

保育者音声対話システムの構築に向けた 遠隔による子ども – 保育者対話のインタラクション分析

山下紗苗¹ 望月翔太¹ 隈有子² 堺礼³ 佐々木綾香⁴ 東中竜一郎¹

¹ 名古屋大学大学院情報学研究科 ² 株式会社マモル

³ 株式会社 Logical Fabrics ⁴ 株式会社七色

yamashita.sanae.w7@es.mail.nagoya-u.ac.jp, mochizuki.shota@nagoya-u.jp

yuko@mamor.jp, ray.sakai@logical-fabrics.com, ayaka-s@nanairo7.jp

higashinaka@i.nagoya-u.ac.jp

概要

我々は、遠隔で子どもを見守りつつ保育できることを目指し、保育者音声対話システムの構築に取り組んでいる。本研究では、その実装方針を得ることを目的とし、見守り対話における子どもと保育者のインタラクションを分析する。我々は、日本語の子ども – 保育者音声対話コーパスを収集し、音声のおよび言語的特徴の観点から既存コーパスと比較した。その結果、保育者音声対話システムでは、ターン制が適していること、パートごとにバックチャンネルを使い分ける必要があること、パートごとに対話の進行を変える制御性が必要であることが示された。

1 はじめに

音声対話システムはスマートスピーカやスマートフォンを通じて家庭に普及し、人と機械との自然な音声インタラクションを可能にしている [1, 2, 3]。音声は子どもにとって自然な入出力モダリティであり、近年では、子どもを対象とした音声対話研究の潮流が形成されつつある [4, 5, 6, 7, 8]。

株式会社七色は、遠隔にいる保育者が音声を通して子どもの様子を見守る保育サービスである NannyME¹⁾を提供している。現在は人間の保育者が子どもの応対をしているが、コスト削減の必要性などから、一部を自動化することが望ましいケースも多い。そのような背景から、我々は、本サービス上で、遠隔で子どもを見守りつつ対話できる保育者音声対話システムの実現を目指している。

本研究では、保育者音声対話システムの実装方針

を得ることを目的とし、子どもと保育者の間のインタラクションを分析する。具体的には、NannyME 上で日本語の子ども – 保育者対話コーパスを収集し、ターンテイキング、バックチャンネル、発話意図の観点から既存コーパスと比較する。

2 関連研究

子どもと大人の対話については、主に対面の親子のやり取りを記録したコーパスが整備されてきた。CHILDES [9] や CHICA [10] などが代表的だが、日本語に特化したものは CEJC-Child [11] や R-JMICC [12] などに限られている。本研究では、対面による親子の音声対話ではなく、遠隔保育サービス上の子どもと保育者の音声対話に着目する。

子どもと大人の音声対話は、ターンテイキング、バックチャンネル、発話意図といった観点での研究が比較的多く、本研究でもこれらの3つに着目する。ターンテイキングでは、子どもは大人よりオーバーラップが少ないこと [13, 14] や応答が遅いこと [15] が報告されている。バックチャンネルでは、子どもは親より内容応答的なものを多く使うことが示されている [16]。発話意図に関しては、発話を意図に基づいて分類したラベルセットである INCA-A が提案され [17]、子どもと大人の対話の発話意図を表現する標準的なものとなっている。

3 対話コーパスの収集

遠隔保育サービスの NannyME 上で、子どもと保育者による対話を収集した。なお、このコーパス収集は所属機関において倫理審査を経ており、対話の参加者および保護者はコーパスの収集、分析、および分析結果の公開に同意している。

1) <https://nannyme.love/>

表1 我々のコーパス (RCDD) と、既存の日本語親子対話コーパスの統計量。括弧内は1対話あたりの値を示す。発話率は実発話時間を対話時間で割った値を示す。

	RCDD (本研究)		CEJC-child [11]		CHILDES [9, 18]	
	子ども	大人 (保育者)	子ども	大人 (親)	子ども	大人 (親)
子どもの年齢	4~6 歳		0~8 歳		1~2 歳	
対話数	50 対話		53 対話		70 対話	
対話時間	23 時間 (28 分)		13 時間 (15 分)		76 時間 (65 分)	
実発話時間	4.5 時間 (5 分)	13.2 時間 (16 分)	3.0 時間 (3 分)	5.1 時間 (6 分)	16.9 時間 (14 分)	32.6 時間 (28 分)
発話率	19.4%	57.0%	22.5%	37.7%	22.1%	42.8%
異なり話者数	27 名	10 名	7 名	11 名	3 名	15 名

