

多人数対話ロボットの实証実験 —ロボットと対話中の親子間会話の分析— Field Test on Multi-party Dialogue Robot -Analysis of Parent-child Conversation in Dialogue with Robot-

坪倉和哉¹ 加藤里菜² 小林邦和^{2*}
Kazuya TSUBOKURA¹, Rina KATO², Kunikazu KOBAYASHI²

¹ 愛知県立大学大学院 情報科学研究科

¹ Graduate School of Information Science and Technology, Aichi Prefectural University

² 愛知県立大学 情報科学部

² School of Information Science and Technology, Aichi Prefectural University

Abstract: 本研究では、雑談対話ロボットシステムを構築し、実環境での対話データの収集を目的としてイベント会場にて実証実験を行った。実証実験には22グループが参加し、同意を得ることができた20グループ(64人)の対話の様子を収録した動画および音声データを収集した。また、対話後にはアンケートを実施した。アンケートの結果、多くの項目で設計したシステムの有効性を示す評価が得られたが、実験に参加したユーザが対話以前からロボットに興味を持っていたことも考えられる。また親子で実験に参加するグループが多かったため、収集した動画データを用いて、ロボットと対話中の親子間会話の分析も行い、親子間の会話の事例をまとめた。その結果、ロボットと対話中の親子間インタラクションの構造の一部を確認できた。

1 はじめに

本稿では、愛・地球博記念公園(モリコロパーク)にて行った対話ロボットの実証実験について報告する。近年、大規模言語モデルの登場により自然な雑談対話文を生成することが可能になりつつある[2]。「対話システムライブコンペティション4」のように、テキストチャットにおいてはユーザによる大規模な評価実験が行われている[2]。同様に対話ロボットに関しても、「対話ロボットコンペティション2022」が開催され、タスク指向対話ロボットに関する実証実験が行われている。しかしながら、雑談対話ロボットについては実証実験を行っている研究が少ない。

そこで本研究では、大規模言語モデルを導入した雑談対話ロボットシステムを構築し、実証実験を行った。実証実験は「モリコロパークなつあそび」の開催イベントとして、2022年7月17日から7月18日に行った。実験では、参加者にロボットと自由に対話を行ってもらった。対話終了後には参加者にアンケートを実施するとともに、同意の得られた場合はビデオと音声の記録も行った。本研究で収集したデータを分析すること

で、現状の雑談対話ロボットシステムの問題点・課題点を明らかにすることが可能となる。

本実験は子供向けのイベント「モリコロパークなつあそび」に併催する形式で行ったため、参加者は子供が多い結果となった。また、対話実験を行う場合、通常ロボット1台対ユーザ1名であることが多いが、本実験ではロボット1台に対して複数人が対話に参加することを許している。そのため、対話参加者のうち親子での参加が多かった。ロボット1台とユーザ複数人の会話を分析した研究もあるが[1]、親子間の会話に着目した研究は見あたらない。しかし、家庭用対話ロボットを検討する上で、ロボットが親子などの複数人と対話を行う状況は考えられる。そこで、本研究では調査の第一段階として、アンケートによる分析に加え、ロボットと対話中の親子間でどのような会話が行われるかについて分析を行うこととした。

以下、2章では構築した対話ロボットの設計と実験デザインについて述べる。3章では実験参加者に行ったアンケート評価について、4章ではロボットと対話中の親子間会話の分析について述べる。そして5章で本稿をまとめる。

*連絡先: 愛知県立大学 情報科学部
〒480-1342 愛知県長久手市茨ケ廻間1522-3
E-mail: kobayashi@ist.aichi-pu.ac.jp

