

特別支援教育における認知モデルベースロボットの位置付けと導入に向けた課題

Positioning of Cognitive Model-Based Robots in Special Needs Education and Issues for Their Introduction

西川 純平^{1*} 佐々木康佑¹ 森田純哉¹
Jumpei Nishikawa¹ Kosuke Sasaki¹ Junya Morita¹

¹ 静岡大学

¹ Shizuoka University

Abstract: AI技術等の発展により、コミュニケーションロボットの多様な場面を想定した開発・導入が進んでいる。人間とロボットの円滑な対話を実現するためには、人同士の対話のように、機械が状況に応じて言語表現や身体表現を変換する仕組みを持つことが重要になる。本研究では、言語と身体情報をつなぐ認知モデルベースのロボットを提案し、特別支援教育の文脈でどのように活用するかを位置付けることを目的とする。具体的には、特別支援教育現場における学習者がロボットに向ける注意や受容性を検討することで、既存の支援手法との補完関係や提案するロボットの新たな応用の可能性を議論したい。本発表では、フィールド実験による調査のための理論的背景や実施方法の枠組みを提示する。

1 はじめに

近年、人工知能（AI）関連技術の飛躍的な発展に伴い、コミュニケーションロボットの研究と実用化が急速に進展している [1]。このような状況下で、ロボットも人間同士のコミュニケーションのように、状況に応じた柔軟な言語表現や身体表現を実現することが重要になる。人間のコミュニケーションは記号的なあるいは数量的な表現に媒介される。記号的表現とは、言語やアイコンといった離散的な表現であり、数量的表現とは、音声や動作といった物理量に対応する。これらふたつの媒体は、人間内部においては異なるシステムにより相補的に処理される [2]。人間とロボットが円滑に対話するためには、機械も状況に応じてふたつの媒体を変換する仕組みを持つことが重要になる。具体的には、ロボットが言語情報と身体情報を相互に補完しながら、対話の文脈に応じた適切な反応を生成する仕組みが求められる。

人間内部での記号と数量の変換は、認知科学および人工知能分野でさかんに議論されてきた。ホーキンスの提唱する座標系理論 [3] は、概念の裏側に連続的な空間を想定することで、身体動作や時空間に埋め込まれた知識構造、言語などの関連を説明する。分布仮説（意味の類似した単語は類似した文脈で使用される、[4]

に基づく単語分散表現 [5] も、このような概念の裏側の連続的な空間を計算機上にモデル化する手法と言える。

人の内部プロセスを理解・予測するために計算機プログラムと対比する認知モデリングの領域においても、記号と数量を扱う工夫がなされてきた。ACT-R [6] がそのひとつの例である。ACT-R は、認知アーキテクチャと呼ばれる、認知モデル（人の認知過程に近似する計算機プログラム）を構築するための基盤（フレームワーク）のひとつである。ACT-R の主要な機能はシンボリックコンポーネントおよびサブシンボリックコンポーネントに大別される。シンボリックコンポーネントでは、モデル構築者が実装する記号が対象タスクのプロセスと構造を記述する。サブシンボリックコンポーネントによって、これらのシンボルの柔軟な処理・振舞いが設定され、連続的な世界と接続される。

これらを踏まえて、著者らは記号的な表現と数量的（身体的）表現を接続する方法を検討してきた [7]。この研究では、ACT-R に基づく機構によって、記号的表現（言語）と身体的表現（ジェスチャ）が相互作用する対話の認知モデルを構築し、コミュニケーションロボットに搭載した。さらに、記号的表現と身体的表現を接続する仕組みをもつ認知モデルベースロボットが、実世界の社会の人々の活動に及ぼす影響について検討するフィールド実験を科学博物館において実施した [8]。

本発表では、この認知モデルベースロボットのより具体的な実用可能性を検討するため、特別支援教育現

*連絡先：静岡大学

〒432-8011 静岡県浜松市中央区城北3丁目5-1
E-mail: nishikawa.jumpei.16@shizuoka.ac.jp

場に注目する。特別支援教育は、学習者の特性や個人の発達の段階に沿った対応が求められる分野であり、学習者一人ひとりの多様なニーズに応じた支援が不可欠である。特に、言語や身体感覚、あるいはその統合など、認知的領域において顕著な遅れを示す生徒が在籍すること、その遅れに多様性が存在することから、個別の支援が求められている

