Bayesian modelling of clause-level constructional knowledge
for Korean-speaking preschool children

Gyu-Ho Shin (University of Hawaii at Manoa; gyuhoshin@gmail.com)
Seongmin Mun (Université Paris Ouest Nanterre La Défense; simon.seongmin.mun@gmail.com)

Usage-based constructionist approaches concern how to appropriately represent developmental trajectories involving clusters of form-function pairings (*constructions*) based on exposure, together with cognitive-psychological factors. We explore how Korean-speaking preschool children develop clause-level constructional knowledge about a transitive event (active transitives; suffixal passives) through Bayesian simulations.

A canonical active transitive (1a) occurs with a nominative-marked actor, followed by an accusative-marked undergoer, with no active verbal morphology. A canonical suffixal passive (2a) occurs with a nominative-marked undergoer, followed by a dative-marked actor, with passive morphology attached to a verb. These patterns can be scrambled (1b; 2b). Oftentimes, omission applies to a marker or an argument and a marker altogether.

(1a)	actor-nominative	undergoer-accusative	V
(1b)	undergoer-accusative	actor-nominative	V
(2a)	undergoer-nominative	actor-dative	V-passive
(2b)	actor-dative	undergoer-nominative	V-passive

For model training, we created schematised input—pairings of morpho-syntactic and semantic-functional properties involving these construction types (with varying degrees of omission of sentential components)—based on characteristics of the CHILDES database. The actual frequency information about the constructional patterns in the corpus data served as initial priors for learning. Our model, adapting the Alishahi and Stevenson's (2008) learning algorithm, learnt (1) probabilities of individual patterns and (2) conditional probabilities of constructional components within each pattern. We measured posterior probabilities of these patterns per learning (1 to 30) to estimate the degree of clustering for these constructions.

Overall, we found dominance of several patterns (e.g., canonical active transitive with no omission) and their inhibitory effects on the growth of the related patterns (e.g., scrambled active transitive; suffixal passives). This largely mirrored distributional nature of child production found in the same corpus data. Our learning model successfully demonstrated the ability to develop clause-level constructional knowledge as a function of input characteristics and statistical learning, which adds to the cross-linguistic evidence for the effectiveness of Bayesian inference on modelling human learning.

References

Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, *32*(5), 789–834.

Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton, NJ: Princeton University Press.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk (3rd edition). Mahwah, NJ: Lawrence Erlbaum.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.