The Language Dynamics Equations of Population-Based Transition – a Scenario for Creolization –

Makoto Nakamura*†, Takashi Hashimoto‡ and Satoshi Tojo†
Graduate School of {†Information, ‡Knowledge} Science,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Tatsunokuchi-machi, Nomi-gun, Ishikawa, 923-1292, Japan

Abstract ***Children will develop their parental languages correctly, since language learners come to obtain the one which they contact most in the community. At the same time, children would be affected by other languages, the influence of which is proportional to the population of those languages. In this paper, we revise the foregoing evolutionary theory of language, that is differential equations of the population dynamics. We propose that the transition rate in languages is sensitive to the distribution of population of each generation. In addition, we introduce the exposure probability that is the measure of influence from other languages. We show experimental results, in which we could observe the emergence of creole. Furthermore, we analysed which language would be dominant, dependent on the initial distribution of population, together with the exposure probability.***

Keywords: language acquisition, population dynamics, creole

1 Introduction

Language acquisition is intrinsic for human intelligence. One of the main purpose of language use for human beings is to communicate with others. Therefore, it is easy to consider that the language learners come to obtain the language which they hear most in the community, i.e., in most cases, children will develop their parental languages correctly. However, because the learning term is limited, language learners need to acquire a plausible grammar from a finite number of sentences uttered in the community during their training periods though they may not have enough samples to determine grammar rules uniquely, as was argued in Gold [5]. Nevertheless, children growing up in the same speech community correctly deduce the underlying grammatical rules and consistently develop the same language.

Studies on simulating the evolution of language have been often reported in recent years, where a community of autonomous and active agents learns a common grammar through the exchange of sentences between them [2, 6, 12]. However, because the number of agents is finite, the results are often hard to be generalized for us to explain the phenomena in the real world. As opposed to this approach, Komarova et al. [7] generalized an evolutionary theory of language with the universal grammar mathematically. They did not pay attention to an ability of each agent but to the whole behaviour of the population, and thus, this approach is different from multi-agent models. In Komarova's work, given the principles in the universal grammar, the search space for candidate grammars was assumed to be finite. From this assumption, defining the similarity and payoff between grammars, they represented the transition of population of grammars as differential equations which they call the *language dynamics equations*.

Because the population of the languages affects language acquisition, observing the transition of population diachronically is an essence for language evolution. However, their model is too simplified for the real situation. For example, the transition ratio between grammars was fixed and they disregarded the influence from other language groups upon children. We propose that the transition rate in languages is sensitive to the distribution of population of each generation. Our purpose of this paper is to modify the language dynamics equations to ones more adaptable to the real world. Furthermore, we will show the process of creolization, as a result of contact of distinct language groups.

In Section 2, we describe the language dynamics equations and our modification of the equations. In Section 3, we give a model with the modified equations. We describe our experiments in Section 4 and conclude in Section 5.

2 Population Dynamics of Grammar Acquisition

In this section, we firstly explain the language dynamics equations proposed by Nowak et al. [11] and by Komarova et al. [7]. Secondly, we mention the problem of Niyogi's model [9], and after that we modify the equations.

2.1 Komarova's Model

Both of Nowak's and Komarova's purposes are to develop a mathematical theory for the evolutionary and population dynamics of grammar acquisition [7,11]. In Komarova's model, given the principles in the universal grammar, the search space for candidate grammars is assumed to be finite, that is $\{G_1, \ldots, G_n\}$. The population dynamics equations are defined from (i) the similarity between grammars as the matrix $S = \{s_{ij}\}$ and (ii) the probability that children fail to acquire their parental language as the matrix $Q = \{q_{ij}\}$. Individuals reproduce children, the number of which is determined by the *fitness* such as: $f_i(t) = \sum_{j=1}^{n}(s_{ij} + s_{ji})x_j(t)/2$, where $x_j(t)$ is the ratio of the population of G_j speakers and $\sum_{j=1}^{n} x_j(t) = 1$. For simplicity, it is assumed that each child has only one parent, namely, learns only one grammar. The language dynamics equations are given by the following differential equations:

$$\frac{dx_j(t)}{dt} = \sum_{i=1}^{n} q_{ij} f_i(t) x_i(t) - \phi(t) x_j(t) \quad (j = 1, \ldots, n), \qquad (1)$$

where $\phi(t) = \sum_{i=1}^{n} f_i(t)x_i(t)$. The total population size keeps constant by '$-\phi(t)x_j(t)$'.

