

# 極座標系を用いた言語進化モデルによる言語の進化軌道に関する研究

\*東垣外 翔<sup>(1)</sup>, 鈴木 麗璽<sup>(1)</sup>, 有田 隆也<sup>(1)</sup>

(1) 名古屋大学 大学院情報科学研究科 複雑系科学専攻

〒 464-0814 愛知県名古屋市千種区不老町

E-mail : azumagakito@alife.cs.is.nagoya-u.ac.jp

Abstract : 現在, 多様に複雑化した言語が世界に広く分布している. 言語はどのように発達してきたのか, 言語発達の過程を解き明かすための試みの一つとして, 構成論的手法を用いた進化模擬実験がある. 本研究では, 学習の概念を導入した言語進化モデルを構築し, 模擬実験を行った. このモデルでは, 言語を極座標で表現するため進化軌道を二次元平面上に表すことができる. 実験により言語が枝分かれしながら複雑化していく軌跡を観察することができた. 分析によって, 言語の多様化には情報共有の価値の頭打ちが不可欠であることや, 言語の複雑化に学習効果が必要であることが示された. また, 学習効果のダイナミクスも明らかとなった.

## 1 はじめに

人間と他動物との大きな違いの一つは, 言語を持つか持たないかである. 人間の持つ言語は他動物の情報伝達手法と比較しても複雑であり, この複雑性が事物に対する意味や概念の表現を可能にしたと言われている. また, 言語を広く後世に残す手段, 多くは文字による知の蓄積によって我々の文明が発達してきたことは明らかである. 人間が文明を築き上げることが出来たのは, 人間が言語と文字を獲得したことが大きな要因ではないだろうか.

言語の起源は何か, 言語はどのようにして進化してきたか, これらの謎を解くことは, 人間の主体性の基盤を探ることである.

言語進化研究が盛んになったのは, Pinker と Bloom らの論文からである [1]. Pinker らは, 人間の言語は, 情報伝達を目的として適応と自然選択により進化したと主張した. 近年では, 計算機を用いた**構成論的手法**が注目されている [2]. 構成論的手法とは, 計算機による模擬実験で個体を相互に作用させ, 個体間の作用によって生まれる高次元の性質一個々だけでは持ち得ない性質を観察・分析する手法である. 言語は体系としての秩序を有しているが, ほとんどの言語は外部からデザインされて作られたものではない. 話者の言語能力と環境と言語自身との相互作用によって徐々に練り上げられ, 複雑化してきた. そのため, 構成論的手法は言語進化を探る上での有効なアプローチである. 言語を構成論的手法で取り扱うということは, 「言語系が社会的な集団内の相互作用によって創発する」という言語観の導入である. この概念の導入により, 言語の成り立ちと本質について単純な要素のみを用いて考察することができる. 更に, 集団がさらされている環境条件の違いが, 言語系にどのような変化をもたらすかを調べることもできる.

言語の複雑化に焦点を当てると, 言語の複雑さが個体の学習能力と共進化するという仮説を示した鈴木らの研究がある [3]. この研究により, 情報伝達による利益と言語学習のコストのトレードオフが言語の複雑化に寄与することがわかってきた. しかし, この研究では, 言語の複雑さの段階が明示的に規定されており, 言語進化が連続的ではないという問題がある. また, 現象は適応度の上下によってのみ表されるため, どのような変遷を経て言語が進化してきたか理解しづらかった.

そこで, 本研究ではこれを土台とし, 言語の複雑化のダイナミクスを言語空間上の個体分布変遷の可視化に基づいて包括的に理解することを目的とし, 連続的な言語進化と学習の相互作用を表現でき, 言語空間を座標系で扱うことで言語の進化を可視化できるモデルを構築する.

