

人間の多次元的な心的表象に基づく 幼児語彙獲得モデルの構築

藤田守太¹ 南泰浩¹¹ 電気通信大学情報理工学研究所

f2131139@edu.cc.uec.ac.jp minami.yasuhiro@is.uec.ac.jp

概要

人間がどのように語彙を獲得しているのかという課題に対して深層強化学習を用いて幼児の語彙獲得機構をモデル化する研究分野がある。これらの研究で扱われる物体の特徴量は画像データから恣意的に決められた次元数が多く種類の少ない特徴量である。しかし、人間の心的表象に関する研究では人間が多くの特徴量を組み合わせて物体を認識していることがわかっている。本研究では人間の心的表象に注目し、49個の特徴量から語彙獲得を行う機構を構築し、実験により、その選択傾向から幼児が自律的に特徴量の選択に制約を課す機構を獲得することを確認した。

1 はじめに

幼児の語彙獲得では、親が幼児に向けて対象の物体または行動（これを以降、実態と呼ぶ）を示し、実態に対する語を示すことで語彙獲得が進むと考えられる。しかし、幼児から見ると親が指し示している実態は多数存在し、具体的に対象が何を示すのか判定することは難しい。また、同じ名称を持つ物体は複数存在し、似ているものでも違う名称を持つ物体がある。それらの物体の境界線を定めることは難しい。このため、幼児の語彙獲得は非常に困難な課題である [1]。

この困難な課題に対して幼児が語彙を獲得する際に、幼児が語の実態の可能性に制限を課す機構の存在を仮定する仮説がある。その中でも主要な理論として Markman らによる制約アプローチ [2] と Tomasello らによる社会-語用論理論 [3] がある。Markman らによる制約アプローチは、幼児には生得的に制約を行う機構が備わっているため、語彙獲得をスムーズに行えるというものである。一方、Tomasello らによる社会-語用論理論は、幼児が持つ認

知的スキルである「共同注視」、「意図の理解」などの母子間の相互行為によって語の実態の可能性に制限を与える。制約アプローチでは、「相互排他性制約」と「事物全体制約」の2つの制約が主張されている。しかし、相互排他性については、語彙獲得の早い段階で、一つの実態に対して2つ以上の語が対応することを幼児が受け入れることが確認されている [4, 5, 6]。また、事物全体制約についても、親の意図を理解している幼児が未知の実態に対して、物体ではなく行為として認識する例が報告されている [7]。このことから、必ずしも、事物全体制約が生得的に存在していない可能性も示唆されている。このように生得的に制約を獲得しているという制約アプローチでは数々の事象を説明できない。一方、語彙獲得を通して制約を獲得するという Tomasello らの社会-語用論理論 [3] では、以上の事象について説明可能であり、幼児の語彙獲得機構のモデル化に適している。

現在、語彙獲得のモデル化に関する先行研究として語彙獲得機構を機械学習を用いてモデル化する研究がある [8, 9]。これらの研究では語彙獲得機構のモデル化を通して、その解明に寄与している。しかし、これらの研究では母子間の相互行為や心理学的要因について生得的な機能として捉えている。一方で母子間の相互行為や心理学的要因について Tomasello の「幼児が持つ認知的スキルである『共同注視』、『意図の理解』などの母子間の相互行為によって語彙獲得における語の意味の可能性に制限が与えられる」という主張に従い、深層強化学習を用いた幼児語彙獲得機構のモデル化を行い、事前知識のない状態から語彙を獲得するモデルを構築する研究がある [10, 11, 12]。これらの研究では、母子間の相互行為を用いることで事前知識のない状態から意図の理解を通して語彙を獲得するとともに、語の意味の可能性に制限をかける機構を獲得することに成功している。しかし、幼児が取得する物体の特徴量に関して

