

Bayesian simulation of Korean-speaking children’s constructional knowledge about transitive events

Usage-based constructionist approaches assume that language emerges and grows through interactions of frequency and domain-general learning capacities (Goldberg, 2019; Tomasello, 2003). The issue is how to appropriately represent developmental trajectories involving clusters of form-function pairings (*constructions*; Goldberg, 1995) based on exposure, together with non-linguistic forces from cognitive-psychological factors. Inspired by recent success of Bayesian simulations in addressing this issue (Alishahi&Stevenson, 2008; Barak.et.al., 2016), we explore how Korean-speaking children formulate knowledge about argument structure constructions involving a transitive event (active transitives and suffixal passives) through Bayesian modelling. Korean, which is typologically different from major languages currently under investigation, provides an interesting testbed due to language-specific properties such as agglutination, scrambling, and omission of sentential components.

A canonical active transitive (1a) occurs with the nominative-marked actor, followed by the accusative-marked undergoer, with no active verbal morphology *per se*. A canonical suffixal passive (2a) occurs with the nominative-marked undergoer, followed by the 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 |

We investigated how a Bayesian-inference-based model learn active transitives and suffixal passives as schematised input—pairings of morpho-syntactic and semantic-functional properties involving these constructions (with varying degrees of omission). There is no Korean corpus of caregiver input paired with semantic-functional information, so we created artificial input based on characteristics of CHILDES Korean (MacWhinney, 2000). We adapted the Alishahi and Stevenson’s (2008) learning algorithm for our simulation. The actual frequency information about the constructional patterns in CHILDES Korean served as initial priors for learning. All the patterns (with omission involved) were used in one learning phase, with the same frequency as the input. Posterior probabilities of the patterns at every learning phase (1 to 30) was measured to estimate the degree of clustering for these constructions.

Overall, we found the dominance of several patterns (e.g., canonical active transitive with no omission) and their inhibitory effects on the growth of the related patterns. This result largely mirrored the distributional properties of child production in CHILDES Korean. Our learning model successfully demonstrates the ability to formulate constructional knowledge as a function of input and statistical learning. The success of our Bayesian learner adds to the cross-linguistic evidence for the effectiveness of Bayesian inference on modelling human learning.

(word count: 397 / 400 words)

Keywords: Argument structure construction; Transitive event; Korean-speaking children; Bayesian simulation

References

- Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, 32(5), 789–834.
- Barak, L., Goldberg, A. E., & Stevenson, S. (2016). Comparing computational cognitive models of generalization in a language acquisition task. In J. Su, K. Duh & X. Carreras (Eds.), *Proceedings of the 2016 conference on Empirical Methods in Natural Language Processing* (pp. 96–106). Association for Computational Linguistics.
- Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- 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.