Language and Cognition* 14/2, 249-274. <https://doi:10.1017/langcog.2022.7>
- Wierzbicka, Anna. 1984. Diminutives and depreciatives: Semantic representation for derivational categories. *Quaderni di semantica* 5(1), 123-130.