A unified computer model for internal and external constraints in language evolution

Christophe Coupé

Laboratoire Dynamique du Langage, CNRS a Université Lyon 2, Lyon, France

Christophe.Coupe@ish-lyon.cnrs.fr

During the last decade, an increasing number of computational approaches have been developed to study human language, which have employed notions such as dynamic complex systems or self-organization. The emergence of linguistic structures has especially received a great deal of attention; whether they focus on the lexicon (Steels 1996), phonetic inventories (de Boer 1999) or sets of syntactic rules (Batali 1997, Kirby 2002, Steels 2000), all studies share the goal of explaining how a coherent global system can emerge from simple local interactions. In parallel, a few computer studies have addressed the role of social parameters in language evolution, such as Nettle's study on social factors (1999), application of Steels' naming games to language contact (Steels 1997, Marsico & al 2000), or Niyogi & Berwick'(1997)s work on the competition of linguistic variants across generations. Such topics can benefit from the advantages of computer models and remain puzzling to most linguists: does a larger population evolve faster or slower? What parameters affect linguistic diversity? Is it possible to refine the glottochronology with additional social settings?

Because of a legitimate difference of focus, models of the emergence of language are usually limited when it comes to study social factors, as models on the evolution of languages often reduce language to an extremely simplified system. While these limitations are helpful to delimitate the role of each parameter, additional phenomena might emerge from the interactions of external (social) and internal (cognitive, production/perception) constraints.

To better investigate such interactions, we propose a model which offers a unified mathematical framework for both types of constraints weighting on a set of linguistic systems, also called agents, which can correspond to either idiolects or communal languages. To this end, we rely on the two key notions of fitness landscape and social network. Internal constraints are defined by a fitness landscape on which linguistic systems draw evolutionary trajectories. All possible states are predefined and no emergence occurs, but the shape of the landscape can be computed from a large variety of situations. Furthermore, a simple model of social network (Milroy 1993) leads to additional attractions or repulsions between agents, following intuitive statements such as Bloomsfield's proposal about the convergence of idiolects of closely and friendly related individuals (Labov 2002). Such a general social model allows investigating a large number of situations, from uniform populations of various sizes, to complex communities with more or less connected sub-networks.

For each agent, the direction of change is determined by the constraints derived from the local slope of the fitness landscape and his social environment. Random draws following probabilistic Gaussian distributions modelling the former constraints lead to the probabilistic evolution of linguistic systems. The global diversity of the population of agents and the mean rate of change can be measured from the trajectories of the systems in the landscape.

We present the conclusions of experiments which first consider the two types of constraints independently, and then combine them to evaluate their respective influence. Among others, Nettle's results are reproduced, and the two types of constraints appear to be operationally independent. We also try to link various topologies of the social network with hypotheses about the prehistory of languages and its characteristic (diversity, rate of evolution). Further enhancements are finally reported, such as the on-going extraction of a fitness landscape from the UPSID database of phonological inventories in world's languages (Maddieson 1984).

References

Batali, J. (1997). Computational Simulations of the Emergence of Grammar. Cambridge University Press, Oxford. p. 405-426

De Boer, B. (1999). Self-organization in vowel systems. Phd thesis, Vrije Universiteit Brussels.

Kirby, S. (2001). Spontaneous Evolution of Linguistic Structure: an iterated learning model of the emergence of regularity and irregularity. Special Issue of IEEE Transactions on Evolutionary Computation on Evolutionary Computation and Cognitive Science 5(2):102-10.

Labov, W. (2002). Driving forces in linguistic change. International Conference on Korean Linguistics, Seoul National University.

Maddieson, I. 1984. Patterns of sound. Cambridge: Cambridge University Press.

Marsico, E., Coupe, C., & Pellegrino, F. (2000). Evaluating the influence of language contact on lexical changes. Proceedings of the third Conference on the Evolution of Language, p. 154-155.

Milroy, J. (1992). Linguistic Variation and Change. Blackwell, Oxford.

Nettle, D. (1999). Is the rate of linguistic change constant? Lingua 108:119-136.

Niyogi, P. & Berwick, R. C. (1997). A Dynamical Systems Model for Language Change. Complex Systems 11:161-204.

Steels, L. (1996). Self-organising vocabularies. MIT Press, Cambridge, MA.

Steels, L. (1997a). Language Learning and Language Contact. In Daelemans, W. (ed.), Proceedings of the workshop on Empirical Approaches to Language Aquisition, p. 11-24.

Steels, L. (2000). A brain for language. In The Ecological. Brain. Sony CSL - Paris - 2000 Symposium.