3.1 話者の募集

本サービスの利用者から、日本語を母国語とする子どもと保育者を募集した。子どもは音声で保育者とやり取り可能な4~6歳に統制し、保育者は本サービスで保育を行う者とし、性別の偏りに配慮した。参加者は子ども27名(男児10名, 女児17名), 保育者10名(男性4名, 女性6名)で, 子どもの年齢は4歳8名, 5歳7名, 6歳12名であった。

3.2 対話の収集

NannyMEでは, 子どもと保育者がクイズや雑談を行うことが多い。本研究では, これらのインタラクションを収集するために, 対話を自己紹介, クイズ, 雑談, クロージングの4パート構成とした。クイズと雑談の組は, 1対話中に連続して3回程度行うものとした。

自己紹介パートでは, 保育者が子どもと保護者に挨拶し, 互いの名前や好きな食べ物などの属性を自由に共有した。クイズパートでは, 保育者が3つのヒントを順に提示するクイズを行い, ヒントを提示するたびに子どもに正解を予想させた。保育者は, 子どもが正解するか, 難易度が高いと判断した時点で正解を開示した。雑談パートでは, クイズの正解に関する話題から自由に対話を展開した。クロージングパートでは, 対話の内容を振り返った。

保育者には, 対話が上記の4パートから構成されるように進行するよう指示した。また, 音声や言語によるやり取りを中心に収集するため, 道具を用いた遊びを控えるように指示した。

3.3 収集の実施

子どもと保育者の1対1の対話を収めた約30分間の映像を50件収集した。対話はTwillio video²⁾ま

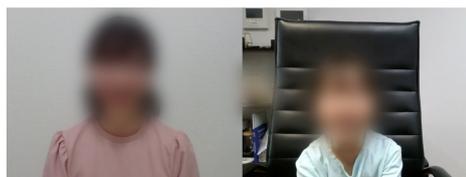


図1 RCDDの対話中の様子。

たはAgora³⁾を用いたビデオ通話で行い, 話者はPC, タブレット, スマートフォンのいずれかから参加した。同じ子どもと保育者の組み合わせによる対話は1回のみとした。子どもの参加回数は1人あたり最大2回, 保育者は1~18回と幅があった。

得られた対話音声は全て人手で書き起こした。図1に対話中の様子を, 付録の図4に対話例を示す。

3.4 対話の統計量

表1に, 我々のコーパス (Remote Childcare Dialogue Data; RCDD) と既存コーパスの統計量を示す。CEJC-childはモニター版のうち2人対話のみを使用し, CHILDESはタイムスタンプ付きの書き起こしがあるMiiProサブセット[18]のうち2人対話のみを使用した。いずれも子どもと大人による2人対話だが, RCDDは子どもと保育者, CEJC-childとCHILDESは親である点が異なる。子どもの年齢も, RCDDは4~6歳, CEJC-childは0~8歳, CHILDESは1~2歳と, それぞれカバー範囲が異なる。子どもの話者数はRCDDが27名と, 既存コーパスより多い。また, CEJC-childとCHILDESは対面による対話であるが, RCDDは遠隔による対話である点が大きく異なる。

4 比較手法

子どもと保育者の間でどのようなやり取りが行われているのかを明らかにするため, ターンテイキング, バックチャンネル, 発話意図の観点から, RCDDを既存の親子対話コーパスと比較する。

2) <https://twillio.com/docs/video/>

3) <https://www.agora.io>

表2 1分間に発生した IPU, Pause, Gap, Overlap の累積時間 (秒) の比較. 括弧内は 1 個あたりの長さの平均を示す. 太字は各行の最大値を示す. 大人対話の IPU と Pause の値は, 大人 2 人分の合計である.

