

話者スタイル抽出と対話フロー生成に基づく 対話データ拡張手法

齊志揚 稲葉 通将

電気通信大学

{qizhiyang,m-inaba}@uec.ac.jp

概要

自然言語処理技術の進展に伴い、対話システムが多様なユーザと対話を行う場面が増加している。しかし、特殊な話者スタイルを示す小規模のユーザグループに対しては、データ不足がシステム性能向上を妨げる要因となっている。本研究では、話者スタイルと対話行為の流れを組み合わせたデータ拡張手法を提案する。大規模言語モデル (Large Language Model; LLM) を用いて話者スタイルを抽出し、事前学習済み言語モデル (Pre-trained Language Model; PLM) を活用して対話行為系列を生成することで、その話者に特化した豊かな対話データを生成する。実験により、提案手法が低リソースユーザグループに適応可能な対話システムの開発に寄与することを示した。

1 はじめに

近年の自然言語処理技術の進歩により、対話システムは大きな関心を集めている [1, 2, 3]。LLM が強力な生成能力を示している一方で、人間同士の対話データを用いてファインチューニングされた対話システムに関する研究は、その優れた制御性と解釈可能性により依然として重要である [4]。しかし、ユーザごとのスタイルの違いは、システムの設計に課題をもたらす。図 1 に示すように、成人 (ユーザ A) と比べて未成年 (ユーザ B) は、意図が明確でない場合や曖昧な応答が多く、円滑な対話の実現にはオペレータによる確認や追加の質問がより多く求められる。このスタイルの違いにより、オペレータは異なるユーザと対話する際、状況に応じて戦略を調整する必要がある。こうした適応的アプローチは、現実世界における対話システムの有効性や使いやすさを高めるうえで不可欠である。

このような特殊な話者スタイルにシステムを適応

ユーザAのスタイル:

- ・目的地やアクティビティに関して明確な意図を持っている。
- ・より詳しい情報や具体的な提案を求められる。(…)

ユーザAの対話:

オペレータ: [方面・県・サブエリアを聞く質問] えっと、ご旅行の予定とゆーか、あの、どちらの方面に行きたいとか、あの、ございますか？

ユーザA: はい、えーと、北海道に行きたいんですけど。

オペレータ: [希望する季節を聞く質問] はい、えーと、季節とか、ご希望ございますか？

ユーザA: えーと、秋頃ですかね。

オペレータ: [旅行人数・人間関係を聞く質問] えーと、お客様は何名様で、ご予約されていらっしゃいますか？

ユーザA: あ、私、私一人だけなんですけれども。

オペレータ: [希望を聞く質問] えーと、特にあの一、何かやりたいこと、とゆーか、あの、ご希望ございますか？

ユーザA: そうですね。えーと、ちょっと紅葉が綺麗な所に行きたいんですけど。

オペレータ: [観光地に対する要望・条件の確認、探す意思の表明] はい、あ、紅葉の綺麗なところ、で、まずじゃあ、調べますね。

ユーザBのスタイル:

- ・目的地やアクティビティに関して明確な意図がない。
- ・質問や提案に対して簡潔に答える傾向がある。(…)

ユーザBの対話:

オペレータ: [方面・県・サブエリアを聞く質問] では早速なんですけど、えー、どちらが行きたい所、えーござい、もうお決まりですか？

ユーザB: はい。

オペレータ: [方面・県・サブエリアを聞く質問] はい、どちらでしょうか？

ユーザB: 広島。

オペレータ: [観光地に対する要望・条件の確認、方面・県・サブエリアを聞く質問] はい、広島ですね。広島県のどこまでとか、って細かい場所まで決まっていますか？

ユーザB: いや、決まってないです。

オペレータ: [希望を聞く質問] うん。例えば何かしたい、とか、遊びたいとか、遊びたい、食べたいとか、何かあったりしますか？

ユーザB: おじいちゃんとおばあちゃんに会いたいです。

オペレータ: [観光地に対する要望・条件の確認] あー、おじいちゃんとおばあちゃん、あー、そら辺に住んでいるんですね。

図 1 実際の人間同士の観光地推薦対話 [5, 6]

