

**科学研究費助成事業 研究成果報告書**

平成 27 年 5 月 21 日現在

機関番号：13901

研究種目：若手研究(B)

研究期間：2011～2014

課題番号：23700310

研究課題名(和文) 言語進化論的アプローチによる急激な環境変化にともなう言語変化のモデル化

研究課題名(英文) Evolutionary linguistic approach to modeling the change in language associated with the rapid change in environment

研究代表者

中村 誠 (Nakamura, Makoto)

名古屋大学・法学(政治学)研究科(研究院)・助教

研究者番号：50377438

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：本研究課題では、自然言語の文法が言語使用者間のコミュニケーションによって動的に形成される過程を計算機シミュレーションで再現し、社会構造と言語変化の関連を示す定量的なモデルを構築した。これにより、言語接触などによる急激な言語変化現象の解明に貢献した。このモデルでは、文法獲得の能力を持ったエージェントを複雑ネットワーク上に配置し、隣接するエージェント間の会話を通じた学習により、通時的、共時的な言語変化を観察した。

研究成果の概要(英文)：We proposed an agent-based model for language contact, where we employed Kirby's iterated learning model (ILM) and complex networks. Totally, we succeeded to implement a linguistic community with learning agents connected with a social network. The network model makes it possible to observe not only diachronic but also synchronic changes in grammar. We achieved implementation of a large-scale, agent-based model where 100 processes of ILM run in parallel, which contributes to the simulation study on language evolution.

研究分野：言語進化

キーワード：言語進化 モデル化 シミュレーション エージェント

### 1. 研究開始当初の背景

人間の言語変化にはさまざまな要因が考えられるが、本研究においては、言語使用者が急激な環境変化にさらされた状況に焦点を当てる。このような例として、ピジンとクレオールが挙げられる。ピジンとは異なる言語を話すコミュニティ間の接触によって現れる一時的な言語であり、それが母語化したものをクレオールと呼ぶ。ピジンやクレオールは、農園におけるオーナーと共通の言語を持たない労働者との支配関係のように、社会的な構造が言語的特徴に強く影響することが知られている。このような言語変化の過程をモデル化することにより、言語接触の構造と言語変化との関係を明らかにすることが可能となる。

このような研究の応用として、インターネット上の言語分析などが期待できる。チャットや掲示板などで話されている言葉は、書き言葉よりも話し言葉に近く、また、非日常的な共通認識や専門知識の共有によって独特の言語表現も使われるようになる。このような現象は日常会話においても見られ、特段珍しいことではないが、この言語の変化に要する時間と伝播の規模は、著しく早く、大きい。インターネットにおける急激な言語変化というのは、ピジンコミュニティと同様に日常生活との社会構造的な乖離、すなわち急激な環境の変化による影響が大きいのではないかと考えることができる。

### 2. 研究の目的

本研究の目的は、自然言語の文法が言語使用者間のコミュニケーションによって動的に形成される過程を計算機シミュレーションで再現し、社会構造と言語変化の関連を示す定量的なモデルを構築することで、急激な言語変化現象の解明に新展開を図ることである。具体的には、仮定のネットワーク上においてエージェント同士のコミュニケーションによって言語獲得をする過程をモデル

化し、言語変化を促す要因を探求する。

### 3. 研究の方法

代表的な文法発達モデルとして、Kirbyによる繰り返し学習モデル (Iterated Learning Model) が挙げられる。これは、親子の2エージェントからなる文法獲得モデルである。基本的な子供の言語発達の能力を仮定し、何世代にもわたる繰り返し学習を行う。このモデルの特徴は、言語獲得期における「刺激の貧困」を解消すべく、文法の抽象度を高めることで、世代を経て複雑な言語構造を獲得するようになることである。

本研究においては、この繰り返し学習モデルを急激な環境変化にともなう言語変化のモデル化に対応するように拡張を行う。具体的には、(1) マルチエージェントのフレームワークへの対応と、(2) 学習機構の拡張である。

(1) では、言語の変化に社会ネットワークがどのように影響しているのかを明らかにするために、マルチエージェントへの拡張を行う。つまり、近隣のエージェントとの会話によって動的に言語コミュニティを構築する過程を再現し、通時的および共時的な両方の側面から文法の変化をモデル化する。

エージェントを結ぶ社会ネットワークの例を図1に示す。ここでは、スケールフリーネットワークを用いている。

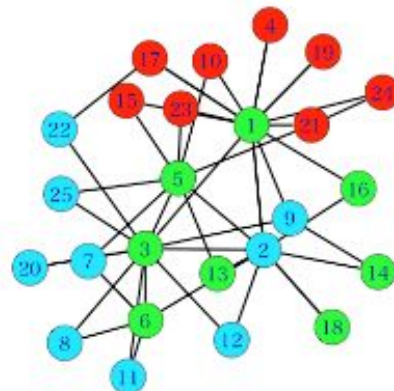


図1. ネットワークの例(N=25)

