

超領域リベラルアーツ (RA) 最終成果物

記号創発システムにおける 第二言語獲得についての考察

2023年7月5日

| | | |
|--------|-----|---------------|
| 総合心理学部 | 4回生 | 竹島 遥貴 |
| 情報理工学部 | 4回生 | 矢野 晃祐 |
| 文学部 | 3回生 | LAU Chun Chit |
| 情報理工学部 | 4回生 | 長友 束咲 |
| 情報理工学部 | 4回生 | 小熊 健太 |
| 経済学部 | 4回生 | 芳村 憲治 |

目次

| | | |
|-----|------------------------|---|
| 1 | 研究背景・目的 | 2 |
| 1.1 | 記号創発システム | 2 |
| 1.2 | 記号創発システム科学における第一言語獲得過程 | 2 |
| 1.3 | 第二言語獲得過程を考える意義 | 2 |
| 1.4 | 研究目的 | 3 |
| 2 | 研究手法 | 3 |
| 3 | 各専門的分野からの考察 | 3 |
| 3.1 | 自然言語処理系 | 3 |
| 3.2 | 機械学習系 | 4 |
| 3.3 | 第一言語に影響しない形での第二言語獲得過程 | 4 |
| 3.4 | 既存のモデルを用いた L2 の獲得過程の考察 | 5 |
| 3.5 | 哲学系 | 8 |
| 4 | 総合的考察 | 8 |

1 研究背景・目的

1.1 記号創発システム

記号創発システムとは、認知システムと社会システムの合体によって成り立つものである [1]。ここで、記号創発システムの記号とは、パースの記号論における記号を指している。パースの記号論においては、〇〇。認知システムとは、個々の人間の対象に対する恣意的なサインの割り当てやその解釈項が、感覚運動情報からボトムアップに形成されつつ、社会からトップダウン的な制約も受けるというループによって生じているというシステムのことである。社会システムとは、社会全体における対象に対するサインの割り当てやその解釈項が、社会における記号の解釈からボトムアップに形成され、トップダウン的な制約を受けるというループによって生じさせるシステムのことである。記号創発システムは、一般的にロボティクスを用いたモデルによって研究される。ロボティクスを用いるのは、実世界における視覚、聴覚、触覚等の複数の感覚情報を利用できるためである。また、モデルを作り実環境の中で適切に動作するかによって妥当性を検証することを、構成論的アプローチと呼ばれる。加えて、認知システムは、「認知的な閉じ」の状態となっていることが重要である。「認知的な閉じ」とは、認知が自らの感覚運動情報に閉じた状態のことを指す。認知が感覚運動情報のみから内的に作り出されるものであるということも言える。そして、認知システムをモデル化する場合は、「認知的な閉じ」を前提とすることができる教師なし学習のモデルが用いられる。教師なし学習のモデルの中でも特に、確率的生成モデルが一般的に用いられている。確率的生成モデルとは、観測変数がいくつかの潜在変数によって確率的に生成されると仮定するモデルである。モデルを学習させる際には、観測変数の情報のみから潜在変数をベイズの定理を用いて推論する。認知システムのモデルの文脈では、感覚運動情報という観測変数のみから、内的な認知（対象に対する恣意的なサインの割り当てやその解釈項）という潜在変数が推論されているということができる。

1.2 記号創発システム科学における第一言語獲得過程

記号創発システム科学の文脈で研究されたモデルは、基本的にはカテゴリや単語などの学習対象の知識有していない。これは、認知的閉じを仮定しているためである。つまり、記号創発システム科学において研究されたモデルは、第 1 言語（以下 L1 とする）獲得過程のモデルであると言える。具体的には、NPYLM-MLDA や HDP-HLM のようなモデルがある。NPYLM-MLDA は、視覚情報、聴覚情報、触覚情報に加えて、言語の音響情報から物体の概念を獲得するモデルである [2]。HDP-HLM は、言語の音響情報から、音素と単語を獲得するモデルである [3]。

