

短時間育児支援のための対話ロボットの研究

Research on Dialogue Robots for Short-Time Childcare Support

豊田 将也[†], 金野 武司[†]

Masaya Toyoda, Takeshi Konno

[†] 金沢工業大学 工学部 電気電子工学科

Kanazawa Institute of Technology, College of Engineering,

[†] Electrical and Electronic Engineering

c1016125@planet.kanazawa-it.ac.jp, konno-tks@neptune.kanazawa-it.ac.jp

概要

本研究では、AIによる短時間育児支援について、幼児が飽きない対話方法および内容を検討するために、ロボットをアバターにした保育士と園児の対話環境を構築し、AIに必要な能力とその水準を明らかにすることを目的とした。対話実験に参加した園児は計7名であり、2名は1回のみ、他5名は日を置いて2回の実験に参加した。各回の園児の様子について、保育士は普段の様子との違いをアンケートで回答した。結果、2回程度のロボットとの対面では、たとえロボットから流れる音声の日頃から接している保育士であってもそれと気づくことはなく、かつ園児たちは普段よりも大きく緊張した状態になり、ほとんど対話できないことが確認された。他方、そういった状態の園児に対して保育士が語りかける方法と内容は、ある一定の傾向を持ちながらも多岐に渡り、AIが再現すべき発話サンプルを十分に確保できることがわかった。保育士の語りかける方法と内容を大規模言語モデルで再現する場合に、こういった難しさがあると考えられるのかを議論する。

キーワード：人工知能，育児支援，対話，ロボットアバター

1. はじめに

近年の日本は少子化や核家族化、地域との関わりの希薄化により、育児における養育者（主に母親）の孤立化が起こっていると考えられる。子育て中の養育者に対して、周りからのサポートがない場合には、育児ストレスを解消できずうつ傾向になる可能性があることも指摘されている [1]。このことから、子育て世帯への社会的な育児支援は急務な状態であると考えられる。阿部ら [2] は、末子に未就学児のいる親を対象にAIの受容性についてのアンケートを行った。そのアンケートの中では、「10分でいいから育児を手伝っ

て欲しいか」という質問に対し、6~7割の親が「ほとんど毎日」、「週3, 4回」と回答した。これは多くの養育者が普段の育児を大変だと感じていることを表した結果だと考えられるが、我々は10分程度の短時間であっても、育児支援が望まれる状況があるのではないかと考えた。例えば養育者が家事（料理や洗濯物の取り込みなど）に取り組む際に、幼児を自分の方に近づけたくない、かつ幼児から目を離さなければならない状況が存在するのではないかと考えられるからである。このような状況で、養育者が安心して幼児から目を離すことができるようにすることは、短時間の育児支援として潜在的に望まれているのではないかと考えられる。

それでは、ロボットやAIが10分程度の短時間でも、幼児の気を引くためにはどのようなことができれば良いのだろうか。この実現に最も重要と考えられることの1つは、幼児を飽きさせないようにすることであろうと思われる。ここでYoutubeのようなデジタルメディアの視聴も1つの手段ではあるが、そういった一方向的なメディアの幼児単独での利用は、幼児の発達や健康に関する観点から不適切であることが指摘されている [3]。この観点から、幼児とのやりとりにおいては相互のコミュニケーションが重要だと考えられる。阿部ら [4] は、遠隔操作型の幼児向け対話ロボットを用いて、遠隔保育を目指した研究を行っている。この研究の場合、ビデオチャットによる対話を行うのは、対象の幼児とは離れた場所にいる幼児の祖父母や親戚、保育士である。この人間が行う部分をAIが代替することができれば、経済的な助けになると考えられる。この観点に加え、ロボットにジェスチャや移動を可能にするような身体を持たせることなく、対話のみでどこまで幼児の気を引き続けることができるのかを検証することは大きな価値があると考えられる。したがって本研究では、対話のみで幼児の気を引き続け

ることの可能性の検証を目指した。

大規模言語モデル (LLM) を利用した対話サービスの発展は近年著しく、音声会話型の AI サービスは成人向けのもの [5] だけでなく、アニメキャラクターを活用した幼児向けの対話サービスも始まろうとしている [6]。しかし、実際にこれら AI との対話が幼児にどのような影響を与えるのかについては、丁寧な検証が必要になるものと考えられる。加えて、AI がどのような話し掛けをすることが適切なのかを検討する必要もあるのではないと思われるが、その知見が十分に積み重ねられたとは言えないのではないだろうか [7]。