これらを踏まえて、本研究では、認知モデルのアプローチに則った（内部プロセスに関する説明可能な表現を持ち、個人に適応可能なパラメータを備える）機構を持つロボットが従来の支援手法とどのように補完的な関係を築き、あるいは新たな支援の可能性を拓くかを検討することを目的とする。具体的には、特別支援教育の現場で実施するフィールド実験を通じて、学習者がロボットに向ける注意や受容性の特徴を明らかにし、それらが学習効果に与える影響を検証する予定である。

2 言語と身体を融合する認知モデルベースロボット

本節では、著者らがこれまで開発してきた言語と身体を融合する認知モデルベースロボットについて概説する。図1は、このロボットの概観である¹。このモデルは簡単な対話の例としてしりとりを実施する。記号と身体性の関連は、しりとり中に登場する単語について、その「大きさ」（単語の指示対象の物理的な大きさ [9]）とロボットのジェスチャが相互に影響し合うこととして表現される。

図中の点線の領域が、認知アーキテクチャACT-R [6] を利用した記号処理システムである。ACT-R では、脳部位に対応づけられる各モジュールを、プロダクションモジュールが制御する。このモジュール構造のなかでは「チャンク」と呼ばれる記号的表現を単位として処理がおこなわれる。しりとりにおける聞き取った単語をもとに自分の回答を選択する処理が、単語や音に関するチャンクを単位として実行される。

単語選択の処理は身体としてのロボットに接続され、発話とジェスチャが影響しあう。このモデルにおける記号から数量への変換は、ロボットが名詞の「大きさのイメージ」に基づくジェスチャをとることに対応する（図1中の橙矢印）。「大きさのイメージ」は、単語分散表現から取り出される人間が単語の指示対象の物理的な大きさに対して持つ一般的なイメージである。数量から記号への影響は、直前のジェスチャの大きさが

モデルの次の単語選択に影響するというかたちで表現される（図1中の緑矢印）。

個人々に応じた支援が求められる特別支援教育の現場においてより高い学習効果を期待するには、このようなモデルが状況に応じて言語表現や身体表現を変換する仕組みを持つことが重要になると考えられる。このことを実現する第一歩として、しり通りの単語系列によって対話の文脈を表現することを考える。今回は、モデルが持つ単語の知識にカテゴリ（りんごに対するフルーツなど）の情報を付与し、ACT-R が持つパラメータのうち連想を表現するもの（活性化拡散と呼ばれる）を設定した（この拡張の詳細は [10] を参照）。これによりロボットは、文脈（カテゴリの連想）の影響を受けた単語系列（たとえば動物ばかり回答）のしりとりや、それらの影響を受けない（単語のカテゴリによらず身近な単語を回答する）しりとりといったバリエーションを持った対話を実現する。

3 実験

本原稿の執筆時点で、特別支援学校をフィールドとした、学習者がロボットに向ける注意や受容性の特徴について調査する実験を実施中である。この実験では、前述のモデルが2体のロボットに実装される。2体のロボットは、2025年2月12日から2月20日のうち5日間、静岡県立浜北特別支援学校のプレイルーム²およびいくつかの教室に時間帯をわけて展示される。実験の対象となるのは、肢体不自由を併せ持つ個人を含めた、小学校1年生から高校3年生の児童・生徒である。児童・生徒は2体のロボットが交互に単語を発話し、ジェスチャをする（しりとりを行う）様子を観察する。

今回の展示では、設計されたモデルの効果を確認するため、前述した仕組みに基づいてロボットがジェスチャをする二つの条件を用意した。

- C 条件：ロボットは、一定の文脈（単語のカテゴリ）を重視してしりとりをする。また、ロボットは単語を発話する際、その単語の「大きさ」を直前の単語の「大きさ」との比較に基づいてジェスチャする。
- NC 条件：ロボットは、文脈（単語のカテゴリ）に基づかずしりとりをする。また、ロボットは単語を発話する際、その単語の「大きさ」を直前の単語の「大きさ」との比較に基づいてジェスチャする。

この2条件に基づいてそれぞれ3つのしりとり単語系列を作成し、各展示時間中には2条件が交互に示され

¹物理的な実体としてのコミュニケーションロボットは Vston 社による Sota を利用した。
<https://www.vstone.co.jp/products/sota/index.html>

