City Research Online

City, University of London Institutional Repository

City Research Online: http://openaccess.city.ac.uk/   publications@city.ac.uk

Modeling is a tool, and data are crucial
A comment on "Modelling language evolution: Examples and predictions"
by Tao gong et al.

Andrea Baronchelli

Department of Mathematics, City University London,
Northampton Square, London EC1V 0HB, UK

"Evolutionary linguistics did not escape the invasion of computer modeling" Tao, Lan and Zhang affirm correctly. In fact, examples are today so numerous that they prefer to offer a detailed analysis of two models rather than a systematic review of the vast literature - a very reasonable approach to illustrate the predictive power of modeling, be it based on equations or multi-agent systems (even though in some lucky case the latter can be understood analytically, too).

A first merit of the paper by Tao and coworkers consists in stressing that modeling is a tool for, and not a subfield of, Language Evolution. Indeed, the subfield misinterpretation has longly plagued research on the origins of language, denouncing a suspect towards the power of models that other disciplines abandoned a long time ago (in this respect, see for instance the "Modeling" session at the "Evolution of Language" - Evolang conferences). Modeling helps verifying hypotheses, showing for example which minimal ingredients are sufficient to account for the emergence of certain language properties. In this respect, simple models have provided important insights into such problems as the emergence of compositionality [1], the possible genetic basis for human language [2, 3], and the categorization of color [4]. However, model can also inform new experiments by identifying possible mechanisms or features responsible for the observed phenomena. Thus, for example, the possible biological responsible for the universality observed in color naming patterns across cultures, namely the Just Noticeable Difference relative to hue perception in humans, has been pointed out by computer simulations [5]

and lately indirectly confirmed in experiments on the the influence of the two cone-opponent channels in the retinogeniculate pathway [6]. Of course, this evidence does not close the debate on color naming universals, but it helps to substantiate the debate around the role of a concrete biological source of universality that, together with the randomness introduced by cultural evolution, provides a quantitative interpretation to the existing data [7].

The second interesting point raised by the paper is the need for (comparisons with) data. Models certainly help identify which points of a theory are superfluous, or test whether proposed mechanisms are viable, but where they radically change the game is in their ability to interpret existing data in a compact way, and to predict new phenomena. Such approaches as experimental semiotics on the one hand and the analyses of existing databases on the other are certainly fundamental in substantiating models (and henceforth theories) of language evolution, and models contribute by raising new questions and suggesting new directions for data production or collection. Here examples are still limited (see the above mentioned case of color naming universals, as an example), but their number is growing and in the future the dialogue between modelers and experimentalists (or data collectors) will certainly become routinary.

References

[1] Smith K, Brighton H, Kirby S. Complex systems in language evolution: the cultural emergence of compositional structure. Advances in Complex Systems 2003;6:537.

[2] Chater N, Reali F, Christiansen MH, Restrictions on biological adaptation in language evolution, Proc. Natl. Acad. Sci. USA 2009;106:1015.

[3] Baronchelli A, Chater N, Pastor-Satorras R and Christiansen MH. The biological origin of linguistic diversity, PloS ONE 2012;7:e48029

[4] Puglisi A, Baronchelli A, Loreto V. Cultural route to the emergence of linguistic categories. Proc. Natl. Acad. Sci. USA 2008;105:7936.

[5] Baronchelli A, Gong T, Puglisi A, Loreto V. Modeling the emergence of universality in color naming patterns. Proc. Natl. Acad. Sci. USA 2010;107:2403

[6] Xiao Y, Kavanau C, Bertin L. The biological basis of a universal constraint on color naming: cone contrasts and the two-way categorization of colors. PloS ONE 2011;6:e24994

[7] Baronchelli A, Loreto V, Puglisi A. Cognitive biases and language universals. Preprint arXiv:1310.7782 (2013).