LLM に学習させるデータとしては、例えば親と幼児が対話する様子を収めたコーパスの利用が考えられる。一例として、NTT 乳幼児音声データベース [8] の内容を参照すると、そこには大きく 2 つの観点で課題があることがわかる。1 つは、当たり前だが幼児の対話相手は親であり、AI やロボットではないという点である。幼児の視点から見て、親という認識による対話内容は、親という認識を前提としていない対話型 AI (LLM) に学習させるデータとしては不適切だと考えられる。そしてもう 1 つは、コーパスに収められた対話内容のほとんどが、絵本やおもちゃを介した三項関係の対話であり、共通の体験やエピソードを基にした二項関係の対話例が少ない点である。我々は、対話のみで幼児の気を引き続けることができるのか検証することを課題とするため、LLM の学習データとしては不十分であると考え、そこで本研究では、ロボットをアバターにした保育士が園児と対話する環境を構築し、そこで交わされる対話内容を収集することで、ロボットや AI に必要とされる幼児への話しかけ方や内容についての知見を得ることを目的とした。

2. 対話事例収集環境の構築

幼児に対して、ロボットや AI からはどのような話しかけが適切なのだろうか。あるいはまた、そういったロボットや AI からの話しかけに対して、幼児はどのように反応するのだろうか。この 2 つを同時に検証できるようにするために、我々は保育士がロボットの役を演じ、そのロボットと対面する幼児とロボット越しに対話する様子を収録する方法を考案した (図 1)。

収録では隣接する 2 つの部屋を用意し、それぞれを部屋 A, B とした。部屋 A にはテーブルの上にロボット (タカラトミー社製 Robi ジュニア) とその背後に円筒形スピーカ (SONY 社製 SRS-XE300) を図 2 のように設置し、同じ部屋の園児からは見えない位置に設置したノート PC (Apple 社製 Macbook) と Bluetooth

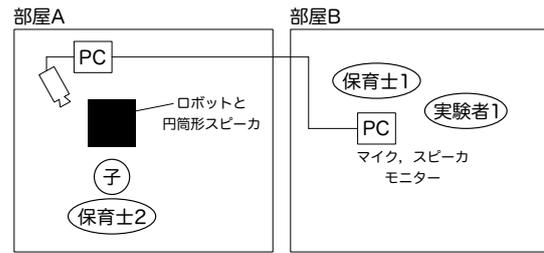


図 1 実験概要図



図 2 収録風景

により無線接続した。この PC は部屋 B に設置するノート PC に無線 LAN で接続し、Zoom でやりとりできるようにした。これにより、部屋 B に設置した PC は、部屋 A と映像および音声を使ってやりとりできるようになっていた。

収録に参加する園児は、保育士 2 の誘導により部屋 A に入り、ロボットの前に設置された椅子に腰掛けた。保育士 2 はそのまま園児の横で待機した。このとき、保育士 2 には対話には極力参加しないことを原則としながら、園児が困った様子を見せる場合には手助けをして構わないこと、および園児がロボットに触れることを静止する必要はないことが説明された。園児の対話相手となる保育士 1 は部屋 B に設置された PC の前で待機し、園児の着席後に話しかけを開始した。保育士 1 には、冒頭のみ「ボクは機械なんだけど、おしゃべりが大好きなんだ！○○くん/ちゃんとお話がしたくて、△△先生に連れて来てもらったんだ！ボクとお話してくれるかな？(○○に入るのは園児の名前、△△に入るのは保育士の名前)」といった内容の発言を行なうことが求められ、それ以降は自由に対話することが依頼された。対話時間の目安は 10 分程度としたが、園児の様子によってはそれ以前に切り上げることも可能とした。時間の経過については部屋 B に待機する実験者が適宜提示した。

保育士 1 は、一人の園児との対話が終わるごとに、普段の様子との違いを尋ねた 3 つの設問 (以下) に 7

件法で答えた。ここで、*は反転項目とした。

1. 園児の発話量は普段と比べてどうでしたか？（少ない～多い）
2. *園児は普段と比べて緊張していましたか？（していなかった～していた）
3. 園児が機械との対話を楽しむ様子は見られましたか？（見られなかった～見られた）

