Building and exploiting resources for derivational morphology: Data-driven and theory-driven approaches

Sebastian Padó
(University of Stuttgart

A central characteristic of derivational morphology is its (semi)regularity, i.e., the existence of regular patterns which are however subject to exceptions, gaps and subregularities. In the first part of my talk, I will argue that the creation of derivational resources can profit from the combination of theory-driven and data-driven methods, and will present evidence for this claim from the construction of DErivBase, a derivational dictionary for German, which combines hand written rules, distributional data, and graph theoretic methods (Zeller et al., 2013, 2014; Papay et al., 2017). In the second part, I will move to the exploitation of such resources and discuss the tension between how the semantic effects of derivation are captured on the theoretical side (transparency, specificity) and how they are captured on the distributional side (Padó et al., 2016; Lapesa et al., 2017; Varvara et al., 2021).

References

Gabriella Lapesa, Sebastian Padó, Tillmann Pross, and Antje Rossdeutscher. 2017. Are doggies cuter than dogs? emotional valence and concreteness in german derivational morphology. In *Proceedings of IWCS*. Montpellier, France.

Sebastian Padó, Aurélie Herbelot, Max Kisselew, and Jan Šnajder. 2016. Predictability of distributional semantics in derivational word formation. In *Proceedings of COLING*. Osaka, Japan, pages 1285–1296.

Sean Papay, Gabriella Lapesa, and Sebastian Padó. 2017. Evaluating and improving a derivational lexicon with graph-theoretical methods. In *Proceedings of DeriMo*. Milan, Italy.

Rossella Varvara, Gabriella Lapesa, and Sebastian Padó. 2021. Grounding semantic transparency in context: A distributional semantic study on german event nominalizations. *Morphology* .

Britta D. Zeller, Padó Sebastian, and Jan Šnajder. 2014. Towards semantic validation of a derivational lexicon. In *Proceedings of COLING*. Dublin, Ireland, pages 1728–1739.

Britta D. Zeller, Jan Šnajder, and Padó Sebastian. 2013. Derivbase: Inducing and evaluating a derivational morphology resource for german. In *Proceedings of ACL*. Sofia, Bulgaria, pages 1201–1211.