²スロープやマットが設置され、児童・生徒が遊ぶことができる室内の多目的スペース

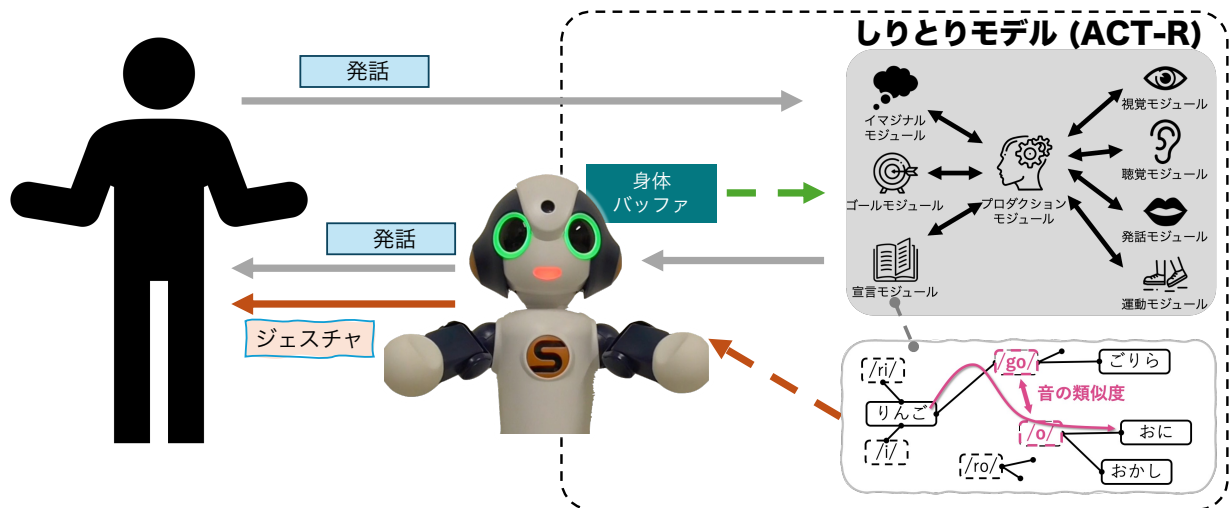


図 1: 提案モデルの概観. 先行研究 [7, 8] の図をもとに著者が再構成. 記号的表現と身体が, 単語選択からジェスチャへの影響 (橙矢印) とジェスチャから単語選択への影響 (緑矢印) として接続される.

る. また, 展示場所を円滑に移動するため, 長机の上にロボット 2 体, 制御用ノート PC, ビデオカメラ 2 台, ポータブル電源が固定された. 児童・生徒のロボットに対する反応は 2 台のビデオカメラによって記録される. また, 児童・生徒が観察する様子を見守る教職員に対して, 以下の項目からなるインタビューを実施する.

- ロボットを観察した教職員の感想
- 児童・生徒たちがロボットを見ている様子から受けた感想
- 特別支援教育の現場で, このようなコミュニケーションロボットが使えるか, どのように使えるか
 - 特別支援教育の現場で, しりとりをするロボットがどう使えると思うか
 - 特別支援教育の現場で, 話す単語の大きさのジェスチャをするロボットがどう使えると思うか
 - 話す内容と身体が連動していることの重要性をどう考えるか

このような条件のもとで, 言語と身体が連動するロボットに対し, 児童・生徒が向ける注意や受容性の特徴を調査する,

3.1 予備的な結果

本稿執筆時点で, 2 日間の展示が完了している. 小学校 1~3 年生および, 四肢不自由の中学生を対象に展示し, 11 名の教職員に対してインタビューを実施した.

ここでは, これらの途中経過において観察された顕著な例を表 1 に示す.

表に列挙した反応から, 児童・生徒に関しては, ロボットのそれぞれの要素を理解し, 反応したと考えられる行動が観察された. ただし, ロボットが「大きさ」を表現していることや, しりどりの文脈 (単語系列におけるカテゴリの連続) の違いに着目した顕著な反応は確認できていない. 教職員に対するインタビューからも, 言葉と身体動作を関連づけることの困難さが指摘されている. このような観点でロボットの振る舞いを把握・理解しているのかに関して, 今後の注意深い分析が必要である.

4 むすび

本稿では, 人間とロボットの円滑な対話を実現するためには, 人同士の対話のように, 機械が状況に応じて言語表現や身体表現を変換する仕組みを持つことが重要になるという考えから, 言語と身体情報をつなぐ認知モデルベースのロボットを提案し, 特別支援教育の文脈でどのように活用しうるかを位置付けることを目的とした. とくに, 既存の支援手法との補完関係や提案するロボットの新たな応用の可能性を議論するために, 特別支援学校において児童・生徒がロボットに向ける注意や受容性を検討する実験を示した. 今後は, 本稿で示した実験を完了する. そして, ロボットの展示に対する児童・生徒の反応および教職員へのインタビュー結果の分析を通して, 特別支援教育現場において, 認知モデルベースロボットが果たすことのできる役割を整理してゆく.

