

科学研究費助成事業 研究成果報告書

令和元年6月19日現在

機関番号：53203

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K16013

研究課題名(和文) 認知バイアスを考慮した言語獲得シミュレーションモデルの構築

研究課題名(英文) Construction of Language Acquisition Model with Cognitive Biases

研究代表者

的場 隆一 (Matoba, Ryuichi)

富山高等専門学校・電子情報工学科・准教授

研究者番号：30592323

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：人間の言語獲得の過程を解明することは大変意義深いですが、頭の中を直接観察することはできない。そこで本研究では言語獲得における繰り返し学習モデルの、意味表現、および、認知バイアスの有効性の測定について検討してきた。意味表現については既存の繰り返し学習モデルに共同注視の枠組みを導入した意味選択型繰り返し学習モデルを提案した。また、認知バイアスを提案したモデルのエージェントに組み込み、その有効性を検証した。

研究成果の学術的意義や社会的意義

これまでの認知バイアスと言語獲得の関係は、主に語彙獲得に焦点が当てられており、人間の言語獲得という大きな問題に関して語彙だけに限定してバイアスの有効性を主張するのでは不十分である。本研究では、言語進化モデルである繰り返し学習モデルに共同注視の枠組みを導入することで意味の多重性を表現するモデルを作成した。また、モデルにおけるエージェントの学習能力として認知バイアスを持たせることで、文法獲得における認知バイアスの効果を計算機シミュレーションによって示した。

研究成果の概要(英文)：It is well known that the cognitive biases much accelerates the vocabulary learning, also some related researches reports that efficacy of them also help to acquire grammar rules faster. The efficacy of the cognitive biases enables infants to connect an utterance to its meaning, even one uttered situation contains many possible meanings. In this study, we constructed Meaning Selection Iterated Learning Model and the method for measuring language distance between generations, and evaluated the efficacy of the cognitive bias in grammar acquisition.

研究分野：情報科学

キーワード：認知バイアス 言語獲得 繰り返し学習モデル

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

(1) 子どもは、生後 10 ヶ月から 18 ヶ月頃では言語の獲得速度は遅く、その定着性も低い。また、語彙の過大汎用による誤用が見られ、親の側の「思いこみ」が相互理解に一定の役割を果たす。これに対し、18 ヶ月以降になると、言語の獲得速度が上がり語彙の過大汎用がなくなっていく。この頃は、1 日平均 7 語から 15 語の語彙を獲得すると言われている。発話された状況から発話の意味に対する可能な仮説すべてを考慮しているのでは、少数の限られた事例から語彙を推論し獲得するのは不可能である。この問題に対し、ある種の制約を利用することで可能性を限定し効率よく推論を行っていると言われている。この制約は認知バイアスとして言語獲得との関係を示す研究成果が報告されている。

(2) 語彙獲得に関しては多くの認知バイアスとの関連が指摘されている。しかしながら、これまでの認知バイアスと言語獲得の関係は、主に語彙獲得に焦点が当てられており、人間の言語獲得という大きな問題に関して語彙だけに限定してバイアスの有効性を主張するのでは不十分である。動物と人間のコミュニケーションの最も大きな違いは文法、特に構文に関する制約であり、これこそが言語獲得と認知バイアスに深くかかわる部分であると考えられる。よって、構文獲得における認知バイアスの効果の検証は重要である。

(3) 言語獲得の研究において、その関連領域は言語学や認知科学、心理学や脳神経科学など多岐にわたる。それと共に近年計算機の精度が急激に発達しており、人工知能や計算機科学など計算機シミュレーションや数理モデルを用いた研究が注目されている。計算機を用いた研究では時間軸や獲得者の特性、獲得環境などを柔軟に設定することが可能となる。そのため実世界では観察できない現象をつくりだすことが可能となる。

(4) 計算機シミュレーションを用いた文法進化の研究として Kirb と Hurford が提唱した繰り返し学習モデル (ILM: Iterated Learning Model) [1] がある。このモデルでは言語獲得におけるボトルネックが起因して、合成性のない言語知識が世代を繰り返すことにより合成性のある言語知識になることを示したモデルである。モデルでは初期世代のエージェントは言語知識を持たず与えられた意味に対してランダムな発話を行う。しかし、文法を一般化する能力をもつエージェントが世代交代を繰り返すと、シミュレーション世界に定義された意味すべてに対して発話できる言語知識を持つようになる。