させるには、膨大な注釈付きデータが必要となる。一方で、特に少数の特殊な話者スタイルを示すユーザに関するデータは収集が難しい場合も多い。その結果、こうしたユーザグループに対応するシステムの性能向上が妨げられている。したがって、リソースが限られたユーザグループに特化した学習データを効率的に生成するデータ拡張手法が求められる。

本研究では、特殊な話者スタイルを持つ低リソースユーザグループに対応するためのデータ拡張フレームワークを提案する。未成年の特殊な話者スタイルやデータ収集の難しさ [7] を考慮し、本研究では未成年向けのデータ拡張に着目した。図 1 に示すように、ユーザの話者スタイルは、相手の対話行為

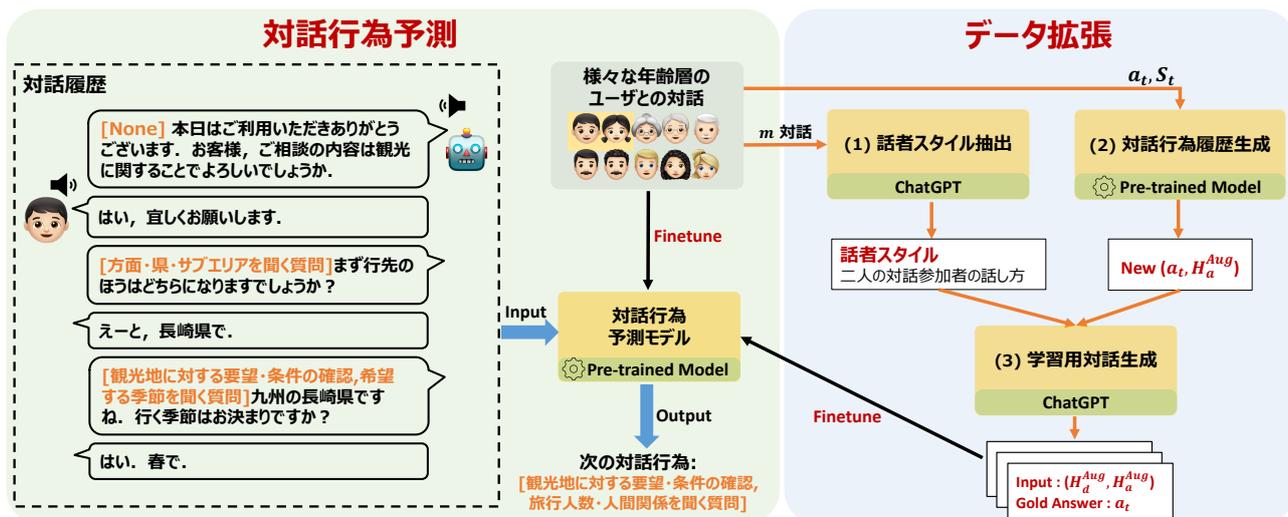


図2 提案したデータ拡張フレームワーク

(Dialogue Act; DA) や応答内容に影響する。そのため、本研究では、話者スタイルと対話フローの両方に注目したデータ拡張フレームワークを構築した。LLM で対象ユーザと相手の話者スタイルを抽出し、PLM を低リソース環境でファインチューニングして多様な DA 履歴を生成する。得られた話者スタイルと DA 履歴を LLM に入力することで、対象ユーザに特化した学習データを生成する。このデータを限られた学習データに追加することで、発話生成の前段階における重要なタスクである DA 予測の性能を向上させることが可能となる。

2 関連研究

注釈付きデータの不足とデータの不均衡は、人工知能分野で依然として問題である [8, 9, 10, 11]。これらの問題に対処するために、データ拡張が使用されており、異なるタスクにおける先行研究でも示されている [12, 13]。例えば、Schick らは、PLM を用いてテキスト類似性データセットを生成した [14]。また、個々の発話操作によるデータ拡張で対話要約タスクの性能を向上させる研究 [15, 16] が提案されているが、いずれも個別の文生成に焦点を当てている。一方で、本研究では複数の発話文からなる一貫した対話の生成を目指している。