表 1: ロボットの各要素に対する児童・生徒および教職員の反応

	児童・生徒	教職員
ロボットの言葉	<ul style="list-style-type: none"> ● ロボットが「ウサギ」と発したとき、耳に見立てた両手を頭の上に立ててジャンプした <p>単語の指示対象の特徴を表現したジェスチャであると考えられる</p>	<ul style="list-style-type: none"> ● 「ゴリラだって」など、ロボットの発した単語を復唱しながら児童・生徒に話しかけた ● やや馴染みのない単語の出現（スカート → トレーニングウェアなど）に対して驚いたような声色で復唱したり、馴染みのある単語の出現（ダンプカー → あさがお）にうなずいた ● インタビューに対し、合成音声は、児童・生徒にとって聞き取りづらいとだろうと回答した <p>児童・生徒と同じ立場で観察しながら、その注意や反応をガイドするような行動であると考えられる</p>
ロボットの身体動作	<ul style="list-style-type: none"> ● ロボットの動きに合わせて両手を肩ほどの高さで広げた <p>ロボットの動作を真似たと考えられる</p>	<ul style="list-style-type: none"> ● インタビューに対し、このロボットが人型である（人を模し、自分と同じように喋る・考えているかもしれない）と理解することも一部の児童・生徒にとっては難しいかもしれないと回答した <p>ロボットがエージェンシーを喚起しない可能性に関する指摘である</p>
ロボットの言葉と身体動作の関連	<ul style="list-style-type: none"> ● ロボットが「ゴリラ」と発しながら両手を広げるジェスチャをしたとき、猿のジェスチャ（片手を頭の上で、もう片方の手を腹の前で拳を握る）をしながら「ゴリラじゃないの？」と問いかけた <p>単語の指示対象の特徴とロボットの動作が異なると指摘していると考えられる</p>	<ul style="list-style-type: none"> ● インタビューに対し、ロボットが児童・生徒と対話するような設定が好ましいと回答した ● インタビューに対し、言葉とジェスチャだけでなく胸部ディスプレイなどによって物を表示することができると児童・生徒にとってわかりやすいだろうと回答した ● インタビューに対し、実際の訓練の中でプラスチック製スプーンとその写真が同じものを指すと理解することが困難な場合があることから、一部の児童・生徒にとっては、言葉とジェスチャを関連づけ、同じものを指し示していると理解することも困難であると回答した <p>補助的な情報の必要性や、言葉とジェスチャを関連づけが困難であるとの懸念である</p>

謝辞

本研究の実施にあたり、静岡県立浜北特別支援学校の皆さまには多大なご協力をいただきました。また、静岡大学情報学部講師の白砂大先生、学生の Nethmee Rasanga Gunasekara さん、新堀耕平さん、佐々木健矢さんに実験実施を補助していただきました。ここに深謝いたします。本研究は JSPS 科研費 23KJ1049 の助成を受けて実施されたものです。

参考文献

- [1] 総務省. 情報通信白書 令和 6 年度版. Technical report, 総務省, 2024.
- [2] Barbara Tversky. *Mind in motion: How action shapes thought*. Hachette UK, Paris, 2019.
- [3] Jeff Hawkins. *A Thousand Brains: A New Theory of Intelligence*. Basic Books, 2021.
- [4] Zellig S Harris. Distributional structure. *Word*, Vol. 10, No. 2-3, pp. 146–162, 1954.
- [5] Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. A neural probabilistic language model. *Advances in neural information processing systems*, Vol. 13, , 2000.
- [6] John R. Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, 2007.
- [7] Jumpei Nishikawa, Kosuke Sasaki, Junya Morita, Alexis Meneses, Kazuki Sakai, and Yuichiro Yoshikawa. Embodied communication model in

act-r based robot. In *Proceedings of ICCM 2023, 21th International Conference on Cognitive Modelling*, 2023.

- [8] 西川純平, 佐々木康佑, 森田純哉. シンボリック表現と身体性を融合する認知モデルベースロボットの社会実装へ向けた評価の試み. 日本認知科学会第41回大会論文集, 2024.
- [9] Kosuke Sasaki, Jumpei Nishikawa, and Junya Morita. Evaluation of co-speech gestures grounded in word-distributed representation. *Frontiers in Robotics and AI*, Vol. 11, , 2024.
- [10] 西川純平, 佐々木康佑, 森田純哉. しりとり単語系列の特徴を制御する認知的要因に関する認知モデルを利用した調査. 第31回年次大会発表論文集, 2025. in press.