MODELING THE EMERGENCE OF CREOLE LANGUAGES

FRANCESCA TRIA
ISI Foundation
Turin, ITALY
fratrig@gmail.com

VITTORIO LORETO
Physics Dept., Sapienza University of Rome
Rome, ITALY
&
ISI Foundation
Turin, ITALY
vittorio.loreto@roma1.infn.it

SALIKOKO MUFWENE
Dept. of Linguistics, University of Chicago
Chicago, IL 60637, USA
s-mufwene@uchicago.edu

VITO D.P. SERVEDIO
Institute for Complex Systems, CNR
Rome, ITALY
Vito.Servedio@roma1.infn.it

Creole languages (Reinecke, 1975; Holm, 2000; Holm & Michaelis, 2009; Michaelis, Maurer, Haspelmath, & Huber, 2013) offer an invaluable opportunity to study the processes leading to the emergence and evolution of Language, thanks to the short - typically a few generations - and reasonably well defined time-scales involved in their emergence. Another well-known case of a very fast emergence of a Language, though referring to a much smaller population size and different ecological conditions, is that of the Nicaraguan Sign Language (Senghas, Kita, & Özyürek, 2004). What these two phenomena have in common is that in both cases what is emerging is a contact language, i.e., a language born out of the non-trivial interaction of two (or more) parent languages. This is a typical case of what is known in biology as horizontal transmission. In many well-documented cases, creoles emerged in large segregated sugarcane or rice plantations on which the

slave laborers were the overwhelming majority. Lacking a common substrate language, slaves were naturally brought to shift to the economically and politically dominant European language (often referred to as the lexifier) to bootstrap an effective communication system among themselves (Chaudenson, 2001; Mufwene, 2001). Here, we focus on the emergence of creole languages originated in the contacts of European colonists and slaves during the 17th and 18th centuries in exogenous plantation colonies of especially the Atlantic and Indian Ocean, where detailed census data are available. Those for several States of USA can be found at *http://www.census.gov/history*, while for Central America and the Caribbean can be found at *http://www.jamaicanfamilysearch.com/Samples/1790al11.htm*. Without entering in the details of the creole formation at a fine-grained linguistic level, we aim at uncovering some of the general mechanisms that determine the emergence of contact languages, and that successfully apply to the case of creole formation.

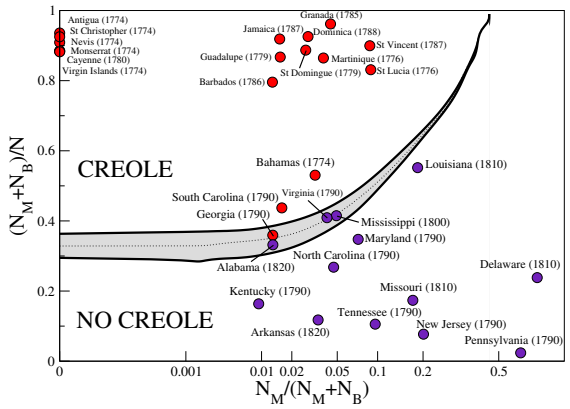

Figure 1. **Census data and model predictions**. We refer to census data about the latest decades of the 18th century and the beginning of the 19th, reporting the number of Free Whites, Free Creole Blacks, and Bozal Blacks B in several States of USA and the Caribbean. Points are the projection of the census data in a bidimensional plane. Red circles mark States where a creole language emerged while purple ones identify States where a creole language historically did not emerge. The gray stripe is the outcome of our modeling scheme and separates the regions where respectively the creole C (above the stripe) and the European E (below the stripe) represent the dominant language (i.e., spread among more than the 80% of the population) in the Mulattos and Bozal populations in the asymptotic states of the model.

We demonstrate a dynamical processes leading to the emergence and stabilization of creole languages, suggesting ways in which modeling can be used as a research tool to clarify accounts of where creoles emerged and what specific ecological factors explain why they did not emerge elsewhere. We judged the

language games (Baronchelli, Felici, Caglioti, Loreto, & Steels, 2006; Loreto & Steels, 2007; Puglisi, Baronchelli, & Loreto, 2008; Loreto, Baronchelli, Mukherjee, Puglisi, & Tria, 2011; Tria, Galantucci, & Loreto, 2012) framework as particularly suitable for this task since it simulates how a population of individuals can bootstrap linguistic consensus–on cultural timescale–out of the local interactions of pairs of individuals. Inspired by the Naming Game (NG) (Baronchelli et al., 2006), our modeling scheme (Tria, Servedio, Mufwene, & Loreto, 2015) incorporates demographic information about the colonial population in the framework of a non-trivial interaction network including three populations (Pucci, Gravino, & Servedio, 2014): Europeans, Mulattos/Creoles, and Bozal slaves. We show how this sole information makes it possible to discriminate territories that produced modern creoles from those that did not, with a surprising accuracy (Fig. 1). We submit that these tools could be relevant to addressing problems related to contact phenomena in many cultural domains (Weinreich, 1963; Thomason, 2001; Bakker & Matras, 2013; Castellano, Fortunato, & Loreto, 2009): e.g., emergence of dialects, language competition and hybridization processes like those undergone by languages with an important number of non-native speakers and more generally all processes where different cultural features come in contact. Thanks to its flexibility, we believe our modeling scheme can be easily modified and adapted to very different contact ecologies, and in this sense it could become an important tool of investigation for scholars interested in testing specific hypotheses about the emergence and evolution of Language.

References

Bakker, P., & Matras, Y. (2013). *Contact languages: A comprehensive guide.* De Gruyter Mouton.

Baronchelli, A., Felici, M., Caglioti, E., Loreto, V., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics*, *P06014*.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, *81*, 591-646.

Chaudenson, R. (2001). *Creolization of language and culture.* New York: Routledge - Taylor and Francis -.

Holm, J., & Michaelis, S. M. (Eds.). (2009). *Contact languages: Critical concepts in linguistics.* Routledge.

Holm, J. A. (2000). *An introduction to pidgins and creoles.* Cambridge University Press.

Loreto, V., Baronchelli, A., Mukherjee, A., Puglisi, A., & Tria, F. (2011). Statistical physics of language dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, *2011*(04), P04006.

Loreto, V., & Steels, L. (2007). Social dynamics: the emergence of language. *Nat. Phys.*, *3*, 758-760.

Michaelis, S. M., Maurer, P., Haspelmath, M., & Huber, M. (Eds.). (2013). *Apics online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Mufwene, S. S. (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.

Pucci, L., Gravino, P., & Servedio, V. D. P. (2014). Modeling the emergence of a new language: Naming game with hybridization. In W. Elmenreich, F. Dressler, & V. Loreto (Eds.), *Self-organizing systems* (Vol. 8221, p. 78-89). Springer Berlin Heidelberg.

Puglisi, A., Baronchelli, A., & Loreto, V. (2008). Cultural route to the emergence of linguistic categories. *Proc. Natl. Acad. Sci. USA (PNAS)*, *105*, 7936-7940.

Reinecke, J. (1975). *A bibliography of pidgin and creole languages.* University Press of Hawaii.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*(5691), 1779–1782.

Thomason, S. (2001). *Language contact.* Edinburgh University Press.

Tria, F., Galantucci, B., & Loreto, V. (2012). Naming a structured world: A cultural route to duality of patterning. *PLoS ONE*, *7*(6), e37744.

Tria, F., Servedio, V. D., Mufwene, S. S., & Loreto, V. (2015). Modeling the emergence of contact languages. *PLoS ONE*, *10*(4), e0120771.

Weinreich, U. (1963). *Languages in contact: findings and problems.* Mouton.