「色」「味」「形」という3つの特徴量を恣意的に決定していること、特徴量の情報量が膨大な代わりにその種類が少ないため、幼児語彙獲得モデルの特徴量選択の傾向が不明瞭であることが挙げられる。実際に Martin らによる研究 [13, 14] では人間は物体を49個の次元で捉えており、それらの特徴量を組み合わせることで物体を判別していることが確認されている。また、従来手法 [10, 11, 12] では母子間の相互作用による語彙獲得を考量している反面、幼児自身が持つ好奇心による語彙獲得への影響も考慮していない。そこで本研究では従来の恣意的な特徴量ではなく、Martin らの研究から判明した49次元の特徴量を用いて、幼児の持つ好奇心と母子の相互作用を組み合わせた語彙獲得モデルを構築する。

2 考慮する心理学的要因

2.1 母子間の相互行為

共同注視 共同注視とは、非言語的コミュニケーションの一つであり、親、あるいは、幼児が何かに注目しているときにもう一方も同じ対象に注目する行動である。語彙獲得を行うには、親と幼児が共通の対象を注視することが必要条件であると言われている [15]。また、Tomasello は6組の母子に共同注視の時間と語彙数について実験を行い、母子間で共同注視を確立し、維持することが言語発達と高い相関を示すことを確認した [16]。このことから、幼児の語彙獲得において親と幼児が相互に同じ対象に注視しようとする性質は語彙獲得において重要な働きをしていると言える。この性質は語彙獲得の初期段階の幼児においても同様に見られる [17]。

意図の理解 母親が幼児の注意を追うことと語彙数の間には強い相関があることが分かっている [18, 19, 20] が、幼児は語彙獲得において受動的な存在であるわけではない。Baldwin の実験 [21] では幼児は親の意図について積極的に理解しようと試みており、親と幼児による相互行為によって語彙獲得は進められていることが示された。また、Tomasello が行った3つの実験では「親の表情から正解の物体を選択できること」 [22, 23]、「単純な親の表情で判断するのではなく、その背景にある意図を理解すること」 [24]、「親が注視している最中でなくても推論から意図を理解することができる」 [25] といったことが確認された。これらのことから幼児は複雑な状況下においても自ら親から意図を汲み取ろうと

し、適切な語彙獲得を行うことができることが分かる。このように語彙獲得は母子間での共同注視だけでなく、親の意図の理解によって進められている。

制約の獲得 幼児の語彙獲得において注目する実態について無限の選択肢から選択することは不可能である。そのため、これまでは生得的な制約を考えてきた [2]。しかし、実際には生得的な制約で説明できない事象も存在する [4, 5, 6, 7]。そこで本研究では、このような制約が母子の相互行為によって後天的に獲得されることを示す。

2.2 人間の心的表象

人間がどのように物体を捉えているかは幼児の語彙獲得において重要な課題の一つである。特に物体をどう心的表象として保持しているかは語彙獲得における物体の特徴量設計に必要な指標である。この人間の心的表象に関する研究として Martin の研究では幼児の言語獲得のモデル化をおこない、物体間の潜在的な類似性構造に注目し、ペア選択タスクを通して、説明可能な49次元の心的表象を提言している [13, 14]。これらの心的表象は再現性が高く、カテゴリ分類、典型性判断、特徴量評価など、特定のタスクだけでなく普遍的な心的表現であることも確認されている。

本研究ではこの心的表象に注目し、Martin の心的表象に基づいて人手で評価した特徴量を幼児の物体に対する特徴量として用いる。

2.3 幼児の内発的な動機付け

多くの生物が報酬がない場合であっても、探索的、遊び的、好奇心主導での行動を取ることが発見されている [27]。それは人間であっても例外ではなく、好奇心による行動を行うことが認知的な発達において重要な要素であり、人間が本来持っている興味に基づいて行動することは知能を成長させることに繋がります [28]。これは幼児の語彙獲得においても確認されており、幼児の語彙獲得は対象となる語を一律に獲得するのではなく、時期により集中的に獲得されている [29]。これは外発的な動機付けでは、説明できない現象であり、幼児自身が内発的な動機付けにより、語彙獲得を行なっていると言える。

