Submitted to:
Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems
Isola di San Servolo (Venice), Italy, October 20-24, 2008

Self-Organization of Creole Community in Spatial Language Dynamics

Makoto Nakamura[1], Takashi Hashimoto[2], and Satoshi Tojo[1]
School of {[1]Information, [2]Knowledge} Science
Japan Advanced Institute of Science and Technology
1-1, Asahidai, Nomi, Ishikawa, 923-1292, Japan
{mnakamur, hash, tojo}@jaist.ac.jp

Abstract

Creolization is a self-organization process of new language community. Thus far, a simulation study of the emergence of creoles has been reported in the mathematical framework. In this paper we introduce a spatial structure to the framework. We show that local creole communities are organized, and creolization may occur when language learners learn often from non-parental language speakers, in contrast to the non-spatial model.

1 Introduction

The emergence of pidgins and creoles is one of the most interesting phenomena in language change. Pidgins are simplified tentative languages spoken in multilingual communities. They come into being where people need to communicate but do not have a language in common. Creoles are full-fledged new languages which children of the pidgin speakers acquire as their native languages. Grammar of a creole is different from any contact languages, although its vocabulary is often borrowed from them [2]. Our goal is to discover specific conditions under which creoles emerge.

Thus far, Nakamura et al. [4] proposed a mathematical framework for the emergence of creoles based on the language dynamics equation by Nowak et al. [5], showing that creoles become dominant under specific conditions of similarity among languages and linguistic environment of language learners. Our purpose in the present study is to introduce a *spatial structure* to Nakamura et al.'s model, in order to observe self-organization process of creole community. Especially, in this paper we compare behaviors of the two models. A related work for introducing a spatial structure into a mathematical model of language change has been done by Castello et al. [3], who have analyzed a bilinguals and spatial version of a mathematical framework by Abrams et al. [1]. Different from Abrams-Strogatz's model, Nakamura et al.'s model [4] is well-defined in terms of learning algorithm and a learning environment.

2 Learning Algorithm and Transition Probability

The most remarkable point in the model of Nakamura et al. [4] is to introduce an *exposure ratio* α, which determines how often language learners are exposed to a variety of language speakers other than their parents. They modified the learning algorithm of Nowak et al., taking the exposure ratio into account in order to model the emergence of creole community. Nakamura et al. [4] have shown that a certain range of α is necessary for a creole to emerge.

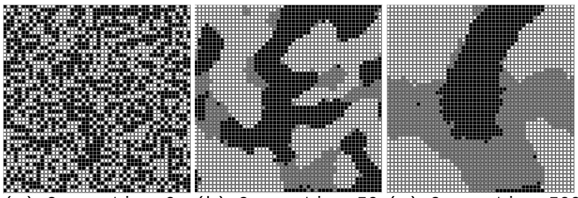
The learning algorithm determines a transition probability $Q = \{q_{ij}\}$ that a language learner whose parent speaks G_i acquires G_j, given the distribution of population $X = \{x_i\}$, similarity among languages $S = \{s_{ij}\}$, the number of input sentences w, and the exposure ratio α (See [4] for detail).

In the spatial model, we use the language distribution in neighbors surrounding each agent to calculate the transition probability Q. Each agent acquires a language through a roulette selection according to the local transition probability.

3 Experiments and Results

The spatial structure is a toroidal 50-by-50 square grid. Each agent has 8 neighbors. Each agent chooses one of three languages every generation, two of which, G_1 and G_2, are pre-existing and randomly distributed with the same total number at the initial state. The remained language, G_3, is a creole, having a certain similarity between two languages. The similarity means the probability that a sentence uttered by a G_i speaker is accepted by G_j. In this paper, we take the following values: $s_{12} = s_{21} = 0$, $s_{13} = s_{31} = 0.3$, $s_{23} = s_{32} = 0.4$, $s_{ii} = 1$, and $w = 10$ for the number of input sentences.

Submitted to:
Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems
Isola di San Servolo (Venice), Italy, October 20-24, 2008



(a) Generation 0 (b) Generation 50 (c) Generation 500

Figure 1. Example of the spatial dynamics (white:G_1, black:G_2, gray:G_3; $\alpha = 0.7$)

We show an examples of the spatial dynamics in Figure 1; (a) Only G_1 and G_2 are distributed at the initial stage. (b) Some local communities (hereafter colonies) of creole are organized at the early stage. (c) Both G_1, G_2 and creole coexist at a quasi-stable stage. In this trial, the creole eventually becomes dominant at Generation 1552. Agents surrounded by both G_1 and G_2 neighbors are likely to acquire the creole. In fact, creole speakers often appear on the border between communities. This is because the large value of α makes the agents to be exposed to both languages, and the creole is the most efficient for accepting input utterances from both languages.