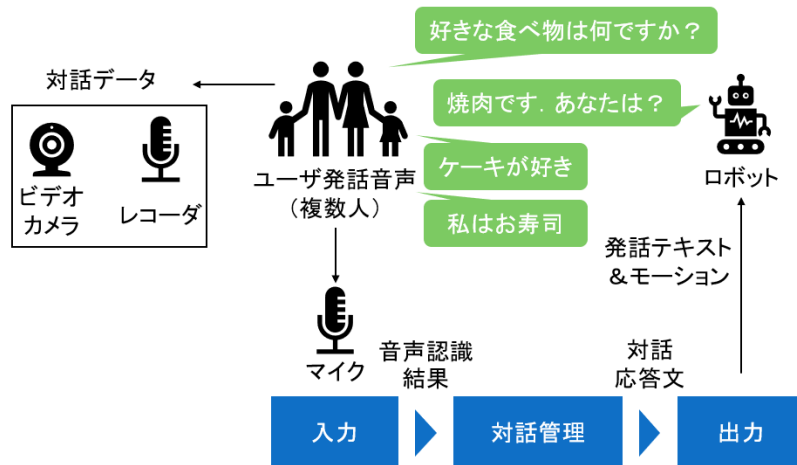


図 1: 対話ロボットシステムの設計

2 対話ロボットの設計と実験デザイン

2.1 対話ロボットの設計

本節では、実験で用いた対話ロボットの設計について述べる。本システムは、「入力部」、「対話管理部」、「出力部」から構成されている（図 1）。

「入力部」ではユーザの音声発話を受け取る。「対話管理部」では、ユーザ発話に対応する応答文を生成する。「出力部」では、対話管理部で生成された応答文を出力する。

対話ロボットとしては、ソフトバンクロボティクス社の NAO を用いる。NAO は高さ約 60cm のヒューマノイドロボットであり、介護や教育の分野でも用いられる親しみやすい設計のロボットである。

以下、システムの各部について詳述する。

2.1.1 入力部

入力部では、マイクによりユーザの発話音声を取得し、Google Cloud の Speech-to-Text API¹を用いて、発話テキストを得る。発話テキストは対話管理部に送られる。

2.1.2 対話管理部

対話管理部では、発話テキストに対応する応答文を生成する。本システムでは応答文生成システムとして生成ベースの雑談対話システムを採用した。具体的には、NTT が提供する日本語 Transformer Encoder-decoder



図 2: 実験場所

対話モデル「japanese-dialog-transformers」[3]を用いた。

2.1.3 出力部

対話管理部によって生成された応答文を音声合成し、NAO のマイクから出力する。音声合成は NAOqi API の ALTextToSpeech を用いる。

2.2 実証実験

実証実験は「モリコロパークなつあそび」の出展イベントとして、2022 年 7 月 17 日から 7 月 18 日に行った。実験場所はモリコロパークの交流センターの多目的スタジオ（室内）である（図 2）。NAO を机上に配置し、実験参加者は椅子に着席した状態でマイクに向かってしゃべりかけることでロボットと対話を行う。マイクの後ろにレコーダを設置し対話中の音声を記録するとともに、ロボットに向かって右側にビデオカメラを設置し参加者の対話の様子を録画した。ロボット背

¹<https://cloud.google.com/speech-to-text/>

後のパーティションの後ろにシステムを動作させるためのワークステーションを配置している。

対話は22グループが参加し、事前に同意を得ることができた20グループ（合計64人、うち36人は子供）に対して対話中の様子を録画・録音を行った。また、対話終了後にアンケートを実施した。機材の不備でアンケートを実施できなかった1グループを除く21グループのアンケートデータを得た。

2.3 対話の流れ

実験参加者には、実験前に実験目的を伝え、同意を得られた場合のみ音声と動画を記録した。対話上の留意点として、以下の4点を対話前に実験参加者に伝えた。

- ロボットへの話しかけはマイクに向かって話す。
- 参加者間の会話も音声入力されてしまう可能性があるため、参加者間の会話はなるべく小さい声で行う。
- （マシンスペックの関係でシステムが即時に返答することが困難であったため、）ロボットにしゃべりかけた後は、少し待つ。
- 対話を終了したいときは、「さようなら」や「バイバイ」と発話する。

2.4 対話後アンケート

対話終了後には実験参加者にアンケートを実施した。アンケートはグループ毎に1度回答してもらった。アンケートでは以下の項目を尋ねた。また、ロボットとの対話について、感想や気づいた点などを自由に記述してもらった。

1. 将来利用したいかどうかについての質問（1項目）
2. ロボットとの会話の自然性に関する質問（4項目）
3. 会話者の対人コミュニケーション認知に関する質問（6項目）

1点目として、「近い将来、今回会話したような、雑談ができるロボットを使ってみたいですか？」という質問に対して、7段階の評点（1が「使いたくない」、7が「使ってみたい」）で回答してもらった。さらに理由や補足があれば自由に記述してもらった。この項目は、マルチモーダル対話コーパスの収集を行っている文献[4]を参考にした。