		RCDD				CEJC-child	CHILDES	大人対話 [19]
		自己紹介	クイズ	雑談	クロージング			
IPU	子ども	11.0 (1.5)	3.7 (1.3)	8.3 (2.0)	0.7 (1.5)	12.6 (1.2)	13.3 (2.1)	–
	大人	37.2 (3.4)	15.2 (3.6)	17.3 (2.7)	2.5 (4.4)	21.7 (1.5)	25.5 (2.9)	59.7
Pause	子ども	3.1 (2.4)	1.3 (3.2)	3.5 (5.9)	0.2 (1.3)	3.8 (1.8)	3.3 (3.0)	–
	大人	7.7 (1.4)	22.0 (11.9)	9.9 (2.9)	0.6 (1.3)	11.9 (1.8)	9.3 (2.4)	3.5
Gap	–	8.6 (1.1)	7.7 (3.1)	10.0 (3.6)	0.5 (1.4)	13.0 (1.3)	8.1 (2.0)	4.0
Overlap	–	3.1 (0.6)	1.2 (0.6)	2.0 (0.6)	0.3 (0.7)	1.3 (0.4)	3.0 (0.5)	8.1

ターンテイキングについては, 先行研究 [20] に基づき, IPU (0.2 秒以上の無音で区切られた発話区間の長さ), Pause (IPU 間の長さ), Gap (異なる話者の発話間の無音区間の長さ), Overlap (異なる話者の発話が重複する区間の長さ) を算出する.

バックチャンネルについては, 書き起こし内のバックチャンネルをカテゴリ別に分類し, 実発話時間あたりの出現頻度を算出する. カテゴリは先行研究 [21] に基づき, 応答系感動詞 R (はい, うん, ああ, ええ, ふん), 感情表出系感動詞 E (お, へえ, ふうん), 語彙的応答 L (そう (ですね), なるほど, たしかに, ね), 評価応答 A (すごい, おもしろい, こわい) とする. カテゴリの分類には, 正規表現を用いて, これらの語と完全一致する場合にそのカテゴリにカウントする.

発話意図については, 各発話の意図をカテゴリに分類し, カテゴリ間の遷移確率をグラフ化することで, やり取りの流れを定性的に検討する. 分類には INCA-A を基にした 19 カテゴリを用いる. 分類は LLM を用いて自動化し, OpenAI の GPT-5 nano⁴⁾ に対してカテゴリ定義と例を付録の図 5 のプロンプトで与え, 最大 50 発話の履歴から各発話に最適なカテゴリを出力させる. なお, LLM の利用規約の観点から, 比較対象は CHILDES に限定する.

5 比較結果

5.1 ターンテイキングの比較結果

表 2 に, ターンテイキングの各指標を示す. 大人対話 (成人同士の対話) の参考値は文献 [19] による.

IPUs については, 子ども - 大人対話は大人対話より累積時間が全体的に少なく, 無音区間が多いことが示された. IPU 長はどのコーパスでも大人のほうが子どもより長い, その差はコーパス間で異なっていた. RCDD では自己紹介, クイズ, クロージ

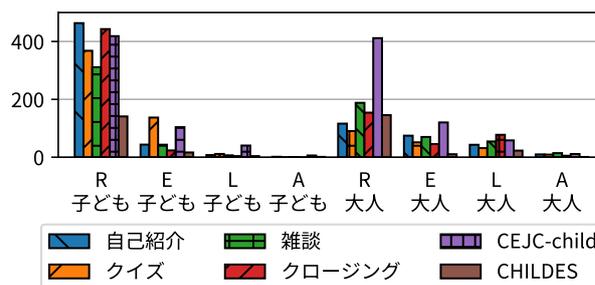


図 2 実発話時間あたりのバックチャンネルの頻度.

グで差が大きく, 雑談で小さかった. CEJC-child と CHILDES でも差は小さかった.

Pause と Gap については, 大人は子どもより累積時間が多かった. ただし, Pause 長には差があり, 通常 1~2 秒程度のところ, RCDD のクイズでは 10 秒を超えることがあった. クイズでは Gap 長が 3.1 秒と比較的長いことから, このパートでは沈黙が長くなりやすいことが分かった.

Overlap については, 子ども - 大人対話は大人対話より累積時間が少なく, RCDD のクイズ, 雑談, クロージングでは 1 分あたり 1 秒や 2 秒程度であった. 最も多い RCDD の自己紹介や CHILDES の場合でも 3 秒程度であった. つまり, 子どもと大人の発話が重なることは少ないことが分かった.

5.2 バックチャンネルの比較結果

図 2 に, 実発話時間 1 時間あたりのバックチャンネルの出現頻度を示す.