The situation is depicted in the following procedure (See also Fig. 1):

1. Individuals leave offsprings proportional to their fitness for each grammar, which is regarded as its communicability. Total distribution of children of G_i speakers is $f_i x_i$.
2. They learn a language from their parents. However, only children whose parents use G_j in the ratio of q_{jj} come to use G_j correctly. The others in the ratio of $1 - q_{jj}$ flow out.
3. Some children whose parents use G_i mistake their target grammar for the other one, namely G_j, in the ratio of q_{ij}.

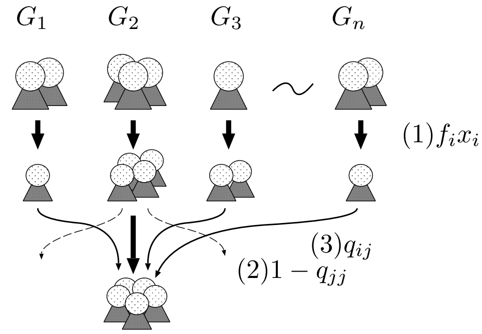


Fig. 1 Flow of population change

From the above interpretation, it is assumed that only adult individuals talk to the other language groups, while children communicate with only their parents. In this circumstance, it may be difficult to consider that the children mistake their parental grammar for another one. In the next subsection, we discuss the possibility of mistake connected with Niyogi's model [9].

2.2 Niyogi's Model

Niyogi [9] proposed a model of language acquisition with linguistically well-grounded grammars together with the trigger learning algorithm (TLA) [4]. He showed the parameter settings in the grammars as a Markov structure through language acquisition of children. Based on his results, the possibilities in which the children do not learn their target language, namely $q_{ij} > 0 (i \neq j)$, can be considered as follows:

A) There are some states in Markov structure, in which children can not escape from the states even if the children receive only correct examples of the target grammar, called an Absorbing State in the Marcov chain.

B) Sufficient quantity of stimuli are not given for children before fully acquiring the target grammar. Any finite number of example sentences is not proved to be enough to determine uniquely the underlying grammar [5].

The matrix $Q = \{q_{ij}\}$ depends on the matrix $S = \{s_{ij}\}$ because the latter concerns the transitivity between grammars [7]. As above, the accuracy of language acquisition also depends on the learning algorithm. Therefore, Niyogi's learning algorithm is possible to cause q_{ij} unnaturally high.

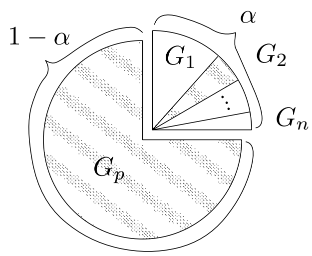


Fig. 2 Exposure probability α ($p = 2$)

2.3 Our Modification

Here, let us consider Niyogi's two issues. With regard to A), it is rather natural to regard that children have the ability to escape from such local traps. It is generally assumed that children acquire their target grammars without error. Therefore children should have a learning algorithm which can attain to an arbitrary target grammar. As to B), if a child has not been given enough example sentences within her critical period of language acquisition, she also cannot learn *another* human language and would be 'Genie' a wolf-child [13]. From the both reasons, the probability that children fail to learn their parental language and learn another language is quite low, i.e., q_{ij} $(i \neq j)$ should be very near to 0.

Thus, we start our experiments with $Q \approx I$ (the unit matrix), admitting the slight possibility of grammar transition. We contended that the Q matrix changes through generations, in regard to the distribution of grammar populations, as we have already shown by computer simulations [8]. Our prime revision is to consider the probability α that children are affected by the other language speakers than their parents. Thus, the probability which the children learn a language from their parents comes to $(1-\alpha)$. Note that α does not exclude children's parental language. We call α the *exposure probability*. In Fig. 2, G_p is the mother language. The children are exposed to other languages at the rate of α. Suppose the parental grammar of a child is G_2, the shaded part of the figure denotes the distribution in which the child is exposed to the mother language. Although G_1 speakers seem to occupy the most part in the distribution of population in the community, the language that is most frequently exposed to the children is the parental one, i.e., $G_p = G_2$, for the value of α is small. However, in general, if α is large enough, the language for the most frequent exposure to a child may differ from her parental one.

