

概念と言語の相互学習における育児語の影響の解析

船田 美雪[†] 中村 友昭[†] 長井 隆行[†] 金子 正秀[†]

[†]電気通信大学大学院情報理工学研究科

1 はじめに

人の赤ちゃんは、外界から得られる知覚情報や養育者からの語りかけによって言語や概念を獲得する。このとき、教示される言語の特徴が母語の獲得に影響を与え、言語や文化毎に異なる概念が構築されると考えられる。また、養育者は幼児に向けては擬音語・擬態語や音韻反復などの形態的特徴を持った育児語を用い、成長と共に成人語を用いる傾向がある。このような変化の中で、人は育児語と成人語の双方を学習する [1].

我々は、自身が取得可能なマルチモーダルな知覚情報をクラスタリングすることで形成されるカテゴリが概念であると考え、さらに他者とのインタラクションを通して言語を学習し、概念と結びつけることで語意を獲得できると考えている。このような概念・言語学習を記号創発システム [2] として捉え、我々はこれまで、ロボットが取得したマルチモーダル情報から概念と言語が相互に影響しあいながら学習するアルゴリズムを提案してきた [3]. しかし、そのような学習に関わる要因と、その要因が学習モデルに及ぼす影響は十分に検証されていない。そこで本稿では、概念と言語の相互学習モデルにおける育児語と成人語の学習の過程と、それぞれの語が概念形成に及ぼす影響を検証することを目的とする。最終的には、我々が提案してきた学習モデル [3] に基づいて、育児語と成人語が混在する教示発話を用いて物体概念の学習を行うことを検討しているが、本稿では育児語と成人語を区別し、それぞれの語について概念形成シミュレーションを行う。

2 概念と言語モデルの相互学習

文献 [3] において、我々は言語と概念が相互に影響しあいながら学習するモデルを提案してきた。図1は、言語モデルと物体概念を統合したグラフィカルモデルであり、灰色で示されたノードは未観測ノードを表している。図中の o が人から教示される音声である。この音声を、 A をパラメータとする音響モデル、 L をパラメータとする言語モデルにより認識した結果が s である。さらに、認識結果 s を言語モデル L を用いて単語へ分割し、Bag of words (BoW) 表現へと変換した

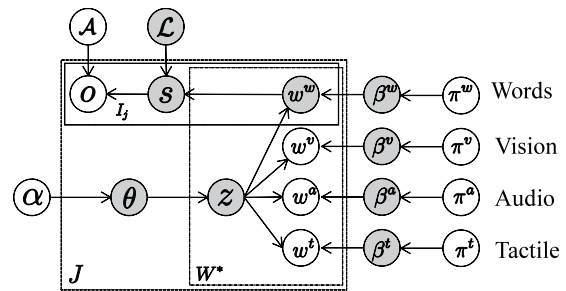


図 1: 物体概念と言語のグラフィカルモデル

ものが単語情報 w^w であり、さらに、 w^v, w^a, w^t はそれぞれ物体から得られる視覚情報、聴覚情報、触覚情報を示している。

また z は物体のカテゴリを表している。さらに、 w^v, w^a, w^t, w^w は、それぞれ $\beta^v, \beta^a, \beta^t, \beta^w$ をパラメータとする多項分布から発生する。これらの多項分布は、それぞれ π^* をパラメータとするディリクレ事前分布に従う。また、カテゴリ z の出現確率分布を表す多項分布のパラメータを θ とする。このパラメータは、ハイパーパラメータ α により決まるディリクレ事前分布に従う。 J, W^*, I_j は、それぞれ物体数、モダリティ*の情報の生起回数、 j 番目の物体への教示発話数を表している。このモデルでは、音声認識結果 s と物体カテゴリ z が単語 w^w によって接続されているため、音声認識と物体概念形成が相互に影響するモデルとなっており、音声認識・単語の接地・概念獲得などが統合されたモデルとなっている。このモデルにより、概念と言語モデルを相互に学習することで、音声認識と物体概念が影響しあい、それぞれの精度が向上する学習が可能となった。

3 概念と言語モデルの相互学習のシミュレーション実験

2章で示した相互学習モデルは非常に複雑であり、言語が概念形成に及ぼす影響を単純に解析することはできない。そのため本稿では、概念と言語の相互学習の近似モデルとして、文献 [4] の色概念と色語の相互学習モデルを用いる。このときの概念と言語の相互学習モデルのグラフィカルモデルは図2のように表すことができ、ロボットは取得した色情報 w^c と教示発話 o をクラスタリングすることで、概念の獲得が可能となる。本稿では、この色概念の学習において、特に「ぐるぐる」や「とんとん」といった音韻反復型の育児語の影響を解析する。

Analysis of the Effect of Childhood Vocabulary by Mutual Learning of Concept and Language

Miyuki FUNADA[†], Tomoaki NAKAMURA[†], Takayuki NAGAI[†] and Masahide KANEKO[†]

[†]The University of Electro-Communications, Department of Mechanical Engineering and Intelligent Systems
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan
{m.funada, nakamura, kaneko}@radish.ee.uec.ac.jp, tna-gai@ee.uec.ac.jp

