

Modelling constructional change with distributional semantics

Florent Perek

University of Birmingham

In the nascent field of diachronic construction grammar (DCxG, Traugott & Trousdale 2013), language change is analysed with reference to the idea that grammar is best described as a structured inventory of form-meaning pairs, aka constructions (Goldberg 1995, 2006). For instance, the English *way*-construction (e.g., *They hacked their way through the jungle*) pairs the form [NP V *one's way* PP] with the notion that the subject referent moves along some path. In DCxG, language change consists either in the creation of new form-meaning pairs (constructionalisation), or in changes in aspects of the form or meaning of existing constructions (constructional change).

As DCxG is a usage-based approach, constructional change is analysed by looking at usage data in historical corpora, and two aspects are especially focused on when examining the diachronic development of constructions: (i) their productivity, i.e., the range of lexical items that can occur in their slots, and (ii) their schematicity, i.e., the degree of openness of a construction's semantics and of its slots. These two aspects are commonly thought to be interrelated: variation in productivity often corresponds to variation in schematicity, but this crucially depends on how attested lexical items are spread in semantic space (cf. Suttle & Goldberg 2011). Therefore, a proper characterization of constructional change requires a semantic description of the lexical items occurring in a given construction, especially in terms of how these items are related.

This talk describes a research programme aimed at applying distributional semantics to the study of constructional change. Distributional semantics offers a way to capture the semantic relatedness of lexical items through their frequent collocates in a large corpus, which eschews the need for manual intervention to characterize lexical meaning. Two methods drawing on distributional semantic representations to track changes in the semantic domain of constructions over time will be discussed. The first method consists in plotting the lexical distribution of a slot of a construction in different time periods, so as to visualize trends of change and identify semantic classes of items that join or leave the distribution of the construction, which would be, at best, harder to discern on the sole basis of semantic intuitions. The second method aims at identifying stages, i.e., times of relative stability vs. turning points, in the semantic history of constructions, by using a customized version of variability-based neighbour clustering (Gries & Hilpert 2008). Several case studies of recent change (i.e., over the last 200 years) in English verbal and nominal constructions are presented to illustrate these methods and show how their results can be interpreted in terms of changes in the grammatical representation of constructions.

References

- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Gries, Stefan & Martin Hilpert. 2008. The Identification of Stages in Diachronic Data: Variability-based Neighbor Clustering. *Corpora* 3. 59–81.
- Suttle, Laura & Adele E. Goldberg. 2011. The partial productivity of constructions as induction. *Linguistics* 49(6). 1237–1269.
- Traugott, Elizabeth C. & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.