Exposure Dependent Creolization in Language Dynamics Equation

Makoto Nakamura[†], Takashi Hashimoto[‡] and Satoshi Tojo[†*]

1 Introduction

Children correctly inherit language from their parents and/or neighbors during their acquisition period, even though it has not yet been clarified how children correctly deduce the underlying grammatical rules and acquire the same language. On the other hand, pidgin and creole are defined as two different stages of language change [1, 3]. Pidgin is a simplified tentative language spoken in multilingual communities. Creole is a full-fledged new language which children of the pidgin speakers obtain as their native language. Some properties of creoles imply the existence of an innate universal grammar.

Linguistic studies are going to have been clarified why and how creoles emerged. Observing actual pidgins and creoles, linguists have argued that creoles would appear under a specific environment like a pidgin community. From the linguistic efforts, it is clear that the emergence of creole is affected by a contact with other languages, the distribution of population for each language, similarities among the languages. On the contrary in population dynamics [8], we could derive boundary conditions from the numerical analyses by parametrizing these elements, and then could contribute to specify the function of the universal grammar.

Thus far, we revised the language dynamics by Nowak et al. [7] in such a way that the transition rates changed according to the distribution of population of each grammar at each generation. In addition, we introduced an *exposure rate* by which a child is exposed to other languages than that of his/her parents. Using this approach, we have shown the emergence of a creole when multiple parental languages are similar in some way [4]. We improved our model to exclude *fitness* that dominated the ratio of offsprings with regard to communicability [5], but we observed such unnatural phenomena that the creole emerged even when children learned language only from their parents. In this paper, we will present a new formalism to remedy this problem.

2 Language Dynamics Equation without Fitness

In this section, we breafly explain our previous model and consider the emergence of creole in population dynamics.

In response to the language dynamics equation by Nowak et al. [7], we assume that any language is classified into one of a certain number of grammars. Thus, the population of language speakers are distributed to grammar $\{G_1 \dots G_n\}$. Let x be a ratio of speakers of each language. Then, the model is expressed as a dynamics which deals with the change of population rate for each language.

In the language dynamics equations, the S matrix and the $\overline{Q}(t)$ matrix play important roles. The similarity matrix $S = \{s_{ij}\}$ denotes the probability that a sentence of G_i is accepted by G_j. The transition matrix $\overline{Q}(t) = \{\overline{q}_{ij}(t)\}$ is defined as the probability that a child of G_i speaker obtains G_j by the exposure to her parental language and to other ones. Being different from the definition

*Graduate School of {†Information, ‡Knowledge} Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, 923-1292, Japan (mnakamur@jaist.ac.jp)

by Nowak et al., our definition of $\overline{Q}(t)$ depends on the generation parameter t, as well as the S matrix and a learning algorithm.

Because Nowak et al. assume that language speakers bear offspring in proportion to their successful communication, they embed a fitness term in their model which determines the birth rate of each language group. Our model excludes the fitness on the assumption that in the real world creoles do not emerge because creole speakers have more offspring than that of other pre-existing languages. We have already shown the difference between the models with and without fitness [5], in which the latter becomes:

$$\frac{dx_j(t)}{dt} = \sum_{i=1}^{n} \overline{q}_{ij}(t) x_i(t) - x_j(t) \quad . \tag{1}$$

3 Language Acquisition and Transition of Population

In this section, we propose a new transition matrix $\overline{\overline{Q}} = \{\overline{\overline{q}}_{ij}(t)\}$. Our approach takes account of a probability distribution of the number of acceptable sentences for each grammar against the number of input sentences during acquisition term, in order to reflect a learning algorithm into the transition matrix. Firstly, we explain the learning algorithm. Secondly, we represent the transition matrix $\overline{\overline{Q}}$.

3.1 Learning Algorithm

We introduce the exposure rate α that a child is affected by the other language speakers than their parents (See Fig. 1). On the other hand, a probability in which the child learns a language from his/her parents comes to $(1-\alpha)$. Note that α does not exclude children's parental language; it is also included in α in proportion to the distribution of population.