Since the distribution of speakers changes in time, and the Q matrix depends on it, the Q matrix should include the time parameter t, that is $\overline{Q}(X(t)) = \{\overline{q}_{ij}(t)\}$, where $X(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. We call $\overline{Q}(X(t))$ the *modified accuracy matrix*. Together with the S matrix and a given α, a learning algorithm determines $\overline{Q}(X(t))$. Thus, the new language dynamics equations are as follows:

$$\frac{dx_j(t)}{dt} = \sum_{i=1}^{n} \overline{q}_{ij}(t) f_i(t) x_i(t) - \phi(t) x_j(t) \qquad (j = 1, \ldots, n). \tag{2}$$

3 The Population-Based Model

In this section, we explain the modified language dynamics equations, eqn (2), by introducing a set of grammars and a learning algorithm.

3.1 The Grammars

We adopt the same set of grammars as Niyogi [9], that is a three-parameter syntactic subsystem described in Gibson et al. [4]. Thus the set has exactly eight grammars, generating languages from $L(G_1)$ to $L(G_8)$. It includes two parameters from X-bar theory. Specifically, they are concerned with specifier-head relations, and head-complement relations in phrase structures. The following production rules correspond to the parameters:

$$\begin{array}{lcl} XP & \rightarrow & \textit{Spec } X' \ (p_1 = 0) \text{ or } X' \textit{ Spec} \ (p_1 = 1), \\ X' & \rightarrow & \textit{Comp } X' \ (p_2 = 0) \text{ or } X' \textit{ Comp} \ (p_2 = 1), \\ X' & \rightarrow & X. \end{array}$$

The third parameter concerns the verb movement. In German and Dutch root declarative clauses, it is observed that the verb occupies exactly the second position. This Verb-Second phenomenon might or might not be present in the world's languages, and this variation is captured by means of the V2 parameter. Table 1 provides the unembedded (degree-0) sentences from each of the eight languages obtained by setting the three parameters.

We do not provide different vocabularies for each grammar, because we suppose that the same words are shared among languages. Individuals would acquire foreign words when necessary in the multilingual community. To make languages simple, each word category in the table, for example, S, V and so on, has only one word. Namely, each category directly stands for the corresponding word.

Table 1 The Languages (cited from [4,9])

Language	[Spec,Comp,V2]	Degree-0 unembedded sentences
$L(G_1)$	[1,1,0]	"V S" "V O S" "V O1 O2 S" "AUX V S" "AUX V O S" "AUX V O1 O2 S" "ADV V S" "ADV V O S" "ADV V O1 O2 S" "ADV AUX V S" "ADV AUX V O S" "ADV AUX V O1 O2 S"
$L(G_2)$	[1,1,1]	"S V" "S V O" "O V S" "S V O1 O2" "O1 V O2 S" "O2 V O1 S" "S AUX V" "S AUX V O" "O AUX V S" "S AUX V O1 O2" "O1 AUX V O2 S" "O2 AUX V O1 S" "ADV V S" "ADV V O S" "ADV V O1 O2 S" "ADV AUX V S" "ADV AUX V O S" "ADV AUX V O1 O2 S"
$L(G_3)$	[1,0,0]	"V S" "O V S" "O2 O1 V S" "V AUX S" "O V AUX S" "O2 O1 V AUX S" "ADV V S" "ADV O V S" "ADV O2 O1 V S" "ADV V AUX S" "ADV O V AUX S" "ADV O2 O1 V AUX S"
$L(G_4)$	[1,0,1]	"S V" "O V S" "S V O" "S V O2 O1" "O1 V O2 S" "O2 V O1 S" "S AUX V" "S AUX O V" "O AUX V S" "S AUX O2 O1 V" "O1 AUX O2 V S" "O2 AUX O1 V S" "ADV V S" "ADV V O S" "ADV V O2 O1 S" "ADV AUX V S" "ADV AUX O V S" "ADV AUX O2 O1 V S"
$L(G_5)$ (English)	[0,1,0]	"S V" "S V O" "S V O1 O2" "S AUX V" "S AUX V O" "S AUX V O1 O2" "ADV S V" "ADV S V O" "ADV S V O1 O2" "ADV S AUX V" "ADV S AUX V O" "ADV S AUX V O1 O2"
$L(G_6)$	[0,1,1]	"S V" "S V O" "O V S" "S V O1 O2" "O1 V S O2" "O2 V S O1" "S AUX V" "S AUX V O" "O AUX S V" "S AUX V O1 O2" "O1 AUX S V O2" "O2 AUX S V O1" "ADV V S" "ADV V S O" "ADV V S O1 O2" "ADV AUX S V" "ADV AUX S V O" "ADV AUX S V O1 O2"
$L(G_7)$ (Bengali)	[0,0,0]	"S V" "S O V" "S O2 O1 V" "S V AUX" "S O2 O1 V AUX" "ADV S V" "ADV S O V" "ADV S O2 O1 V" "ADV S V AUX" "ADV S O V AUX" "ADV S O2 O1 V AUX"
$L(G_8)$ (German, Dutch)	[0,0,1]	"S V" "S V O" "O V S" "S V O2 O1" "O1 V S O2" "O2 V S O1" "S AUX V" "S AUX O V" "O AUX S V" "O1 AUX S O2 V" "O2 AUX S O1 V" "ADV V S" "ADV V S O" "ADV V S O2 O1" "ADV AUX S V" "ADV AUX S O V" "S AUX O2 O1 V" "S O V AUX" "ADV AUX S O2 O1 V"

