

言語動力学によるクレオール化のシミュレーション —言語獲得環境への依存性—

中村 誠[†], 橋本 敬^{††}, 東条 敏[†]

北陸先端科学技術大学院大学 { [†] 情報, ^{††} 知識 } 科学研究科

言語進化の分野において、クレオールといわれる言語現象が取り扱われている。本研究の目的は、クレオールが発現するための言語能力と社会、言語環境についての条件を数理科学的な側面から一般化された形で定式化することである。我々はこれまでに M.Nowak らが提唱する言語動力学を修正し、他言語話者との接触に関するパラメータを追加した。これにより、言語動力学上でクレオールの発現を観察し、そのために必要とされる言語間の類似性に関する条件を導き出している。本稿では言語学習者に焦点を当て、クレオール化の学習環境への依存性を検証する。実験により言語獲得期における言語入力が少ないとクレオールが発現しやすいという結果が得られた。

The Simulation of Creolization on Language Dynamics — Dependency on the Environment on Language Acquisition —

Makoto NAKAMURA[†], Takashi HASHIMOTO^{††} and Satoshi TOJO[†]

Graduate School of { [†] Information, ^{††} Knowledge } Science,
Japan Advanced Institute of Science and Technology

Creole is one of the main topics in the fields concerning the language change and evolution. The purpose of this paper is to develop a new formalism of language dynamics so that creole may emerge. Thus far, we modified the language dynamics equation by Nowak et al. so as to change in accordance with the distribution of population of each language at each generation. Thus, we could observe creolization under limited conditions. In this paper, focusing on language learners, we investigate the relationship between the environment on language acquisition and creolization. In the experiments, we show that creolization is easy to occur when language learners receive a small amount of language input.

1 はじめに

社会言語学や発達言語心理学などの分野において、ピジンやクレオールといわれる言語現象について盛んに研究が行われている [1]。ピジンとは共通の言語を持っていないが、通商その他の目的で互いにコミュニケーションをとる必要がある人々の間に発達した伝達システムである。その語彙や文法は単純化されており、構造や言語使用の点ではるかに発達している土地の言語と並んで、補助言語として習得される。その後ピジンが発展し、ある共同体の母語となったものはクレオールと呼ばれる。クレオールは、元となったどの言語とも文法的に異なり、それ自身が文法的にも表現能力としても充実した土地の言語となる。ピジンやクレオールは世界中で発見されており、それぞれが独

自に発達した言語体系であるにも関わらず、非常に似通った特徴と文法構造を持っていることが大きな特徴である。

ピジンとクレオールは言語進化の観点からも興味を持たれている。言語の進化とは、生物進化、個人の学習、文化進化という3つの時間スケールが相互作用し、共進化するシステムである。また、集団で使用されている言語ルールを獲得し、そのルールに制約されて言語を使用するが、その獲得や使用自体が社会での言語的ルールを改変するというルールダイナミクスも考えられ、図1のようなダブルループ・ダイナミクスとして概念化することができる [2]。図中の3つの要素に囲まれた外側のループと、生物進化を除いた内側のループの間ではタイムスケールが大きく異なる。クレオール化の問題は、内側のループで論じることができる問

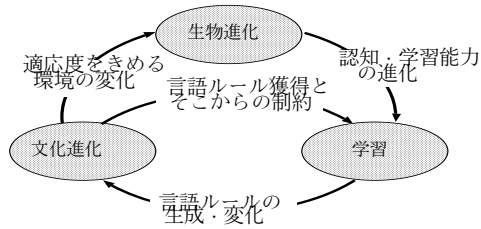


図 1: ダブルループ・ダイナミクス [2]

題である。

本研究の目的は、クレオールが発現するための言語獲得能力と社会、言語環境についての条件を数理科学的に導出し、言語変化の過程を通時的な側面から一般化された形で定式化することである。これまで言語学的な側面からビジンやクレオールは調査されてきたが、これらが発現する条件を実際のクレオールから厳密に求めることは現実的に不可能であった。しかし近年における言語進化に関する研究は、構成論的手法によってめざましい発展を遂げている [3]。これらの研究成果に基づいて、これまでのクレオールの実地調査からは得られなかった、クレオールが発現するための言語話者の人口構成比や、言語獲得環境に関する条件等を計算機上のシミュレーションによって導き出すことが可能となる。我々はこれまでに Nowak et al. [4] が提案した言語動力学モデルを修正し、現実のクレオールが発現する環境をモデル化し、実験によってその過程を観察した。また、クレオールの発現に必要なとされる言語間の類似性に関する条件を導き出している [5]。さらに、子供の言語獲得期における言語入力量に関するパラメータを動力学モデルに導入し、言語入力十分である環境のもとでクレオールが発現する過程を定量的に分析した [6]。本稿では、言語獲得に必要な言語入力量について再検証し、言語入力とクレオールの関係について言及する。