子どもについては, R が最頻出で, 特に RCDD の自己紹介とクロージング, CEJC-child で多かった. E は RCDD のクイズと CEJC-child に多く, その中でも「え」「あ」が頻出した. これは, 子どもがバックチャンネルをフィードバックだけでなく, 思考状態の表出としても用いているためだと考えられる.

大人については, R が最頻出で, 特に RCDD の雑談, CEJC-child で多かった. これは, 子どもの発話

4) <https://platform.openai.com/docs/models/gpt-5-nano>

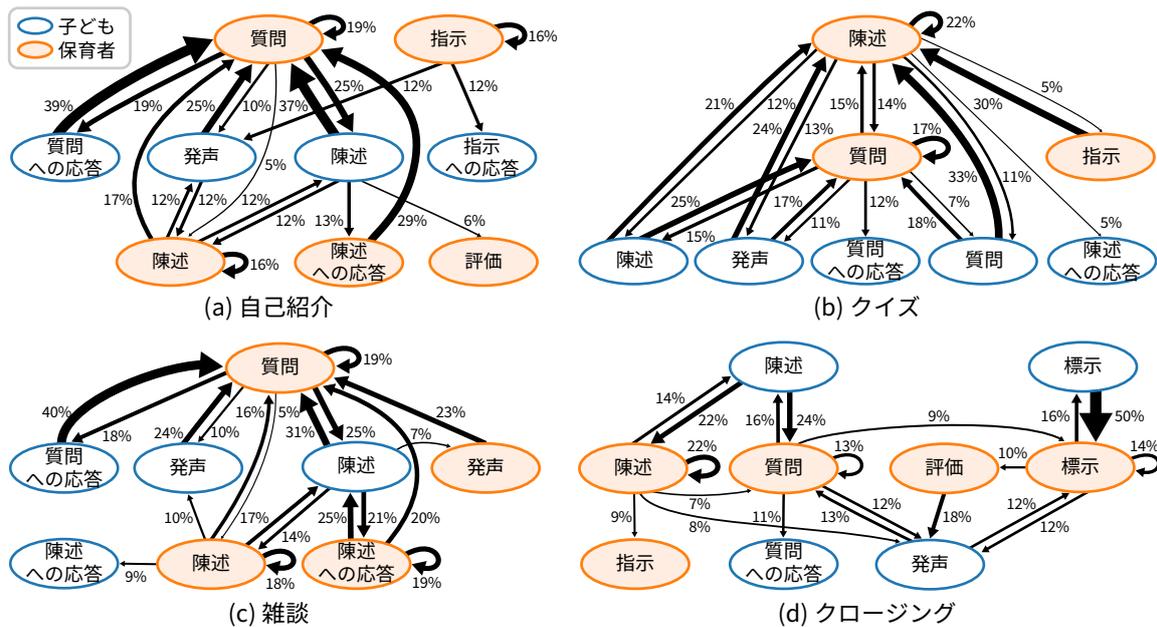


図3 RCDDの発話意図の遷移。エッジ上の数値は遷移確率を示す。

を積極的に受け止めていることを示す。EとLはRCDDのクイズやCHILDESで少なかった。これは、大人が対話を主導しているためだと考えられる。これらのことから、大人はパートに応じてバックチャンネルを使い分けしているといえる。

5.3 発話意図の比較結果

図3に、RCDDのパートごとの発話意図の遷移を示す。グラフの概形から、自己紹介と雑談は比較的類似していた。これらのパートでは、2ターンからなる隣接ペア [22, 23] だけでなく、まず保育者の質問、次に子どもの陳述、最後に評価やコメントがあるという、3ターンの Question-Response-Follow-up 構造 [24] のやり取りが多かった。一方で、クイズは異なる概形であった。このパートでは、まず保育者の陳述や質問、次に子どもの質問や質問への応答、最後に再び保育者の質問があるという、ヒントの開示、回答、次のヒントの開示というサイクルが形成されていた。クロージングは他のパートとは異なり、子どもの標示（感謝や挨拶など）から保育者の標示が最頻出であった。図は省略するが、CHILDESは、保育者の陳述や質問、子どもの陳述が中心で、そこから様々な発話意図と双方向に接続しており、比較的クイズに近い遷移であった。