3.2 The Similarity Matrix

Generally, the S matrix is uniquely calculated when the grammars and the probability for each sentence are given. Suppose that an individual who uses G_i utters a sentence in $L(G_i)$ with a uniform probability, then s_{ij} is the number of common sentences between $L(G_i)$ and $L(G_j)$ divided by the number of sentences in $L(G_i)$. In this sense, every diagonal element, that is s_{ii}, is always 1, and s_{12} is 6/12 in Table 1.

3.3 The Learning Algorithm

We give the learning algorithm with the following restrictions:

- New born children do not have any particular grammar, thus we do not give any initial values for the parameters.
- Children who contact only with their parents must obtain the parental (target) grammar. It is not guaranteed in TLA; it depends on the transitivity of the grammars.
- The learning period should be given enough for the estimation of the target grammar.

We introduce the simplest learning algorithm which satisfies the above restrictions. Children are required to accumulate the number of acceptable sentences from each grammar. The learning algorithm becomes as follows (See also Fig. 3):

1) A child receives a sentence uttered by an adult.
2) For each grammar, if a sentence is acceptable for the child, the grammar scores a point in her memory.

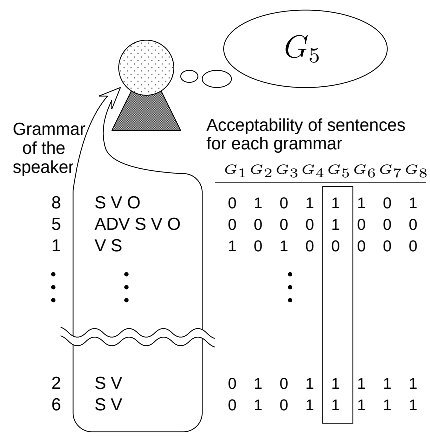


Fig. 3 The simplest learning algorithm

3) 1) and 2) are repeated until the child receives a fixed number of sentences that is regarded as enough for the estimation of the grammar.
4) The child adopts the grammar with the counter of the highest score.

In case a child talks only with her parents, the grammar of the child would be G_{j^*} by the second restriction of the learning algorithm, as:

$$\begin{aligned} j^* &= \operatorname*{argmax}_j \{\eta s_{pj}\} \\ &= \operatorname*{argmax}_j s_{pj} \quad (= p), \end{aligned}$$

where η denotes the number of sentences the child receives, and p is the index of the parental grammar. Because s_{pj} is regarded as the acceptability of G_j when the number of sentences of G_p are given,

ηs_{pj} is the expected value of the score for G_j when η is sufficiently large.