以降、2 節で Nowak et al. の言語動力学の説明をし、我々のモデルの導入と修正点について述べる。3 節では学習アルゴリズムについて述べる。4 節で言語動力学上でクレオールがどう扱われるか説明し、5 節で実験を行なう。最後に 6 節でまとめを行なう。

2 言語動力学

Nowak et al. [4] は言語獲得に関する理論を発展させることを目的として、数理生態学的な人口動

力学 [7] を応用したモデルを提案している。これを言語動力学 (Language Dynamics Equations) と呼ぶ。彼らは子供の第一言語獲得が系全体で話される言語にどのように影響を与えるかについてモデル化を行なった。そこでは大人を言語話者、子供を言語学習者とし、子供は親からのみ言葉を聞いて文法を推定する。実験により、言語学習者の学習精度と集団内で使用される言語が支配/均衡する関係を導き出している。

言語動力学では、普遍文法、特に原理とパラメータ理論を仮定している。原理とパラメータ理論とは、普遍文法はすべての人間言語に共通な原則の体系、すなわち原理 (principle) とそれに付随するパラメータ (parameter) からなると仮定し、子供の文法獲得は普遍文法の原理に組み込まれたパラメータの値を言語経験により固定する過程であるとする考え方である [8]。すなわち、原理によって与えられた文法の探索空間は有限であると仮定し、言語話者が用いる文法は $\{G_1, \dots, G_n\}$ としてあらかじめ定義される。すると言語の変化とは、言語話者が所有するパラメータの変化を示しており、その変化はそれぞれのパラメータ値に対応する言語間の人口遷移によって表現される。ここで、ある言語集団において文法 G_i を持つ言語話者の人口比率を x_i 、すなわち $\sum_{i=1}^n x_i = 1$ であるとする、 x_i の時間変化は動力学系として表わされる。

また、Nowak et al. は、言語動力学において言語が持つ性質を次のように与えている：

類似性 言語間の類似度を表わす。任意の 2 つの言語間の類似度を示す確率行列 $S = \{s_{ij}\}$ で表わされる。要素 s_{ij} は、 G_i 話者がランダムに文を発話したときの文法 G_j を持つ聞き手に理解される確率として求められる。

遷移性 言語間の遷移確率を表わす。 G_i 話者である親の子供が文法 G_j を獲得する確率は行列 $Q = \{q_{ij}\}$ で表わされる。

我々はこれまでに、この言語動力学を修正し、実際にクレオールが発現する環境をモデル化した。具体的には、Nowak et al. のモデルが子供は親からしか言語を学習しないと仮定していたのに対して、集団内の親以外からも言語入力を受けるようにした。我々が提案する動的遷移行列モデルは次の式で表わされる：

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

Nowak et al. のモデルとの違いは次の通りである :

1. 集団内において言語が通じる尺度を示す適度関数を排除した [9]. すなわち, 言語話者である大人は子供を産むが, 生まれる子供の数は使用言語に依存せず一定である. これにより, 図 1 のダブルループのうち, 内部ループについて言及しているものであると考えることができる.

2. 定数として与えられていた遷移確率 $Q = \{q_{ij}\}$ を, 集団内の人口構成比に依存し時間変化する関数 $\bar{Q}(t) = \{\bar{q}_{ij}(t)\}$ とした. これを動的遷移行列と呼ぶ.

動的遷移行列 $\bar{Q}(t)$ は, 学習アルゴリズムに依存する. 次節において学習アルゴリズムと動的遷移行列について述べる.

3 学習アルゴリズムと動的遷移行列

本節では, 新たな遷移行列 $\bar{Q} = \{\bar{q}_{ij}(t)\}$ を提案する. これは Nowak et al. のモデルでは定数として与えられていたものである. 子供の言語獲得期において, さまざまな言語話者と接触する環境と, それに影響を受けて文法を獲得するような学習アルゴリズムを与え, これらを反映した動的遷移行列 \bar{Q} を提案する.

