

The productivity of constructions: beyond type frequency

Florent Perek¹ & Adele Goldberg²

¹University of Birmingham, f.b.perek@bham.ac.uk ²Princeton University

Keywords: productivity, type frequency, usage-based model, constructions, schematization, artificial language learning

When do speakers use constructions in creative ways? Studies on this topic typically designate type frequency, i.e., the number of different items attested in a construction, as the driving factor of syntactic productivity (e.g., Bybee & Thompson 1997). There is intuitive appeal to his idea: speakers should be more confident that a construction can be extended to new items if they have witnessed it with many items than if they have seen it used with only a few. A less common but equally sound account instead posits *variability* as the critical factor, i.e., the diversity (especially semantic) of the items witnessed in a construction. On this view, type frequency is a mere proxy for variability. Since type frequency and variability are typically correlated in natural data (more diversity requires more types, and typically more types implies greater diversity), the two factors have proven difficult to tease apart.

In this study, we attempt to separately test the role of type frequency and variability on syntactic productivity, using an artificial language learning experiment (e.g., Perek & Goldberg 2015, 2017), which allows us more control over the input provided to language users. Over two sessions, participants are exposed to two nonce constructions, “Verb Agent Patient-po” (e.g. *Mooped the cat the monkey-po*) and “Verb Patient-po Agent” (e.g. *Glimmed the rabbit-po the wolf*), through video clips paired with sentence descriptions. Each construction is attested with two different sets of nonce verbs with transitive meanings. In one condition, one construction has higher type frequency than the other, but both have low variability (i.e., each is attested with highly similar variants of the same verb meaning). In another condition, both constructions are matched in type frequency but one of them has higher variability than the other (i.e., it is attested with a more diverse range of verb meanings). After exposure in one or the other condition, participants are asked to produce new sentences in the artificial language with either a verb attested in the input, a new verb similar in meaning to those in the distributions of the constructions, or a new verb that is semantically distinct from witnessed verbs.

Participants successfully learn the distribution of each construction in both conditions, in that they use attested and similar verbs in the relevant construction. Critically, they treat the higher variability *or* the higher type frequency construction as more productive; i.e., participants strongly prefer this construction over the low variability or low type frequency construction when the verb has a novel meaning. In other words, we find independent effects of type frequency *and* semantic variability on syntactic productivity, when the other factor is held constant. While the effect of type frequency replicates earlier studies, the fact that higher variability alone can drive productivity independently of type frequency in an experimental setting is a novel finding. We interpret our results in a schema-based model of grammatical generalisations, in which schematisation can take more than one route.

References

- Bybee, Joan & Sandra Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistics Society* 23. 65-85.
- Perek, Florent & Adele Goldberg. 2015. Generalizing beyond the input: the functions of the constructions matter. *Journal of Memory and Language* 84. 108-127.
- Perek, Florent & Adele Goldberg. 2017. Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition* 168. 276-293.