6 おわりに

本研究では、日本語の子ども - 保育者音声対話コーパスを収集し、既存コーパスと比較した。その

結果、保育者音声対話システムを構築するにあたって次の知見が得られた。

- ターンテイキングの分析から、発話のオーバーラップが少ないことが分かった。したがって、近年研究が盛んである Full-duplex [19] のような双方向性は必要なく、堅実に対話を進められるターン制でも十分な可能性がある。
- バックチャンネルの分析から、パートごとの使い分けが必要であることが分かった。特に、クイズにおいては、感情表出系感動詞 E を多用する必要がある。
- 発話意図の分析から、パートごとに発話意図の遷移が異なることが分かった。このことは、対話の進行を意図的に制御する仕組みが不可欠であることを示している。

今後の課題として、ピッチなどの韻律的観点から RCDD の特徴を明らかにすることが挙げられる。さらに、本サービス向けの保育者音声言語モデルの構築を目指す上では、制御性の確保と十分なデータ収集が不可欠である。特にデータ収集においては、より広い年齢層や性格特性を網羅する必要がある。今後は、これらの対応を目指したい。

なお、子ども向け音声対話システムは多くのリスクを伴うため倫理的問題が生じやすい。実装にあたっては、子どもに対する安全性、プライバシー保護、および、問題発生時の責任について検討を重ねつつ開発を進めていく必要があると考えている。

謝辞

本研究は、JST ムーンショット型研究開発事業、JPMJMS2011 の支援を受けたものである。また、対話の収集において多大なご協力を賜った NannyME の利用者の皆様に感謝する。