文法進化のモデルは文法を獲得するエージェントが存在し、文法を獲得したエージェントが次世代のエージェントに文法を伝達するという点では文法獲得のモデルであるといえる。繰り返し学習モデルはシンプルであるにもかかわらず構文獲得の本質に着目した優れたモデルであるため、進化言語学の分野や多くの言語学者が比較研究をしている。

2. 研究の目的

本研究では繰り返し学習モデルによる計算機シミュレーションを用いて認知バイアスが言語獲得に及ぼす有効性について考察する。また、繰り返し学習モデルは文法の発達・進化モデルであるため、認知バイアスの有効性を観察できるモデルに改変し以下について検討する。すなわち、言語獲得における繰り返し学習モデルの意味表現、および認知バイアスの有効性の測定における課題について認知モデルとして妥当であり、かつ計算モデルとして高速である文法獲得のパラダイムを提示する。

3. 研究の方法

ベースとした繰り返し学習モデルについて以下の改変を行うことで認知バイアスの有効性を測定した。

(1) ベースとした繰り返し学習モデルでは、親エージェントはランダムに意味を与えられ、意味を表現する発話を生成する。一方で、子エージェントは発話と共に親エージェントの発話の意味を知ることができる。現実世界では、発話の意味は見るできないため、親の発話内容を子が間違っ て解釈することがある。そこで本研究では、意味の伝達を行わず発話状況を親子で共有する共同注視の枠組みをモデルに導入した。また、認知バイアスの効果を調査するためには、1 つの発話に対して複数の意味候補が存在する発話状況から発話の意味を選択しなければいけない環境を作らなければならない。そこで意味の多重性をモデルに導入した。

(2) 文法進化モデルとしての繰り返し学習モデルでは、エージェントが獲得した文法を評価する際、表現度 (モデルで設定した意味の数に対して獲得した文法が表現できる意味の数) と文法を構成する規則数を用いていた。本研究では言語間距離を用いて認知バイアスの有無により子が獲得した文法が親のそれとどれだけ近いものであるかという評価する。

4. 研究成果

(1) 共同注視環境: 本研究では、文法進化モデルである繰り返し学習モデル (図 1) を改変し共同注視を導入した意味選択型繰り返し学習モデル (Meaning Selected Iterated Learning Model: MSILM) (図 2) を構築した。従来の繰り返し学習モデルでは発話と意味が一対一に結び

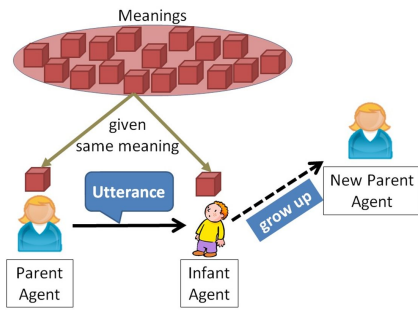


図1 繰り返し学習モデル

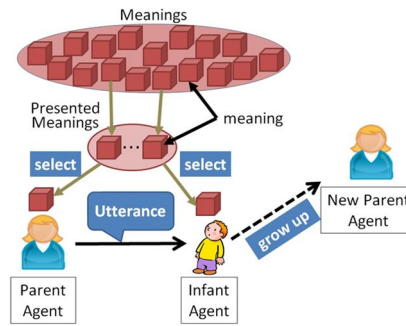


図2 意味選択型繰り返し学習モデル

ついたものを親子間で伝達していたのに対し、MSILM では1つの状況から任意の数の意味が読み取れる環境を作っている。これにより子が意味を特定できない場合を設けることで認知バイアスが働く環境を用意している。すなわち、従来の繰り返し学習モデルにおいては、子は親から発話と共に発話意味を受け取ることができるのに対し、MSILM では親は複数の意味を読み取ることができる状況に置かれ、その状況から意味を1つ読み取り発話を作成する。子は親の発話と譲許を受け取る。子は状況から意味を1つ読み取り、受け取った発話と読み取った意味のペアで学習を行う。これにより、親子が同じ状況を共有しながら親が発話をしている、いわゆる共同注視の環境をモデルに組み込んだ。

