

CONCEPTUAL STRUCTURE IS SHAPED BY COMPETING PRESSURES FOR SIMPLICITY AND INFORMATIVENESS

Jon W. Carr^{*1}, Kenny Smith¹, Jennifer Culbertson¹, and Simon Kirby¹

^{*}Corresponding Author: j.w.carr@ed.ac.uk

¹Centre for Language Evolution, University of Edinburgh, UK

Languages are shaped by competing pressures from learning and communication. Learning favours simple languages, while communication favours informative ones, giving rise to the simplicity–informativeness tradeoff. Languages that evolve under this tradeoff are both maximally simple (learnable) and maximally informative (communicatively useful). This has been shown in natural language and in experimental settings. For example, Kemp and Regier (2012) showed that kinship systems exist at the optimal frontier of simplicity and informativeness. In a separate line of experimental work, Kirby, Tamariz, Cornish, and Smith (2015) showed that when artificial languages evolve under a learning pressure alone, they become simple and uninformative, and when languages evolve during communication, they become complex and informative; it is only when both pressures are at play that we find languages at the optimal frontier.

However, a recent iterated learning experiment by Carstensen, Xu, Smith, and Regier (2015) showed that artificial languages expressing spatial relationships tended to become more informative when subjected to a pressure from learning. This is a surprising result given the previous work briefly reviewed above, which says that informativeness is driven by the pressure from communication, not from learning. One potential explanation for this result lies in their measure of informativeness, *communicative cost*, which is sensitive to (a) the number of words that the language is comprised of (expressivity) and (b) the extent to which similar meanings are expressed by the same word (which we will term convexity). In their experiment expressivity was fixed at four words. As a result, the reduction in communicative cost they found must be due to categories evolving to become more convex, i.e. picking out increasingly tightly-clustered sets of meanings.

To demonstrate that learning favours convex categories, we conducted two experiments in which participants learned and produced a category system for stimuli varying on two dimensions, size and angle. In Experiment 1 participants were trained on one of three systems: One marking a distinction in angle, one marking a distinction in size, and one marking a distinction on both dimensions (see Fig. 1). The results indicated that the Angle-only system was easiest to learn,

followed by the Size-only system; the Angle & Size system was hardest to learn, despite having the lowest communicative cost.

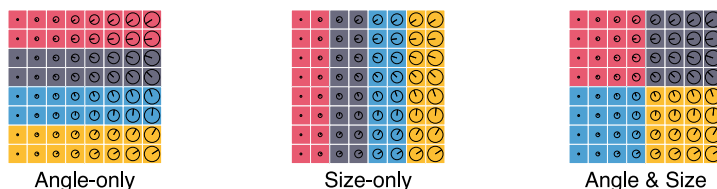


Figure 1. The three systems participants were trained on in Experiment 1. Each colour indicates a category (a group of meanings labelled by the same word).

In Experiment 2, the output of one participant became the input to the following participant in a standard iterated learning design. An example chain is shown in Fig. 2. Over 12 generations the category system became increasingly easy to learn, as indicated by decreasing intergenerational transmission error. Furthermore, in the majority of chains, the language converged on a system marking only a distinction on the angle dimension, which participants found easiest in Experiment 1. This increase in simplicity is driven by increasing convexity.

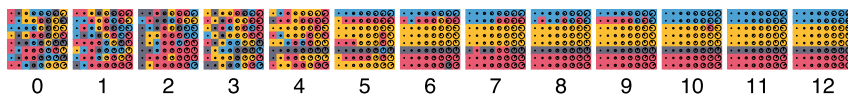


Figure 2. An example iterated learning chain from Experiment 2, showing 12 generations of change in the category system. The initial state of the language at generation 0 was generated randomly. By generation 12, the language has evolved to mark an angle distinction.

We also found that most chains converged on fewer than four categories. This suggests that iterated learning acts as a pressure for simplicity by simultaneously decreasing expressivity and increasing convexity. However, if, as in Carstensen et al. (2015), expressivity is held constant, the learning pressure can only act through convexity: Although languages may become more informative under iterated learning, they do so not because of a pressure to be more communicatively useful, which in Carstensen et al.'s study necessarily decreases communicative cost as a side-effect of increasing convexity. This therefore suggests that, contra Carstensen et al. (2015), languages which are both simple yet informative will only emerge when pressures from learning and communication are at play. We support these conclusions with a Bayesian iterated learning model that displays strikingly similar results.

Acknowledgements

JWC was funded by the Economic and Social Research Council (grant number ES/J500136/1).

References

- Carstensen, A., Xu, J., Smith, C. T., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 303–308). Austin, TX: Cognitive Science Society.
- Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*, 1049–1054.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.