対話のデータ拡張に関する研究も盛んに行われている、Mohapatra らは、GPT-2 を用いてユーザとエージェントのボットを開発し、低リソースシナリオでの性能改善を示した [17]。最近では、LLM を活用したデータ拡張が注目され、Kim らは知識グラフを用いて社会的対話を生成し [11]、Pan らはドメイン外

の会話から抽出した対話パスを基にタスク指向の対話を生成した [18]。しかし、Pan らが既存データから対話パスを抽出するのに対し、本研究はターゲットユーザに特化した DA 履歴を新たに生成する。

3 提案フレームワーク

対話の進行中に直近の対話履歴を基に次のターンの DA を予測するタスクは、適切な応答生成を可能にし、対話の一貫性と精度を高める。また、これにより誤応答の発生が抑えられ、ユーザ体験およびシステムの信頼性向上に寄与する [19]。本研究では、特殊な話者スタイルを持つ低リソースユーザグループに対する DA 予測性能の向上のための新たなデータ拡張フレームワークを提案する。

図2の左側に示すように、DA 予測タスクは、次の対話ターン t に対し、直近 n ターンの対話履歴 $H_d = (S_{t-n}, U_{t-n}, \dots, S_{t-1}, U_{t-1})$ とシステムの DA 履歴 $H_a = (a_{t-n}, \dots, a_{t-1})$ を入力し、次のターンのシステムの DA a_t を出力するものとして定義される。ここで、 S はシステム、 U はユーザの発話を表す。

対話と DA 履歴に基づいて次のターンの DA を予測するためには、ターゲットユーザの話者スタイルや対話フローを正確に反映したデータが必要である。本研究では、対話参加者の話者スタイルを捉え、ターゲットユーザグループとの対話フローを模倣することで、多様な学習データを生成する。図2の右側に示すように、提案するデータ拡張フレームワークは以下の3つのコンポーネントで構成される：(1) ChatGPT¹⁾を用いて話者スタイルを抽

1) <https://openai.com/blog/chatgpt>

出する, (2) PLM をファインチューニングしてシステムの DA 履歴 $H_a^{Aug} = (a_{t-n}^{Aug}, \dots, a_{t-1}^{Aug})$ を生成する, (3) 抽出した話者スタイルと生成したシステムの DA 履歴 H_a^{Aug} を ChatGPT に入力し, 学習対話データ $H_d^{Aug} = (S_{t-n}^{Aug}, U_{t-n}^{Aug}, \dots, S_{t-1}^{Aug}, U_{t-1}^{Aug})$ を生成する.

3.1 話者スタイル抽出

対象ユーザグループ特有の話者スタイルは会話内容に大きな影響を与えるため, それを正確に捉えることが重要である. 本研究では, 対象ユーザグループと非対象ユーザグループをそれぞれ半数ずつ含む m 個の対話セットを比較し, 対象ユーザグループ特有の話者スタイルを抽出する手法を採用する. 具体的には, ChatGPT を用いて対話セットを入力し, 比較を通じて対象ユーザグループの顕著な話者スタイルを識別し, 対象ユーザグループおよびその対話相手の話者スタイルを抽出する.

特に, 対象ユーザが「目的地や活動に対して曖昧な意図を示す」などの抽象的なスタイルの抽出に注力する. このスタイルは会話の方向性に影響を与え, 生成される対話を実際の人間の会話に近づけるうえで重要である. プロンプトと抽出結果の例は付録 A に示す.

3.2 対話行為履歴生成

図 1 に示すように, 対象ユーザ特有の話者スタイルは, 対話フローに影響を与える. そこで本研究では, 対象ユーザグループを模した多様な DA 履歴 H_a^{Aug} を生成することを目的とし, 図 3 に示す PLM のファインチューニング手法を提案する.

まず, 次のターン t の DA a_t および発話 S_t を入力し, 直近 n ターンの DA 履歴 H_a を出力とする学習データを構築する. 学習フェーズでは, 全学習データで PLM をファインチューニングし, DA 履歴生成を学習させる. 次に, 対象ユーザグループのデータに限定した二次ファインチューニングを行い, モデルを対象ユーザグループ特有の話者スタイルに適応させる. この二段階のファインチューニングにより, モデルは人間らしい DA 履歴を生成しつつ, 対象ユーザグループの特徴に対応可能となる.