In general, a child has chances to talk with other adult individuals in proportion to the populations. The expected grammar of the child should be G_{j^*} by such learning algorithm that:

$$\begin{aligned} j^* &= \underset{j}{\text{argmax}}\{\sum_k \eta s_{kj} x_k(t)\} \\ &= \underset{j}{\text{argmax}}\{\sum_k s_{kj} x_k(t)\}. \end{aligned}$$

Similarly, because the child listens to sentences from all the individuals in the community, the expected value of the counter for G_j becomes $\sum_k \eta s_{kj} x_k(t)$. Then the child would adopt a grammar of the highest score. Here, we introduce the exposure probability α, that prescribes the ratio a child talks to people other than her parents. Thus, the choice of grammar would depend on the linear combination of the above two equations, i.e., the estimated grammar of the child is G_{j^*} such that:

$$j^* = \underset{j}{\text{argmax}}\{\alpha \sum_k s_{kj} x_k(t) + (1-\alpha) s_{pj}\}. \quad (3)$$

3.4 The Modified Accuracy Matrix

Suppose first that there are only two grammars G_1 and G_2 and a child whose parental grammar is G_1, the child acquires G_1 when

$$\begin{aligned} &\{\alpha(s_{11}x_1(t) + s_{21}x_1(t)) + (1-\alpha)s_{11}\} \\ \geq &\{\alpha(s_{12}x_2(t) + s_{22}x_2(t)) + (1-\alpha)s_{12}\}. \end{aligned}$$

Because the right-hand side of the above is unknown to an individual and simply regarded uniform between 0 and 1, the probability of the choice of G_1 would be the left-hand side itself. In the similar way, the probability that the child i chooses G_j in n candidate grammars is supposed to be:

$$P_{ij} = \{\alpha \sum_k s_{kj} x_k(t) + (1-\alpha) s_{ij}\}^{n-1}. \quad (4)$$

Hence, the probability in which a child whose parents have G_i learns G_j, that is $\overline{q}_{ij}(t)$, is as follows:

$$\overline{q}_{ij}(X(t)) = \frac{(\alpha \sum_k s_{kj} x_k(t) + (1-\alpha) s_{ij})^{n-1}}{\sum_l (\alpha \sum_k s_{kl} x_k(t) + (1-\alpha) s_{il})^{n-1}}. \quad (5)$$

4 Experiments

In this section, we show the experimental result of the language dynamics equations of population-based transition mentioned in Section 2.3 and 3. Variables in eqn (2) are the distribution of populations $X(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ and the parameter is the exposure probability α.

4.1 Population Dynamics toward Creolization

In Fig. 4, we show the result of the population dynamics. We gave initial distributions of populations as $x_1(0) = 0.32, x_3(0) = 0.32, x_5(0) = 0.36$, and other $x_i(0)$'s were set to 0. The exposure probability α is examined at the range from 0 to 1.

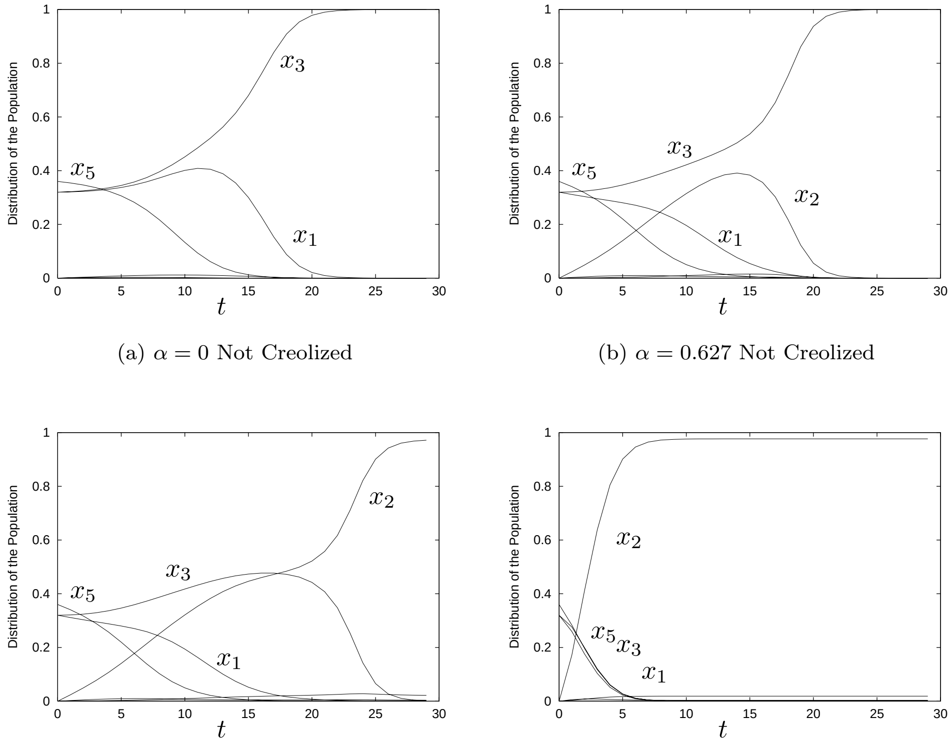
In case $\alpha = 0$, children learn a language only from their parents. In fact, Fig. 4(a) represents that a population of G_3 given at the initial state eventually occupies the distribution in the community. Because the modified accuracy matrix $\overline{Q} = \{\overline{q}_{ij}\}$ is constant in eqn (2), the system behaves the same as eqn (1), that is proportional to the similarity between grammars and to the distribution of population. In consequence, both of x_1 and x_5 decrease in accordance with the diminishing fitness.

