Exploring individual variation in constructional schematicity with mixed models

Svetlana Vetchinnikova, University of Helsinki, svetlana.vetchinnikova@helsinki.fi

This paper takes a usage-based constructionist approach to individual variation and assumes that: 1) language is an inventory of constructions at different levels of schematicity and 2) personal construct-i-cons vary as a function of usage. With repeated use, constructions move along the continuum from more schematic to more lexically specified through the process of chunking and can undergo reanalysis. A well-known example is *I don't know,* which is phonologically reduced and conveys an additional pragmatic function of mitigated disagreement when used as a unit (Bybee & Scheibman 1999). Reduction is typical property of chunks and can serve as a diagnostic of a change in the internal structure of an expression. Given individuality of language usage, to what extent do different instantiations of constructions vary in schematicity in personal construct-i-cons?

As a case study, I used a 1.75-million-word corpus of comments posted by one blogger over 8 years. As a dependent variable, I chose the alternation between contracted (reduced) and uncontracted forms of *it is* hypothesizing that it was more likely to be reduced in chunks. *It is* occurs in a wide variety of syntactic structures including clefts, progressives, passives, extraposed and copular structures: altogether 10,000 corpus occurrences of *it is/it's* were categorized into 14 frequent constructions. For each lexical item filling the open slot in each construction, I used delta P statistic to compute the degree to which a lexical item associates with a construction and the degree to which a construction associates with a lexical item (Gries & Ellis 2015). In addition, I calculated normalized entropy for each construction as a measure of dispersion (Gries & Ellis 2015; Gries 2021). In a logistic regression model predicting the reduced form, I included the two delta P measures, normalized entropy, possible priming and temporal order of occurrence as fixed effects and lexically specified instantiations of constructions as random effects ($R^2_{\text{conditional}} = 0.21$, $R^2_{\text{marginal}} = 0.33$, AUC = 0.81). Variance in random intercepts showed variation of lexically specified instantiations in schematicity while variance in random slopes for the effect of temporal order showed change in schematicity over time.

Keywords: chunking, reduction, individual variation, constructions, mixed models

References

Bybee, Joan & Joanne Scheibman. 1999. The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37(4). 575–596. https://doi.org/10.1515/ling.37.4.575.

Gries, Stefan Th. 2021. *Statistics for Linguistics with R: A Practical Introduction*. *Statistics for Linguistics with R*. De Gruyter Mouton. https://doi.org/10.1515/9783110718256.

Gries, Stefan Th. & Nick C. Ellis. 2015. Statistical measures for usage-based linguistics. *Language Learning* 65(S1). 228–255. https://doi.org/10.1111/lang.12119.