2点目として、ロボットとの会話の自然性に関する以下の4項目に対して、5段階の評点（1が「不自然、できていない」、5が「自然、できた」）で回答してもらった。

- 発話内容の自然性：発話内容が自然かどうか
- 話題追従：ロボットは適切に応答できたかどうか
- 話題提供：ロボットは新たな情報を提供できたかどうか
- 話し方の自然性：音声やジェスチャー、表情などで示される話し方が自然かどうか

この4項目は、オーディエンスの前で実際に対話システムを動作させ、評価を行うイベントである「対話システムライブコンペティション5」²の評価基準を参考にした。さらに理由や補足があれば自由に記述してもらった。

3点目として、会話者の対人コミュニケーション認知に関する6項目に対して、8段階の評点（1が「全くそうでない」、8が「全くその通りである」）で回答してもらった。この6項目は会話者の対人コミュニケーション認知に関する測定項目18項目[5]から選定した。18項目のアンケートは実験参加者を時間的に拘束してしまうと考えられた。そのため実験参加者の負担を低減するため、18項目から雑談対話ロボットに関する評価として適切であると考えられるものを優先して以下の6項目とした。

- 会話に退屈していた
- 不満足な会話であった
- テンポの悪い会話であった
- 好意的に会話ができた
- 活発に会話した
- 会話は価値あるものだった

3 アンケートによる評価

本章では、対話終了後に実験参加者に実施した、対話ロボットに関するアンケート評価について分析を行う。アンケートは対話者本人または付添者が回答した。アンケート項目は2.4節で述べた11項目である。アンケートは21グループから回答を得た。

1点目の「近い将来、今回会話したような、雑談ができるロボットを使ってみたいですか？」に対する質問については、平均で5.10点であった（図3）。1が「使いたくない」、7が「使ってみたい」であることも考慮すると、使ってみたくないユーザより使ってみたいとする実験参加者が多いと言える。しかし、家庭内で使用することを考えると、ロボットの価格や安全性の面での課題があり[6]、長期的な使用における印象の変化も検討しなければならない。

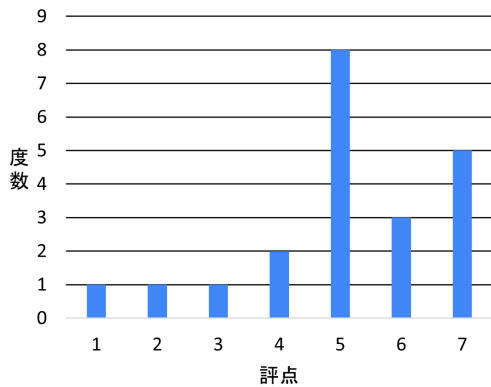


図 3: 将来利用したいかどうか (1が「使いたくない」、7が「使ってみたい」)

表 1: ロボットとの会話の自然性の平均

項目	発話内容	話題追隨	話題提供	話し方
平均	2.95	2.62	3.52	3.43

2点目のロボットとの会話の自然性に関する4項目については、項目毎の平均を表1に、度数分布図を図4に示す。アンケートは1(不自然, できていない)から5(自然, できた)の5段階であるため、「話題提供」と「話し方」はやや高く、一方、「話題追隨」はやや低い評価となった。実験参加者は子供が多く、また対話者だけでなく、付添者も参加した対話も多かったため、同時に複数人がロボットに話しかける状況や実験参加者同士の会話が認識されてしまう場合が見られた。そのため、複数話者の音声認識が正しく行われなかったことが、対話者の話す話題に適切に追隨できなかった要因の一つであると考えられる。

3点目の会話者の対人コミュニケーション認知に関する6項目については、項目毎の平均を表2に、度数分布図を図5に示す。「退屈」、「不満足」、「テンポが悪い」の3項目は反転項目である。これら3項目の中でも、対話のテンポの評価が低い結果となった。実験に用いたワークステーションの性能の関係で対話応答に10秒以上かかることがあったため、このような結果になったと考えられる。ロボットが対話者に反応するまでに2秒以上かかってしまうと印象が悪化することが報告されており[7]、音声認識にかかる時間も含めて2秒以内に抑えるべきであった。「テンポが悪い」を除く5項目については、評価が高い結果となったため、ロボットの反応に遅延が発生していたことを除くと、実験参加者におおむね良い印象を与えられたと考えられる。