## 2 目的

本研究の目的は, 言語の複雑化のダイナミクスを言語空間上の個体分布変遷の可視化に基づいて包括的に理解することである. 特に本研究では, 他個体と情報交換に基づいた言語能力の適応進化と, 言語能力の可塑性の適応進化の2者が共進化することに焦点を当て, 計算機上で言語能力の進化プロセスを表現し, 包括的に分析する. 本研究は言語進化に

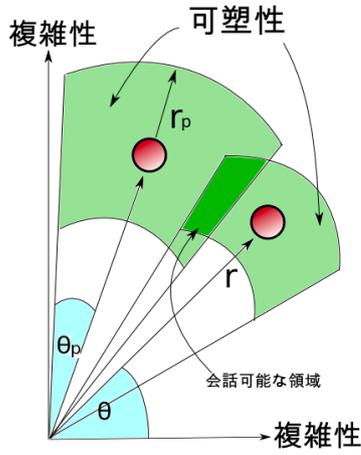


図 1: 言語表現の概念図

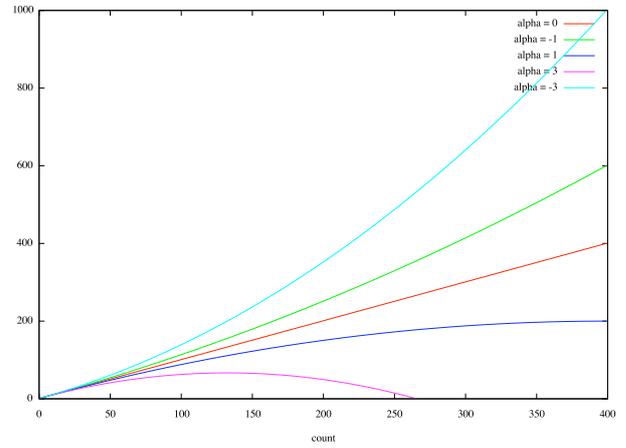


図 2:  $\alpha$  による利益増減の設定例

おける難問である「言語はどのように進化するか」へのアプローチの一つである。

### 3 極座標系を用いた言語進化模型

本模型では、個体は他個体との情報伝達を目的として活動し、活動の結果得た利益をもとに世代交代で進化していく。また、情報伝達の利益のみに焦点をあて、言語を使うことのメリットとする。以下にこの模型の詳細を説明する。

#### 3.1 個体の表現

個体は極座標で表す。極座標による個体表現の概念図を図 1 に示す。個体の先天的な言語能力として、扱う言語の複雑さを動径  $r$ 、扱える言語の特性を偏角  $\theta$  に対応付けた点で表現する。また、個体の後天的な言語能力の可塑性として、学習で扱えるようになる言語の複雑さ、特性をそれぞれ可塑的な動径  $r_p$  と偏角  $\theta_p$  として対応付け、個体を中心とした領域で表現する。個体は学習によって、領域の重なっている別の個体と会話できることとする。個体を座標で表すことで進化の軌道を二次元平面上の軌跡として可視化することができ、適応度や可塑性の増減による振る舞いと合わせた観察が可能となる。極座標系を使うのは、扱う言語が複雑化するほど、情報伝達量は増すが、他者との会話を通じ合いにくくなるという仮定を自然にモデル化するためである。すなわち動径が長いほど、偏角が同じなら個体間の言語空間上の距離が離れるので、会話を通じにくくなる。

#### 3.2 言語進化の表現

各世代において、個体は他個体と会話し、会話を通じ合った個体数に応じて利益を得る。このとき得られる利益は、適応度評価式のパラメータ  $\alpha$  によって、1) 会話を通じ合えば通じ合うだけ利益を得るか、2) 利益に一定の頭打ちが存在するか、3) 通じ合うほど利益が指数関数的に増加するかの設定が可能である。適応度評価式については後述する。また、言語が複雑なほど意味をより正確に伝えることができると仮定し、個体が扱える言語の複雑さに応じて利益を得る。言語の複雑性の利益についても、情報共有の利益と同様の設定が可能である。これら利益の掛け合わせから学習コストを引いたものをその個体の環境における適応度とする。次世代に残る個体にはそれぞれ突然変異をかけ、可塑性と複雑性を変化させる。適応度が高いものほど次世代に残り易いため、進化が起こる。