3.1 言語話者との接触状況

子供は言語獲得期において, 親やその他の大人の言語話者と接することによって文法を獲得する. 使用言語がひとつしかないコミュニティでは, 親以外の言語話者も親と同じ言語を使用するため, 目標文法はひとつに定まる. しかし, さまざまな言語を母語とする話者が混在するコミュニティでは, 子供は多くの言語と接することとなり, 結果として最も効率的に入力文を受理するような文法を獲得することが考えられる.

他の言語と比べて親が話す言語に子供がどれだけ接したかということ推量するため, 言語入力を次のような 2 つのカテゴリに分けて考えることとする. ひとつは親からの入力, もうひとつは親以外の話者からの言語入力である. 特に後者の比率を接触比率 α (exposure rate α) と呼ぶこととする. 親以外の話者との接触に関して, それぞれの言語と接触する比率はそれぞれの言語話者の人口比に比例すると仮定することにより, この α から子供が接触するそれぞれの言語の比率を導き出すことができる. ある子供が接触する言語の分布を

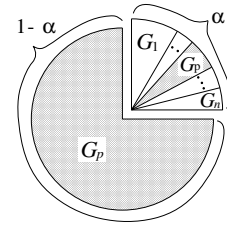


図 2: 接触比率 α

図 2 に示す. 親が G_p 話者であるこの子供は, 影がかかった部分の割合で言語 $L(G_p)$ と接することとなる. すなわち $\alpha x_p + (1 - \alpha)$ である. そのほかの母語以外の言語 G_i と接する比率は αx_i である.

3.2 学習アルゴリズム

本モデルは Nowak et al. のモデルと同様, 普遍文法を仮定している. すなわち, 人間が話すことができる文法の数は有限であり, 言語学習とはその中から獲得するのに最もふさわしいと考えられるものをひとつ選択することに過ぎない. 本モデルにおける学習アルゴリズムは Niyogi [10] が指摘した問題を解消したものである. Niyogi はトリガー学習アルゴリズムによる学習過程をマルコフ遷移図で示し, 言語学習を行なう子供は親からしか言語入力を受け取らないにも関わらず, 学習開始時の初期状態によって親の言語を正常に獲得できない子供がいることを示した. これを受けて, 我々は学習アルゴリズムに次のような制約を設けた.

- 言語入力として親の言語のみを受け取った場合, 親の言語を獲得する.
- 複数の言語と接触した場合においても, 入力文に応じて最もふさわしい文法を採用する.

次にあげるアルゴリズムがこの学習のひとつであると考えられる :

1. 子供の頭の中に, それぞれの文法ごとにスコア表があると考え.
2. 子供は大人が発話した文を受け取る.
3. 全ての文法について, その文が受理可能かどうか識別し, その文が受理できた文法は, スコアに 1 ポイント加算される.
4. 子供が言語獲得に十分な量と考えられる入力文を受け取るまで (w 文) 2. と 3. を繰り返す.

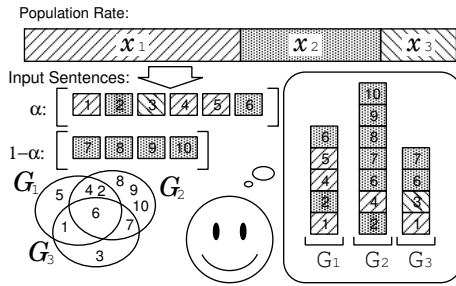


図 3: 学習アルゴリズム

5. その子供は最も高いスコアを獲得した文法を採用する。

各言語に対する話者の人口比率と接触比率 α により、子供が接触する言語の比率が求まる。また、類似行列 S は、それぞれの文法について入力文の受理能力を決定するものである。図 3 は G_2 話者の子供が学習の結果 G_2 を獲得した例である。学習期間において、子供は、図中 1 から 10 と番号がふられた文を受け取る。入力文は接触比率 α にしたがって、2 つのカテゴリに振り分けられる。ひとつは親以外の言語話者が発話した文であり、もう一つは親が発話したものである。前者において、それぞれの文は特定の文法を持った言語話者から発話されたものであるが、どの言語話者の発話であるかはそれぞれの言語話者の人口比率によって決定される。図 3 の例は、子供は文 1, 4, 5 を G_1 話者から、文 3 を G_3 話者から聞いていることを表わしている。その他の文は全て G_2 話者からである。したがって、これらの文は特定の文法で必ず受理できるものである。言語学習者である子供は、入力文のうち、それぞれの文法で受理できる文の数をカウントし、最も多くの文を受理した文法を採用する。例えば、文 1 は G_1 話者が発話したものであるが、 G_3 でも受理することができるため、 G_1 と G_3 のカウンタにそれぞれ 1 を加える。どの文がどの文法で受理することができるかということが図中のベン図で示されており、 S 行列はこの関係を言語間の類似度として表わしたものである。

