Re-explaining the universal patterns of color categorization

April 10, 2014

Ever since Berlin and Kay's *Basic Color Terms* (1969), linguistic categorization of color, with its apparent universal tendencies, constitutes a widely discussed topic at the interface of linguistics, philosophy and cognitive psychology. For a long time, the discussion had been polarized between the universalists (such as Paul Kay) and the relativists (such as Barbara Saunders), who differ in their opinions both as to how strong the reported cross-linguistic tendencies of color categorization actually are, and as to how such tendencies are to be explained. Some color categorization models proposed more recently (*e.g.* Jameson and Komarova, 2009; Baronchelli et al., 2010; Loreto et al., 2012) bridge the opposition by incorporating both universalist and relativist insights: the role of the shared color space on one hand, emergence of categories from linguistic interaction on the other. We believe, however, these models to be problematic as regards the choice of the perceptual and interactional principles employed. Moreover, the authors do not provide a transparent and satisfactory method of evaluation. In this article, we adopt the general game-theoretic approach, but improve substantively on both aspects. From this perspective, we also aim to examine the claim by Regier et al. (2007) that universal patterns of color categorization can be explained in terms of optimal partitions of the irregular space of colors.

We model the linguistic categorization of color by means of dynamic game-theoretic interaction of agents in a population, using similarity-maximization signaling games in the spirit of Jäger and van Rooij (2007). The agents exchange messages about points within the human visible spectrum represented in the CIELAB color space, which is supposed to capture human-like perceived similarity and difference between particular colors. Furthermore, we enhance the perceptual space with quasi-realistic probabilities based on the frequency of occurrence of each color in a natural environment. We project the learned signaling strategies onto the 330 Munsell chips and evaluate the variants of our model against the empirical data provided by the World Color Survey (Kay et al., 2009; Cook et al.). Our assumption is that a model which, based on sufficiently realistic principles, generates realistic categorical partitions is relevant with respect to explaining the patterns of color categorization in the languages of the world.

The preliminary results suggest that our model has a predictive power comparable to that of Regier et al. (2007), while having more realistic motivations in several respects. On a more general level, we establish a, so far missing, methodology of evaluating explanatory claims regarding the universal patterns of color categorization.

References

A. Baronchelli, T. Gong, A. Puglisi, and V. Loreto. Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Sciences*, 107:2403–2407, 2010.

R. Cook, P. Kay, and T. Regier. WCS Data Archives. Available at http://www.icsi.berkeley.edu/wcs/data.html (approached January 2013).

G. Jäger and R. van Rooij. Language structure: psychological and social constraints. *Synthese*, 159:99–130, 2007.

K. J. Jameson and N. L. Komarova. Evolutionary models of color categorization. II. Realistic observer models and population heterogeneity. *J. Opt. Soc. Am.*, 26:1424–1436, 2009.

P. Kay, B. Berlin, L. Maffi, W. R. Merrifield, and R. Cook. *The World Color Survey.* Center for the Study of Language and Information, Stanford, 2009.

V. Loreto, A. Mukherjee, and F. Tria. On the origin of the hierarchy of color names. *Proceedings of the National Academy of Sciences*, 109:6819–6824, 2012.

T. Regier, P. Kay, and N. Khetarpal. Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104:1436–1441, 2007.