参考文献

- [1] Ashwin Ram, Rohit Prasad, Chandra Khatri, Anu Venkatesh, Raefer Gabriel, Qing Liu, Jeff Nunn, Behnam Hedayatnia, Ming Cheng, Ashish Nagar, Eric King, Kate Bland, Amanda Wartick, Yi Pan, Han Song, Sk Jayadevan, Gene Hwang, and Art Pettigrue. Conversational AI: The science behind the Alexa Prize. **arXiv preprint arXiv:1801.03604**, 2018.
- [2] Li Zhou, Jianfeng Gao, Di Li, and Heung-Yeung Shum. The design and implementation of XiaoIce, an empathetic social chatbot. **arXiv preprint arXiv:1812.08989**, 2019.
- [3] OpenAI, et al. Gpt-4 technical report. **arXiv preprint arXiv:2506.02979**, 2024.
- [4] Jack Mostow, Greg Aist, Paul Burkhead, Albert Corbett, Andrew Cuneo, Susan Eitelman, Cathy Huang, Brian Junker, Mary Beth Sklar, and Brian Tobin. Evaluation of an automated reading tutor that listens: Comparison to human tutoring and classroom instruction. **Journal of Educational Computing Research**, Vol. 29, No. 1, pp. 61–117, 2003.
- [5] Jacqueline M Kory Westlund, Sooyeon Jeong, Hae W Park, Samuel Ronfard, Aradhana Adhikari, Paul L Harris, David DeSteno, and Cynthia L Breazeal. Flat vs. expressive storytelling: Young children’s learning and retention of a social robot’s narrative. **Frontiers in human neuroscience**, Vol. 11, p. 295, 2017.
- [6] Jacqueline M Kory-Westlund and Cynthia Breazeal. A long-term study of young children’s rapport, social emulation, and language learning with a peer-like robot playmate in preschool. **Frontiers in Robotics and AI**, Vol. 6, p. 81, 2019.
- [7] Alireza Esfandbod, Zeynab Rokhi, Ali F Meghdari, Alireza Taheri, Minoo Alemi, and Mahdiah Karimi. Utilizing an emotional robot capable of lip-syncing in robot-assisted speech therapy sessions for children with language disorders. **International journal of social robotics**, Vol. 15, No. 2, pp. 165–183, 2023.
- [8] Nazerke Rakhymbayeva, Aida Amirova, and Anara Sandygulova. A long-term engagement with a social robot for autism therapy. **Frontiers in robotics and AI**, Vol. 8, p. 669972, 2021.
- [9] Brian Macwhinney. **The CHILDES Project: Tools for Analyzing Talk**. Lawrence Erlbaum Associates, 2000.
- [10] Dhia Elhak Goumri, Abhishek Agrawal, Mitja Nikolaus, Hong Duc Thang Vu, Kübra Bodur, Elias Emmar, Cassandre Armand, Chiara Mazzocconi, Shreejata Gupta, Laurent Prévot, Benoit Favre, Leonor Becerra-Bonache, and Abdellah Fourtassi. CHICA: A developmental corpus of child-caregiver’s face-to-face vs. video call conversations in middle childhood. In **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation**, pp. 3153–3164, 2024.
- [11] 小磯花絵, 石本祐一, 居關友里子, 江口典子, 柏野和佳子, 川端良子, 田中真理子, 田中弥生, 西川賢哉. 『子ども版日本語日常会話コーパス』モニター版の構築. 言語処理学会第 31 回年次大会発表論文集, pp. 3525–3528, 2025.
- [12] 西海枝洋子, 渡辺和希, 小西隆之, 伊藤直子, 金礪愛, 五十嵐陽介, 宮澤幸希, 西川賢哉, 馬塚れい子. 『理研母子会話コーパス (R-JMICC)』構築の試みと研究成果-対乳児自発音声における日本語特有の韻律的・分節的特徴の解明を目指して-. 第 3 回コーパス日本語学ワークショップ予稿集, pp. 383–392, 2013.
- [13] Tanya M Gallagher and Holly K Craig. An investigation of overlap in children’s speech. **Journal of Psycholinguistic Research**, Vol. 11, No. 1, pp. 63–75, 1982.
- [14] Viktória Horváth and Valéria Krespsz. Temporal characteristics of child-adult conversations: Utterances and turn-taking. **Taikomoji kalbotyra**, No. 19, pp. 3–13, 2023.
- [15] Marisa Casillas, Susan C Bobb, and Eve V Clark. Turn-taking, timing, and planning in early language acquisition. **Journal of child language**, Vol. 43, No. 6, pp. 1310–1337, 2016.
- [16] Kübra Bodur, Mitja Nikolaus, Laurent Prévot, and Abdellah Fourtassi. Using video calls to study children’s conversational development: The case of backchannel signaling. **Frontiers in Computer Science**, Vol. 5, p. 1088752, 2023.
- [17] Anat Ninio, Catherine E. Snow, Barbara A. Pan, and Pamela R. Rollins. Classifying communicative acts in children’s interactions. **Journal of Communication Disorders**, Vol. 27, No. 2, pp. 157–187, 1994.
- [18] Sayo Miyata. **Japanese CHILDES: The 2012 CHILDES manual for Japanese**, 2012.
- [19] Atsumoto Ohashi, Shinya Iizuka, Jingjing Jiang, and Ryuichiro Higashinaka. Towards a Japanese full-duplex spoken dialogue system. In **Proceedings of the 26th Interspeech Conference**, pp. 1783–1787, 2025.
- [20] Tu Anh Nguyen, Eugene Kharitonov, Jade Copet, Yossi Adi, Wei-Ning Hsu, Ali Elkahky, Paden Tomasello, Robin Algayres, Benoit Sagot, Abdelrahman Mohamed, and Emmanuel Dupoux. Generative spoken dialogue language modeling. **Transactions of the Association for Computational Linguistics**, Vol. 11, pp. 250–266, 2023.
- [21] Yasuharu Den, Nao Yoshida, Katsuya Takanashi, and Hanae Koiso. Annotation of Japanese response tokens and preliminary analysis on their distribution in three-party conversations. In **Proceedings of the 2011 International Conference on Speech Database and Assessments**, pp. 168–173, 2011.
- [22] Emanuel A. Schegloff and Harvey Sacks. Opening up closings. **Semiotica**, Vol. 8, No. 4, pp. 289–327, 1973.
- [23] Emanuel A. Schegloff. **Sequence Organization in Interaction: A Primer in Conversation Analysis**. Cambridge University Press, 2007.
- [24] Malcolm Coulthard. **An introduction to discourse analysis**. Routledge, 2014.