適応度評価には以下の式を使用する。

$$Fitness = (L - \alpha \frac{L^2}{2G})^{w_1} * (r - \beta \frac{r^2}{2A})^{w_2} - (\theta_p r_p)^{w_3} \quad (1)$$

第一項は学習によって可能な情報共有による利益を表し、第二項は言語の複雑性による利益を表す。第三項は学習コストである。  $L$  は学習で言語を理解し情報伝達が成功した他個体の数、  $r$  は動径（原点からの距離）であり、  $G$  は総個体数、  $A$  は動径の限界、  $w_1, w_2, w_3$  は重みである。  $\alpha, \beta$  はそれぞれ  $L, r$  の増加分に応じた減衰量をきめる定数であり、情報共有の頭打ちを表現できる。図 2 に例を示す。  $\alpha$  が正なら、情報を共有する個体数が一定の閾値を超えることで適応度

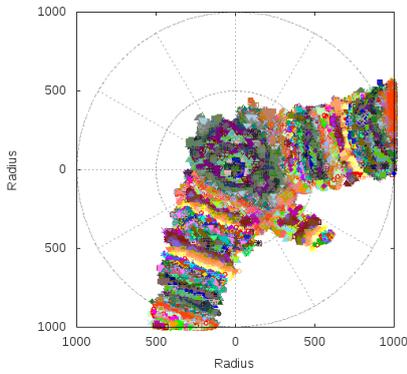


図 3:  $\alpha = 10$  での言語進化の軌跡

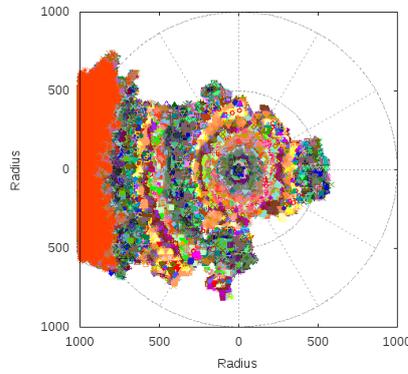


図 4:  $\alpha = 0$  での言語進化軌跡

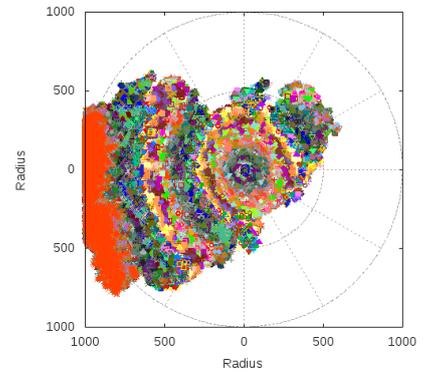


図 5:  $\alpha = -10$  での言語進化軌跡

が下がる状況を表現できる。0 なら線形に利益が増加し、負であれば情報を共有する個体が増加するほど適応度が指数関数的に増加する。同様に  $\beta$  が正なら、言語の複雑さが閾値を超えると適応度が下がり、負であれば複雑なほど適応度が増加する。

## 4 実験結果と考察

以上のモデルを用いて実験を行った。実験では、 $\alpha$  の値を 10, 0, -10 と変え、通じ合う言語数が利益に及ぼす影響を調べた。図 3 は、 $\alpha$  を 10 にし、通じ合う言語が多すぎても個体の適応度を下げるという設定の実験結果で、250 世代の個体座標を 1 世代ごとに色分けして表した言語進化軌跡である。個体が枝分かれしながら複雑化している様子を見ることができる。図 6 は世代ごとの適応度グラフである。横軸が世代、縦軸が世代ごとの適応度評価値を表す。赤線が世代における各個体の平均値であり、緑線は各世代での最高値である。世代が進むにつれ適応度が上がり、進化していることがわかる。図 7 は世代ごとの通じ合う言語数のグラフである。横軸が世代、縦軸が世代ごとの通じ合った言語数である。赤線は世代における各個体の平均値であり、緑線は各世代での最高値である。同図より、通じ合う言語数が一定の数に収束していくことがわかる。これは情報共有による利益に頭打ちがあるためと考えられる。