本研究では従来手法が外発的な動機付けに基づく語彙獲得がなされていたのに対して、このような幼児の内発的な動機付けも考慮した語彙獲得モデルを構築する。

3 幼児語彙獲得モデル

3.1 モデルの概要

モデルの構造 幼児語彙獲得機構モデルに深層強化学習を用いる。その理由として、強化学習とは環境の中で試行錯誤を繰り返しながら報酬を最大化して学習する手法であり、幼児の語彙獲得のプロセスと酷似していることが挙げられる。また、語彙獲得における状態数は無限に近いので、深層学習を用いる必要があることも挙げられる。モデルは LSTM（緑）と二つの DQN（青）からなる。LSTM は幼児の心的状態を表し、二つの DQN はその心的状態に基づき、幼児の行動を出力する。DQN1 は物体の特徴量を選択する行動を出力し、DQN2 は幼児が思い描く物体の名前を答える行動を出力する。

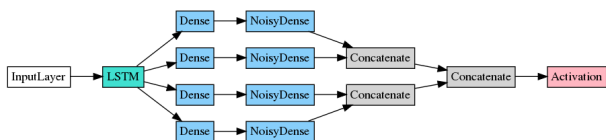


図 1: 幼児語彙獲得モデルの構造

報酬 従来の幼児語彙獲得モデルでは母子間の相互作用の中で正解か不正解かのみを報酬とする DRQN[30] を用いていた。

$$r_t^e = \begin{cases} 1 & \text{正解のとき} \\ -1 & \text{不正解のとき} \end{cases} \quad (1)$$

しかし、本研究では物体の特徴量の種類が従来研究の 4 個から 49 個と増えており、学習の初期段階で物体の名称を正解するのに必要な特徴量を選択することが難しく、学習初期の段階では物体の名称を正しく答えられないため、DRQN モデルでは正の報酬を与えることが困難であった。また、実際の幼児の語彙獲得は母子間の相互作用のみではなく、幼児自身が未知のものに対して関心を持つことで進む。

そこで本研究では新たな特徴量を探索した際に報酬を与えるという内部報酬を導入した NGU[31] を基にしたモデルを構築する。具体的には従来の報酬 r_t^e に加えて内部報酬 r_t^i を考慮した以下の報酬 r_t を従来の報酬の代わりとする。ただし、本実験では同エピソード内で状態の遷移が不可逆に進むため、同エピソードで同じ特徴量の選択状態になる可能性は前ステップの状態と同一な時のみである。そこで現在の状態が前ステップと同一かを判定し、その状態を

訪れた回数 n の逆数をエピソード報酬とする。また、 L は任意の定数、 α_t は RND[32] で提案されている値を用いる。

$$r_t^i = \frac{1}{n} \times \min(\max(\alpha_t, 1), L) \quad (2)$$

$$r_t = r_t^e + r_t^i \quad (3)$$

3.2 学習の流れ

語彙獲得は母子間の意図の理解を介して行われる。図 2 は幼児語彙獲得モデルの一連の語彙獲得の流れであり、語彙獲得のアルゴリズムを以下に示す。まず、共同注視と意図の理解により母子間で共通の物体に対する注視が起こり、語彙獲得が始まる。次に、物体の特徴量と幼児の行動を持った初期状態 s_0 を心的状態を表す LSTM に入力する。その出力結果（心的状態）を DQN1 に入力として与える。これにより、DQN1 はどの特徴を選択するかを出力した後、選択した行動の種類とその選択に対応する特徴量を s_1 に追加する。この s_1 を次の LSTM に入力し、その出力結果（心的状態）を DQN2 の入力として与えることで、DQN2 は語の選択を行い、親へと伝える。選択された語が親の意図した物体と一致していれば正の報酬を、不正解ならば負の報酬を与える。この処理を語が一致するか定められた回数まで繰り返す。

