

‘Over reference’: a comparative study on German prefix-verbs

Introduction In this paper, we approach the integration of distributional and conceptual perspectives on word meaning into referential approaches. We report on a case study on the meaning of German prefix-verb constructions with *über* (‘over’). Consider the data in (1).

- | | | | | |
|-----|----|---|----|--|
| (1) | a. | <i>über-bringen</i>
over.PREFIX-bring
‘to deliver’ (TRANSFER) | c. | <i>über-springen</i>
over.PREFIX-skip
‘to skip over’ (ACROSS) |
| | b. | <i>über-kleben</i>
over.PREFIX-paste
‘to paste over’ (APPLICATION) | d. | <i>über-be-werten</i>
over.PREFIX-PREFIX-value
‘to overvalue’ (SCALE) |

From the perspective of truth-conditions and reference, the denotation of prepositions like *über* has been modelled as a directed set of vectors in a 3D model of space [Zwarts and Winter, 2000]. Consequently, when *über* is combined with a verb, the resulting denotation can be identified as the movement of the direct object of the verb along the direction (from A to B as in (1-a)) or into the region (the ‘above’-region of the direct object as in (1-b)) identified by *über*. The spatial contribution in *über* constructions can be further refined by taking into account the conceptual interpretation of the *über*+verb construction, where we adopt the approach to the conceptualization of spatial prefix and particle verbs of Pross and Roßdeutscher [2015] to *über*. We distinguish four conceptual classes based on their syntactic and grammatical behaviour with respect to case assignment and the licensing of argument structure and prepositional phrases: TRANSFER of an object from A to B as in (1-a), APPLICATION of an object to another object as in (1-b), movement ACROSS some obstacle as in (1-c) and exceeding a certain threshold on a SCALE that the base verb provides as in (1-d).

Experiment We conducted an experiment in which we computed distributional [Miller and Charles, 1991] vectors of 40 base verbs and their *über*-derivations (10 for each class), using a state-of-the-art model [Mikolov et al., 2013]. Then we applied a hierarchical clustering algorithm (cosine similarity, average linkage). Rather than to reproduce the gold standard, our goal was to learn about the relationship between distributional and conceptual classifications and thus about what kind and amount of conceptual information distributional representations encode. For our qualitative analysis, we characterized each base and derived verb by its 10 nearest neighbors (e.g., the most related words to *rennen* and *überrennen*, respectively). The nearest neighbors replace the typically reported “most salient dimensions” of the distributional vector, since these are unanalyzable in the Mikolov et al. approach.

Analysis Since we do not have the space to discuss the clustering results, this abstract concentrates on a small number of exemplary analyses (a final presentation would cover both). As the first example, consider the nearest neighbors for the APPLICATION verb *kleben* in (2-a) and *überkleben* in (2-b), respectively.

- | | | |
|-----|----|---|
| (2) | a. | <i>kleben.V</i> : aufkleben.V ausschneiden.V festkleben.V bekleben.V
to glue to affix sth. to cut sth. out to tape sth. to stick sth. all over sth. |
| | b. | <i>überkleben.V</i> : Aufkleber.N Aufschrift.N kleben.V Schriftzug.N
to paste over sticker label to glue logo |

The neighbors in (2) correspond fairly well with standard interpretations of Distributional Semantics like the one by Baroni et al. [2014]: they provide a set of linguistic expressions that are conceptually similar, connected by narrative chains or derivationally related. The neighbors also link nicely to our conceptual classification: as expected for the APPLICATION class, the direct object which is

pasted over (e.g. a sticker or a logo) figures prominently in the distributional characterization of the *über*-construction, but not in that of the base verb. This good fit might seem disappointing, since the distributional information does not seem to provide the novel insights for the conceptual modelling of word meaning that have been hoped for and which, for example, form the motivation for this workshop. Reassuringly, *überkleben* is not representative for all *über*-constructions. For a considerable number of instances, we encountered aspects of meaning that find expression only in the distributional characterization. Consider, as an example, the verb *überrennen* (3) and neighbors of the base and target verb in (4-a) and (4-b), respectively.