3.3 動的遷移行列

子供は、学習期間中に受け取った入力文を最も多く受理することができる文法を $\{G_1, \dots, G_n\}$ の中から選択する。動的遷移行列 $\bar{Q} = \{\bar{q}_{ij}(t)\}$ は、 G_i 話者の子供が G_j を獲得する確率であり、この学習アルゴリズムから導き出すことができる。

子供は、接触比率 α と言語話者比率に比例してそれぞれの言語話者である大人から入力文を聞くと仮定する。 G_i 話者を親に持つ子供がある一文を聞いたとき、それが G_j で受理できる確率は次のように表すことができる：

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

子供が聞く入力文の数を w とすると、 G_i 話者の子供が G_j を仮定したときに、そのうち r 文受理できる確率は次のような二項分布を用いて表すことができる：

$$g_{ij}(r) = \binom{w}{r} (U_{ij})^r (1 - U_{ij})^{w-r} . \quad (3)$$

また、 G_j で受理できた文の数が r 未満である確率は次の通りである：

$$h_{ij}(r) = \sum_{k=0}^{r-1} \binom{w}{k} (U_{ij})^k (1 - U_{ij})^{w-k} . \quad (4)$$

これらの確率分布を用いることにより、 G_i 話者の子供が G_j を仮定したときに w 文中 k 文受理し、かつ、その他の文法を仮定したときは $k-1$ 文以下であった確率は $g_{ij}(k) \prod_{l=1, l \neq j}^n h_{il}(k)$ で表わされる。 G_i 話者の子供が w 文聞いたあとに G_j を獲得するためには、 n 個ある文法のうちで G_j が最も効率的な文法でなければならない。すなわち、 G_j は少なくとも w 文ある入力のうち少なくとも $\lceil \frac{w}{n} \rceil$ 文は受理できなければならない。なぜなら、受け取った文のそれぞれが、ある言語話者から発話されたものであることから、少なくともその文法で受理することができるため、 w 文の入力のうちで、 $\lceil \frac{w}{n} \rceil$ 文以上受理できた文法が必ず存在するからである。いいかえると、もし G_j が受理した文の数が $\lceil \frac{w}{n} \rceil$ に満たなかった場合、その子供は G_j を採用しない。ここから、遷移確率 \bar{q}_{ij} は G_j が $w, w-1, \dots, \lceil \frac{w}{n} \rceil$ 文受理できたときの確率の和として求めることができる。したがって、動的遷移行列 $\bar{Q}(t) = \{\bar{q}_{ij}(t)\}$ は次のような式で表わされる：

$$\begin{aligned} \bar{q}_{ij}(t) &= \frac{\sum_{k=\lceil \frac{w}{n} \rceil}^w \left\{ g_{ij}(k) \prod_{\substack{l=1 \\ l \neq j}}^n h_{il}(k) + R_{ij}(k, n) \right\}}{\sum_{m=1}^n \left[\sum_{k=\lceil \frac{w}{n} \rceil}^w \left\{ g_{im}(k) \prod_{\substack{l=1 \\ l \neq m}}^n h_{il}(k) + R_{ij}(k, n) \right\} \right]} , \end{aligned} \quad (5)$$

ここで $R_{ij}(k, n)$ は 2 つ以上の文法が同じ数の入力文を受理し, その結果 G_j を獲得する確率の総和である. m 個の文法が同じ数だけ受理した場合, G_j を獲得する確率は, その出現確率を m で割った値として求められる. 文法の数 $n = 3$ であったとき, $R_{ij}(k, n)$ は次のように表わされる:

$$\begin{aligned} R_{ij}(k, 3) &= \frac{1}{2}\{g_{ij}(k)g_{ij_2}(k)h_{ij_3}(k) + g_{ij}(k)h_{ij_2}(k)g_{ij_3}(k)\} \\ &+ \frac{1}{3}\{g_{ij}(k)g_{ij_2}(k)g_{ij_3}(k)\} \\ &\quad (j_2, j_3 \in \{1, 2, 3\}, j \neq j_2, j_3, j_2 \neq j_3). \end{aligned} \quad (6)$$

4 言語獲得モデルとクレオール

本節では, 人口動力学上のクレオールの定義と, 言語入力と言語獲得の関係について述べる.