According to the increase of α, x_2 rises gradually though $x_2(0) = 0$ (See Fig. 4(b)). This means that some individuals come to speak a grammar that no one spoke at the initial state. Because the value of q_{12} in eqn (2) especially rises for large α, x_1 comes to flow into x_2. This is the emergence of creole. The result shows that the individuals tend to choose a grammar of the highest communicability in the community rather than the parental (target) grammar. In case Fig. 4(b), that is $\alpha = 0.627$, because x_2 did not increase any longer after x_1 and x_5 mostly flowed out, the language G_2 declined and eventually disappeared.

However, at this sheer boundary of $\alpha = 0.628$, the language G_2 comes to be dominant (See Fig. 4(c)). In a series of the experiments, G_2 keeps dominant and stabler while α comes near to 1 (See Fig. 4(d)). Moreover, these two figures represent that the higher α makes the convergence shorter. From the definition of α, it is admitted the condition of $\alpha = 1$ is the easiest for creole to emerge. From the viewpoint of population dynamics, creole is defined as such a grammar G_c that:

$$x_c(0) = 0, \; x_c(t) > \theta_c, \quad (6)$$

where x_c denotes the distribution of the population of G_c, and θ_c denotes a certain threshold to be regarded as creole [1, 3]. Thus, we can regard our experimental result is a creolization.



(a) $\alpha = 0$ Not Creolized

(b) $\alpha = 0.627$ Not Creolized

(c) $\alpha = 0.628$ Creolized

(d) $\alpha = 1$ Creolized

Fig. 4 Result of population-based transition

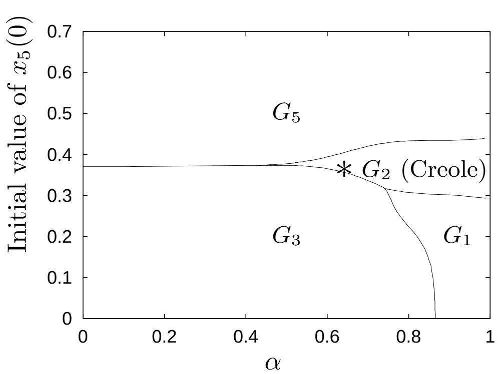


Fig. 5 Regions of the dominant grammar

4.2 Regions of the Dominant Grammar

In Section 4.1, we could represent an example of creolization. The occurrence of creole depends on the initial value of the distribution and the exposure probability α. Actually, the α value and the distribution of population decide which language would be dominant, regardless whether creolization occurs or not. The next experiment aims at drawing a map as to which language would be dominant in various initial values and the parameter values.

In order to show the boundary conditions clearly, we parameterize the population rate of G_5, that is $x_5(t)$ from 0 to 1. The other two groups $x_1(t)$ and $x_3(t)$ are set to be equal as follows:

$$x_1(0) = x_3(0) = (1 - x_5(0))/2.$$

In this α-$x_5(0)$-space, we could detect boundaries between which of two different languages become dominant. The result is shown in Fig. 5. In the figure, solid lines denote the boundaries between would-be dominant languages. Creole(G_2) appears while the other grammars are existent at the initial stages. The asterisk in Fig. 5 corresponds to the case of Fig. 4(c).

Looking the map closely, we could see the creolization in the right-middle area as the value of α becomes higher. In the upper region of the figure, G_5 comes to dominant because $x_5(0)$ is high. When the $x_5(0)$ value is low, G_1 and G_3 share the region, because the $x_1(0)$ and $x_3(0)$ are equivalently given at the initial stage. Because G_1 is more similar to G_2 (creole) than G_3, that is $s_{12} = s_{21} = 6/12$

against $s_{13} = s_{31} = 2/12$, x_1 often comes to high together with x_2, for G_1 earns more fitness than G_3 when x_2 increases. Once creole emerges at an α value on a fixed $x_5(0)$, it has never disappeared while α is larger than the value. Thus, the condition of $\alpha = 1$ is the easiest for creole to emerge.