図 4 と図 5 はそれぞれ  $\alpha$  が 0, -10 の時の言語進化軌跡である。図 3 に比べ、集団の枝分かれが少なく、本模型においては、それぞれの個体が可塑性を広げるよりも個体同士が位置的に近いほうが学習のコストの面で有利である。そのため、通じ合えば合うだけ利益を得るという設定では、個体の位置は近ければ近いほど適応度が高くなり枝分かれが発生しにくい。

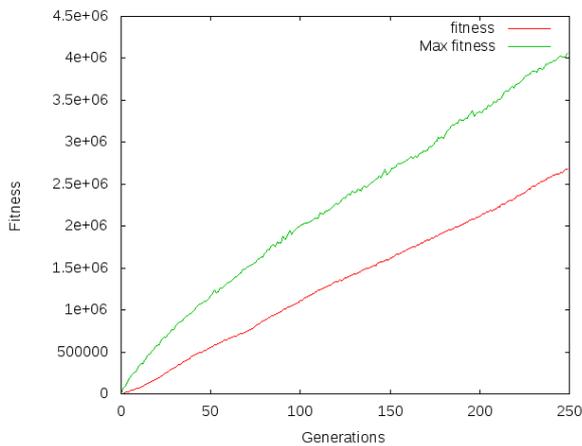


図 6: 適応度

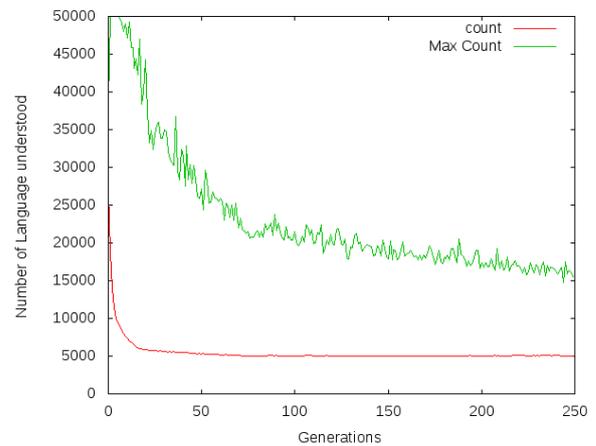


図 7: 通じ合う言語数

図 8 は  $\alpha = 10$  における可塑性のコストである。横軸が世代、縦軸が世代ごとの可塑性のコストの平均である。図 9 は  $\alpha = 10$  における情報共有による利益である。横軸が世代、縦軸が世代ごとの利益の平均である。これらの図を用いて、枝分かれの理由を考察することができる。図 8 では、20 世代ごろから可塑性が減りはじめ、それ以降減少しているのがわかる。図 9 では同時に 20 世代ごろから利益が一次的に減り、その後回復し増加している。このことから、枝分かれの

