

二重ネットワークを用いた言語変化

Change of linguistic structure using a complex doubly structural network model

高橋 聡^{*1} 山田 隆志^{*1} 吉川 厚^{*1} 寺野 隆雄^{*1}
 Satoshi Takahashi Takashi Yamada Atsushi Yoshikawa Takao Terano

^{*1} 東京工業大学大学院総合理工学研究科知能システム科学専攻
 Department of Computational Intelligence and Systems Science,
 Interdisciplinary Graduate School of Science and Engineering,
 Tokyo Institute of Technology

In this paper, we analyze the influence on linguistic drift by social structure using agent-based simulation. We propose the new model by applying linguistic structure to the complex doubly structure network model. The new model represents linguistic structure and social structure. We run two types simulations using the new model and show that social structure affect linguistic drift.

1.はじめに

言葉や文字は表現が部分的に変化することによって、新たな言葉を生み出すことがある。例えば“(^^)”や“(;_;)”のように”顔文字”という表現形式は複数の記号が組み合わさって表現が構成され、その一部を変化させていくことによって様々な表現を生み出している。このような変化で生み出された表現は使用されることによって集団内で共有されたり、新たな表現の出現によって消滅したりする

集団内で共有語彙が発生していく過程や、言葉が進化していく過程に関して様々な研究が行われてきた。このような現象が実世界の中で変化する過程を追うことは難しく、シミュレーションがさかんに用いられてきた[橋本 04]。

そのシミュレーション手法の一つとしてエージェントベースシミュレーション(以下,ABS)がある[寺野 04]。ABSではエージェント同士が情報の伝播・学習などの相互作用することによって創発する共有知識の発生過程を解析することができる。

ABS を用いて集団内で共有語彙が獲得されて過程に着目した研究として言語ゲーム[Steels 96]がある。言語ゲームではエージェントはいくつかの単語と各語が指し示す対象の対のリストを持つ。そして、エージェントはコミュニケーション、学習によって対象に対して新たな単語を割り当て、他のエージェントのリストの対を学ぶ。複数のエージェントがコミュニケーション、学習を重ねた結果、エージェント共有語彙、すなわち、単語と対象の組み合わせの集合を獲得する。

一方、集団内での情報や文化の伝播過程に関しては社会ネットワーク構造が大きな影響を与えることがわかっている[松山 08]。[國上 08]では複雑二重ネットワークモデルを用いて、複数のエージェントと複数のモノが存在する世界を仮定し、エージェント同士が物々交換できるモノとモノの情報を交換することによって、複数のモノと交換することができるモノ、すなわち貨幣が創発する現象に着目し、ネットワーク構造によって創発する貨幣の数が変化することを示している。

そこで、本研究では複雑二重ネットワークモデルを用いて[Steels 96]の言語ゲームに社会ネットワークを取り入れ、既存の表現の変化によって生み出された表現が、集団内で共有かされたり、使用されなくなって消滅されたりする過程に対して社会ネットワーク構造が与える影響を考察する。

2. 研究手法・モデル

本研究は複雑二重ネットワークを拡張して用いる。社会構造をエージェント間のネットワーク構造で表現し、表現と意味の対応関係を内部ネットワークによって表現する。エージェントは会話プロセスによって新たな意味の表現を生み出したり、表現と意味との対応関係を交換・学習したりする。

2.1. 複雑二重ネットワークモデルとは

複雑二重ネットワークモデル[寺野 08]、[國上 08] (図 1) は社会を構成するエージェント間のネットワーク構造(外部ネットワーク)と、エージェントが認識・学習等の対象とする客体間のネットワーク構造(内部ネットワーク)によって構成されるモデルである。エージェント間の学習/伝播等の相互作用により内部ネットワークの構造が変化する。このモデルを用いることにより、社会構造をエージェント間のネットワークで、言語構造を客体間のネットワークで、社会構造から言語構造への影響を学習/伝播過程で表現することが