We examine the probability of dominance for each language (Figure 2). Note that the spatial model is based on a stochastic dynamics. This graph is the result of 100 runs for 1,000,000 generations at each α value. The corresponding result in the non-spatial model is the population distribution at the stable generation, shown in Figure 3, since the non-spatial model is based on the deterministic dynamics. This parameter set makes creole dominant at the range $0.1 \lesssim \alpha \lesssim 0.8$ (See [4]). In the spatial model, the probability that the creole is dominant gradually decreases from $\alpha > 0.3$, and it becomes 0.3 around $\alpha > 0.8$.

These differences can be understood by considering local interaction and stochastic dynamics. The pre-existing language may be able to form a colony due to stochasticity. Once a colony with certain size is formed, agents in the colony are surrounded by the same language and the exposure ratio effectively comes to $\alpha = 0$. This situation is hard for the creole speakers to organize a colony. Thus, the probability to be dominant is restrained by the pre-existing language at the middle-high range of α. At the higher range of α, the creole can organize a colony at the early stage with certain ratio through random drift. The colony can grow to the whole space.

4 Conclusion

We show that in the spatial language dynamics, creole can be dominant even in the high exposure ratio, different from the non-spatial model. The analysis of the result tells

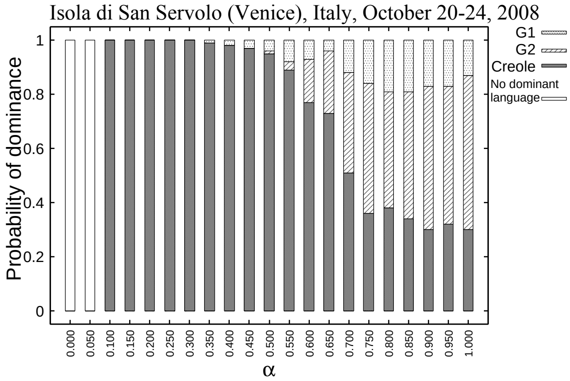


Figure 2. Probability of dominant language in the spatial model

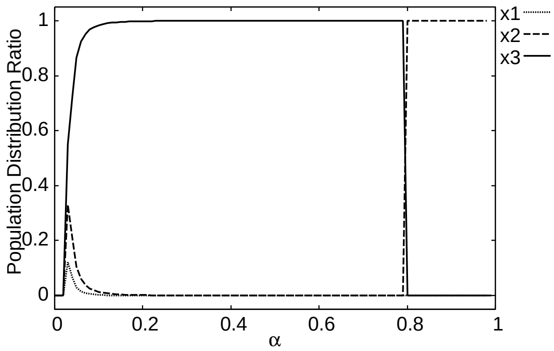


Figure 3. Stable population distribution in the non-spatial model

us that emergence of local colonies at the early stage tends to induce the full creolization.

References

[1] D. M. Abrams and S. H. Strogatz. Modelling the dynamics of language death. *Nature*, 424:900, August 2003.

[2] J. Arends, P. Muysken, and N. Smith, editors. *Pidgins and Creoles*. John Benjamins Publishing Co., Amsterdam, 1994.

[3] X. Castelló, V. M. Eguíluz, M. S. Miguel, L. Loureiro-Porto, R. Toivonen, J. Saramäki, and K. Kaski. Modelling language competition: bilingualism and complex social networks. In A. Smith et al. (Eds), *The Evolution of Language (EVOLANG7)*, pages 59–66. World Scientific, 2008.

[4] M. Nakamura, T. Hashimoto, and S. Tojo. Exposure dependent creolization in language dynamics equation. In A. Sakurai et al. (Eds), *New Frontiers in Artificial Intelligence (EELC2004)*, volume 3609 of *Lecture Notes in Artificial Intelligence*, pages 295–304. Springer, 2004.

[5] M. A. Nowak, N. L. Komarova, and P. Niyogi. Evolution of universal grammar. *Science*, 291:114–118, 2001.