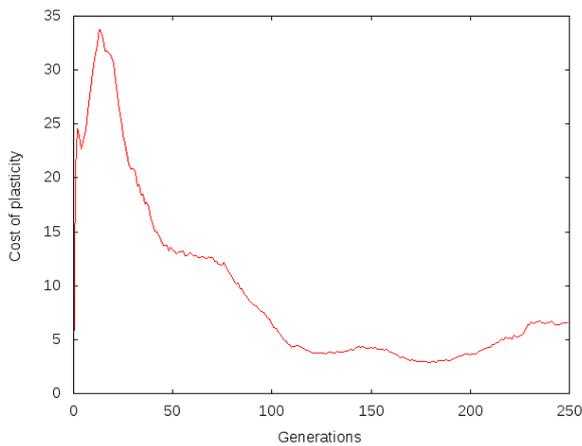


図 8:  $\alpha = 10$  可塑性のコスト

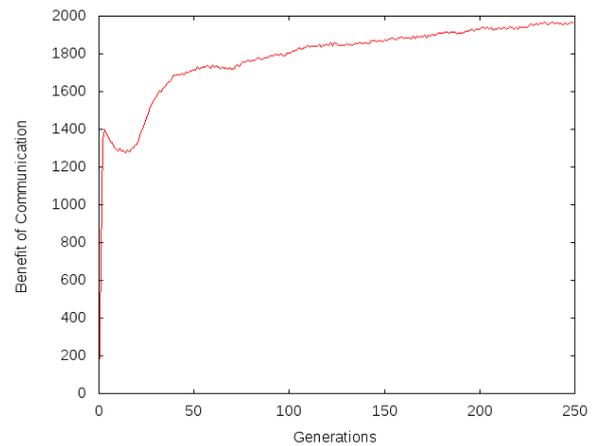


図 9:  $\alpha = 10$  情報共有による利益

ダイナミクスは、1) 情報共有の利益を得るために可塑性が多い個体が増える。2) 学習のコストにより、可塑性が少なくても通じ合うような個体（近くにいる個体）が増える。これが言語を共有する小集団が複数生じるきっかけになる。3)  $\alpha$  が正であれば、情報の利益に頭打ちがあるため集団数が一定に収束するということが言える。これは、可塑性がはじめはメリットとして増加するが、コストがあるために減少し、その特徴が生得的なものとなる助けとなるという、典型的なボールドウィン効果そのものである [4]。

## 5 おわりに

本研究では、計算機を用いた人工生命的な手法を用いて、言語の複雑化のダイナミクスを包括的に理解するためのモデルを構築し、模擬実験を行った。結果、言語の複雑化、多様化には情報共有の頭打ちが重要な影響を及ぼすこと [5]、後天的な学習能力が先天的な言語能力として定着していく段階（ボールドウィン効果）が言語進化にも重要な働きを及ぼすということをサポートする結果が得られた。さらに、言語が枝分かれする際の共進化ダイナミクスについての知見が得られた。

本研究で用いたモデルでは、個体の言語能力の進化のみに注目しているため、言語進化における原初的な言語能力の進化までしか表現出来ていない。言語進化の全体的なプロセスは言語能力と言語自体の共進化である [2]。言語自体は他言語との相互作用によって進化すると考えられ、文化に大きく影響されている。今後の展開として、文化進化的な言語進化を本モデルの上で表現し、言語能力の発達だけの世界と文化進化する言語自体の世界とを結びつけることが必要となる。

言語の枝分かれについて、生物種の分岐と同様、言語においても地理的な隔離が枝分かれに影響すると考えられるが、今回はそのような構造がない前提で言語の多様化に焦点を当てている。今後はこの点も考慮していきたい。

また、コミュニケーションが成立することによって得られる利益の問題がある。これまでの言語進化研究においては、コミュニケーションの利益は聞き手と話し手が同等に得られるものとするものが多かった。実際には、聞き手に有利な情報、話し手に有利な情報など、様々な形態の情報伝達があるはずであり、言語進化を包括的に理解するにはこれらの利益の違いを考慮する必要がある。

## 参考文献

- [1] Steven Pinker and Paul Bloom. Natural language and natural selection. *Behavioral and Brain Sciences*, Vol. 13, No. 4, pp. 707–784, 1990.
- [2] 有田隆也. コミュニケーションの創発. 計測と制御, Vol. 48, No. 1, pp. 39–46, 2009.
- [3] Reiji Suzuki and Takaya Arita. How learning can guide evolution of communication. *Proceedings of Artificial Life XI*, pp. 608–615, 2008.
- [4] James Mark Baldwin. A new factor in evolution. *American Naturalist*, Vol. 30, pp. 441–451, 1896.
- [5] Takaya Arita and Yuhji Koyama. Evolution of linguistic diversity in a simple communication system. *Artificial Life*, Vol. 4, No. 1, pp. 109–124, Winter 1998.