できる。

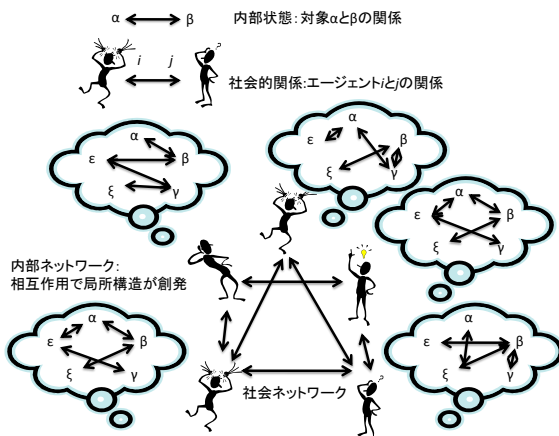


図 1: 複雑二重ネットワーク

2.2. 本モデルの内部ネットワーク

エージェントは意味と文字列によって構成される表現を持つ。この意味と複数の文字との対応構造を複雑二重ネットワークモデルの内部ネットワークとする。例えば、エージェントが意味 1 として”AAA”を持ち、意味 4 として”CCC”を持つとき内部ネットワークは図 2 のようになる。

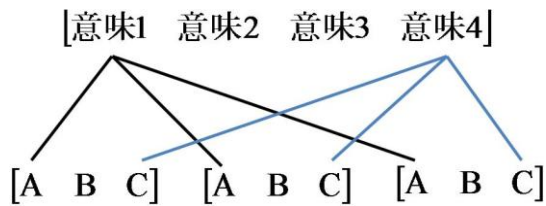


図 2: 内部ネットワーク

2.3. 会話プロセス

会話プロセスではまずエージェントの中からエッジで繋がったエージェントを一組ランダム選び、片方を表現エージェント、もう片方を観察エージェントとする。そして表現エージェントが表現する意味をランダムに選択する。

表現エージェントは表現したい意味を持っているときにはそれを表現し、観察エージェントに伝える。持っていないときには元々持っている表現の中から表現の一つを選択しその一部を変化させて新たな表現を作り出す。

例えば、表現エージェントが意味 1 として”AAA”の表現を持っていて、まだ表現を持たない意味 3 を表現した時には”AAC”や”BAA”等を作り出し意味 3 の表現とする。

観察エージェントは選択された意味の表現を元々持っていない場合は表現エージェントの意味と表現のセットを獲得する。持っている場合は、その表現が表現エージェントの表現と異なるときには表現の一部を表現エージェントの表現に変更し、新たな表現を作り出す。

例えば、表現エージェントが意味 1 として”AAA”を表現し、観察エージェントが意味 1 として”BBB”を持っている時には、観察エージェントは意味 1 の表現として”BBA”や”BAB”を獲得する。

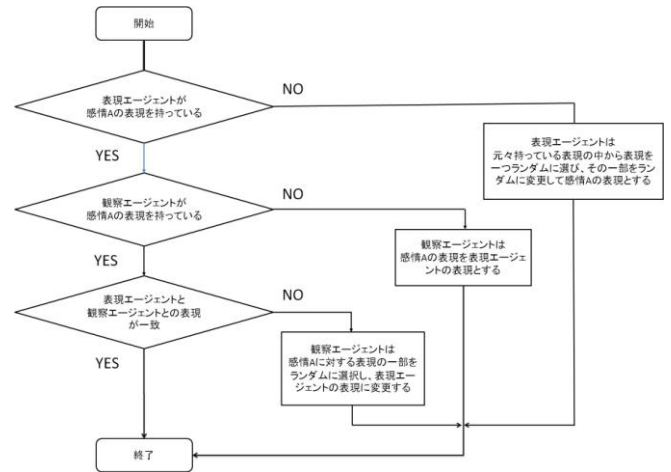


図 3: 会話プロセス

3. 実験

実験 1 では社会ネットワーク構造を変化させ、集団内で語彙が共有・消滅する過程に対して社会ネットワーク構造が与える考察する。実験 2 では異なる表現を持つ文化が接触した時、接触の仕方の違いによってどのような表現が生まれるのか調べる。