(2) 認知バイアスの導入：本研究では認知バイアスの中でも対称性バイアスおよび相互排他性バイアスをエージェントの能力として導入した。

対称性バイアスとは $A \rightarrow B$ ならば $B \rightarrow A$ を認識する認知の偏りである。これにより、発話とその意味の対応づけを学習する際には、発話 U という刺激に対してある意味 M を既に対応づけしているなら、同じ発話 U という刺激を受けた時に意味 M を想起することができる。本研究では、発話意味があり発話を行うというプロセスに対し、発話から発話意味を推論するプロセスを対称性バイアスの効果としている。

相互排他性バイアスとは $A \rightarrow B$ ならば $\neg A \rightarrow \neg B$ と認識する認知の偏りである。つまり、発話 U_1 に意味 M を対応付けしているならば、発話 U_2 に対して意味 M を対応付けしない。本研究では、子の持つ文法では入力発話に対応する意味を得られない場合、この持つ文法規則以外の文法規則を相互排他性バイアスにより生成する。

これらのバイアスは1つの発話に対して1つの意味のみ(対称性バイアス)、1つの意味に対して1つの発話のみ(相互排他性バイアス)と認知するように作用するため、意味と発話を一対一で結びつける効果がある。

これらがMSILM でどのように作用するかについて図3、図4で示す。子が状況と発話を受け取った際に、状況から認識できる意味の候補からどの意味を選択するかについて認知バイアスが影響する。認知バイアスが働かない場合、子は状況が与える意味の候補からランダムで1つを選択する。一方で、認知バイアスが働く場合、親の発話意味を推論により選択する。

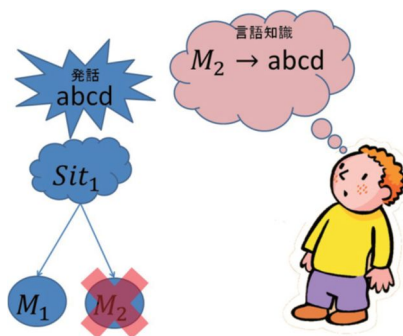


図3 対称性バイアスによる意味の制約

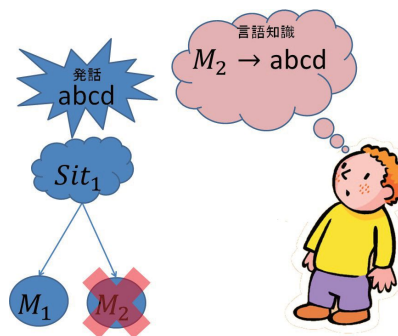


図4 相互排他性バイアスによる意味の制約

(3) 言語間距離：言語獲得における認知バイアスの有効性は、親の意図した意味を子が獲得できるか否かである。表現可能な意味数と言語知識を構成する規則数という評価法では、親子間にどれだけの認識の差があるのかを評価できない。よって、認知バイアスの影響を確認するには親子間における言語間類似度を知る必要がある。そこで本研究ではMSILMにおける言語間距離の測定法を考案した。以下に距離の測定手法を述べる。

意味 (m_c) と発話 (u_c) の対で構成されている子供の文法知識 (G_c) から文法規則 (g_c) を1つ選択する。親の文法規則 (G_p) から m_c とのハミング距離として一番近い意味 ($m_p^{m_c}$) を持

つ規則 ($g_p^{m_c}$) を選択する .

規則 $g_p^{m_c}$ の発話部分 ($u_p^{m_c}$) と u_c においてレーベンシュタイン距離 ($d(u_c, u_p^{m_c})$) を計算する .
 レーベンシュタイン距離は発話の長さで最大値が変わるので 0 から 1 に正規化する .

G_c を構成する全ての規則について から を行う .これらの平均値を親子間の言語距離とする .すなわち ,

$$Dist_{C \text{ to } G_p} = \frac{1}{|G_c|} \left(\sum_{i=0}^{|G_c|} \frac{d(u_{c_i}, u_p^{m_{c_i}})}{|u_{c_i}| + |u_p^{m_{c_i}}|} \right)$$

(4) 認知バイアスの有効性:MSILM を利用して認知バイアスの有効性を表現度と言語間距離を用いて評価した . バイアスを持たないエージェント Randomly MSILM, バイアスによる効果が意味推論時に完全な形で適用される Perfect Matching, Perfect Matching で候補を得られない場合に発話のレーベンシュタイン距離の近い候補の集合から意味のハミング距離が一番近いものを選択する Imperfect Matching の 3 種類で比較を行った .実験設定は, 発話回数を 1 世代 50 回とし, 最大 100 世代まで行った .また, 意味空間の大きさを 100 とした . 1 つの状況が与える意味数を 2 個とした .これにより, 意味空間から作られる状況は 4950 種類となる . 100 試行の平均を図 5 および図 6 に示す .

