

日本語日常会話における他者開始修復の分析

坪倉 和哉¹ 入部 百合絵¹ 北岡 教英²

¹ 愛知県立大学大学院 情報科学研究科 ² 豊橋技術科学大学 情報・知能工学系
id231001@cis.aichi-pu.ac.jp, iribe@ist.aichi-pu.ac.jp, kitaoka@tut.jp

概要

LLMの登場により対話システムはより自然な発話が生成できるようになったが、依然として対話の破綻が生じているため、破綻を修復する手法が必要となる。本研究では対話システムの自然な修復を実現するため、人同士の会話における修復開始発話の言語パターンを分析することにより、人がどのように修復を開始するかを明らかにする。トラブルの発生をどのように伝えるかによって他者開始修復(OIR)発話のタイプが3つ提案されているが、日本語日常会話コーパスを用いて分析を行った結果、OIR発話タイプ毎に品詞の出現パターンが異なることが明らかになった。具体的には、OIR発話タイプ毎に形態素数が異なることや、品詞の種類や出現頻度が異なることが示された。本結果は、対話システムが自然に修復を実行するための基礎的な知見となり得る。

1 はじめに

大規模言語モデルの登場により、対話システムは自然な応答が生成できるようになった。しかし、人間同士の会話でも対話中に理解や聞き取り上のトラブルが発生し修復が必要となるため、人とシステムとの対話中においても対話中のトラブル(いわゆる対話破綻)の発生を検出し、適切に対話を修復する必要がある。これまで、対話破綻を検出する研究は様々行われてきた。例えば、「対話破綻検出チャレンジ」[1]はユーザとのテキスト対話における対話破綻検出精度を競うワークショップであり、破綻検出のための手法が研究開発されてきた。一方で、破綻を修復する方法については、いくつか提案されているが[2, 3]、確立された修復手法は存在しない。

修復は人間同士の会話でも発生するため、人-人コミュニケーションにおける修復方法を分析することで、その知見を人-対話システム間の修復に活用することが期待できる。Ngoらの研究[4]では会話

エージェントにおける修復開始の検出と修復発話生成のために、他者開始修復発話の特徴付ける明確なパターンを明らかにした。しかし、Ngoらの研究[4]ではタスク指向対話に着目しており、非タスク指向対話における修復開始パターンについては分析されていない。

本研究では、システムが破綻した際に、ユーザが修復を開始し、システムまたはユーザが修復を実行することを想定した「他者開始修復(Other-Initiation-Repair; OIR)」に着目し、非タスク指向対話における修復開始発話の分析を行う。OIR発話は先行研究[5]において3つのタイプに分類されており、本研究ではOIR発話のタイプ毎の品詞の出現パターンの特徴を探る。これにより、人間が自然に行っている修復の開始を対話システムに実装するための知見を得ることを目的とする。また、システムがユーザの修復開始発話を認識することを想定して、OIR発話タイプの識別実験も行い、自動でOIR発話タイプを分類することが可能かを示す。本研究の貢献は次の2点である。

- 日本語日常会話における他者開始修復発話の品詞パターンをそのタイプ毎に明らかにした。
- 他者開始修復発話タイプの自動分類を行い、79.9%の精度を得た。

2 データセット

日本語の非タスク指向対話における他者修復開始発話を分析するため、本研究では日本語日常会話コーパス(Corpus of Everyday Japanese Conversation; CEJC)[6]を用いる。CEJCから修復事例を抽出し(2.2節)、他者開始修復発話タイプのアノテーションを行う(2.3節)。

2.1 日本語日常会話コーパス(CEJC)

日本語日常会話コーパス(Corpus of Everyday Japanese Conversation; CEJC)[6]とは、日常生活の中

で自然に生じる会話を収録したコーパスである。多様な場面における多様な話者との会話がバランスよく収集されている。本研究では、人間同士の会話における修復事例を分析するため、CEJCのうち、談話行為情報が提供されている「コア」(コーパス全体のうち複数のアノテーションが提供されている20時間、52セッションのデータ)と呼ばれるサブセットを分析対象とする。

