

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

An Attempt to Evaluate the Social Implementation of Cognitive Model-Based Robots for the Fusion of Symbol and Embodiment

西川 純平[†], 森田 純哉[†]

Jumpei Nishikawa, Junya Morita

[†] 静岡大学

Shizuoka University

nishikawa.jumpei.16@shizuoka.ac.jp

概要

人間は複数の表現を相補的に処理しコミュニケーションを行う。人間と機械の円滑な対話のためには、機械が状況に応じてこれらの媒体を変換する仕組みを持つことが重要になる。本研究では、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。ロボットの印象を調査するフィールド実験において、ロボットの発話する単語とジェスチャが同期しているとき、いくつかの項目が高得点となった。

キーワード：シンボリック表現, 身体性, 認知モデリング, ACT-R

1. はじめに

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 (Tversky, 2019)。人間と機械の円滑な対話のためには、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。このことを踏まえて、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 (Nishikawa et al., 2023)。本研究では、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。本稿では、この目的のために実施したフィールド実験について報告する。この実験を通してロボットの言語と身体が同期していると違和感なく受け入れられるという可能性を検討する。

2. 記号表現と身体表現を接続する認知モデル

図1は記号と数量 (身体) の接続を表現したモデル (Nishikawa et al., 2023) の概観である。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」 (Sasaki et al., 2024) 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 (Anderson, 2007) を利用した記号処理システムである。脳部位に対応づけられるそれぞれのモジュールを、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ー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に対して持つ一般的な大きさのイメージである [4]。数量から記号への影響は、直前のジェスチャー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)。

3. フィールド実験

この認知モデルを組み入れたロボットが、実社会の人々の活動に及ぼす影響を検討するフィールド実験を実施した。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、2024年2月3日 (土) 10:00-16:45 にかけて、科学博物館 (日本科学未来館) のハブスペースに展示された (図2)。

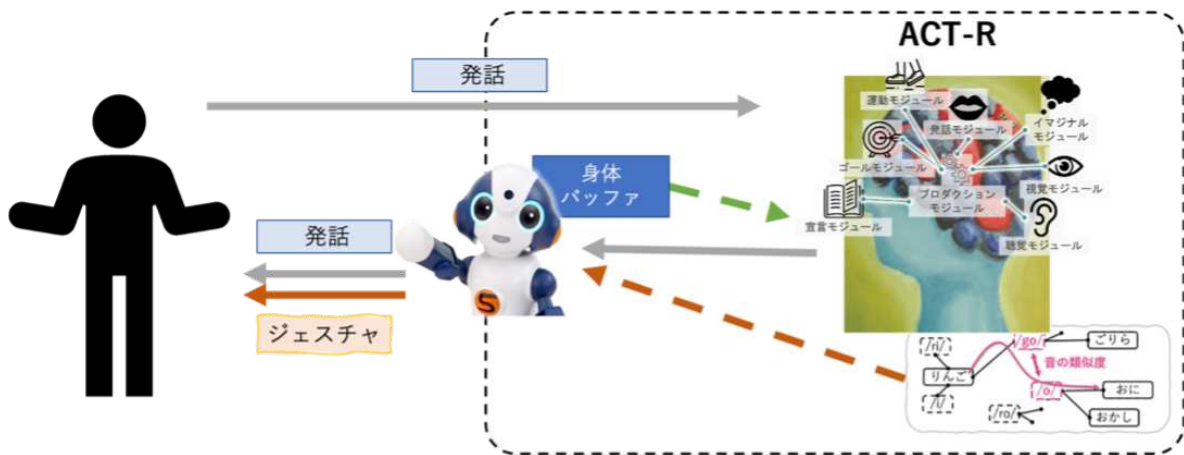


図1 提案モデルの概観. 先行研究 (Nishikawa et al., 2023) の図を著者が和訳. 記号的表現と身体が, 単語選択からジェスチャへの影響 (橙矢印) とジェスチャから単語選択への影響 (緑矢印) として接続される.