生成フェーズでは, 次のターン t の DA a_t および発話 S_t を入力し, 複数の DA 履歴 H_a^{Aug} を生成する. 多様性を確保するため, 元のデータセットで観測されていない (a_t, H_a^{Aug}) の組み合わせを選択する.

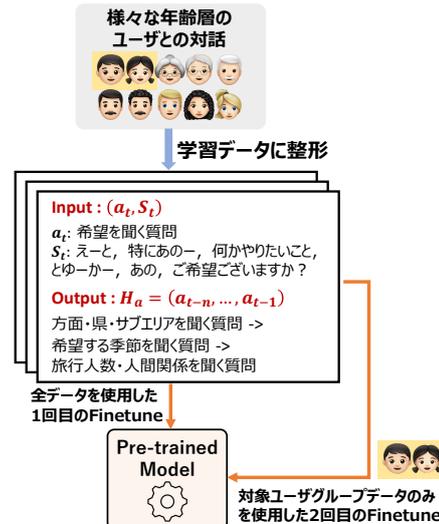


図 3 DA 履歴生成のための 2 段階ファインチューニング

3.3 学習用対話生成

対象ユーザグループに合わせた話者スタイルと DA 履歴を取得した最終目標は, これらを基に対話を生成し, DA 予測の学習データを充実させることである. 本研究では, ChatGPT の生成能力を活用し, Few-shot プロンプトに抽出した話者スタイルと DA 履歴 H_a^{Aug} を入力して, 対象ユーザグループの話者スタイルを反映した対話 H_d^{Aug} を生成する. 生成した H_d^{Aug} と H_a^{Aug} を用い, a_t を正解ラベルとして学習データを構築する. プロンプトの詳細は付録 B に記載している.

4 評価実験

提案したデータ拡張フレームワークの有効性を評価するため, 未成年者の会話スタイルを反映したデータを用い, 低リソース環境下で異なるデータ規模のモデル性能を比較した.

4.1 データセット

本研究では, 日本語データセット「旅行代理店タスク対話コーパス」を使用した [5, 6]. このデータセットには, 様々な年齢層のユーザとの 330 件の対話 (合計 115 時間, 平均 20 分) が含まれ, DA の注釈が施されている. 対話は Zoom を介して実施され, 6 人のオペレータと 55 人のユーザ (未成年 20 名, 成人 25 名, 高齢者 10 名) が参加した. 先行研究 [5] では, 未成年が特殊な話者スタイルを用い, 独立した意見を表明する頻度が少ないことが確認されている. DA の注釈は 29 の事前定義されたラベル

表 1 実験結果

モデル	日本語 GPT-NeoX		日本語 T5-base	
	完全一致率	部分一致率	完全一致率	部分一致率
Minors-Only	0.2352	0.3438	0.2477	0.3452
Zero-Shot	0.2950	0.4054	0.2984	0.4069
Low-Resource	0.3012	0.4199	0.3038	0.4207
Low-Resource + Aug (Ours)	<u>0.3085</u>	<u>0.4278</u>	<u>0.3071</u>	<u>0.4240</u>
- w/o DA History Generation	0.3052	0.4263	-	-
- w/o Second Finetune	0.3072	0.4269	-	-
- w/o Speaker Style	0.3027	0.4274	-	-
Full-Resource	0.3117	0.4431	0.3136	0.4378

で行われ、セグメント単位²⁾で注釈付けされている。DAの種類については付録Cに示す。

4.2 低リソース設定

低リソース環境をシミュレーションするため、未成年3人の18対話を学習、10人の60対話を評価に使用した。実験では、規模が異なるデータで5種類のDA予測モデルをファインチューニングし、各モデルの学習データの設定は以下である：

- **Minors-Only**：未成年3人からの18対話のみを使用した。
- **Zero-Shot**：成人および高齢者からの210対話を使用した。
- **Low-Resource**：Minors-Onlyの18対話に、成人および高齢者からの210対話を加えた計228対話を使用した。
- **Full-Resource**：未成年10人からの60対話に成人および高齢者の210対話を加えた計270対話を使用した。
- **Low-Resource+Aug (Ours)**：Low-Resourceの228対話に、提案手法で生成したデータを追加した。直接比較のため、データ規模をFull-Resourceと同等になるまで拡張した。