また、1日の収録が終わったときに、保育士1,2は園児に対する気付きを尋ねる以下の設問に自由記述で回答した。

1. 機械を演じるにあたって、意識したことがあれば教えてください。
2. 機械が園児とやりとりするにあたって、園児がより対話を楽しむには、どのような内容が望ましいと思いますか？振り返って思いつくことがあれば、教えてください。また、対話に限るとそれが難しいと思われる場合には、他にどのような手段があれば良いと思いますか？
3. 対話調査全体を通じて、何か他に気づいたことがあれば教えてください。

収録には、金沢工業大学に併設された扇が丘保育園に通う園児7名（女児4名、男児3名、平均月齢30.71ヶ月、 $SD = 5.34$ ）が参加した。このうち、2名は1回のみ、5名は2回の収録で、計12回の収録が実施された。園児にとっての2回目の収録は、おおよそ1～2週間の間が開けられた。保育士は計6名の協力を得た。実験は金沢工業大学倫理審査委員会の承認を受け実施された。

3. 実験結果

7件法によるアンケートへの回答結果と対話時間を表1に示す。表1では、数字が表す意味合いを揃えるため、反転項目は数値を逆転して記載した（緊張度について、1が「していた」、7が「していなかった」とした）。これにより、回答結果は数字が小さいほど普段に比べて発話量が少なく緊張し、楽しむ様子も見られなかったことを表す。また、対話時間はロボット役の保育士が冒頭文を読み上げ始めたとき、または「こんにちは」といった挨拶を行なったときを開始時点とし、「バイバイ」などといった別れの挨拶を行ったときを終了時点とした。

発話量、緊張度、および楽しんでいる度合いについての保育士の回答を見ると、平均的には中立の4より低い評価となった。これは保育士の印象として、園児の様子が普段と比べて全体的には発話量が少なく緊張しており、楽しんでいる様子が見られなかったこと

を示している。ただし、園児によっては、発話量や緊張する様子は見られながらも、1回目から楽しんで見せる様子も見せることもあったようである。また、2回目においてその評価が大きく変化する園児もいたようであり、これは回数を重ねることで異なる反応を見せるようになる可能性があることを示す結果だと考えられる。

対話時間については、園児A、Bの1回目の収録では、当初の想定として10分間（600秒）を設定したが、対話を行う保育士側に手詰まり感が見られ、7分程度で対話を終えることになった。この結果を踏まえ、その他の収録では、保育士に依頼する対話時間を5分間（300秒）に設定した。また、対話の内容を確認すると、ほぼ全ての時間にわたって保育士が一方的に園児に語りかける状況になっており、園児からの応答や発話は極めて少ない状態だった。

4. 議論

実験結果より、2回程度のロボットとのインタラクションでは、園児たちは緊張し、対話を行うことが困難であることが確認できた。このような結果になった要因について、園児側と保育士側に分けて考察する。園児側における要因は主に2つあると考えられる。1つは、ロボットへの慣れが不十分であったことである。ここで、園児Fの発話量や緊張度に注目すると、1回目と2回目とで数値が大きく変化していることが確認できる。このことから、インタラクションの回数を重ねることでロボットに慣れ、対話のみで気を引き続けることが可能になることも十分に考えられる。

もう1つは、実験環境によるものである。本研究で行った収録では、参加する園児を保育士が集団から1人だけ連れ出し、収録を行う部屋に移動させた。これが園児の不安を高め、消極的な姿勢を生んだのではないかと考えられる。これに対しては、園児が複数人いる集団の中にロボットを配置して対話を行うなど、対話の状況を変化させることが考えられる。

保育士側における要因についても主に2つが考えられる。1つは、園児がほとんど応答を返さないことで、対話が一方的になったことである。園児が緊張した様子であることに加え、表情やジェスチャーなどを使えない対話のみでのやりとりでは、保育士であっても園児からの反応がほとんど得られず、一方的な対話になってしまい、保育士が手詰まり感を感じている様子が見られた。もう1つは、収録におけるシステムの不備によるものである。隣接する部屋で遠隔対話を行ったことで、実音声と隣の部屋からの音声がズレを持

表1 アンケートの回答(7件法)と対話時間(秒). ハイフンは不参加を表す.
緊張度の数値は逆転して表記した.