1.3 第二言語獲得過程を考える意義

記号創発システム科学における研究で扱われているモデルの大半は、L1 獲得過程について考察するモデルである。第 2 言語（以下 L2）獲得は、L1 獲得と異なり、事前の知識があるという特徴

がある。事前の知識を仮定した記号創発は、これまで明確に研究されていない。また、言語圏によって異なる記号創発システムが存在しているという考え方もできると考えられる。日本語圏には日本語的な記号解釈を生み出す記号創発システムがあり、英語圏には英語的な記号解釈を生み出す記号創発システムがあるということである。そのため、L2 獲得は、記号創発システム間の交互作用について考えることにつながると考えられる。加えて、記号創発システム科学以外の分野にも波及する可能性がある。具体的には、心理学の分野が考えられる。特に、神経・生理心理学においては、L1 と L2 の処理の違いについて、生理反応を指標とした実験で研究されている [4]。しかし、L1 と L2 で処理機構にどのような違いがあるのかまでは明らかにできていない。記号創発システム科学的な L2 獲得過程の研究は、人間がどのように L2 を獲得しているのかを明らかにする研究ではない。しかし、人間がどのように L2 を獲得するのかの示唆を与えることはできると考えられる。そのため、L2 に関する心理学の分野の研究を支援するものとなると予想される。以上の観点から、記号創発システム科学的に、L2 獲得について考えることは意義があると考えられる。

1.4 研究目的

本研究では、記号創発システムの観点から、第 2 言語の獲得プロセスを考察する。具体的には、記号創発システムの観点から、多分野の知見も踏まえつつ、第 2 言語の獲得プロセスについての仮説を立てることを目的とする。

2 研究手法

まず、文献調査を行い、記号創発システムについての理解を深めると共に、既存の議論や考察について、情報を収集する。次に、収集した情報から、各自の専門分野の観点で考察を行い、それらの統合により L2 獲得プロセスについての仮説を立てる。そして、その仮説に基づいたグラフィカルモデルの記述を行い、論文形式にてまとめる。

3 各専門的分野からの考察

3.1 自然言語処理系

本節では、自然言語処理分野の観点から記号創発システムにおける第二言語獲得過程を考察する。

まず、自然言語処理の過程について、大きく分けて『形態素解析』『構文解析』『より高次の解析』の三段階に分類することが出来る [5]。形態素解析は与えられた文字列について単語への分割、品詞情報の付与、単語の原型の復元を行う。構文解析においては文中の各単語間における修飾・被修飾関係が有向グラフの形で抽出される。より高次の解析においては評判分析や意味解析等、より特定の分析目的に特化した、複雑な解析が行われる。本節では記号創発システムにおいて単一の言語のみを学習したモデルが既に存在し、概念が形成されていることを前提として第二言語獲得過程に

おける概念の獲得に注目し、

3.2 機械学習系

まず機械学習とは何かというところに触れていく。機械学習とは簡単に言えば蓄積したデータを利用して学習するアルゴリズム全般を指すものである。記号創発システムの言語獲得過程でも教師なし学習や自然言語処理などという形で多く取り入れられている。

今回は、現在言語獲得の過程で使われている技術の一つである「SHRDLURN」に着目した。SHRDLURN は言葉を使ったゲームによる学習のことで教師なし学習の一種である。[6] まず、簡単に概要を説明する。言語の学習方法として積み木を使用しており、人間がロボットに対して積み木の操作について指示をする過程でロボットが言語を獲得していくものとなっている。指示の方法に何も制約がなく、これによって指示を出す言語が人によって様々になったり、独自のコマンドを作成して伝える人もいるということが特徴である。

この技術では積み木を動かすといった極めて少ない操作しか行っておらず、完全な言語獲得というところまでには至っていない。しかし、この技術で作り出している獲得方法が記号創発システムでの言語獲得に似ており、記号創発システムでの第二言語獲得にも活かすことができるのではないかと考えた。現状、マルチモーダルなモデルでは第一言語獲得過程における観測変数としては一つの言語のみを使用している。しかし、先に挙げた技術で多数の言語を観測変数としてもロボットが学習できることがわかったため、第一言語獲得過程で獲得したい第二言語も使用することもできるのではないと思う。それによって第二言語が第一言語獲得過程ですでに知っている言語になるのではないかと考える。

