

# 概念・言語獲得モデルに対する構成論的解析 - 概念形成過程における育児語の影響 -

船田 美雪<sup>†</sup> 中村 友昭<sup>†</sup> 長井 隆行<sup>†</sup> 金子 正秀<sup>†</sup>

<sup>†</sup>電気通信大学大学院情報理工学研究科

## 1 はじめに

人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。特に日本においては、養育者は幼児に向けて擬音語・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る傾向がある。一方、英語圏での育児語の使用頻度は日本語圏よりも少ない。また、日本語児は英語児に比べて語彙の獲得が緩やかだが、新しい語を正確に学習する能力を早期から持つことが分かっており、育児語が言語の学習に影響を及ぼしていると考えられる [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定義し、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。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 [2]。しかし、この学習過程における言語と概念の相互作用とその要因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シミュレーションを行い、それぞれの語が概念形成に及ぼす影響を構成論的アプローチに基づき検証することを目的とする。

## 2 概念と言語モデルの相互学習

文献 [2] において、我々はこれまで物体概念を扱ってきた。しかし、物体概念と言語の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。そこで、本稿では、文献 [4] の色の概念学習を対象とする。図 1 は、言語モデルと色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

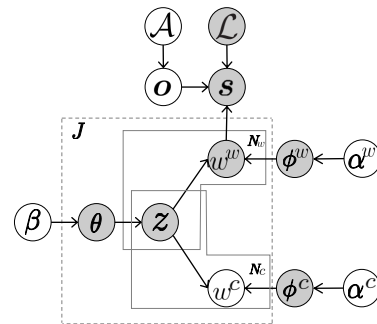


図 1: 色概念と言語のグラフィカルモデル

結果  $s$  を言語モデル  $L$  を用いて単語へ分割し、Bag of words (BoW) 表現へと変換したものが単語情報  $w^w$  であり、さらに、 $w^c$  はそれぞれの色から得られる知覚情報を示している。これらの情報  $w^*$  は、 $\phi^*$  をパラメータとする多項分布より生成される。また  $z$  は  $\theta$  をパラメータとする多項分布より生成される色のカテゴリを表している。 $\alpha^*$  と  $\beta$  は、それぞれ  $\phi^*$  と  $\theta$  の事前分布であるディリクレ分布のパラメータである。 $N_c$ 、 $N_w$  は、モダリティ  $c$ 、 $w$  の情報の生起回数を表している。ロボットが取得した複数の情報  $w^*$  をマルチモーダル LDA [3] によりクラスタリングすることにより、ロボットはセンサ情報を範疇化した概念  $z$  を教師なしで獲得することができる。

## 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。

### 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、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

Constructive Analysis of Acquisition Model of Concept and Language  
- The Effect of Infant-Directed Speech on Concept Formation Process -

Miyuki FUNADA<sup>†</sup>, Tomoaki NAKAMURA<sup>†</sup>, Takayuki NAGAI<sup>†</sup> and Masahide KANEKO<sup>†</sup>

<sup>†</sup>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

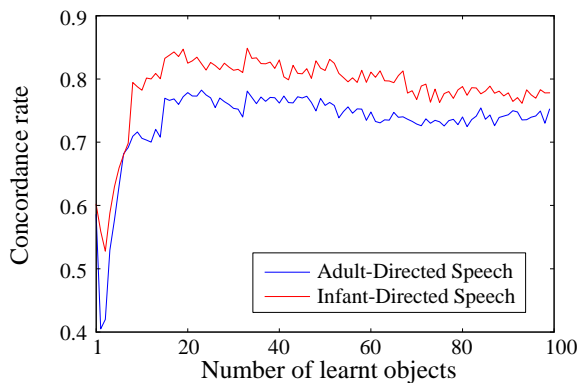


図 2: 学習データ数に対する育児語と成人語を含む教示発話全体の音声認識精度

り当てて、「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、学習にどういった影響を与えるのかを検証した。