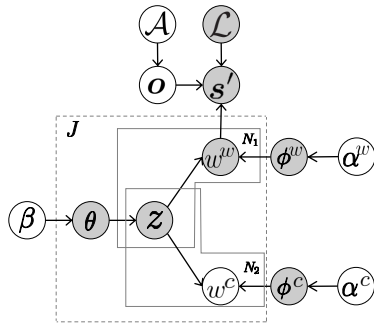


図 2: 色概念と言語のグラフィカルモデル

3.1 実験設定

まず、ランダムな色の単色画像とその画像の特徴を表す教示発話の組を 160 組生成した。次に、このうち 100 組を学習データとし、1 個から 100 個まで変化させながら学習を行い、各データ数毎に学習されたモデルを用いて 60 組のデータの認識を行った。このとき成人語での教示の場合には、「あか」や「あお」などの色語を用い、「これはあかだよ」といった教示発話を与えた。一方、育児語では、必ずしも全ての色語と対応した音韻反復型の育児語が存在するわけではない。そのため、各色に対して擬似的に音韻反復語型の単語を割り当て、育児語を用いた学習をシミュレーションした。例えば、赤色に対しては「ぐるぐる」といった単語を割り当て、「これはぐるぐるだよ」といった教示発話を与えた。すなわち、育児語での教示の場合は、成人語での「あか」や「あお」といった色語の代わりに、それらの色語と対応させた音韻反復語を用いている点のみが異なり、他の条件は全て同じである。これにより、本実験では音韻反復型の育児語が、学習にどのような影響をあたえるのかを検証した。

3.2 音声認識一致率

学習データ数を増加させたとき、教示発話全体から音声認識された文字列と正解文字列の一致率は図 3 のようになった。このとき、認識文字列 S_r と正解文字列 S_c の一致率 C は式 (1) によって表すことができる。すなわち、 C は比較する文字列の編集距離 $LD(S_r, S_c)$ を文字列長で割り、それを 1 から引いた値として定義される。一致率の値が大きいほど正解に近い音声認識ができていたことを意味する。

$$C = 1 - \frac{LD(S_r, S_c)}{\max(S_r, S_c)} \quad (1)$$

図 3 において、実線が成人語、破線が育児語を用いて学習した結果を示す。さらに、それぞれの認識文字列において、成人語と育児語に相当する単語のみに注目した場合、図 4 のようになった。このとき、実線が成人語、破線が育児語の音声認識一致率を示す。図 4 から、学習前期には成人語の方が育児語よりも正解に近い音声認識をすることができ、学習後期には育児語の方が成人語よりも正解に近い音声認識をすることができていることが分かる。このような音声認識の一致率の差は学習において無視できるものではなく、言語の学習だけでなく、概念の学習にも影響を及ぼすと考

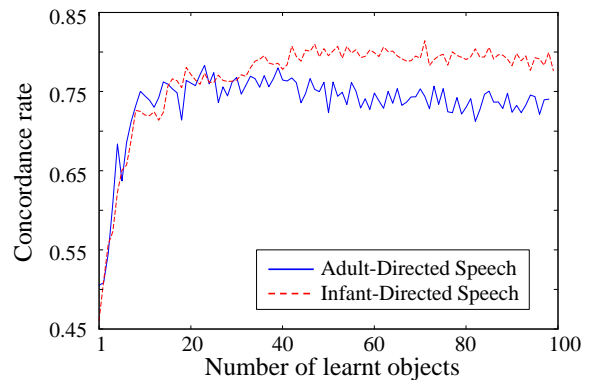


図 3: 学習データ数に対する育児語と成人語を含む教示発話全体の音声認識一致率の推移

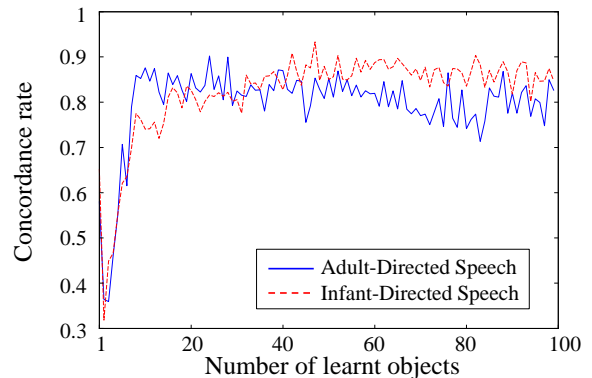


図 4: 学習データ数に対する育児語と成人語の音声認識一致率の推移

えられる。

4 おわりに

本稿では、ロボットの概念と言語の相互学習において教示発話が学習に及ぼす影響の解析を目的として、成人語と育児語を用いた概念形成シミュレーション実験を行った。今後、それぞれの語が学習に及ぼす影響を検証する。さらに、物体概念と言語を相互に学習するモデルを用いた研究に発展させることを検討している。

謝辞

本研究は JST CREST, JSPS 科研費 JP16H02835 の助成を受け実施したものである。

参考文献

- [1] 小林 哲生, 奥村 優子, 服部 正嗣: 幼児における育児語と成人語の学習しやすさの違いを探る, NTT 技術ジャーナル, 2015
- [2] 谷口 忠大: 記号創発ロボティクス 知能のメカニズム入門, 講談社, 2014
- [3] 中村 友昭, 長井 隆行, 船越 孝太郎, 谷口 忠大, 岩橋 直人, 金子 正秀: マルチモーダル LDA と NPYLEM を用いたロボットによる物体概念と言語モデルの相互学習, 人工知能学会誌, Vol.30, No.3, pp.498-509, 2015
- [4] 船田 美雪, 中村 友昭, 長井 隆行, 金子 正秀: MLDA と教示なし単語分割に基づく概念と言語モデルの学習過程の解析, 計測自動制御学会 システム・情報部門学術講演会, GS13-5, 2016