In summary, the transition is strongly affected by the similarity between grammars, the distribution of population, and the exposure probability α.

5 Conclusions

In this paper, we have argued that language acquisition is strongly affected by the population of the languages in the community. First, we discussed that the language dynamics equations proposed by Komarova et al. [7] has the problem that the probability transition matrix Q is given as constant through generations. Next in the study of Niyogi [9], we mentioned what phenomena are reflected in $q_{ij} > 0$ $(i \neq j)$.

Considering the above, we proposed that (i) $Q = \{q_{ij}\}$ must be sensitive to the distribution of population at each generation, and also (ii) Q must be quite near to the unit matrix, that is, $q_{ii} \simeq 1$ and $q_{ij} \simeq 0$ where $i \neq j$. From these observations, we revised the language dynamics equations, introducing $x_i(t)$'s that is the population rate at each generation and the exposure probability α. In order to show the adequateness of our equations, we have shown experimental results. That is, when α is rather large, a new language appears that is most communicative for every individual. We regarded this phenomena as the emergence of creole. Next, we searched for the boundary conditions in which creole emerges, in the two-dimensional space of α and the ratio of the most populous language.

The relationship among languages is expressed by the S matrix in the present population dynamics. If we assumed that the value for each element of the S matrix is continuous, we could further parameterize $7 \times 7 = 49$ variables which consist of elements of the S matrix except to the diagonal elements. Namely, Fig. 5 would merely denote a two-dimensional surface of a body in the 51-dimensional space. Furthermore, the other 7 candidate parameters instead of $x_5(0)$ are regarded as the vertical axis like Fig. 5. Incidentally, we have found some combinations of elements of the S matrix in which creole never occur with any distribution of population, although they have not been well explained yet. We need to study in further generalized conditions, to clarify the boundary conditions of creolization.

Acknowledgement

We give our thanks for Jim Hurford and Hajime Yamauchi for their helpful comments.

References

[1] Bickerton, D.: Language and Species, The University of Chicago Press (1990)

[2] Briscoe, E.J.: Grammatical Acquisition and Linguistic Selection, In: Briscoe, T. (ed.): Linguistic Evolution through Language Acquisition: Formal and Computational Models, Cambridge University Press (2002)

[3] DeGraff, M. (ed.): Language Creation and Language Change, The MIT Press (1999)

[4] Gibson, E., Wexler, K.: Triggers, Linguistic Inquiry, 25 (1994)

[5] Gold, E. M.: Language identification in the limit, Inform. Control 10, pp.447-474 (1967)

[6] Hashimoto, T., Ikegami, T.: Evolution of symbolic grammar systems, In: Moran, F., Moreno, A., Merelo, J.J., Chacon, P. (eds.): Advances in Artificial Life. Springer, Berlin (1995) pp.812-823

[7] Komarova, N.L., Niyogi, P., Nowak, M.A.: The Evolutionary Dynamics of Grammar Acquisition, J.Theor.Biol. **209** (2001) 43-59

[8] Nakamura, M., Tojo, S.: The Emergence of Artificial Creole by the EM Algorithm, Proceedings of the Fifth International Conference on Discovery Science (DS2002), Lecture Notes in Computer Science 2534, Springer (2002) pp.374-381

[9] Niyogi, P.: The Informational Complexity of Learning from Examples, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA (1994)

[10] Niyogi, P., Berwick, R.: The logical problem of language change. Technical Report AI Memo 1516 / CBCL Paper 115, MIT AI Laboratory and Center for Biological and Computational Learning, Department of Brain and Cognitive Sciences.

[11] Nowak, M.A., Komarova, N.L.: Towards an evolutionary theory of language, Trends in Cognitive Sciences **5**(7) (2001) pp.288-295.

[12] Ono, T., Tojo, S., Sato, S.: Common Language Acquisition by Multi-Agents, International Computer Symposium (ICS'96), Proceedings on Artificial Intelligence (1996) pp.218-223

[13] Pinker, S.: The Language Instinct, W. Morrow & Co. (1994)