4.1 人口動力学上のクレオールの定義

我々は普遍文法の立場から, 言語獲得のメカニズムとしてあらかじめ原理とパラメータを仮定している. そして, 普遍文法にはクレオールも含まれるという立場をとり [11], パラメータセッティングのレベルでのクレオールの出現を想定している. すなわち, クレオールの文法を G_c とすると, 原理によって制限される文法集合は $\{G_1, \dots, G_c, \dots, G_n\}$ となる.

ここで我々は, これまでの言語学的な定義 [12, 13] とは大きく異なるが, 人口動力学の視点からみたクレオールの定義をする. クレオールとは, 新しい言語の発現である. すなわち, あるコミュニティで初期に存在しなかった言語が, 後に存在するようになる現象と考えることができる. よって, 優勢クレオールとは, 次のような言語 $L(G_c)$ であると定義される:

$$x_c(0) = 0, \quad x_c(t) > \theta_d. \quad (7)$$

ここで $x_c(t)$ は人口動態が収束し, 安定した t 世代における G_c 話者の人口比率, θ_d は優勢 (dominant) 言語であるとみなすための人口比率の閾値を示し, 本稿では $\theta_d = 0.9$ としている. これらの定義は, 初期状態では誰も話していなかった言語が, 最終的にはそのコミュニティでほとんどの人に使用される状態を表わしている.

4.2 言語入力と言語獲得の関係

これまでに述べたモデル ((1) 式から (6) 式まで) は, 定数として $S = \{s_{ij}\}$, α , w を持ち, これらをパラメータ化することによって動力学の振

る舞いを観察することができる. このうち, 類似行列 $S = \{s_{ij}\}$ は言語の特性, 接触比率 α および言語入力量 w は子供の学習環境をそれぞれ表わしている.

本モデルでは, 子供の文法獲得とは有限の文法集合 $\{G_1, \dots, G_n\}$ の中から選択することに等しい. また, 言語の特性として特定の文法ルール等を設定することなく, 言語間の類似性を S 行列で表わしている. つまり, 文法構造を形成していく過程を動力学上で表現することが原理的に不可能であり, したがって, 言語入力がどんなに少なくても, 子供は完全な言語をひとつ獲得することができる. この問題は, 普遍文法を仮定した他の多くの言語獲得モデル [14, 15] が共通して抱える問題である.

現実の言語獲得に関しては, Genie の例 [16] のように, 生まれてから言語入力がほとんどないまま成長すると, 人間言語のような複雑な文法を持った言語を正常に獲得できないだろう. この問題は, 完全な言語を選択するというを前提とした言語動力学で扱うことはできない. また, 言語獲得の途中段階にある子供が話すことばは, 二重否定の使用や活用の省略など, クレオールに見られる特徴を持っているといわれており, Bickerton は, もしその子供がそれ以降言語入力が与えられずに成長すれば, クレオールに非常に近い言語を獲得するだろうと述べている [17]. しかし, クレオールは言語獲得期の途中段階において獲得されるような劣った言語ではないため, クレオールの発現についても適切な言語入力量を考慮する必要がある. 本モデルで言語入力が極端に少ない状況下での言語獲得問題を扱うことは, 現実の問題に則していないため, たとえそのときに新言語が現れたとしても, 現実のクレオール化の過程をモデル化したとはいえない. そのため, 子供の言語獲得に関して, 言語入力量は十分にあるという前提のもとで実験を行なう必要がある. そのために妥当と考えられる言語入力量 w を求める必要がある. したがって, 4.1 節で人口比率の変化によって発現した新言語をクレオールと定義したが, それに加えて, 言語学習者の言語入力に関してもクレオールであるとみなすための条件が必要であると思われる.