4.3 モデルと評価指標

DA履歴の生成には日本語T5-Large³⁾を使用した。また、DA予測にはT5-Base⁴⁾とGPT-NeoX⁵⁾を用い、それぞれ異なる未成年からの対話を学習・評価データとして用いて4回の実験を行い、完全一致率と部分一致率に基づく評価を行った。これらの指標を採用した理由は、一つの発話が複数セグメントに分かれ、複数タグが付く可能性があるためである。

2) コミュニケーション機能を持つ最小単位の対話行動
 3) <https://huggingface.co/retrieva-jp/t5-large-long>
 4) <https://huggingface.co/retrieva-jp/t5-base-long>
 5) <https://huggingface.co/stockmark/gpt-neox-japanese-1.4b>

4.4 実験結果とアブレーション

表1は、異なる未成年データを使用して4回実行し、シード1~5を用いた平均結果を示す。Minors-Onlyは未成年データのみを使用するため、学習データが限定的で、Zero-Shotモデルより性能が劣った。提案したLow-Resource + Aug (Ours)は、Low-Resourceを上回る性能を示し、未成年特有の話者スタイルを効果的に捉えたデータを生成できたことを示している。一方で、データ量をFull-Resourceと同等に拡張しても、Full-Resourceには一貫して劣る結果となった。生成データの質の管理不足や、実際の対話に含まれるフィラーなどの口語的な要素に対し、ChatGPT生成データがより構造化され過度に流暢であることが要因となった可能性がある。

また、提案フレームワークの有効性を評価するため、日本語GPT-NeoXを用いたアブレーション実験を実施した。既存の (a_t, H_a^{Aug}) を用いてデータ拡張するw/o DA History Generationと、拡張時にスタイルを利用しないw/o Speaker Styleは、いずれもLow-Resourceを上回り、スタイル抽出やDA履歴生成の単独利用でも性能向上が可能であることを示した。w/o Second Finetuneは、2回目のファインチューニングなしでも性能向上に寄与する結果を示し、新しい (a_t, H_a^{Aug}) の組み合わせの有用性が確認された。提案手法(Ours)は完全一致率と部分一致率で最良の性能を達成し、話者スタイル抽出とDA履歴生成の統合、および対象ユーザグループに特化した2段階ファインチューニングの有効性を示した。

5 おわりに

本研究では、データが限られ、特殊な話者スタイルを持つユーザグループに対するDA予測モデルの性能向上を目的としたデータ拡張手法を提案した。実験結果により、提案手法の信頼性とその構成要素の有効性が確認された。