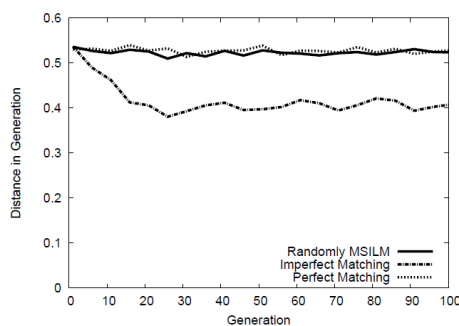


図 5 各世代間の言語間距離

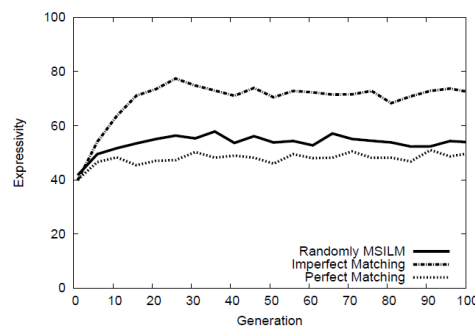


図 6 各世代の表現語

完全にバイアスを適用させるとバイアスを持たない場合とほぼ同等な文法を獲得するのに対し, 不完全な形で (ゆるく) バイアスかけるエージェントは表現度が最も高く, また言語間距離においても親子間で最も近い文法を獲得できていることが示された .

< 引用文献 >

[1] Kirby S., Hurford J.R., The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model, In: Cangelosi A., Parisi D. (eds) Simulating the Evolution of Language, Springer, London, pp121-147, 2002.

5 . 主な発表論文等

[雑誌論文](計 1 件)

Hiroki Sudo, Ryuichi Matoba, Todd Cooper, and Akira Tsukada, Effect of the symmetry bias on linguistic evolution, Artificial Life and Robotics, 査読有, 276, 2016, 1-8.

[学会発表](計 11 件)

米澤 智貴, 的場 隆一, 実世界における繰り返し学習実験, 平成 30 年度北陸地区学生による研究発表会, 2019.

Asuka Hashimoto, Kyo Takata, Nao Ito, Ryuichi Matoba, Kentaro Tani, Yoshinobu Maeda, Study of the Influence of an Obstacle on Evacuation Behavior Using Multi-Agent Simulation Where the Intimate Space Around Each Agent Is Considered, The 11th International Conference on Computer Modeling and Simulation, 2019.

Ryuichi Matoba, Yuya Hayashi, and Shingo Hagiwara, Applying Overextension to First Language Acquisition in a Joint Attention Frame, 22nd International Symposium on Artificial Life and Robotics, 2017.

Ryuichi Matoba, Yusaku Funane, and Todd Cooper, Automatic Construction of Emotion Corpus with Emotion Estimation by Naive Bayes, 22nd International Symposium on Artificial Life and Robotics, 2017.

Makoto Nakamura, Yuya Hayashi, and Ryuichi Matoba, Simulation of Language Evolution Based on Actual Diachronic Change Extracted from Legal Terminology, 9th International Conference on Agents and Artificial Intelligence, 2017.

細川 誠介, 的場 隆一, ディレクトリ型検索エンジンにおける Web ページのカテゴリ自動分類, 言語処理学会第 23 回年次大会, 2017.

Nann Kawdavid, Ryuichi Matoba, Extractive Text Summarization Implemented with SKG Formula, 言語処理学会第 23 回年次大会, 2017.

林 裕也, 萩原 信吾, 的場 隆一, 拡大解釈を導入した共同注視環境における第一言語獲得, 電子情報通信学会ソサエティ大会, 2016.

Yuya Hahashi, Ryuichi Matoba, Applying Extended Interpretation to First Language Acquisition under Joint Attention Frame, 2016 International Conference on Engineering and Technology(ICET2016), 2016.

Yusaku Funane, Ryuichi Matoba, Emotion Estimation and Automatic Construction of Emotion Corpus by Naive Bayes, 2016 International Conference on Engineering and Technology(ICET2016), 2016.

舟根 優作, 的場 隆一, ナイーブベイズ分類による文の感情推定と感情コーパスの自動生成, とやま産学官金交流会 2016, 2016.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。