5 実験

クレオールが発現するための最も単純なモデルとして 3 言語の場合を考える. すなわち 2 つの言語集団が接触した結果, 第 3 の言語としてクレオー

ルが発現する可能性がある環境である。3つの言語の初期の人口配分を $(x_1, x_2, x_3) = (0.5, 0.5, 0)$ とすることにより, G_1 と G_2 を既存言語, G_3 をクレオールとみなし, これが優勢言語となるための条件を考える。ここでは S 行列を次のような対称行列とした:

$$S = \begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \quad (8)$$

要素 a は既存の2言語間の類似度であり, b, c はそれぞれ G_1 とクレオール, G_2 とクレオールの間の類似度を表わしている。ここではこの類似行列の値を $(b, c) = (0.45, 0.35)$ と恣意的に割り当てた。以降, 既存の言語間の類似度 a , 接触比率 α , 言語入力量 w をパラメータとして, 動力学的振る舞いを観察する。

5.1 言語間の人口遷移

世代の変化に対するそれぞれの言語話者の人口変化を図4に示す。ここでは $a = 0$ とした, すなわち既存の2言語の間では, 話者同士で会話が全く成立しない状況を表わしている。

図4(a)および図4(b)は $\alpha = 0$ とし, 言語話者である子供が, 親からしか言語を学習しないときの人口変化を表わしている。このとき, 子供は学習期間中にそれぞれ5個と30個の文を受け取り, 文法を推定する。子供は学習期間中に親からしか入力文を受け取らず, さらに受け取った文はもう一方の既存の言語の文法では受理できないため, G_1 話者の子供が G_2 を獲得することはない。しかし, 第3の言語である G_3 はそれぞれの言語と類似した言語であるため, G_1 話者の子供は G_3 を獲得する可能性がある。その条件は, G_1 話者の子供が受け取った入力文の全てが G_1 で受理できるために, 受け取った入力文の全てが G_3 でも受理できた場合に限られ, その場合1/2の確率で G_1 話者の子供は G_3 を獲得する。同様に G_2 話者の子供も G_3 を獲得する可能性がある。また G_3 を獲得した子供は次世代では親となり, その子供は G_3 だけではなく, G_1, G_2 を獲得する可能性を持っているのである。その可能性は入力文の数が少ないほど高まり, 言語間の人口の遷移は活発になることが考えられる。

図4(a)と図4(b)は, それぞれ入力文の数が $w = 5$, $w = 30$ のときの結果である。 $w = 5$ のときは各言語の人口比率がほぼ同じになり, 最終的

に3言語が共存することを示している。このとき, G_1 と G_3 の類似度 (b) が, G_2 と G_3 のそれ (c) よりも高いため, G_1 話者の子供は比較的 G_3 を獲得しやすい。そのため, 初期において G_1 話者の人口が G_2 に比べて急激に G_3 に移行しているのがわかる。 $w = 30$ のときは, 入力文数の多さから子供が親の言語を獲得する確率が高まるため, 人口が遷移しにくい。したがって, $w = 30$ のときも人口が遷移し, 3言語が共存することが考えられるが, 100万世代追跡しても, ほとんど人口の変化が見られなかった。

入力文の数を $w = 30$ とし, 接触比率 α を増加させたときの結果を図4(c)に示す。接触比率は子供の言語獲得期において, 親以外の言語話者と接触する割合を表わしたものである。この割合を増加させると, 子供は親が持っている文法では受理できない文を受け取る可能性がでてくるため, 入力文を最も多く受理できる文法を採用することとなる。したがって, α の増加は人口遷移をより活発にするものであると考えられる。図は $\alpha = 0.668$ のときの結果であり, 初期において誰も話していなかった言語 G_3 を, 時間が経過するにつれてほとんどの話者が話すようになったことを示している。このときの G_3 がクレオールである。クレオールが発現するかどうかは類似行列 S , 接触比率 α , および言語入力量 w に依存し, その条件は限られている [6]。

この実験により, 初期条件によって複数の言語が共存するか, ひとつの言語が優勢となるかが異なるということを示した。特に言語入力量が少ないときは, 多いときと比べて親の言語が子供に継承されにくいことを示した。

5.2 言語入力量と優勢言語

言語動力学モデルは, 複数の言語が共存するか, ひとつの言語が優勢言語となって収束する。次の実験では, 接触比率 α と既存の言語間の類似度 a をパラメータとして, 最終的にどの言語が優勢となるかを調べた。結果を図5に示す。それぞれの図に優勢言語とその領域が示されている。アスタリスク (*) は, 優勢言語が現れずに複数の言語が共存するか, 100万世代まで計算しても収束しなかった領域である。このときの言語入力量は $w = 30, 20, 10$ である。以降, 言語入力量の差が, 優勢言語の分布にどのように影響を与えるかについて考察する。