Our learning algorithm resolves Niyogi [6]'s problem that there is an unrealistic Markov structure which implies that some children cannot learn certain kinds of language. From the viewpoint of a universal grammar that all conceivable grammars of human beings are restricted to

Fig. 1 The exposure rate α

Fig. 2 The learning algorithm

a finite set, language learning is considered as a choice of a plausible grammar from them. Thus, the learning algorithm is given as: 1) In a child's memory, there supposed to be a score table of grammars. 2) The child receives a sentence uttered by an adult. 3) For each grammar, if a sentence is acceptable for the child, the grammar scores a point in his/her memory. 4) 2) and 3) are repeated until the child receives a fixed number (w) of sentences that is regarded as enough for the estimation of the grammar. 5) The child adopts the grammar with the highest score.

The distribution of population and the exposure rate α determine the rate of the adult speaker to which the child is exposed for each grammar, while the S matrix determines the acceptability of a sentence. Fig. 2 shows an example that a child of G_2 speaker obtains G_2 after the exposure to a variety of languages. The child receives sentences, that are numbered boxes from 1 to 10. The input sentences are divided into two sets by the exposure rate α. One of the sets consists of sentences of a variety of grammars, dependent on the rate of population. For example, the child hears sentences 1, 4 and 5 uttered by G_1 speakers. The other is that of her parents. Therefore, these sentences are acceptable by a particular grammar. Because her parental grammar is G_2, for example, the sentences 7 to 10 are randomly chosen from the language of G_2. The child counts acceptable sentences for each grammar. The sentence 1 can be accepted by G_3 other than G_1, while it is uttered by a G_1 speaker. The Venn diagram in Fig. 2 represents that each language shares sentences with others. In this case, because the sentence 1 is acceptable both by G_1 and by G_3, the child adds 1 to both of the counters in his/her mind.

3.2 Revised Transition Probability

Suppose that children hear sentences from adult speakers depending on the exposure rate and on the distribution of population. A probability that a child whose parents speak G_i accepts a sentence by G_j is expressed by

$$U_{ij} = \alpha \sum_{k=1}^{n} s_{kj} x_k + (1-\alpha) s_{ij} \quad . \tag{2}$$

After sufficient time, the child will adopt the most plausible grammar, which is estimated by counting a number of accepted sentences by each grammar. This learning algorithm is simply represented in the following equation. Exposed to a variety of languages in proportion to the ratio of adult speakers, children whose parents speak G_i will adopt G_{j^*} in the following manner:

$$j^* = \underset{j}{\operatorname{argmax}} \{U_{ij}\} \quad . \tag{3}$$

When children hear w sentences, a probability that a child of G_i speaker accepts r sentences with G_j is given by a binomial distribution,

$$g_{ij}(r) = {}_wC_r (U_{ij})^r (1-U_{ij})^{w-r} \quad . \tag{4}$$

On the other hand, a probability that the child accepts less than r sentences with G_j is

$$h_{ij}(r) = \sum_{k=0}^{r-1} {}_wC_k (U_{ij})^k (1-U_{ij})^{w-k} \quad . \tag{5}$$

Using these two probability distributions, the probability that a child of G_i speaker accepts k sentences with G_j, while less than $k-1$ sentences with the other grammars, comes to $g_{ij}(k) \prod_{l=1, l\neq j}^{n} h_{il}(k)$. Because the candidate grammar must accept at least $\lceil \frac{w}{n} \rceil$ sentences, the probability that the grammar G_j is adopted as the most plausible one by the child of G_i speaker is

$$\bar{\bar{q}}_{ij}(t) = \frac{\displaystyle\sum_{k=\lceil \frac{w}{n} \rceil}^{w} \{ g_{ij}(k) \prod_{\substack{l=1 \\ l\neq j}}^{n} h_{il}(k) + R \}}{\displaystyle\sum_{m=1}^{n} [\sum_{k=\lceil \frac{w}{n} \rceil}^{w} \{ g_{im}(k) \prod_{\substack{l=1 \\ l\neq m}}^{n} h_{il}(k) + R \}]} \quad , \tag{6}$$

where R is an offset term that one or more other grammars accept the same number of sentences as G_j. This function is the revised transition probability.

4 Creole in Population Dynamics

We presuppose that the emergence of creole strictly depends on the population distribution, as opposed to traditional linguistic explanations [2, 3]. From the viewpoint, a creole is considered as such a grammar G_c that: $x_c(0) = 0, x_c(t) > \theta_c$, where $x_c(t)$ denotes the rate of the population of G_c at a convergent time t, and θ_c is a certain threshold to be regarded as a dominant language. In this paper, we put $\theta_c = 0.9$ through the experiments. This definition represents that some individuals come to speak a language that no one spoke at the initial state, and consequently dominates the community.