A 付録

図 4 に、我々が収集した RCDD の対話例を示す。図 5 に、発話意図の推定に用いたプロンプトを示す。

<p>自己紹介</p> <p>保育者: こんにちは. 子ども: こんにちはー. 保育者: あ. ありがとう. 保育者: お名前教えてください. 子ども: ○○○○でーすー. 保育者: ○○○ちゃん. ○○○ちゃん何歳ですか? 子ども: 四歳. (中略) 保育者: ○○○ちゃん何が好き? 子ども: ミニー. 保育者: あミニーちゃん. 一緒だ. やったー. ミニーちゃん可愛いよね.</p>
<p>クイズ</p> <p>保育者: ジャヒント二問目聞いてみよう. 保育者: 髭があります. 保育者: あこれ○○ちゃんかな? 子ども: ライオン. (中略) 保育者: でも, お髭ある動物っていっぱいあるよね. ヤギさんとかも, お髭あるかな? (中略) 保育者: 私は, ニャーっと, 鳴きます. 子ども: ネコ? 保育者: ネーコ. 保育者: せいーかい.</p>
<p>雑談</p> <p>保育者: ネコさん触ったことある? 子ども: 見たことある. 保育者: あ見たことあるのー? どこで見たことある? (中略) 子ども: ちゃんはーミルクとかー, 魚とかが好き. 保育者: うん. 保育者: うん. 保育者: よく知ってるねえー. 子ども: だって本で見たことあるから. 保育者: 本とかで読んだの? 保育者: あっ, すっごい. 勉強家だねえー.</p>
<p>クロージング</p> <p>保育者: 眠くなっちゃうね. ほっぺも落ちちゃったし, お目目も閉じちゃうし. 保育者: あー○○ちゃんそろそろさバイバイの時間になっちゃったね. 保育者: 楽しかった? 保育者: え, ○○○先生も楽しかったー. (中略) 保育者: あっ, バイバイの時間だ. バイバイ. 子ども: バッバー. 保育者: ありがと.</p>

図 4 RCDD の対話例.

<p># 指示</p> <p>あなたのタスクは、与えられた対話の各発話にラベルを付与することです。以下の制約と出力形式に従ってください。</p> <p># 制約</p> <p>与えられた対話の各発話に対して、文脈を踏まえ、以下の19個の選択肢の中から最も適切なラベルを1つ付与すること。 (以下、発話意図の定義と例)</p> <ol style="list-style-type: none"> 1. 指示 (Directives): 聞き手に対して、ある行為を行うよう求める、提案する、命じる発話。 2. 指示に対する応答 (Responses_for_Directives) 3. 発話誘発 (Speech_Elicitations): 相手の発話を引き出すための行為、またはそれに応じる行為。単語や文の模倣、完成、音声模倣の要請などが含まれる。 4. 発話誘発に対する応答 (Responses_for_Speech_Elicitations) 5. 約束 (Commitments): 話し手自身の行為に関する意図、許可、禁止を表す発話。自分の行動を宣言する。 6. 約束に対する応答 (Responses_for_Commitments) 7. 宣言 (Declarations): 発話によって新しい社会的、事実的状态を成立させる行為。 8. 宣言に対する応答 (Responses_for_Declarations) 9. 標示 (Markings): 出来事の発生を社会的に示したり、感情的反応を表す行為。感謝、挨拶、謝罪、祝いなどを含む。 10. 標示に対する応答 (Responses_for_Markings) 11. 陳述 (Statements): 事実、意見、希望を述べる発話。 12. 陳述に対する応答 (Responses_for_Statements) 13. 質問 (Questions): 情報を求める発話。wh質問、Yes/No質問、選択質問、確認質問などが含まれる。 14. 質問に対する応答 (Responses_for_Questions) 15. 行為遂行 (Performances): ゲームや活動の中でルールに従って行う発話。 16. 評価 (Evaluations): 相手または自分の行動に対する肯定的、否定的な評価を表す発話。 17. 確認要求 (Demands_for_Clarification): 相手の発話を聞き返したり、繰り返しを求めたりする行為。 18. 言語形式の修正 (Text_Editing): 他者の誤った発話を訂正する行為。正しい言語形式を提示する。 19. 発声 (Vocalizations): 明確な機能をもたない語様発声や、理解不能な音声を含む。 <p># 出力形式</p> <p>以下の形式で出力すること。</p> <pre><Format> {"utterance_id":0,"label":"Markings"} {"utterance_id":1,"label":"Responses_for_Markings"} {"utterance_id":2,"label":"Questions"} ... </Format></pre> <p># タスク</p> <p>## 与えられた対話</p> <p>\$dialogue</p>
--

図 5 発話意図の推定に用いたプロンプト.