3.3 第一言語に影響しない形での第二言語獲得過程

本節では、第一言語と第二言語を同じ獲得方法で習得した場合の第二言語獲得過程を考察する。

今回は、第一言語獲得の獲得過程で使われている「NPYLM-MLDA」の方法に着目した。これは 1.2 でも説明した通り、視覚情報、聴覚情報、触覚情報に加えて、言語の音響情報から物体の概念を獲得するモデルである。

これは全ての情報から、カテゴリーを定め、収納しているので、視覚と触覚の情報は同じだが、聴覚情報のみが違うと推測され、別のカテゴリーに収納されるのではないだろうか。この場合、同じものでも別のカテゴリーに入れられるため、第一言語と第二言語の関係性は存在しない。

しかし、この関係を結び付けるためには、言語を獲得したい対象物が全く同じものを指しているという状態が、言語情報、聴覚情報以外が同じ場合のみ、同じカテゴリーに含めるという事が必要であり、この際でも、第一言語の獲得が、第二言語の獲得に影響する要素はないと考える。

3.4 既存のモデルを用いた L2 の獲得過程の考察

3.4.1 考察の前提

記号創発システム的に L2 獲得過程を表すモデルを考えるにあたって、既存の記号創発システムのモデルの拡張によって考えられると仮説を立てた。ここでは、単一の言語を学習した記号創発システムのモデルを、L1 を獲得したモデルとして仮定する。これを、L1 獲得済みモデルと呼ぶ。また、L2 は L1 の知識に基づいて習得されるものである [7]。そのため、L2 の獲得は、L1 学習を獲得したモデルの潜在変数の分布から得られる情報も利用して行っていくこととする。具体的に、L1 獲得済みモデルを事前学習済みモデルとして、L2 の学習を行わせることで、L2 獲得獲得過程のモデルを考える。

3.4.2 モデルの定性的な妥当性の検討方法

L1 学習済みモデルを用いた、L2 の獲得のモデルの妥当性を保証するために、人間が L2 の獲得を行うにあたってどのような現象が起こるのかに基づいて考える。ここで、記号創発システム科学においては、人間が構成要素となっている記号創発システムがどのような仕組みで成り立っているのかを調べる。そのために、人間がどのような記号創発を行っているのかを、ロボティクスによるシミュレーションによって構成論的にしらべていく [1]。これを踏まえると、人間がどのように L2 を獲得しているのかについて考えることは、記号創発システム科学的意義のあることであると考えられる。人間が L2 の獲得の際に起こる現象として、L2 の獲得のしやすさが年齢によって異なることを取り上げる。ここで、年齢が上昇するにつれて増える L1 の知識が、L2 の獲得を促進する側面もあれば、抑制する側面もあるといわれている [8]。促進する側面としては、語彙や意味の獲得があげられる。抑制する側面としては、発声や音声、統語的な要素があげられる。年齢の上昇によって L1 の知識が多いことは、より L1 に対応するように潜在変数が推論されていると言い換えることができる。モデルの妥当性は、L1 に対応するように推論された潜在変数が、L2 を適切に生成できるように必要と予想されるベイズ更新の定性的な量を用いて判定する。L1 の知識が L2 を獲得するのを促進するということは、L1 に対応して推論された潜在変数が、L2 を適切に生成できるようになるまでに必要なベイズ更新が少ないということであるといえる。そのため、L1 の知識が L2 を獲得するのを促進する語彙や意味や意味の獲得のモデルは、L2 を適切に生成できるようになるまでに必要な定性的なベイズ更新の量が少ないことをもって妥当であると判定する。一方、L1 の知識が L2 の獲得を抑制するということは、L1 に対応して推論された潜在変数が、L2 を適切に生成できるようになるまでに必要なベイズ更新が多いということであるといえる。そのため、L1 の知識が L2 を獲得するのを抑制する発音の獲得のモデルは、L2 を適切に生成できるようになるまでに必要な定性的なベイズ更新の量が多いことをもって妥当であると判定する。