We have mainly observed the behavior of the model of three grammars. Suppose the size of language is the same and each sentence of the language is chosen in a uniform probability, the similarity matrix can be expressed as such a symmetric matrix that:

$$S = \begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \quad . \tag{7}$$

We regard G_3 as a creole grammar, giving the initial condition as $(x_1(0), x_2(0), x_3(0)) = (0.5, 0.5, 0)$. Therefore, the element a denotes the similarities between two pre-existing languages, and b and c are the similarity between G_1 and the creole, and between G_2 and the creole, respectively.

5 Experiments

In this section, we show the experimental result of our model. We examine the conditions that creole appears and comes to be dominant in combinations of the S matrix and α.

5.1 Emergence of Creole

In Fig. 3, we show the result of our model. We arbitrarily set the S matrix to $(a, b, c) = (0, 0.45, 0.35)$, in which the pre-existing grammars G_1 and G_2 do not share any sentence. We gave the number of input sentences $w = 30$ that was found to be large enough for language acquisition in three grammars. The exposure rate α is examined at the range from 0 to 1.

In case $\alpha = 0$, children learn a language only from their parents. In fact, Fig. 3(a) shows that both populations of G_1 and G_2 hardly transmit. In the previous model [5], we found a problem that a creole coexists with other languages at $\alpha = 0$. However, we come to resolve the problem.

According to the increase of α, x_3 rises gradually though $x_3(0) = 0$, while x_1 falls down to 0 in Fig. 3(b). However, x_3 declined in further generations and eventually disappeared. Because the transition of population depends on Eqn 2, we can approximately compare the directions of population transition among grammars with Eqn 2. Eqn 8 expresses an expansion of Eqn 2 at $a = 0$. Although the population is shared among only G_1 and G_2 at the initial generation, the increase of α makes the transition from G_1 and G_2 to G_3 active. Because U_{13} is greater than U_{31} and $U_{23} > U_{32}$ at early generations, x_1 and x_2 starts flowing out to x_3. Moreover, because $U_{13} > U_{23}$, x_1 is easier to flow into x_3 than x_2. Once x_3 has earned a certain rate of population, U_{13} becomes further greater than U_{31} and the outflow of x_1 to x_3 accelerates, while $U_{23} \simeq U_{32}$. After x_1 mostly diminished, the difference between U_{23} and U_{32} concerns that between $(cx_2 + x_3)$ and $(x_2 + cx_3)$, that is x_3 and x_2. Because x_2 is barely more than x_3 at the point of generation in $\alpha = 0.3$, G_2 finally dominates the community. Fig. 4(a) shows these flows of the population. We can see that the larger α, the dynamics converges at the earlier generation in Fig. 3(b) and Fig. 3(c). On the contrary, we have encountered that the equation did not converge in realistic time at a small α.

In case of $(cx_2 + x_3) > (x_2 + cx_3)$ at the point, x_3 rises to 1, and G_3 becomes dominant. Fig. 3(d) shows the emergence of creole in $\alpha = 0.668$. Fig. 4(b) depicts the process of creolization that in addition to the inflow of population of G_1 the transition from G_2 to G_3 outstrips the outflow.

Further increasing α, we can observe G_1 becomes dominant although it loses the population at the very first in small α. Also, let us pay attention to U_{12} and U_{21} in Eqn 8. At the early generation, x_3 has not earned enough population yet. When α is large enough like Fig. 3(e), U_{12} is larger than U_{13}. Large α represents that children of G_1 speakers grow up, hearing sentences of G_2 in the almost same rate, with that of G_1. Therefore, the direct transition between G_1 and G_2 occurs at large α. Fig. 4(a) shows the flows of the population between G_1 and G_2. In other words, such an open communication may develop the power game between the pre-existing languages. Thus, we can regard our experimental result is that creoles are not the easiest to emerge at $\alpha = 1$. This result adequately remedied our fallacious expectation [4].

5.2 Dominant Language and Creole

In the previous section, we showed the emergence of a creole, and quantitatively considered the process of creolization. We observed that a creole appears within a certain area of α which must be large enough but less than the value which affects the direct transition between pre-existing languages. The value of a in Eqn 7 concerns the transition, while it is fixed to 0 in the previous ex-

(a) G_1 and G_2 coexist at $\alpha = 0$.