言語入力量を $w = 30$ としたときの分布図を図5(a)に示す。図4(c)に示したクレオール化は

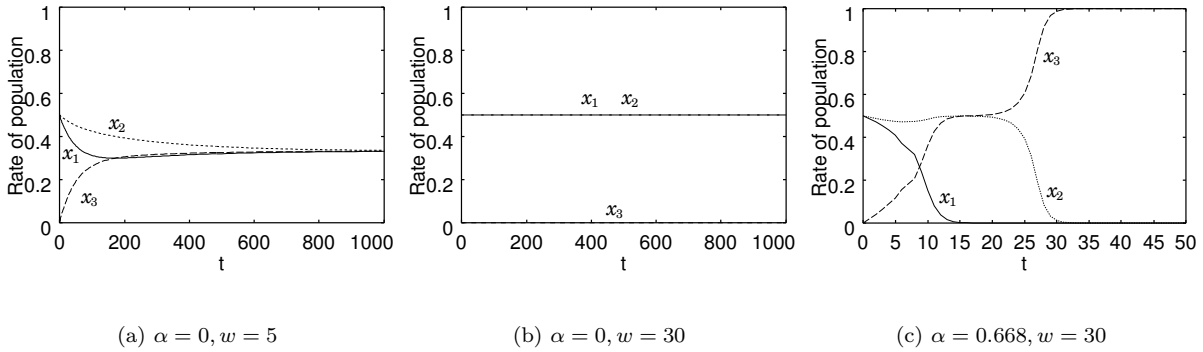


図 4: 3 言語での人口の変化 ((a, b, c) = (0, 0.45, 0.35))

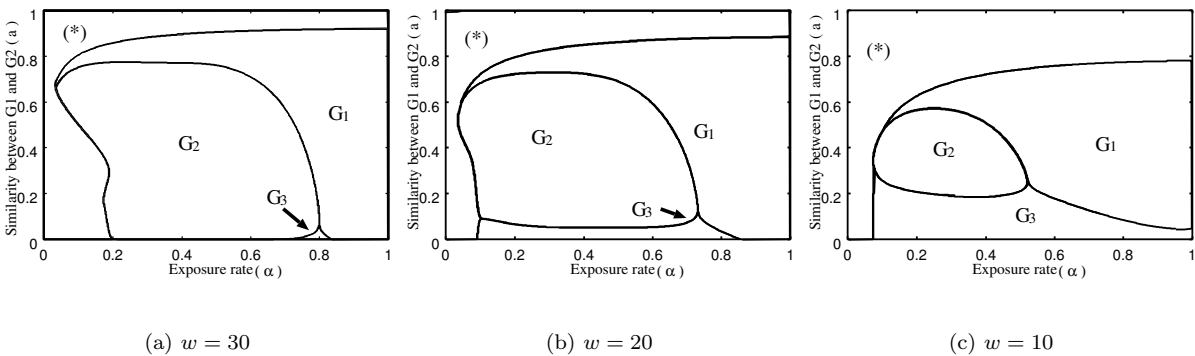


図 5: 接触比率と言語の類似性に対する優勢言語の分布図 ((a, b, c) = (a, 0.45, 0.35))

$(\alpha, a) = (0.668, 0)$ の点で追跡したものである。既存の言語間の類似度 (a) が増加すると、これら 2 言語の間で人口の遷移が活発になり、新言語が発現する可能性は少なくなる。また、 $a \geq 0.9$ の領域は、既存の 2 言語はほとんど同じ言語であり、この 2 言語が共存するため優勢言語は存在しない。

言語入力が少なくなるとクレオール (G_3) の領域が増加することが図 5(b) と図 5(c) を見るとわかる。 a の値が b, c と比べて十分小さいと、 G_3 は既存の 2 言語と類似性がある ($b, c > 0$) のに対し、既存の言語は G_3 としか類似性がないことを意味する。このため、 w が少なく、人口遷移が活発であるときは言語としての優位性が保てなくなり、 w の減少と共に G_3 の領域が増えていく。また、前節の実験結果から見てわかるとおり、 w の値が小さくなると複数の言語が共存しやすくなる傾向がある。つまり、4.2 節で述べたような入力文に関する条件が満たされていないと、不完全な言語入力になっていることと対応し、よってクレオールが

生じやすくなる。したがって、図 5(b) と図 5(c) のように広域でクレオールが発現すると考えられる。

6 おわりに

言語学分野におけるクレオールの主な調査対象は、これまでに発見されてきたクレオール間に見られる言語としての類似性と、そこから推論される普遍文法が存在する可能性についてである。一方、我々はクレオールが発現するための条件を数理的に導き出すことを目的としており、言語の定義を各言語間の類似性から、クレオールの定義を言語話者の人口比率からそれぞれ与え、これまでの言語学的な視点と全く異なるアプローチで調査を行なっている。

