

言語動力学によるクレオール化のシミュレーション

○中村 誠¹, 橋本 敬², 東条 敏¹

北陸先端科学技術大学院大学 {¹ 情報, ² 知識} 科学研究科

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、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、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といわれている。

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、文化進化という三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 [4]。クレオール化の問題は、そのうちの内側のループで論じることができる問題である。

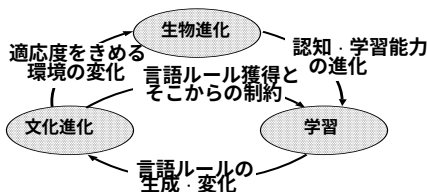


図 1 ダブルループ・ダイナミクス [4]

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ピ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。これらの研究成果を基に、上で述べた条件を導き出すことが可能となる。

2 言語動力学

Nowak et al. [2] の言語動力学 (Language Dynamics Equation) では、普遍文法を仮定することにより、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。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 とすると ($\sum_{i=1}^n x_i = 1$)、 x_i の時間変化は動力学系として表わされる。我々はこれまでに、この言語動力学を修正し、言語の学習環境に関するパラメータを追加した [3]。これは、Nowak et al. のモデルが子供は親からしか言語を学習しなかったのに対して、集団内の親以外からもパラメータに応じて言語入力を受けるようにしたものである。これを接触比率 α と呼び、子供の言語獲得期において親以外の言語話者から言語入力を受け取る比率を決定する。本稿では接触比率 α に加え、言語獲得期における言語入力の頻度に対するクレオール発現の条件について言及する。

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応度関数を排除した。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) 定数であった遷移確率 $Q = \{q_{ij}\}$ 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。この変化率を調整するパラメータが接触比率 α である。

3 動的遷移行列

言語学習者である子供は、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 の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：

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

ここで $S = \{s_{ij}\}$ は、任意の 2 つの言語間の類似度を表す確率行列である。要素 s_{ij} は、 G_i 話者がランダムに文を発話したとき、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。我々は (2) 式をもとに $\bar{Q}(t)$ 行列を提案した [5]。この $\bar{Q}(t)$ 行列は、定数として接触比率 α および学習期間中の入力文の数を示す w を持ち、これらの値の変化に対する動力学的振る舞いを観察することができる。

4 実験とまとめ

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。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_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導き出している。ここでは対称な S 行列を与え、対角要素を 1 としている。図 2 は接触比率と既存の言語間の類似度をパラメータとしたときの、優勢言語の分布を出力したものである。このときの入力文の数は $w = 30$ であり、3 言語のケースでは文法の推定に十分な数である。図中のアスタリスクは優勢言語が存在しなかった領域を表している。既存の言語が互いに似ていると、それぞれの言語話者

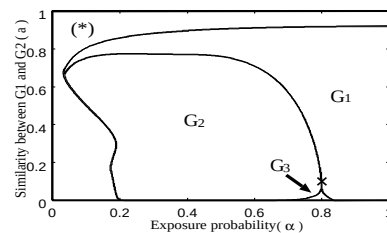


図 2 優勢言語の分布図の例

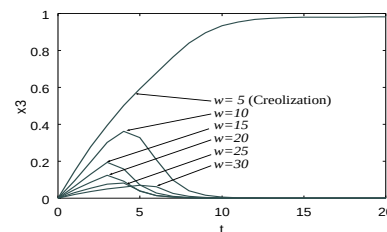


図 3 入力文の数に対するクレオール話者の人口変化

がコミュニケーションをとることができるため、新言語であるクレオールの必要性がなくなる。また、十分な入力を受け取ると、既存の言語がそれぞれ優勢言語となるため、図 2 においてクレオールが発現する可能性は非常に小さいことがわかる。

通常の言語獲得と比較して、文法を推定するためのトリガがないような、著しく貧弱な言語入力しか受け取らなかったときにクレオールは発現するといわれている [1]。図 3 は、図 2 の中の × 印の点に着目し、言語入力の数 w に対する x_3 の時間変化を示したものである。 $w = 30$ のとき、 G_3 は話者がほとんどいないまま消滅したのに対し、入力文の数が減るごとに人口のピークが増加する。そして $w = 5$ で集団内の優勢言語となる。ここから、 w が小さくなると、図 2 の G_3 の領域が広がることが想像できる。

我々は、これまでにクレオールが発現するために必要とされる使用言語そのものの条件を求めてきた。今後は言語学習者に焦点を当て、学習環境によるクレオール化の条件を求めていく。

参考文献

- [1] M. DeGraff (ed). *Language Creation and Language Change*. The MIT Press, Cambridge, MA, 1999.
- [2] M. A. Nowak, N. L. Komarova, and P. Niyogi. Evolution of universal grammar. *Science*, 291:114–118, 2001.
- [3] 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*, 11(3), pages 282–298, 2004.
- [4] 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. In 第 18 回人工知能学会全国大会講演論文集, 金沢, 2004.
- [5] M. Nakamura, T. Hashimoto, and S. Tojo. Exposure dependent creolization in language dynamics equation. In *Proc of EELC*, pages 9–14, 2004.