(b) G_2 is dominant at $\alpha = 0.3$.

(c) G_2 is dominant at $\alpha = 0.667$.

(d) G_3 is dominant at $\alpha = 0.668$.

(e) G_1 is dominant at $\alpha = 0.835$.

(f) G_3 is dominant at $\alpha = 1$.

Fig. 3 The transition of dominant language by changing α $((a, b, c) = (0, 0.45, 0.35))$.

Fig. 5: Distribution of dominant grammars $((a, b, c) = (a, 0.45, 0.35))$

periment. The next experiment aims at drawing a diagram as to which language would be dominant in various values of the similarity between the pre-existing languages.

We parametrized a in the S matrix and α. The result is shown in Fig. 5. In the figure, the region of asterisk $(*)$ denotes none of the languages becomes dominant. Namely, either the dynamics converged to the coexistence of a few languages, or the dynamics could not converge at over a million generations. As we mentioned in the previous experiment, with small values of α the dynamics hardly converges. On the contrary, in the upper side of the asterisk region, the pre-existing languages coexist because those languages are regarded as an almost identical language by very high value of a.

The previous experiment was examined along with the horizontal axis at $a = 0$. At the bottom of the figure around $\alpha = 0.8$, G_3 (creole) becomes dominant. Thus, the lower value of a, the easier creoles emerge. In other words, the value of a implies the degree of the power game between the pre-existing languages. The extent of the region of creole depends on the similarity among pre-existing languages. This result is consistent with that of our previous model [4] that a creole may emerge if the pre-existing languages are not similar to each other, but to the newly appeared language.

(a) G_2 becomes dominant. (b) G_3 becomes creole. (c) G_1 becomes dominant.

Fig. 4 Flow of the population by changing α value ($(a, b, c) = (0, 0.45, 0.35)$)

$$U_{ij} = \begin{pmatrix} (1-\alpha) + \alpha(x_1 + bx_3) & \alpha(x_2 + cx_3) & (1-\alpha)b + \alpha(bx_1 + cx_2 + x_3) \\ \alpha(x_1 + bx_3) & (1-\alpha) + \alpha(x_2 + cx_3) & (1-\alpha)c + \alpha(bx_1 + cx_2 + x_3) \\ (1-\alpha)b + \alpha(x_1 + bx_3) & (1-\alpha)c + \alpha(x_2 + cx_3) & (1-\alpha) + \alpha(bx_1 + cx_2 + x_3) \end{pmatrix} \tag{8}$$

6 Conclusion

In this study, we proposed the modified Q matrix of the population dynamics equation [7], where children may migrate to non-parental languages, counting the number of sentences with probable grammars. As a result, we could show that creolization rarely occur in a high value of *exposure rate* α, no less in a small value of it. In a high value of α, children tend to select a pre-existing dominant language, and in a low value they certainly learn parental language; thus we could contend that the creolization emerges just between the influence of mother tongues and the socially dominant language. As a future work, we need to consider refining the learning algorithm; and in addition, we need to establish the more reliable theory on language similarity.

References

[1] Jacques Arends, Pieter Muysken, and Norval Smith, editors. *Pidgins and Creoles*. John Benjamins Publishing Co., Amsterdam, 1994.

[2] D. Bickerton. *Language and Species*. University of Chicago Press, 1990.

[3] Michel DeGraff, editor. *Language Creation and Language Change*. The MIT Press, Cambridge, MA, 1999.

[4] Makoto Nakamura, Takashi Hashimoto, and Satoshi Tojo. Creole viewed from population dynamics. In *Proceedings of the Workshop/Course on Language Evolution and Computation in ESSLLI*, pages 95–104, 2003.

[5] Makoto Nakamura, Takashi Hashimoto, Satoshi Tojo, and Kenny Smith. The effect of fitness in the emergence of creole. In *Proceedings of the 5th International Conference on the Evolution of Language*, page 85, 2004.

[6] Partha Niyogi. *The Informational Complexity of Learning from Examples*. PhD thesis, Massachusetts Institute of Technology, , 1996.

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

[8] J.W. Weibull. *Evolutionary Game Theory*. The MIT Press, Cambridge, MA, 1995.