- (3) *über-rennen* over.PREFIX-run ‘to overrun’
- (4) a. **rennen.V:** schnappen.V springen.V zurennen.V hüpfen.V
to run to snatch to jump to run towards to hop
- b. **überrennen.V:** Horde.N belagern.V stürmen.V Hunne.N erobern.V
to overrun horde to besiege to storm Hun to conquer

We interpret the neighbors of *überrennen* as expressing a clear negative *evaluation*, describing an event that is perceived as uncontrolled or uncontrollable (an “EXCESS” reading). At the same time, nothing in the referential-conceptual compositional account of constructing the meaning of *überrennen* from *rennen* and *über* indicates that such a reading might arise. Thus, in this case, there is a tension between the compositional account (which describes what meaning we would *expect*) and the distributional account (which describes how the word is actually *used*). It seems fruitful to consider this situation from the point of view of a language user seeking to describe a certain situation. They are faced with the task of choosing among different possible conceptualizations of the situation which they can use as the basis of their verbalization. For *überrennen*, the distributional characterization captures two intertwined consequences of this conceptual choice. The first one is the emergence of a metaphorical shift for *über* to describe situations over which control is lost or can be lost from a salient point of view. The second one is the emergence of the negative evaluative component of *über*. Therefore, what the distributional analysis contributes is that *über* functions like an expressive element that gains the development of a certain disposition over the maximum of what could be controlled or what could be measured on a scale. Both facets of the distributional semantics of *über* are a direct reflection of the kind of data that is used for the construction of distributional model. Distributional models are derived from texts written by authors with a certain communicative intention. A communicative intention, however, is often not explicitly realized linguistically but appears in the form of proffered content, implicatures and associative collocations and thus constitutes part of the discourse strategy of the author [van Dijk and Kintsch, 1983].

Conclusion The use-based meaning components that specifically distributional representations identify with *über* fall outside the usual scope of conceptual or referential semantics and constitutes an additional dimension of word meaning. We believe that this observation is highly relevant for the three approaches of meaning that the workshop addresses. From a formal semantic point of view, the prefixation of verbs as a discourse strategy may be a decisive factor in the analysis of the productivity of prefixation and the border area between compositional and idiosyncratic word meaning. From a conceptual point of view, the metaphorical and expressive dimension of the meaning of verb prefixes like *über* may indicate a conceptual space the formation and organization of which is not only a matter of cognition but rests on the use of language as a strategic tool to realize communicative intentions. From a computational point of view, our case study suggests that the interpretation of distributional representations is more complex than is usually assumed in approaches that envisage a close parallel of the composition of truth-conditions and the composition of distributional representations [Herbelot and Copestake, 2013].

References

- Marco Baroni, Raffaella Bernardi, and Roberto Zamparelli. Frege in space: A program of compositional distributional semantics. *LiLT*, 9:241 – 346, 2014.
- Aurelie Herbelot and Ann Copestake. Lexicalized Compositionality. Unpublished manuscript, 2013.
- Tomas Mikolov and Ilya Sutskever and Kai Chen and Greg S Corrado and Jeff Dean. Distributed representations of words and phrases and their compositionality. *Proceedings of NIPS*, 3111–3119, 2013.
- George A. Miller and Walter G. Charles. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6:1 – 28, 1991.
- Tillmann Pross and Antje Roßdeutscher. Measuring out the relation between conceptual structures and truth-conditional semantics. In Kata Balogh, editor, *Selected papers from the Workshop "Bridging Formal and Conceptual Semantics"*, Düsseldorf, 2015. Düsseldorf University Press.
- T. A. van Dijk and W. Kintsch. *Strategies of Discourse Comprehension*. Academic Press, Orlando, Florida, 1983.
- Joost Zwarts and Yoad Winter. Vector space semantics: A model-theoretic analysis of locative prepositions. *Journal of Logic, Language, and Information*, 9:169 – 211, 2000.