3.4.3 L2 の語彙や意味の獲得について

L1 の語彙や意味の獲得済みモデルとして、NPYLM-MLDA に単一の言語を学習させたモデルを用いる [2]。このモデルは、マルチモーダル情報と単語情報の両方から物体概念生成を行うモデ

ルである。NPYLM が音響情報から単語分割を行うモデルである。MLDA が視覚、聴覚、触覚というマルチモーダル情報から物体概念生成を行うモデルである。MLDA は、クロスモーダル推論を行うことができる。クロスモーダル推論とは、ある感覚情報から、他の感覚情報を推論することである [1]。視覚情報を w^v 、聴覚情報を w^h とした時、視覚情報から聴覚情報のクロスモーダル推論は、 $p(w^h|w^v)$ と表すことができる。図 1 は NPYLM-MLDA のグラフィカルモデルである。以下、議論に必要な変数を説明する。 w^w は単語情報、 w^v は視覚情報、 w^a は聴覚情報、 w^h は触覚情報を表す。 w^w は、NPYLM によって、音響情報から推論された潜在変数である。それぞれの β は、 w を生成するためのパラメータであり、それぞれの π によって生成される。そして、 Z はカテゴリを表し、 Θ によって生成される。 Θ は α をパラメータとしたディリクレ分布によって生成される。

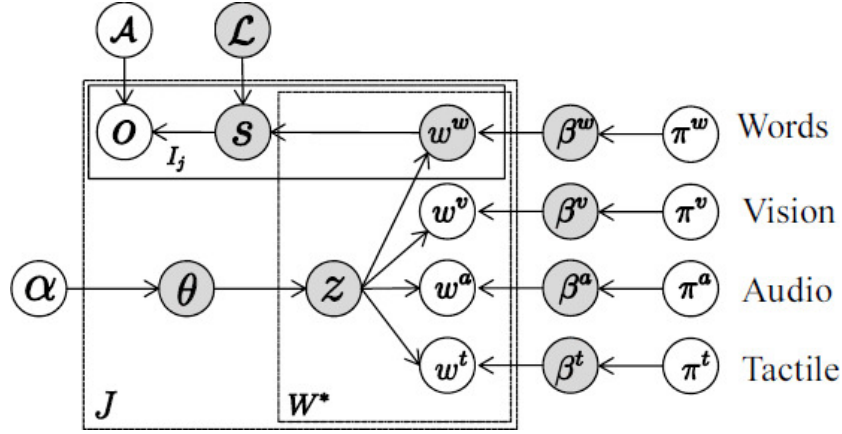


図 1 NPYLM-MLDA のグラフィカルモデル

次に、L1 獲得済み NPYLM-MLDA を事前学習済みモデルとした L2 獲得を考える。観測変数としては、L1 の単語情報 (W^{w11}) とそれに対応する L2 の単語情報 (W^{w12}) を用いる。まず、 W^{w11} から、 w^v 、 w^a 、 w^h をクロスモーダル推論し、MAP 推定等でもっともらしい \hat{w}^v 、 \hat{w}^a 、 \hat{w}^h を求める。次に、 W^{w12} 、 \hat{w}^v 、 \hat{w}^a 、 \hat{w}^h から、 z を推論する。これを複数の単語組について行う。この時、 Z はすでに、L1 の単語情報を適切に推論できるようになっている。また、L1 の単語情報は、それに対応する L2 の単語情報と共起関係にある。よって、 Z が示すカテゴリと L2 の単語情報がすでに対応しているといえる。そのため、 $p(Z)$ と適切に L2 の単語情報を生成する z の確率分布 $q(Z)$ は、ほとんど違いが無いと考えられる。つまり、 $p(Z)$ をベイズ更新する量は少ないと考えられる。よって、L1 獲得済み NPYLM-MLDA モデルは、L2 の語彙や意味の獲得の事前学習済みモデルとして、妥当であると考えられる。