謝辞

本研究は JSPS 科研費 19H05692 の助成を受けたものです。

参考文献

- [1] Shuzheng Si, Wentao Ma, Haoyu Gao, Yuchuan Wu, Ting-En Lin, Yinpei Dai, Hangyu Li, Rui Yan, Fei Huang, and Yongbin Li. SpokenWOZ: A large-scale speech-text benchmark for spoken task-oriented dialogue agents. In **Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track**, 2023.
- [2] Sameera A. Abdul-Kader and Dr. John Woods. Survey on chatbot design techniques in speech conversation systems. **International Journal of Advanced Computer Science and Applications**, Vol. 6, No. 7, 2015.
- [3] Seokhwan Kim, Yang Liu, Di Jin, Alexandros Papanagelis, Karthik Gopalakrishnan, Behnam Hedayatnia, and Dilek Z. Hakkani-Tür. How robust r u?: Evaluating task-oriented dialogue systems on spoken conversations. **2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)**, pp. 1147–1154, 2021.
- [4] Chandan Singh, Jeevana Priya Inala, Michel Galley, Rich Caruana, and Jianfeng Gao. Rethinking interpretability in the era of large language models, 2024.
- [5] Michimasa Inaba, Yuya Chiba, Ryuichiro Higashinaka, Kazunori Komatani, Yusuke Miyao, and Takayuki Nagai. Collection and analysis of travel agency task dialogues with age-diverse speakers. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 5759–5767, Marseille, France, June 2022. European Language Resources Association.
- [6] Michimasa Inaba, Yuya Chiba, Zhiyang Qi, Ryuichiro Higashinaka, Kazunori Komatani, Yusuke Miyao, and Takayuki Nagai. Travel agency task dialogue corpus: A multimodal dataset with age-diverse speakers. **ACM Trans. Asian Low-Resour. Lang. Inf. Process.**, jun 2024.
- [7] Selami Aydin, Leyla Harputlu, Özgehan Uştuk, Şeyda Savran Çelik, and Serhat Güzel. Difficulties in collecting data from children aged 7–12. **International Journal of Teacher Education and Professional Development (IJTEPD)**, Vol. 4, No. 1, pp. 89–101, 2021.
- [8] Connor Shorten and Taghi M. Khoshgoftaar. A survey on image data augmentation for deep learning. **Journal of Big Data**, Vol. 6, pp. 1–48, 2019.
- [9] Wenzhong Shi, Min Zhang, Rui Zhang, Shanxiong Chen, and Zhao Zhan. Change detection based on artificial intelligence: State-of-the-art and challenges. **Remote Sensing**, Vol. 12, No. 10, 2020.
- [10] Michael A. Hedderich, Lukas Lange, Heike Adel, Jannik Strötgen, and Dietrich Klakow. A survey on recent approaches for natural language processing in low-resource scenarios. In **Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, 2021.
- [11] Hyunwoo Kim, Jack Hessel, Liwei Jiang, Peter West, Ximing Lu, Youngjae Yu, Pei Zhou, Ronan Bras, Malihe Alikhani, Gunhee Kim, Maarten Sap, and Yejin Choi. SODA: Million-scale dialogue distillation with social commonsense contextualization. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 12930–12949, 2023.
- [12] Steven Y. Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. A survey of data augmentation approaches for NLP. In **Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021**, 2021.
- [13] Markus Bayer, Marc-André Kauffhold, and Christian Reuter. A survey on data augmentation for text classification. **ACM Comput. Surv.**, Vol. 55, No. 7, dec 2022.
- [14] Timo Schick and Hinrich Schütze. Generating datasets with pretrained language models. In **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, 2021.
- [15] Yongtai Liu, Joshua Maynez, Gonçalo Simões, and Shashi Narayan. Data augmentation for low-resource dialogue summarization. In **Findings of the Association for Computational Linguistics: NAACL 2022**, 2022.
- [16] Jiaao Chen and Diyi Yang. Simple conversational data augmentation for semi-supervised abstractive dialogue summarization. In **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, 2021.
- [17] Biswesh Mohapatra, Gaurav Pandey, Danish Contractor, and Sachindra Joshi. Simulated chats for building dialog systems: Learning to generate conversations from instructions. In **Findings of the Association for Computational Linguistics: EMNLP 2021**, 2021.
- [18] Yan Pan, Davide Cadamuro, and Georg Groh. Data-augmented task-oriented dialogue response generation with domain adaptation. In **Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation**, 2023.
- [19] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. A survey on dialogue systems: Recent advances and new frontiers. **SIGKDD Explor. Newsl.**, Vol. 19, No. 2, p. 25–35, nov 2017.
- [20] Zhiyang Qi and Michimasa Inaba. Data augmentation integrating dialogue flow and style to adapt spoken dialogue systems to low-resource user groups. In **Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue**, 2024.

A 話者スタイル抽出のためのプロンプトと抽出結果

プロンプト

==タスク説明==

三つの未成年の客に対する観光地案内対話と三つの一般の客に対する観光地案内対話が与えられる。与えられた対話に比べて、未成年話者の特徴と、未成年話者に対する店員の特徴をまとめて出力してください。
箇条書きで、できるだけ多く、細かくまとめてください。

==未成年対話1==

[Minor_Dialogue_1]

==未成年対話2==

[Minor_Dialogue_2]

==未成年対話3==

[Minor_Dialogue_3]

==一般対話1==

[Non-minor_Dialogue_1]

==一般対話2==

[Non-minor_Dialogue_2]

==一般対話3==

[Non-minor_Dialogue_3]