	回数	A	B	C	D	E	F	G	平均	SD
月齢		28	38	38	24	33	25	29	30.71	5.34
発話量	1回目	2	1	2	1	1	2	1	1.43	0.49
	2回目	2	-	1	-	1	5	1	2.00	1.55
緊張度*	1回目	1	2	1	1	1	2	1	1.29	0.45
	2回目	5	-	2	-	1	6	1	3.00	2.10
楽しみ度	1回目	5	1	2	1	6	6	1	3.14	2.13
	2回目	4	-	4	-	1	7	1	3.40	2.24
対話時間(s)	1回目	430	429	343	290	183	295	164	304.86	98.22
	2回目	287	-	317	-	272	282	294	290.40	15.11

ちながら同時に聞こえる収録状況だった。隣の部屋が目視できるようにすることを狙った配置であったが、保育士にとっては対話しづらい状態になってしまい、思ったように対話が行えなかったのではないかと考えられる。

園児とロボットによる対話収録では、このような問題点がありながらも、我々が収録したデータは十分に有効な側面があるといえる。なぜなら、収録データには緊張した園児に対する保育士のノウハウが表れているためである。保育士の工夫として、自由記述式のアンケートの回答から例を挙げると、「園児が好きなものについて話すように心がけた」、「身近なもの(日常的なもの)を話題として話した」、「身振り手振りで答えられる問いかけをした」などという意見があった。このような内容が含まれた対話データは、幼児の気を引き続けることを目指す対話型AIが学習・再現すべきものだといえるのではないだろうか。したがって、保育士の対話内容や方法がLLMに再現できるのかを検証し、そのLLMが行う対話内容の有効性について検討していくことを今後の課題としたい。

5. 結論

本研究では、AI(LLM)による幼児が飽きない対話方法・内容を検討するために、ロボットをアバターにした保育士と園児の対話収録実験を行った。実験の結果、2回程度のロボットとの対話では、園児の緊張を緩和することができず、園児の発話を促すことが困難であることが確認できた。その一方で、園児に対する保育士の対話内容や方法は、園児の発話を促すノウハウが詰まっており、AIが学習すべき対話データを十分に確保することができた。今後の課題としては、LLMによる保育士の対話内容の再現性を検証していくことが考えられる。

謝辞

本研究は、JSPS 科研費基盤研究(C)「ことばへの意味づけ過程に見られる解釈的循環メカニズムの脳科学・計算論的解明」/課題番号22K12756の助成を受けました。国立情報学研究所 音声資源コンソーシアムから提供を受けた「NTT 乳幼児音声データベース(INFANT)」を利用しました。また、実験では金沢工業大学扇が丘保育園の保育士の方々および園児たちの多大な協力を得ました。ここに記し謝意を表します。

参考文献

- [1] 高木健志. (2006). うつ病患者に対するソーシャルワーク介入の可能性. 川崎医療福祉学会誌, 16(1), 129-132.
- [2] 阿部香澄, 塩見昌裕, 粕谷美里, & 長井隆行. (2021). 未就学児育児におけるロボット・AIの社会受容性と期待. 日本ロボット学会誌, 39(4), 371-374.
- [3] Hill, D., Ameenuddin, N., Reid Chassiakos, Y. L., Cross, C., Hutchinson, J., Levine, A., ... & Swanson, W. S. (2016). Media and young minds. *Pediatrics*, 138(5).
- [4] 阿部香澄. (2022). ポストコロナ時代に活躍する遠隔保育技術. 計測と制御, 61(3), 227-230.
- [5] 音声会話型おしゃべり AI アプリ Cotomo, <https://cotomo.ai/>, アクセス 2024.4.21.
- [6] Benesse and Softbank Robotics, こどもちゃれんじ 35周年特別企画 AI「しまじろう」, <https://shimajiro.benesse.ne.jp/cp/35th/>, アクセス 2024.4.21.
- [7] 白井宏美. (2022). 談話研究の知見に基づく対話システムの現状と課題. 日本音響学会誌, 78(5), 277-282.
- [8] 天野成昭, 近藤公久, 加藤和美 (2008): NTT 乳幼児音声データベース (INFANT). 国立情報学研究所 音声資源コンソーシアム. (データセット). <https://doi.org/10.32130/src.INFANT>