3.4.4 L2 の発音の獲得について

L1 の発音の獲得済みモデルとしては、単一の言語を学習した HDP-HLM を用いる [3]。HDP-HLM モデルは、文を構成するモデルと単語を構成するモデル、音の情報を持つモデルの 3 つから

なるモデルである。音響情報のみから、音素 (文字) と単語を同時に推論するという特徴がある。音声を獲得することは、単語とそれに対応する音素を獲得することととらえられる。そのため、音声の獲得モデルとして、HDP-HLM モデルを用いる。図 2 は HDP-HLM のグラフィカルモデルである。 y は観測データである。 w は潜在音素、 z は潜在単語である。

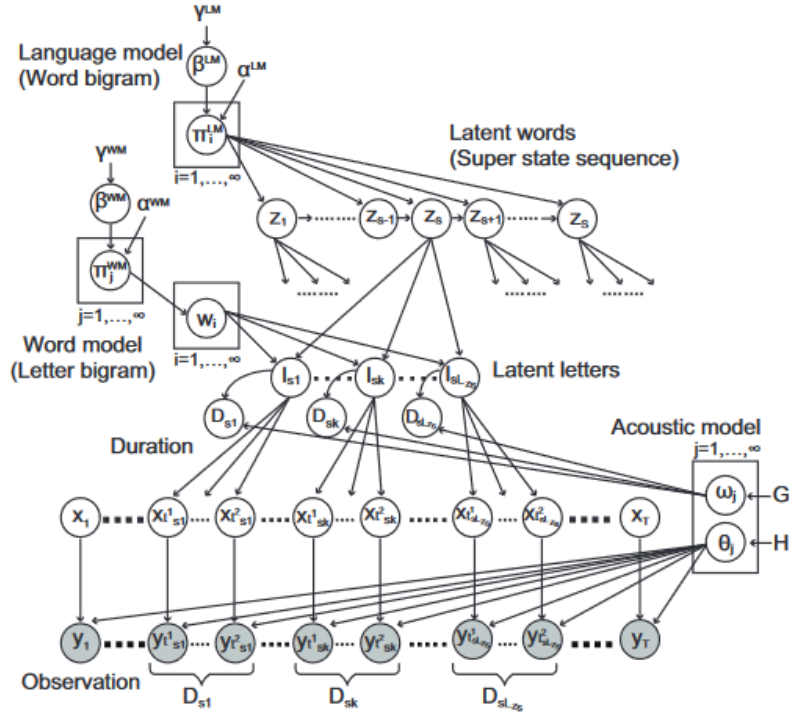


図 2 HDP-HLM のグラフィカルモデル

次に、L1 獲得済み HDP-HLM モデルを用いた L2 獲得を考える。観測変数として、L2 の音響情報を提示するものとする。ここで、L1 獲得済みモデルは音響情報から、L1 の文字と単語の潜在変数を推論するように学習されている。そのため、L2 の音響情報が与えられた時も、L1 の文字と単語を推論する。よって、L1 獲得済みモデルが L2 の音響情報から L2 の文字と単語を推論できるようになるまでには、新たに潜在変数をベイズ更新する必要がある。そのため、潜在変数のベイズ更新は多く必要であると考えられる。このことから、L1 獲得済み HDP-HLM モデルは、L2 の音声の獲得モデルの事前学習済みモデルとして妥当であると考えられる。

3.4.5 結論と 2 つのモデルの統合

以上のことから、記号創発システム科学的な L2 獲得過程のモデルの事前学習済みモデルとして、語彙獲得のモデルとしては NPYLM-MLDA、発音獲得のモデルとしては HDP-HLM が妥当であると考えられる。ただし、これらのモデルは、どちらも音響の情報から、単語を推論するモデルを含んでいる。NPYLM は、音響情報から文字 (音素) を推定できるようになった後に、文字から単語を推定するモデルとなっている。一方、HDP-HLM モデルは、文字 (音素) と単語を同時に学習

