

Using iterated learning to reveal biases for well-structured meanings in language

Jon W. Carr

Centre for Language Evolution

University of Edinburgh

j.w.carr@ed.ac.uk

Our continuous sensory perception of the world is discretized into arbitrary categories. By aligning on a particular system of categorical meaning distinctions, members of a language population can rely on their shared understanding of the observable world to successfully communicate. For example, the items referred to by the English word *cup* form a conceptual category that has fuzzy, English-specific boundaries with neighbouring concepts, such as *bowl*, *glass*, and *pitcher* (Labov, 1973; Malt, Sloman, & Gennari, 2003). What is not well understood, however, is how such categorical meaning distinctions are shaped in conceptual space to optimize for language learning and language use.

Fig. 1 shows three kinds of categorization of a two-dimensional meaning space. In the case of the *random* system, the category membership of every item must be learned independently. In the case of the *connected* system, each category forms a contiguous region of the space, allowing for a more economical cognitive representation. Finally, the *convex* system — consistent with prototype theory — offers the most economical representation; each category can be represented by

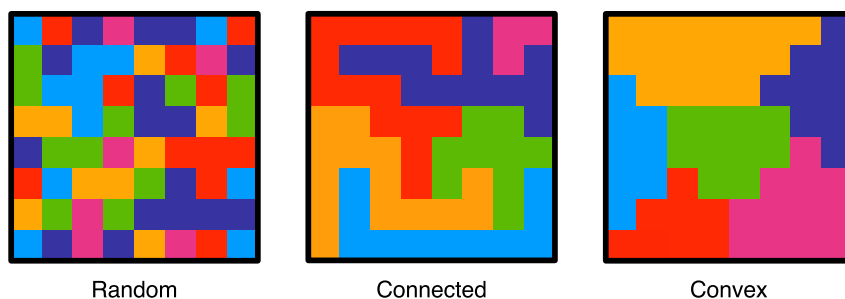


Figure 1: Three kinds of categorization of a two-dimensional meaning space.

a single point in the space. For this reason, Gärdenfors (2000) has argued that categories should be expected to have convex structures; however, there has been little empirical work demonstrating this.

Kirby, Tamariz, Cornish, and Smith (2015) use the *iterated learning* framework to show that pressures for compressibility (i.e., the simplicity of the cognitive representation) and expressivity (i.e., the capacity of the system to convey a useful range of meanings) shape the kinds of structure we find in language. The pressure for compressibility is derived from repeated episodes of learning over generational time, while the pressure for expressivity is inherent to successful communication (cf. Regier, Kemp, & Kay, 2015). These competing pressures may provide an explanation for the kind of categorical structure that arises in language, and the iterated learning paradigm provides a method for teasing apart the contributions of communication and language transmission.

I will present data from a previous study (Carr, Smith, Cornish, & Kirby, 2016) reanalysed in terms of convexity, and I will discuss ongoing experimental work to tease apart the effects of learning and communication on well-formed meaning structures in language. Our predictions are that communication acts as a pressure to move from a random system to a connected system, but that repeated episodes of learning are required to shift emergent languages towards convexity.

References

- Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2016). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive Science*, 1–32. doi: 10.1111/cogs.12371
- Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, MA: MIT Press.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102. doi: 10.1016/j.cognition.2015.03.016
- Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English* (pp. 340–373). Washington, DC: Georgetown University Press.
- Malt, B. C., Sloman, S. A., & Gennari, S. P. (2003). Universality and language specificity in object naming. *Journal of Memory and Language*, 49, 20–42. doi: 10.1016/S0749-596X(03)00021-4
- Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O’Grady (Eds.), *The handbook of language emergence* (pp. 237–263). Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118346136.ch11