親子エージェントは各ノードに配置され

ており、隣人同士で発話と学習を繰り返すことで文法を獲得する。各世代で文法獲得をしたのち、それぞれのエージェントとの言語間距離を計算し、距離行列を作成する。それを元にしてクラスタリングを行った結果をデンドログラムで描画したものを図2に示す。横軸はエージェントのインデックス、縦軸は言語間距離を表している。図中の赤い破線は、言語とみなす閾値を表し、これに交差している縦線の数、は、集団で使用される言語の数に対応している。逆に、この線よりも下にあるサブツリーは同じ言語であるとみなす。図1と図2は対応しており、各エージェントの言語がどこに属しているのかがわかる。このデンドログラムは世代ごとに変化し、コミュニティの言語が収束する過程を観察することができる。

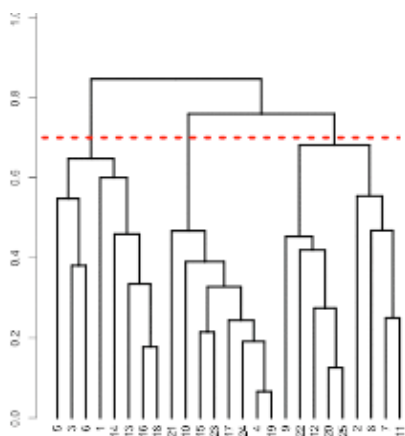


図2 . デンドログラムによる言語分布 (N=25)

また、(2)においては、シミュレーションを円滑に行うための問題解決を行う。

繰り返し学習モデルにおける親子間の文法獲得では、複数のエージェントから言語入力を受け取り学習する場合、世代を経るごとに発話が長くなるという問題が指摘されている。これは、学習において効率的な汎化ができなくなるためであり、組み合わせ爆発の原因となる。したがって、エージェント数を(1)の25エージェント以上に増やすこと

が困難であった。そこで、学習過程において、文法の表層形のうち冗長な部分を刈り込む手法を提案した。これをクリッピング処理という。これは、hamburger を burger, influenza を flu, examination を exam と省略するような、人間の実際の行動を模倣したものである。

#### 4 . 研究成果

(1) 一定世代後に集団で話されている言語の数とネットワークの形状との関連を示すために、格子状ネットワークとの比較をした。エージェント数は  $N=25$  である。実験結果を図3に示す。横軸 ( $0 \leq \alpha \leq 1$ ) は、隣人との接触割合を表す。隣人が複数いる場合は  $1/(\text{隣人の人数})$  の割合でそれぞれ接触し、親との接触割合は  $(1 - \alpha)$  である。縦軸は100世代後における集団で話されている言語の数である。それぞれ100回の試行の平均で表されている。青い線と赤い線は、それぞれスケールフリーネットワークと格子状ネットワークを用いた結果を表している。

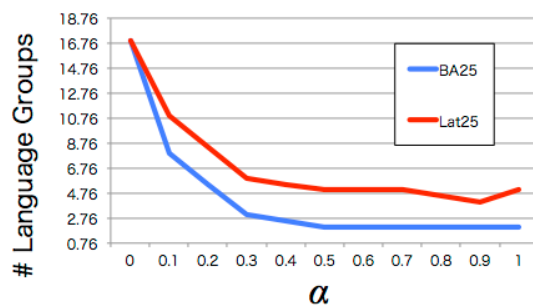


図3 . 実験結果

実験では、スケールフリーネットワークの集団は、 $\alpha$  によらず言語の数が格子状ネットワークの集団よりも少ないことがわかる。これは、ネットワーク形状によって言語の伝播のしやすさに差があり、スケールフリーネットワークに関しては、格子状ネットワークよりも共通言語を獲得しやすい環境であることを示している。したがって、このシミュレーションにより、社会ネットワークの構造の

違いが言語変化に影響を及ぼすことを示した。

(2) 繰り返し学習において、一組の親子エージェントによる文法獲得を行った。親からまとまった量の文を聞いて文法の推論を行うが、その際、その表層形についてクリッピング処理を行う。これにより、表層の文字列が短くなり、冗長な処理がなくなると期待される。

実験は、世代ごとの発話文字列の長さの平均を出力し、クリッピング処理がある場合とない場合との比較を行った。図4にクリッピング処理のない従来手法、図5にクリッピング処理を加えたときの結果を示す。これらは100回の思考の平均である。