していく。人間が単語を学習する際は、L1であろうとL2であろうと、音響情報からは、音素と単語を同時に獲得している。そのため、音響から単語を推論するモデルとしては、HDP-HLMを採用する。

3.4.6 モデルの拡張

HDP-HLMにおいて、複数話者から同時に単一言語の音響情報を与えられた際に、文字や単語の推論の精度が悪くなるということが示されている [9]。その原因として、音響情報を発する話者を区別できていないためであることをあげている。そのため、L2の獲得過程についても、話者の区別と同様に、話されている言語がL1かL2かを表すカテゴリカルな潜在変数が必要であると考えられる。

3.5 哲学系

本章では、主にチョムスキーの言語生得説を取り上げて、合理論的な立場を土台として経験論である記号創発システムに対する批判を行うことによって、記号創発システムにおける合理論としての可能性を吟味し、記号創発システムにおける合理論な側面の可能性を考察する。そして、前述の考察に基づいて言語生得説における第二言語獲得過程について考察する。まず、哲学的な考察を行う前に、この考察における二つの前提を説明しなければならない。一つは、言語獲得において記号創発システムを、経験論的な立場として位置付けることである。おのれを記号論として位置付けている記号創発システムは、認知主体からのボトムアップな秩序形成と社会からのトップダウン的な制約との相互作用によって、創発的記号システムが形成される。このシステムにおける言語習得は、この考えは、システムを常に変化し続けている動的なものとして捉える点においては、確かに経験主義的であるが、しかしこれで認知主体における生得的、先天的、あるいはアプリアリのなものへの否定であるわけではない。チョムスキーの言語生得説は、人間は普遍文法といういかなる言語の習得も可能とする普遍的な言語機能が生得的に備えている、という合理主義的な考えである。本章における考察は、まさに言語生得説は創発的記号システムに受容することについての可能性を考察することである。

4 総合的考察

マルチモーダルで意味を
文字のみでは

参考文献

- [1] 谷口忠太. 心を知るための人工知能——認知科学としての記号創発ロボティクス——. 共立出版, 2020.

- [2] T. Nakamura, T. Nagai, K. Funakoshi, S. Nagasaka, T. Taniguchi, and N. Iwahashi. Mutual learning of an object concept and language model based on mlda and npylm. *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 600–607, 2014.
- [3] T. Taniguchi, S. Nagasaka, and R. Nakashima. Nonparametric bayesian double articulation analyzer for direct language acquisition from continuous speech signals. *IEEE Transactions on Cognitive and Developmental Systems*, Vol. 8, pp. 171–185, 2016.
- [4] M.-L. Vives, V. Costumero, C. Ávila, and A. Costa. Foreign language processing undermines affect labeling. *Affective Science*, Vol. 2, pp. 199–206, 2021.
- [5] 小林雄一郎. R によるやさしいテキストマイニング. オーム社, 2017.
- [6] Sida I Wang, Percy Liang, and Christopher D Manning. Learning language games through interaction. *arXiv preprint arXiv:1606.02447*, 2016.
- [7] B. Pajak, A. B. Fine, D. F. Kleinschmidt, and T. F. Jaeger. Learning additional languages as hierarchical probabilistic inference: Insights from first language processing: Learning languages as hierarchical inference. *Language Learning*, Vol. 66, pp. 900–944, 2016.
- [8] 岩立志津夫, 小椋たみ子. よくわかる言語発達. ミネルヴァ書房, 2017.
- [9] 林楓, 中島諒, 長坂翔吾, 谷口忠太. 階層ディリクレ過程隠れ言語モデルと深層学習を用いた語彙獲得過程の計算論. 日本認知科学会大会, Vol. 33, pp. 522–525, 2016.