抽出結果

==未成年話者の特徴==

- ・行き先ややりたいことについて具体的なイメージがない、曖昧な表現を使う
- ・具体的なプランよりも感覚的な要望を持ちやすい。例えば「自然を楽しみたい」「人が少ない静かな場所がいい」など
- ・詳細な計画を立てるよりも大まかな希望を述べることが多い
- ・アクティビティよりも自然の景色や少し変わった体験に興味を示すことがある
- ・「なんか」「うーん」などと考えながら話すことが多い
- ・「はい」だけで応答する場合が非常に多い
- ・未成年話者の発言は、短く、ためらいがちで、何を言いたいのか明確でないことがある
- ・観光地の名前や地理的な位置にあまり詳しくない
- ・食べ物や好みについて曖昧な回答をすることがある（「肉もいいけど、海鮮もいいかも」等）

==未成年話者に対する店員の特徴==

- ・親しみやすい言葉遣いを使う
- ・レジャーやアクティビティに焦点を当てた提案をすることが多い。景色や体験を重視する提案をすることが多い
- ・動機や興味に合った提案を心がけ、具体的なスポット名を挙げる
- ・観光地の特色や見どころを詳細に説明する
- ・話者の反応を確認しながら提案を行う
- ・未成年話者に対しては、店員が複数の選択肢を提示して、興味に合ったものを選ぶように促すことが多い
- ・店員は、未成年話者の興味やニーズを汲み取るうとする姿勢で、より多くの情報を提供し、質問を繰り返して確認することが多い
- ・質問を多くして話者の興味や希望を引き出そうとする
- ・若い話者の興味を想定した活動（例：体験型アトラクション、写真撮影スポットなど）を提案する
- ・季節や時間帯に合わせた提案をすることを心掛ける
- ・話者が不明確な要望を出した時にも、柔軟に対応し提案を行う

図4 プロンプトと抽出結果

B 対話生成のためのプロンプト

==タスク説明==

与えられた未成年話者特徴に基づき、対話行為から観光地推薦対話を生成してください。

==話者スタイル==

未成年話者の話者スタイル:

- ・行き先ややりたいことについて具体的なイメージがない、曖昧な表現を使う
- ・具体的なプランよりも感覚的な要望を持ちやすい。例えば「自然を楽しみたい」「人が少ない静かな場所がいい」など（…）

未成年話者に対する店員の話者スタイル:

- ・親しみやすい言葉遣いを使う
- ・レジャーやアクティビティに焦点を当てた提案をすることが多い。景色や体験を重視する提案をすることが多い（…）

例1

==対話行為==

希望する季節を聞く質問 -> 観光地に対する要望・条件の確認 -> 旅行人数・人間関係を聞く質問

==対話==

店員: [希望する季節を聞く質問]ご旅行になる季節はもうお決まりでしょうか？

お客: 季節は冬がいいですね。

店員: [観光地に対する要望・条件の確認]冬でございますね。

お客: はい。

店員: [旅行人数・人間関係を聞く質問]かしこまりました。では何名様でご旅行になりますか？

お客: えっと、兄弟2人で旅行したいです。

例2~7

(…)

生成目標

==対話行為==

a_{t-n}, \dots, a_{t-1}

==対話==

C データセット対話行為タグ

旅行代理店対話コーパス [5, 6] では、対話中のオペレータとユーザに対してそれぞれ DA タグが注釈されているが、本研究では、表 2 に示すオペレータの DA タグのみを使用した [20]。

表2 オペレータ役タグ一覧

None
方面・県・サブエリアを聞く質問
希望する季節を聞く質問
旅行人数・人間関係を聞く質問
年齢を聞く質問
経験を聞く質問
希望を聞く質問
検索に関連する質問・提案
観光地に対する要望・条件の確認
目的地の確認
チェックリストへの追加
最終的な旅行計画のまとめ
共有画面上の観光地の提案
共有画面上の観光地に関する質問
探す意思の表明
写真や地図の情報提供
検索条件の情報提供
観光地名の情報提供
検索結果の説明に基づく観光地の説明
営業時間と閉店日の情報提供
料金の情報提供
特徴条件の情報提供
アクセスの情報提供
電話番号の情報提供
駐車場の情報提供
検索結果がない意思表示
観光地の情報の誤りの訂正
特定の観光地についての主観評価や推測
検索結果全体の報告

図5 対話生成のためのプロンプト