COMPRESSIBLE KINSHIP TERMINOLOGIES ARE MORE LEARNABLE THAN LESS COMPRESSIBLE ALTERNATIVES

SARA ROLANDO[1], SIMON KIRBY[1], and KENNY SMITH*[1]
*Corresponding Author: kenny.smith@ed.ac.uk
[1] Centre for Language Evolution, University of Edinburgh, UK

Different languages partition meanings into different semantic categories, labelled with words or morphemes. The scope of variation in these partitions is wide, as systems of semantic categories can differ in both the number of labels used and in the strategies used to group meanings into one category. However, this variation is constrained – not all theoretically-possible partitions are found in natural languages, and similar meanings are encountered in unrelated languages.

This pattern of constrained variation (Regier et al., 2015) has been attested in several domains, such as colour (Berlin and Kay, 1969), spatial terms (Ketharpal et al., 2009), numbers (Greenberg, 1978), and kin classification systems (Murdock, 1970), and clearly stands in need of explanation. Kemp and Regier (2012) propose that this constrained variation is a consequence of pressures for efficient communication (see also Regier et al., 2007; Xu and Regier, 2014; Carstensen et al., 2015). According to this view, category systems are shaped by two competing forces – the need for simplicity, as a learnable or efficient category system minimises cognitive load, and the need for accurate communication, as an efficient category system allows listeners to reliably reconstruct the meanings intended by the speaker.

Suggestive evidence supporting this hypothesis is provided by Kemp and Regier's (2012), who shows that kin categorization systems in natural languages exhibit a near-optimal trade-off between informativeness (ability to uniquely specify individuals in a family tree, weighted by the need probability of being required to refer to each individual) and simplicity (quantified by the length of the grammar underlying the kinship system). Here, we provide a mechanistic test of one of the key assumptions in Kemp and Regier's argument: we focus on

DOI:10.12775/3991-1.102

the pressure arising from language learning, and investigate whether simpler, more compressible kin categorization systems are indeed more learnable than less compressible alternatives.

To test this hypothesis, we ran an artificial language learning experiment in which participants were divided into two conditions and asked to learn two kinship terminologies that differed in their complexity but were matched in their expressivity and their similarity to English. Participants were familiarised with the family tree by introducing 3 members at a time and stating their relationship to each other by only using the primitives "child" and "parent". Once they had learned the set of family relationships to criterion, they were trained and tested on either the Simple or Complex kinship system (see Figure 1). Both kinship systems used 12 labels to individuate 16 possible relationships between relatives but differed in their complexity, where complexity was quantified using the metric proposed by Kemp and Regier (2012), specifically counting the number of rules necessary to describe the system (i.e. the length of the grammar).

As predicted under accounts which assume that simpler kinship systems are easier to learn, we found that participants in the Simple condition learned significantly faster than participants in the Complex condition. We also found that participants in both conditions tended to produce final languages with fewer labels than their input language, collapsing some distinctions encoded in their input, as predicted under accounts of learning where communicative functionality is sacrificed in favour of simplicity (e.g. Kirby et al., 2015).

These results support the claim that the learnability of a kin categorisation system (and by extension, other linguistic systems) depends on the complexity of the mental representation it requires, and that more compressible systems are more learnable than less compressible alternatives. In ongoing work we are testing whether the trade-off between simplicity and informativeness seen in natural languages can emerge from learning or use in isolation or whether (as predicted by Kirby et al., 2015) both learning and use are required to explain the structure of these linguistic systems.

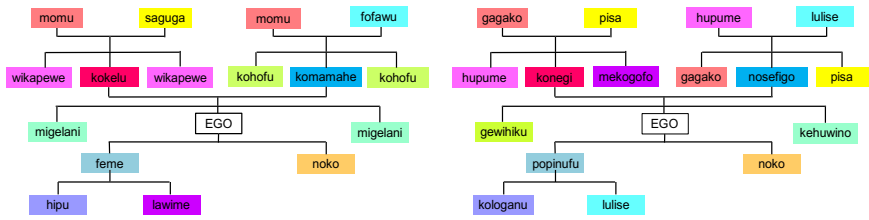

Figure 1: Example Simple (left) and Complex (right) kin categorization systems. Generation is given vertically, relatedness is indicated by lines, males are on the left, females are on the right. While the Simple kinship system underspecifies gender (e.g. using the same label for "brother" and "sister"), the Complex system assigns the same label to very different meanings (e.g. paternal grandfather and maternal uncle).

References

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

Carstensen, A., Xu, J., Smith, C. T., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Greenberg, J. H. (1978). Generalizations about numeral systems. In J. H. Greenberg (Ed.), *Universals of human language, volume 3: Word structure* (pp. 249–295). Stanford: Stanford University Press

Kemp, C., & Regier, T. (2012). Kinship Categories Across Languages Reflect General Communicative Principles. *Science*, *336*, 1049–1054.

Khetarpal, N., Majid, A., & Regier, T. (2009). Spatial terms reflect near-optimal spatial categories. In N. Taatgen et al. (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *PNAS*, *104*, 1436–41.

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.

Murdock, G. P. (1970). Kin Term Patterns and Their Distribution. *Ethnology*, *9*, 165–207.

Xu, J., Dowman, M., & Griffiths, T. L. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B Biological Sciences*, *280*, 20123073.