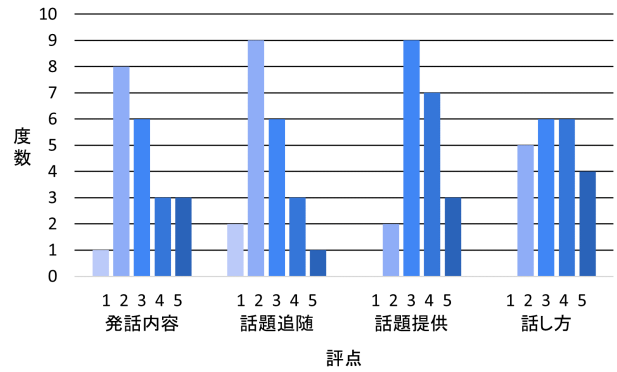


図 4: ロボットとの会話の自然性 (1が「不自然, できていない」、5が「自然, できた」)

4 親子間会話の分析

本研究ではロボット1台と複数人の対話者の会話の中でも、親子間の会話に着目する。家庭用対話ロボットを考える上で、対話者1人の状況だけではなく、対話相手が複数である場合も考慮する必要がある。対話者が複数の場合、対話者同士の会話も考慮して対話を行う必要があるため、ロボット1台対対話者1人の対話と比べ、より複雑な問題となる。そこで、本研究では対話者同士の関係性を親子関係に限定し、分析を行う。具体的には、ロボットと対話中の親子間会話の事例を観察し、分類を行う。

分析対象となる対話は、実証実験で収集した20対話のうち、親子間会話が確認された14対話である。親子間対話を「親から子供に対する働きかけ」と「子供から親に対する働きかけ」の2つに大きく分けた。

親から子供に対する働きかけとしては、表3に示すような事例が確認された。ロボットからの質問に対して困っている子供に対して、親は答えを提案したり質問内容を子供にわかりやすく説明するなど、子供とロボットの会話を円滑に進むように支援するような働きかけが見られた。

子供から親に対する働きかけとしては、表4に示すような事例が確認された。ロボットから質問をされた場合、子供は親に確認したり親に回答を頼むような働きかけが見られた。このように、子供はロボットに直接発話することを躊躇い、親に対して会話に関与するように働きかけるような発話が確認された。

以上のように、本研究で収集したデータには、ロボットと対話中の親子間会話には、親が子供とロボットとの会話をサポートする働きかけと、子供が親にロボットとの対話において支援を求める働きかけが確認された。今後は実証実験を引き続き行い対話データを収集することで、親子間会話の構造を詳細に分析する。これにより、ロボットが親子間会話を理解し、ロボット

²<https://sites.google.com/view/ds1c5>

表 2: 会話者の対人コミュニケーション認知の平均

項目	退屈	不満足	テンポが悪い	好意的	活発	価値がある
平均	3.67	3.62	5.19	5.76	4.52	5.19

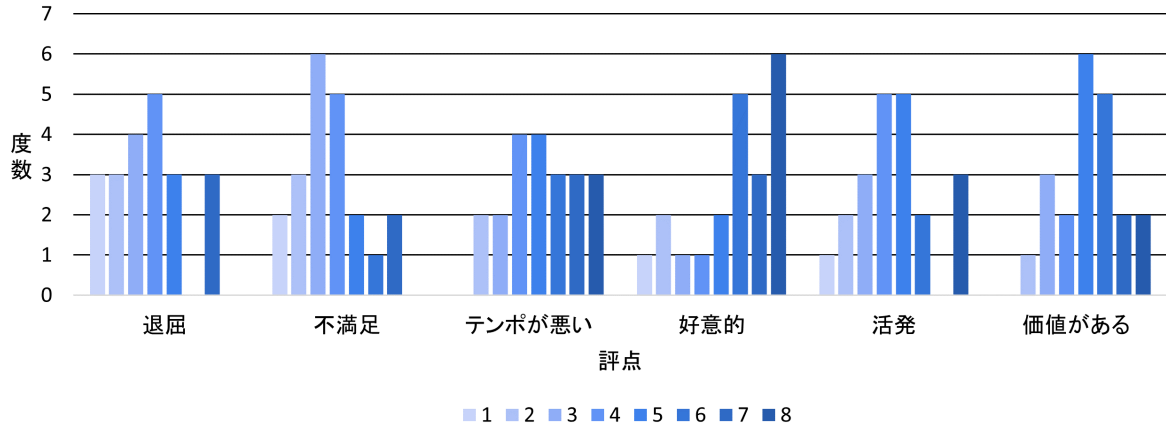


図 5: 会話者の対人コミュニケーション認知 (1 が「全くそうでない」, 8 が「全くその通りである」)