3.1. 実験 1

単純に共通表現の獲得過程に対して社会ネットワークの形状がどのような影響を与えるか考察する。初期値としてすべてのエージェントに意味 1 の表現として AAA を持たせた状態で会話させる。WS ネットワークにおける張り替え率を 0.05 から 0.5 まで変化させたネットワークと、BA ネットワーク、ランダムネットワーク、レギュラーネットワークにおいて実験を行い、ネットワーク全体での表現の種類の変化を見た。エージェント数は 1000、平均次数は 20 でネットワークを作成した。

3.2. 結果 1

結果を図 4 に示す。レギュラーネットワークや張り替え率が低い WS ネットワークは、始めは表現の種類が収束が早い後半は収束しにくくなっている。これはネットワークのクラスタリング係数が高いため、最初は局所的に共有語彙が発生し表現の種類が減少するが、その後は逆に複数の共有語彙から一つの共有語彙への収束が阻まれるためだと考えられる。

一方 BA ネットワークや張り替え率が高い WS ネットワークは、初めは表現の種類が増加するが、その後の収束が早

い、これはネットワークのクラスタリング係数が低いため、最初は共有語彙が発生しにくく表現の種類が増加し、その後一つの表現が増加すると一気に全体でその表現に収束するためだと考えられる。

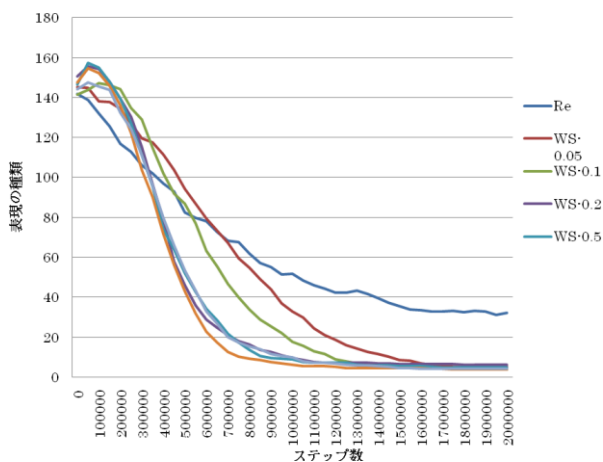


図 4:共有語彙の発生過程の比較

3.3. 実験 2

異なる表現を持つ文化が接触した時にどのような表現が生まれるのか調べる.AAA から発生した表現を持つネットワーク 1 と BBB から発生した表現を持つネットワーク 2 の二つの WS ネットワークを繋ぎ、繋ぐエージェントの割合を変更して比較する.エージェント数は 100,平均次数は 4,張り替え率は 0.1 で二つのネットワークを作成した.

全体の 20%から 100%まで 20%刻みで接触するエージェント数を変化させ、接触させた.例えば 20%の場合は両方のネットワークからランダムに 20 人ずつエージェントを選ぶ.そして、ネットワーク 1 から選ばれたエージェントはネットワーク 2 から選ばれたエージェントか 4 人をランダムに選びエッジを張る.

両方の表現が維持される場合を“分離”とし、一方の表現に全体が染まる場合を“同化”とし、両方の表現が合わさった表現を“混合”とする.混合には A と B の組み合わせで 6 種類存在するが、特に区別しない.

3.4. 結果 2

実験結果を表 1 に示す.また、接触率 20%混合、接触率 20%同化、接触率 100%混合の結果になった実験を一つ取り出し、持っている表現毎にエージェント数の時間変化をグラフ化すると図 5,6,7,8,9 のようになった.

表 1 より、接触の割合が非常に低い場合には“分離”することがあり.接触の割合が増加させると“同化”が減少し“混合”が増加することがわかる.

