
Correlations in acceptability as a window on the grammar

Vsevolod Kapatsinski

University of Oregon

vkapatsi@uoregon.edu

Grammatical theories often differ in which constructs correspond to the same structure. For example, in morphology, Bybee's (1985) product-oriented schemas group together forms that share a paradigm cell and some recurrent formal characteristics regardless of how they are derived from base forms. Thus, the English past tense forms *hit* and *bit* instantiate the same schema even though the base of *hit* is *hit* while the base of *bit* is *bite*. In contrast, under a rule-based theory, these two forms instantiate two different rules, 'do nothing' and $a\bar{1} \rightarrow \bar{1}$. Similarly, in syntax, researchers disagree on whether constructional alternations share the same underlying structure (deep structure in generative grammar, or a 'mother construction' in construction grammar; e.g., Diessel 2023).

How can we tell? If we assume that participants differ in the acceptability they assign to a particular construction / schema / underlying structure, a window on this question is provided by examining individual differences in judgments.

Specifically, if two constructs (sentences, words, etc.) exemplify the same structure, the judged acceptabilities of the two constructs should correlate positively across participants (Kapatsinski 2012). While this may also be observed with production probabilities, acceptability judgments are a better way to assess this prediction because probabilities must sum to 1. For example, if two constructs are alternative outputs for the same input (e.g., *burned* vs. *burnt*), it is difficult to observe whether they share structure with production data, as speakers who choose to produce *burned* more will necessarily produce *burnt* less because the two compete for the same probability mass. In judgment, this is not a foregone conclusion.

If we treat each participant as a dimension, and the participant's judgment of a construct as that construct's score on that dimension, we can situate the constructs relative to each other in a multidimensional similarity space. This allows us to cluster the constructs, e.g., using hierarchical clustering (Kapatsinski 2012). The clustering solution then allows us to identify which constructs might correspond to the same underlying structure.

References: Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins. • Diessel, H. (2023). *The constructionicon: Taxonomies and networks*. Cambridge University Press. • Kapatsinski, V. (2012). What statistics do learners track? Rules, constraints and schemas in (artificial) grammar learning. In S. Th. Gries & D. Divjak (eds.), *Frequency effects in language learning and processing*. Berlin: de Gruyter, 53–82.