本研究で構築する幼児語彙獲得モデルは初期状態では事前知識を持たず、特徴の選択をランダムに行うが、語彙獲得が進むにつれて、語彙を決定するのに適した特徴の選択すなわち制約を獲得する。



図 2: 語彙獲得の一連の流れ

4 実験

4.1 実験設定

データセット データセットには「apple, orange, ball, book, dog, cat, horse」の 7 つの語について Martin の研究で [13, 14] で言われている 49 次元について人手でのラベルの付与を行い、ランダムなノイズを加え、各語あたり 100 物体を作成、80% を学習用、20% を検証用に分割して実験のシミュレーションを行なった。

実験の流れ 「幼児モデルが物体に対して事前に与えられた語候補の内のいずれかを答える」を1ステップ、「モデルが正解の語を答えるか、ステップ数が打ち切り回数を超える」を1エピソードとした。本実験ではエピソード数を100000回、打ち切り回数を10回とした。また、1000エピソード学習する毎にテスト用のデータセットで正答率を測定した。

比較モデル ベースラインは田口らによるDRQNモデル [11] と学習を行わないランダムモデルをベースラインとして用いた。

4.2 実験結果

語彙獲得 ベースラインとの比較結果を図3に示す。本研究では、従来研究とは異なり一つの特徴量が持つ情報量が極めて少ない情報量空間で特徴をいかに組み合わせるかが重要であったが、ベースラインとの比較から、提案手法が顕著に語彙獲得に成功していることがわかる。

特に、DRQNモデルとの差から従来の研究と異なり、幼児語彙獲得モデルが報酬を得られるまでに時間がかかる状況の中でも幼児の内発的な動機付けである好奇心を表す内部報酬を用いることで適切な探索を行っており、学習に成功していることがわかる。これらから幼児の持つ好奇心が語彙獲得に有効に働くことがわかる。

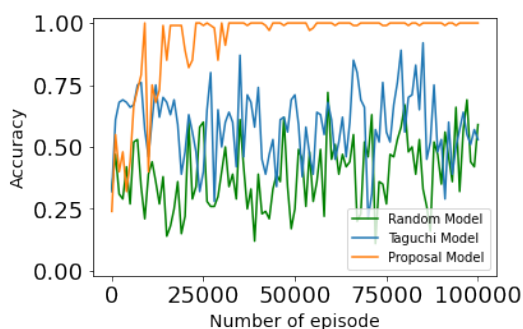


図3: 語彙獲得の結果

特徴量の選択回数 幼児語彙獲得モデルが正解するまでに必要な特徴選択回数を表1に示す。まず、特徴選択回数についてだが、学習の初期段階では平均回数が6.15回と多く、多くの特徴量を取得する必要がある。しかし、学習終盤では、2.35回と少ない回数で理想値である1.94回に近い回数で答えることが可能となっており、物体を識別するのに必要な特徴量を適切に選択できていることがわかる。

表1: 特徴量の選択回数

	Mean	25th percentile	Median	75th percentile
学習前	6.15	3	6	10
学習後	2.35	2	2	3

特徴量の木構造化 選ばれている特徴量について、10000回あたりの学習の中で選択回数が10回以下、あるいは、親ノードの1/10以下の選択回数である特徴量を枝切りし、木構造化したものを図4に示す。学習途中での特徴量選択が複雑な形であるのに対して、学習後の特徴量選択はシンプルな形になっていること、木構造の葉の数と物体の数が一致していることから最適化されていることがわかる。これらのことから幼児語彙獲得モデルは母子間の相互行為を通じた語彙獲得によって自律的に特徴量に優先度を付け、語彙獲得を行う機構を獲得したことがわかる。