### 3.2 音声認識精度

学習データ数を増加させたとき、教示発話全体の音声認識精度は認識文字列と正解文字列の一致率によって求められ、図2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。

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

図2において、青実線が成人語、赤実線が育児語を用いて学習した結果を示す。図2から、学習全体で一貫して育児語の方が成人語よりも音声認識精度が高いことが分かる。これは、育児語での教示の場合では、ほとんどの教示音声を正しく認識できているのに対して、成人語での教示の場合では音声認識を正しく行うことが難しい教示音声が存在するためだと考えられる。つまり、育児語はより正解に近い言語モデルを学習する傾向があるのに対して、成人語は正解とは異なる言語モデルを学習する要因を孕んでいることを示唆している。

### 3.3 学習用文字列のパープレキシティ

次に、学習データから取得された文字列の文字バイグラムパープレキシティを計算した結果、図3のようになった。パープレキシティはエントロピー  $H$  を用いて式(2)のように表され、文字列の複雑さを意味する。ただし、 $M$  は言語モデルの学習に用いた文字列から計算した文字バイグラム、 $\hat{W}$  は評価文字列である。本稿では評価文字列に3.2節の正解文字列を用いた。

$$\begin{aligned} \text{Perplexity} &= 2^H \\ H(\hat{W}|M) &= \frac{1}{|\hat{W}|} \sum_{\hat{w} \in \hat{W}} -\log(P(\hat{w}|M)) \quad (2) \end{aligned}$$

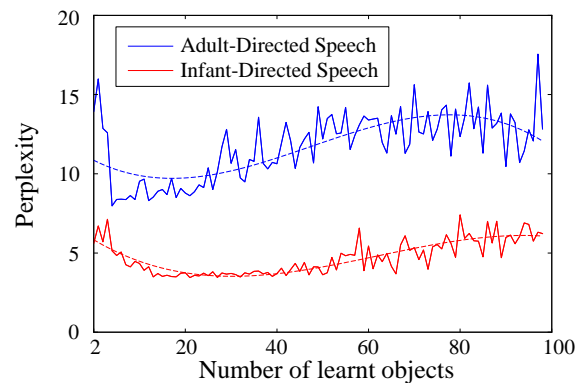


図 3: 学習用文字列の文字バイグラムパープレキシティ

図3において、青実線が成人語、赤実線が育児語を用いて学習した結果であり、破線がそれぞれの近似曲線を示す。図3より、成人語のパープレキシティは学習初期に急激に減少してしていることが分かる。このことから、成人語の学習において過学習が起こっていると考えられる。一方、育児語のパープレキシティは学習初期から中期にかけて緩やかに減少しており、過学習が抑制されていると考えられる。

## 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構成論的解析を目的として、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。育児語の方が成人語よりも音声認識精度が高く、明らかな差があることを示した。その要因として、成人語には過学習の傾向があり、育児語には過学習を抑制する特徴があることが考えられる。この結果は、音韻反復型の育児語による学習は進行が緩やかであるが、言語の学習精度は高いことを示しており、文献[1]における日本語児の学習過程に相当すると考えられる。今後の課題として、成人語と育児語が混在する教示発話への拡張や、より複雑な学習モデルにおける概念と言語の相互作用の解析などが挙げられる。

## 謝辞

本研究は JST CREST JPMJCR15E3, 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 哲生ほか：“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] 中村 友昭ほか：“マルチモーダル LDA と NPYLM 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
- [3] 中村 友昭ほか：“Gibbs Sampling による物体のマルチモーダルカテゴリゼーション,” 第 28 回日本ロボット学会学術講演会, 2I1-3, 2010.
- [4] Miyuki Funada et al., “Analysis of the effect of infant-directed speech on mutual learning of concepts and language based on MLDA and unsupervised word segmentation,” IROS2016: Workshop on ML-HLCR, 2017.