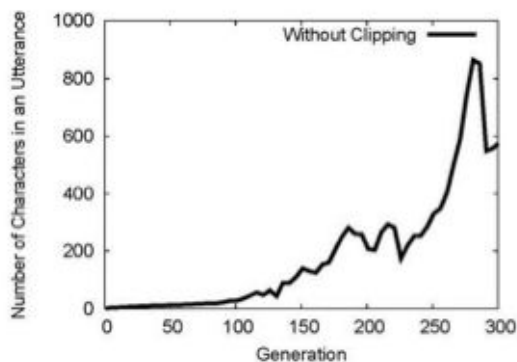


図4．各世代におけるクリッピング処理なしの学習後の発話の平均文字数

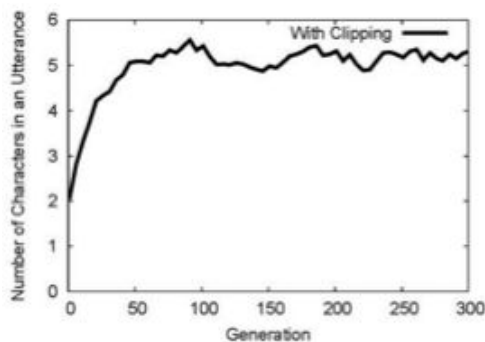


図5．各世代におけるクリッピング処理あり学習後の発話の平均文字数

クリッピング処理がない場合、平均文字

列長が世代を経るごとに増加するのに対し、クリッピング処理がある場合は、平均文字列長は世代を通じて変化がなく、また、縦軸のスケールを見れば明らかなように、総じて短い。また、獲得した文法の表現性や、前世代との言語の差はクリッピング処理がある場合とない場合で大差はなく、純粋に処理速度を向上させることに貢献していることを確認した。

この学習機構をエージェントに持たせ、(1)の実験を行った。N=25 および N=100 で実行し、実行時間の大幅な改善が確認できたが、処理結果に大きな違いが表れた。これは、言語間類似度の計算方法が(1)と(2)で異なっていたことが原因であることが考えられており、これらを加味したモデルを作成して再実験を行うことが今後の課題としてあげられる。

#### 5．主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 2 件)

1 Ryuichi Matoba, Hiroki Sudo, Makoto Nakamura, Shingo Hagiwara, Satoshi Tojo, Process Acceleration in the Iterated Learning Model by String Clipping, International Journal of Computer and Communication Engineering, 査読有, Vol.4, No.2, 2015, pp.100-106.  
DOI: 10.7763/IJCCE.2015.V4.388

2 Makoto Nakamura, Tomohiro Ohno, Yasuhiro Ogawa, Katsuhiko Toyama, Acquisition of Hyponymy Relations for Agricultural Terms from a Japanese Statutory Corpus, Journal of Information Processing in Agriculture, 査読有, Vol.1, No.2, 2014, pp.95-104.  
DOI: 10.1016/j.inpa.2014.11.001

〔学会発表〕(計 9 件)

1 Makoto Nakamura, Ryuichi Matoba, Satoshi Tojo, Simulation of the Emergence of Language Groups Using the Iterated Learning Model on Social Networks, COGNITIVE2015, Nice, France, 2015, pp.175-180.

2 Ryuichi Matoba, Hiroki Sudo, Makoto Nakamura, Satoshi Tojo, Application of Loose Symmetry Bias to Multiple Meaning Environment, COGNITIVE2015, Nice, France, 2015, pp.62-65.

3 中村 誠, 萩原信吾, 的場隆一, 東条敏,

社会ネットワークと流行語の伝播の関係を  
示すシミュレーションモデルの構築 ,2014 年  
度日本認知科学会第 31 回大会 ,名古屋大学 ,  
pp.342-343, 2014.

4 須藤洸基 ,的場隆一 ,萩原信吾 ,中村 誠 ,  
繰り返し学習モデルにおける高速化のため  
の文字省略プロセスの提案 ,2014 年度日本認  
知科学会第 31 回大会 ,名古屋大学 ,pp.255-260,  
2014.

5 須藤洸基 ,的場隆一 ,萩原信吾 ,中村 誠 ,  
東条 敏 ,第一言語獲得における認知バイア  
スに基づいた言語知識修正 ,2013 年度日本認  
知科学会第 30 回大会 ,玉川大学 , pp.343-349,  
2013.

6 Makoto Nakamura, Ryusei Kobayashi,  
Yasuhiro Ogawa, Katsuhiko Toyama, A  
Pattern-Based Approach to Hyponymy Relation  
Acquisition for the Agricultural Thesaurus,  
AOS2012, Kuching, Malaysia, 2012, pp.2-9.

7 Ryuichi Matoba, Makoto Nakamura, Shingo  
Hagiwara, Satoshi Tojo, CogSci2012, Sapporo,  
Japan, 2012, p.2774.

8 Makoto Nakamura, Shingo Hagiwara, Satoshi  
Tojo, An Agent-Based Model for the Emergence  
of Creoles, CogSci2012, Sapporo, Japan, 2012,  
p.2798.

9 Makoto Nakamura, Shingo Hagiwara, Satoshi  
Tojo, Multilayered Formalisms for Language  
Contact, WS on the Constructive Approaches to  
the Language Evolution, Kyoto, Japan, 2012,  
pp.145-147.

## 6 . 研究組織

### (1)研究代表者

中村 誠 (NAKAMURA, Makoto)

名古屋大学・法学(政治学)研究科(研究  
院)・助教

研究者番号 : 50377438