これは幼児が名詞を提示し続けることで世の中には名詞が多いことを認識し、名詞を判別するための制約、例えば、形状を最初に注目することと等しい。すなわち、形状バイアスのような制約を獲得したと考えられる [33]。

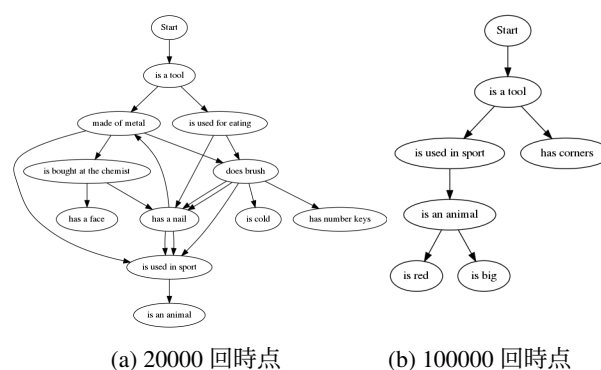


図4: 選択された特徴量の木構造化

5 おわりに

本研究では人間が物体を認識する際に用いられる心的表象に注目すると共に幼児の内発的な動機付けである好奇心を導入することで幼児の語彙獲得機構のモデル化を行った。実験結果により、提案手法が49個の特徴量の組み合わせで語彙獲得を行えることが確認された。

また、特徴量の選択規則を扱う物体の数に最適な木構造化することから生得的とされていた制約も、語彙獲得の過程で、幼児語彙獲得モデルが自律的に獲得できる可能性を示すことが出来た。