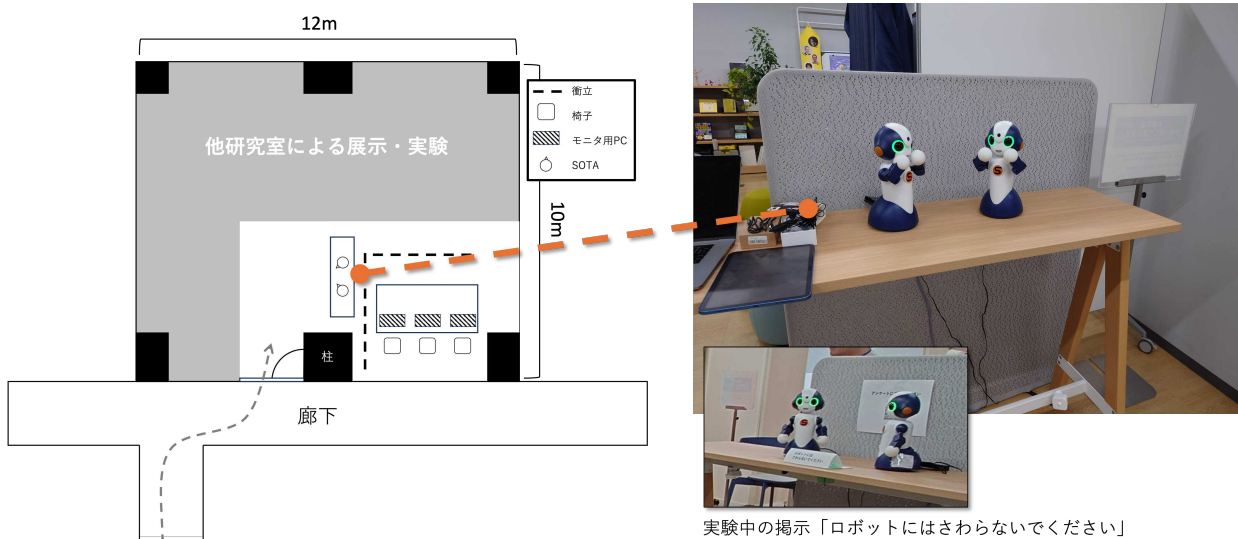


図2 展示の様子. ロボットの配置と動作中の写真.

来場者は2体のロボットが交互に単語を発話し, ジェスチャーをとる (しりとりを行う) 様子を観察した.

今回の展示では, 設計されたモデルの効果を確かめるため, 前述した仕組みに基づいてロボットがジェスチャーをする同期条件に加えて, 二つの条件を用意した.

- 同期条件: ロボットは単語を発話する際に, その単語の「大きさ」に基づいてジェスチャーする.
- 非同期条件: ロボットは単語を発話する際に, ランダムな大きさのジェスチャーする.
- 不動条件: ロボットは単語を発話する際に, ジェスチャーをしない (常に一定の姿勢を保つ).

展示時間中には3条件が繰り返し示された. ある条件のしりとりが終了する (しりとりが20回続く) と, 20秒程度のインターバルののち, 次の条件のしりとりが開始された. また, 展示会場には, 人感センサ, カメラ, ICレコーダが設置され, 来場者のロボットに対

する反応が記録された. また, ロボットの展示に興味を示した来場者に対して, アンケートへの回答を依頼した. アンケートには, 来場者の属性に関する設問, 観察した条件を区別するための設問, ロボットへの主観的な印象を評価する Godspeed 質問紙の設問が含まれた. このような条件のもとで, ロボットの言語と身体が同期していると違和感なく受け入れられるという可能性を検討する. つまり, Godspeed 質問紙の尺度において同期条件は他の条件よりも高く評価されるかどうかをたしかめる.

3.1 結果と考察

図3は得られたアンケート結果から, 来場者が自身の観察したしりどりの試行から受けた印象を Godspeed 質問紙に基づいて評価した結果を取り出したものである ($n=8$, 同期=3, 非同期=3, 不動=2). 3色のバーは, それぞれの来場者が観察した認知モデルとロボットの動作を対応付けた各条件を示す. 図の

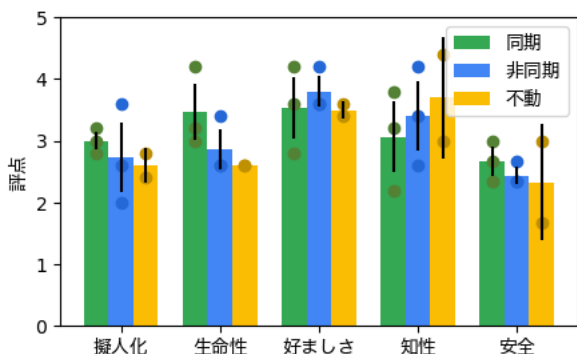


図3 アンケート結果. Godspeed の下位項目ごとに整理. エラーバーは SE. 各点は参加者.

横軸には Godspeed 質問紙の下位項目が並び、縦軸はそれぞれの評点を示す。図より、生命性の項目において、ロボットの身体が動作する2条件（同期・非同期）の評価が不動条件よりも高く、擬人化の項目でも同期条件が不動条件よりもやや高いことが観察される。これらは、ロボットの身体が動作すること、単語の意味と同期していることの効果を示す可能性がある。

4. まとめと今後

本稿では、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ことを目的として実施したフィールド実験について述べた。しりとり中の単語とジェスチャーの関係を変化させた3条件の実験の結果として、ロボットの発話する単語とジェスチャーが同期した条件においてロボットの印象評価でいくつかの項目が高得点となった。これは単語の意味とロボットの身体動作が同期していることの効果を示唆する。

今後、滞在時間など他の記録されたデータを分析し、認知モデルベースのロボットを社会実装する効果に関するさらなる知見を得る。

文献

- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?:* Oxford University Press.
- Nishikawa, J., Sasaki, K., Morita, J., Meneses, A., Sakai, K., & Yoshikawa, Y. (2023). "Embodied communication model in ACT-R based robot," in *Proceedings of ICCM 2023, 21th International Conference on Cognitive Modelling*, URL: <https://mathpsych.org/presentation/1167>.
- Sasaki, K., Nishikawa, J., & Morita, J. (2024). "Evaluation of co-speech gestures grounded in word-distributed representation," *Frontiers in Robotics and AI*, 11, DOI: 10.3389/frobt.2024.1362463.
- Tversky, B. (2019). *Mind in motion: How action shapes thought*, Paris: Hachette UK.