図 5,6 から 20%の接触率の時に “混合”が共有化されるま

で非常に時間がかかっていることがわかる.また、ネットワーク 1 において、“AAB”が全体で共有化された時期と、ネットワーク 2 において“AAB”が増加し始めた時期が一致していることと、その後、ネットワーク 1 とネットワーク 2 の“AAB”の減少と“BBB”の増加の時期が一致していることから、接触初期においてネットワークの接触点付近は様々な表現が入り混じっていたが、ネットワーク 1 で“AAB”が共有化されると同時期に接触点付近においても“AAB”共有化され、その後、接触点付近において“AAB”と“BBB”が接触し、最終的には“AAB”が共有化されたのではないかと考えられる.

図 7,図 8 より、20%の接触率の時の“同化”では“混合”に比べ短い時間で共有語彙が発生していることがわかる.また、ネットワーク 1 において“AAA”が共有化されると同時期にネットワーク 2 でも“AAA”が増加していることから、その時期に接触点周辺で“AAA”が共有化されたと考えられる.そして、その後ネットワーク 2 においても“AAA”は共有化された.

図 9 から 100%の接触率の時の “混合”は非常に短時間で共有化されていることがわかる.接触の初期において“AAA”と“BBB”が急速に減少し、その後、急激に“AAB”が増加し全体で共有化されている.

以上より、接触率が低い場合には全体で共有化された表現は、まず片方のネットワークで共有化された後、さらに接触点付近で共有化され、最後に全体で共有化されるという段階を踏むのではないかと考えられる.このため、接触点が少ないと片方のネットワーク内では“AAA”や“BBB”が有利なため、同化が起きやすくなると考えられる.それに対し、接触率が高い場合は段階的な共有化がなく、さらに接触の初期で“AAA”や“BBB”が急激に減少してしまうため、“混合”が共有化されやすいのだと考えられる.

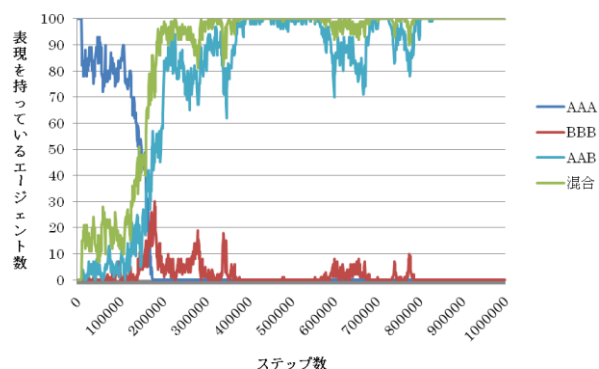


図:5 接触率 20%混合ネットワーク 1 (“AAB”が共有化)

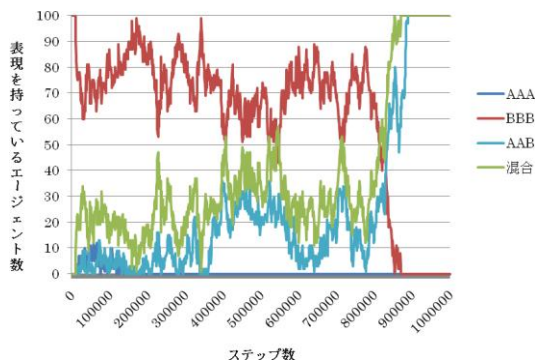


図 6:接触率 20%混合ネットワーク 2
("AAB"が共有化)