参考文献

- [1] W.V.O.Quine. *Word and Object*. MIT Press, 1960.
- [2] E.M.Markman. Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M.R.Gunnar & M.Maratsos (Eds.), *The Minnesota symposia on child psychology*, Vol.25.Modularity and constraints in language and cognition pp.59–101, 1992.
- [3] M.Tomasello, A.C.Kruger and H.H.Ratner. Cultural learning. *BEHAVIORAL AND BRAIN SCIENCES* vol.16, pp. 495–552, 1993.
- [4] C.B.Mervis, R.M.Golinkoff and J.Bertrand. Two-Year-Olds Readily Learn Multiple Labels for the Same Basic-Level Category. *CHILD DEVELOPMENT*, vol.65, Issue4, pp. 1163–1177, 1994.
- [5] G.O.Deak, and M.Maratsos. On having complex representations of things: Preschoolers use multiple words for objects and people. *Developmental Psychology*, vol.34 no.2, pp. 224–240, 1998.
- [6] E.V.Clark and T.A.Svaib. Speaker perspective and reference in young children. *First Language*, Vol 17, Issue 49, 1997.
- [7] M.Tomasello and N.Akhtar. Two-Year-Olds Use Pragmatic Cues to Differentiate Reference to Objects and Actions. *Cognitive Development*, vo. 10, pp. 201–224, 1995.
- [8] Joe Nishihara, Tomoaki Nakamura. Online Algorithm for Robots to Learn Object Concepts and Language.*IEEE Transactions on Cognitive and Developmental Systems* vol. 9, Issue: 3, September 2017.
- [9] O.Rasanen and K Khorrami. A computational model of early language acquisition from audiovisual experiences of young infants. In *Proceedings of Interspeech*, vol. 2019-September, Interspeech-Annual Conference of the International Speech Communication Association, International Speech Communication Association, pp. 3594–3598, Interspeech, Graz, Austria, 2019.
- [10] 野口輝, 南泰浩. DRQN による幼児の語彙獲得のモデル化. *言語処理学会 第 24 回年次大会 発表論文集* pp. 1260–1263, 2017.
- [11] 田口真輝, 南泰浩. ニューラルネットワーク強化学習を用いた幼児語彙獲得のモデル化. *電子情報通信学会. HCS2019–73*, 2020.
- [12] 藤田守太, 南泰浩. 母子間の相互行為に注目した幼児語彙獲得機構のモデル化. *電子情報通信学会 信学技報*, vol. 122, no. 103, TL2022–7, pp.36–40, 2022.
- [13] Martin N. Hebart, Charles Y. Zheng, Francisco Pereira, Chris I. Baker. Revealing the multidimensional mental representations of natural objects underlying human similarity judgments. *Nature Human Behaviour*. 4, pp. 1173–1185, 2020.
- [14] Charles Y. Zheng, Francisco Pereira, Chris I. Baker, Martin N. Hebart. Revealing interpretable object representations from human behavior. *ICLR 2019*.
- [15] J.Bruner. *Child’s Talk: Learning to Use Language*. Oxford University Press, vol. 1, Issue. 1, 1985
- [16] M.Tomasello and J.Todd. Joint attention and lexical acquisition style. *First Language*, vol 4, Issue 12, 1983.
- [17] M.Carpenter, K.Nagell, M.Tomasello, G.Butterworth and C.Moore. Social Cognition. Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, Vol.63, No.4, Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age, pp.1–174, 1998.
- [18] M.Tomasello and M.J.Farrar. Joint Attention and Early Language. *Child Development*, vol. 57, no. 6 December, pp.1454–1463, 1986.
- [19] P.J.Dunham, F.Dunham, A.Curwin. Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology*, vol.29 no.5, pp.827–831, 1993.
- [20] N.Akhtar, F.Dunham and P.Dunham. Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language*, vol. 18 no. 1, pp. 41–49, 1991.
- [21] D.A.Baldwin. Infants’ contribution to the achievement of joint reference. *Child Development*, vol.62, no.5, pp.875–890, 1991.
- [22] M.Tomasello, R.Strosberg, and N.Akhtar. Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, vol. 23, no.1, pp.157–176, 1996.
- [23] M.Tomasello, and M.E.Barton. Learning words in nonostensive contexts.*Developmental Psychology*, vol. 30, no.5, pp.639–650, 1994.
- [24] N.Akhtar and M.Tomasello. Two-year-olds learn words for absent objects and actions. *Developmental Psychology*, vol.14, Issue.1, pp.79–93, 1996.
- [25] N.Akhtar, M.Carpenter and M.Tomasello. The Role of Discourse Novelty in Early Word Learning. *CHILD DEVELOPMENT*, vol.67, Issue.2, pp.635–645, 1996.
- [26] White, R. W. Motivation reconsidered. *Psychological Review*, 66, pp.297–333, 1959.
- [27] Sandra O Mollenauer. Shifts in deprivation level: Different effects depending on amount of preshift training. *Learning and Motivation*, Volume 2, Issue 1, pp.58–66, 1971.
- [28] Richard M. Ryan, Edward L. Deci. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, Volume 25, Issue 1, pp.54–67, 2000.
- [29] 小林哲生, 南泰浩. 日本語習得児における語彙カテゴリ構成の発達の変遷. *電子情報通信学会技術研究報告 信学技報* vol. 114 no. 68, pp.67–72, 2014.
- [30] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fiedjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. Human-level control through deep reinforcement learning. *Nature*, vol. 518, no.7540, pp.529–533, 2015.
- [31] Adrià Puigdomènech Badia, Pablo Sprechmann, Alex Vitvitskiy, Daniel Guo, Bilal Piot, Steven Kapturovski, Olivier Tieleman, Martín Arjovsky, Alexander Pritzel, Andrew Bolt, Charles Blundell. Never Give Up: Learning Directed Exploration Strategies. Published as a conference paper at *ICLR 2020*.
- [32] Yuri Burda, Harrison Edwards, Amos Storkey, Oleg Klimov. Exploration by random network distillation. *ICLR 2019*.
- [33] Susan S. Jones, Linda B. Smith. How children know the relevant properties for generalizing object names. *Developmental Science*, vol.5, no.2, pp.219–232, 2002.