表 3: 親から子供に対する働きかけ

分類	事例
ロボットからの質問に対する答えを提案する	ロボット「どこに住んでいますか？」 親→子「xx だね (xx は都道府県)」
ロボットからの質問に対する詳しい答えを聞く	ロボット「苦手な食べ物がありますか？」 親→子「苦手な食べ物の名前言ってみたら？」
ロボットからの質問内容を噛み砕いて説明する	ロボット「趣味はなんですか？」 親→子「好きなことは何？」
ロボットへの質問内容を提案する	親→子「何歳ですか？と聞いてみたら」

表 4: 子供から親に対する働きかけ

分類	事例
子供がロボットに答える前に親に確認する	ロボット「どこに住んでいますか？」 子→親「xx (xx は都道府県)？」
子供が恥ずかしがってロボットからの質問を親に答えさせようとする	ロボット「好きな食べ物はなんですか？」 子→親「お母さんが言って」

に向けられた発話か否かを判断する，親子間会話に参加することで対話を活性化させる，といったことが可能になると考えている。

5 まとめ

本研究では，雑談対話ロボットシステムを構築し，実証実験を行った．アンケートでは高い評価を得た項目も確認できたが，本実証実験には対話ロボットに元々興味のあるユーザが参加したと想定されるため，対話前から対話ロボットに良い印象を持っていたことも考えられる．そのため，今後は対話前後の印象評価アンケートも行い，ロボットとの対話を体験することによる印象の変化を見ることも検討する．

ロボット対話中の親子間会話の分析では，データ数の関係から事例を挙げるのみとなったが，ロボットと会話中の親子のインタラクションの構造の一部を確認できた．引き続き実証実験を行い，親子間会話の構造を分析する．

参考文献

- [1] 石川 真也, 船越 孝太郎, 篠田 浩一, 中野 幹生, 多人数対話ロボットの実現にむけたマルチモーダル対話データの収集と分析, 第 27 回 (2013) 人工知能学会全国大会論文集, 2013.
- [2] 東中 竜一郎, 船越 孝太郎, 高橋 哲朗, 稲葉 通将, 赤間 怜奈, 佐藤 志貴, 堀内 颯太, ドルサ テヨルス, 小室 允人, 西川 寛之, 宇佐美 まゆみ, 対話システムライブコンペティション 4, 人工知能学会研究会資料 SIG-SLUD-093-17, pp 92-100, 2021.
- [3] Hiroaki Sugiyama, Masahiro Mizukami, Tsunehiro Arimoto, Hiromi Narimatsu, Yuya Chiba, Hideharu Nakajima and Toyomi Meguro, Empirical Analysis of Training Strategies of Transformer-based Japanese Chat-chat Systems, arXiv:2109.05217, 2021.
- [4] 駒谷 和範, 岡田 将吾, 堅田 俊, マルチモーダル対話コーパス Hazumi 公開と生体信号を含む新規データ収集, 人工知能学会研究会資料 SIG-SLUD-C002-35, pp.170-177, 2020.
- [5] 木村 昌紀, 余語 真夫, 大坊 郁夫, 感情エピソードの会話場面における表出性ハロー効果の検討, 感情心理学研究, Vol.12, No.1, pp.12-23, 2005.
- [6] 入鹿山 剛堂, 暮らしの中で活躍する AI とロボット : 1. 暮らしの中のロボット・AI -家庭用ロボットの現状と将来-, 情報処理, Vol.59, No.8, pp.686-691, 2018.
- [7] 志和 敏之, 神田 崇行, 今井 倫太, 石黒 浩, 萩田 紀博, 安西 祐一郎, 対話ロボットの反応時間と反応遅延時における問投詞の効果, 日本ロボット学会誌, Vol.27, No.1, pp.87-95, 2009.