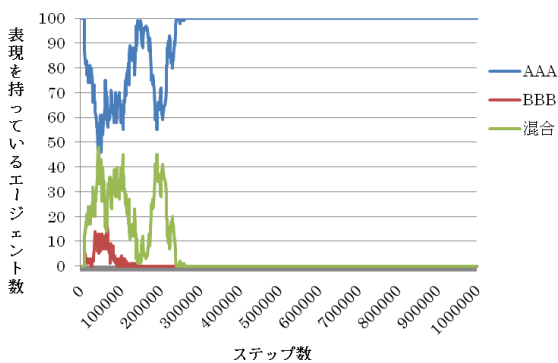


図 7:接触 20%同化ネットワーク 1

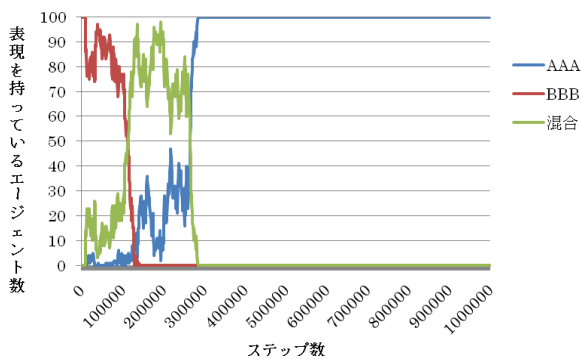


図 8:接触 20%同化ネットワーク 2

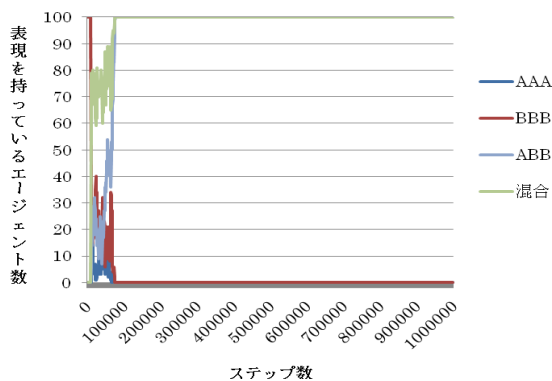


図 9:接触 100%混合ネットワーク 1

表 1:実験結果

	結合割合(%)				
	20	40	60	80	100
分離	169	1	0	0	0
同化	608	534	371	253	247
混合	223	465	629	743	753

4. 結論

本研究では複雑二重ネットワークモデルを拡張したモデルを用いて実験を行い,集団内で語彙が共有・消滅する過程に対して社会ネットワーク構造が与える影響が確認できた. 実験結果から, 1)社会ネットワーク構造が共有語彙への収束過程へ影響を与えること,2)社会ネットワーク同士の接触の仕方が,二つの異なる表現が接触した結果,共有化される表現に影響を与えること,がわかった.

参考文献

[Axelrod 97] Axelrod, R: The Dissemination of Culture ,A Model with Local Convergence and Global Polarization, Jurnal of Conflict Resolution, vol.41, pp203-326, 1997.

[橋本 04] 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, The 18th Annual Conference of the Japanese Society for Artificial Intelligence (2004).

[國上 08] 國上真章, 小林正人, 山寺智, 寺野隆雄: 複雑 2 重ネットワークモデルによる貨幣の創発現象の分析, 情報処理学会論文誌, 数理モデル化と応用(TOM)(07-Special-15), 2008.

[Masuda&Konno 06] Masuda, N., Konno, N.: Multi-state Epidemic Processes on Complex Networks, Journal of Theoretical Biology, vol.243, pp.64-75, 2006.

[松山 07] 松山科子,國上真章,寺野隆雄: ABS によるコンテンツ流通メカニズムの解析, 日本ソフトウェア科学会ネットワークが創発する知能研究会, JWEIN2007 予稿集, 2007.

[Steels 96] Steels, L:Self-Organizing Vocabularies, In:Langton, C and Shimohara, T (Eds.), Artificial LifeV, MIT Press (1996).

[寺野 08] 寺野隆雄: 複雑二重ネットワークモデル-知識と人のネットワークで社会を観る-,オペレーションズ・リサーチ, 2008年12月号, pp661-666,2008.

[寺野 04]寺野隆雄: エージェント・ベース・モデリング: その楽しさと難しさ, 計測自動制御学会, 計測と制御 Vol.43 No.12, pp.927-931, 2004.