我々は Nowak et al. [4] の言語動力学モデルをより現実的なモデルに改良し、新たなモデルを提案した。このとき、現実的であるための条件として次の点を主張した。

- 子供の言語獲得は周りの言語話者が話す言語

に影響を受ける。

- 親としか会話を行なわなかったとき、子供は親の言語を身につける。

このモデルでは、言語間の遷移を特徴づけるパラメータが人口比率に依存するので、遷移行列が動的に変化するものとなる。また、言語学習者である子供が親の言語を継承する精度は言語入力量に比例する。

我々はこれまでに、子供の言語獲得に関して、言語入力量が十分であるという前提の元でクレオール調査を行ってきたが、言語獲得に必要なとされる言語入力量について十分に議論してこなかった。そこで本稿では子供の言語獲得環境に焦点を当て、言語獲得期の子供の言語入力量とクレオール化の関係について実験を行なった。実験では、言語入力が少ないと、たとえ親からのみ言語を学習しても、子供は他の言語に移行しやすい傾向があることを示した。特に新言語への移行が、Bickerton [17] が主張するような、言語獲得過程にある子供がクレオールに近い発話を行なうといった現象を示唆していると考えることができる。

しかし、提案したモデルは抽象的であり、親の言語を正常に獲得するためには、どの程度の言語入力が必要であるかという問いに関して、解答を提示することはできない。そのためには、文法等の具体的な言語に関わる知識を導入する必要があり、数理モデルの上に文法の情報をいかに表現するかということが今後の課題となる。

参考文献

- [1] DeGraff, M.(ed.): *Language Creation and Language Change*, The MIT Press, Cambridge, MA (1999).
- [2] 橋本敬: 言語進化とはどのような問題か? ~構成論的な立場から, 第18回人工知能学会全国大会講演論文集, 金沢 (2004).
- [3] Cangelosi, A. and Parisi, D.(eds.): *Simulating the Evolution of Language*, Springer, London (2002).
- [4] Nowak, M. A., Komarova, N. L. and Niyogi, P.: Evolution of universal grammar, *Science*, Vol. 291, pp. 114-118 (2001).
- [5] 中村誠, 橋本敬, 東条敏: 言語動力学におけるクレオールの創発, *認知科学*, Vol. 11, No. 3, pp. 282-298 (2004).
- [6] Nakamura, M., Hashimoto, T. and Tojo, S.: Exposure Dependent Creolization in Language Dynamics Equation, *Proceedings of the First International Workshop on the Emergence and Evolution of Linguistic Communication*, pp. 9-14 (2004).
- [7] Hofbauer, J. and Sigmund, K.: *Evolutionary Games and Population Dynamics*, Cambridge University Press (1998).
- [8] 井上和子, 原田かづ子, 阿部泰明: 生成言語学入門, 大修館書店, 東京 (1999).
- [9] Nakamura, M., Hashimoto, T., Tojo, S. and Smith, K.: The effect of fitness in the emergence of creole, *Proceedings of EVOLANG5*, p. 85 (2004).
- [10] Niyogi, P.: *The Informational Complexity of Learning from Examples*, PhD Thesis, Massachusetts Institute of Technology (1996).
- [11] Bickerton, D.: *Roots of language*, Karoma, Ann Arbor, MI (1981).
- [12] 亀井孝, 河野六郎, 千野栄一 (編): 言語学大辞典, 三省堂, 東京 (1996).
- [13] 風間喜代三, 長谷川欣也 (編): 言語学百科事典, 大修館書店, 東京 (1992).
- [14] Niyogi, P. and Berwick, R. C.: The Logical Problem of Language Change, Technical Report AIM-1516, AI Lab, MIT (1995).
- [15] Briscoe, E. J.: Grammatical Acquisition and Linguistic Selection, *Linguistic Evolution through Language Acquisition: Formal and Computational Models* (Briscoe, T.(ed.)), Cambridge University Press, chapter 9 (2002).
- [16] Pinker, S.: *The Language Instinct: How the Mind Creates Language*, William Morrow and Company, New York (1994).
- [17] Bickerton, D.: Creole languages, *Scientific American*, pp. 108-115 (July, 1983).