Title

The key property of frequency distributions that facilitates linguistic rule generalisation is long-tailedness

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Pankratz, Elizabeth
Kirby, Simon
Culbertson, Jennifer

Publication Date

2024

Peer reviewed

The key property of frequency distributions that facilitates linguistic rule generalisation is long-tailedness

Elizabeth Pankratz
University of Edinburgh, Edinburgh, United Kingdom

Simon Kirby
The University of Edinburgh, Edinburgh, United Kingdom

Jennifer Culbertson
University of Edinburgh, Edinburgh, United Kingdom

Abstract

Generalisation of a linguistic rule can be facilitated by certain distributional characteristics. Previous work has shown that a rule is better generalised if it applies to items that (i) follow a skewed frequency distribution, or (ii) follow a uniform frequency distribution over many distinct item types. These two observations cannot be unified under explanations of rule generalisation that are based on entropy of the frequency distributions (since skewed distributions have low entropy, while a greater type count increases the entropy), nor explanations that focus on one highly-frequent type providing a basis for analogical extension (since all types in uniform distributions are equally frequent). Using an artificial language learning experiment and an agent-based model, we show that participants' generalisation behaviour is best matched by a model encoding preferential generalisation of rules containing long-tailed distributions—that is, containing a greater number of low-frequency types.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*