2.2 修復開始発話の抽出

CEJCのコアに含まれる対話には、ISO 24617-2をベースに発話単位毎に談話行為情報が付与されている[7]¹⁾。談話行為情報には修復開始を表すRepairInitiationのタグが含まれているため、まずはRepairInitiationタグの集計を行った。その結果、CEJCのコアに含まれる59,324発話のうち、relationType2列²⁾がRepairInitiationである発話は510発話あった。修復開始のRepairInitiationが付与されている発話にはrelation2列³⁾に遡及的依存関係(Retrospective Dependence)または外部遡及的依存関係(External Retrospective Dependence)のタグが付与されている。依存関係のタグは、関連する談話機能を担う二つの発話の関係を示すものであり、二つ目の発話にタグが付与されている。遡及的依存関係は二つ目の発話一つ目の談話機能に対する任意の反応を担うタイプのつながりである。外部遡及的依存関係は周囲で生じた物音に対してフィードバック発話を行う等の発話以外で生じた出来事に対する任意の反応を担うタイプのつながりであるため、トラブルの原因となる発話(Trouble Source)が存在しないことから、分析からは除外した。その結果、508件の修復開始発話を得た。

次に、修復開始発話に対するTrouble Sourceとなる発話を特定する。RepairInitiationタグが付与された発話と遡及的依存関係にある発話は修復開始発話に対する依存先であるため、その発話はTrouble Sourceと見なすことができる。そのため、RepairInitiationタグが付与された発話のrelation2列の発話idを参照することで、Trouble Source発話を特定する。ここで、RepairInitiation発話自身やRepairInitiation発話

より時系列が後の発話にTrouble Sourceが指定されているケースが4件存在したため、これらを分析から除外した。また、遡及的依存関係に複数の発話が指定されているケースが20件存在したが、Trouble Sourceとなる発話が複数あると、分析が複雑になるためここではこれらのケースについても除外した。以上の処理により、484発話のRepairInitiation発話を得た。

最後に、484件のRepairInitiation発話に対して、自己開始または他者開始のタグを付与した。Trouble Source発話とRepairInitiation発話が同一話者により行われている場合、自己開始タグを、それ以外の発話に他者開始タグを付与した。その結果、自己開始修復発話は3件、他者開始修復(Other-Initiation-Repair; OIR)発話は481件であった。人と対話システムとの対話における修復を考える場合にも、ユーザ(他者)からの修復開始発話に基づき、システムが修復を行うことが求められるため、本研究ではOIR発話に焦点を当てる。

2.3 他者開始修復発話タイプのアノテーション

他者開始修復(OIR)発話を分析するため、タスク指向対話のOIR発話の言語構造を分析している先行研究[4]に従い、まずはOIR発話を3つのタイプ[5]に分類する。これによりOIR発話のタイプ毎の言語パターンの特徴を探る。

文献[5]ではOIR発話をOpen Request, Restricted Request, Restricted Offerの3つに分類している(表1)。本研究ではこの分類を採用し、人手でアノテーションを行う。まず、3名のアノテータがそれぞれ481件のOIR発話に対してOpen Request, Restricted Request, Restricted Offer, Other(その他, 不明)のいずれかのラベルを付与した。アノテーションの結果、Fleiss' Kappaは0.667(Substantial agreement: かなり一致)であった。次に、3名のアノテータにより付与されたラベルを用いて多数決により最終的なラベルを決定した。多数決により決定できなかった発話(8件)と多数決の結果Otherに割り当てられた発話(11件)についてはデータから除いた。以上の結果、Open Requestが249発話(53.9%)、Restricted Requestが149発話(32.3%)、Restricted Offerが64発話(13.9%)となった。タスク指向対話における先行研究[4]では、Open Requestが20発話(6.5%)、Restricted Requestが32発話(10.4%)、Restricted Offerが255発話(83.1%)であり、タスク指向対話と非タ

1) 談話行為情報のマニュアル: https://www2.ninjal.ac.jp/conversation/cejc/doc/dialogAct_manual.pdf

2) CEJCのコアデータ中のrelationType2列には談話の展開や会話の調整に関わる情報が付与されている。

3) relation2列はrelationType2のタグの依存関係を結ぶ発話のidが付与されている。

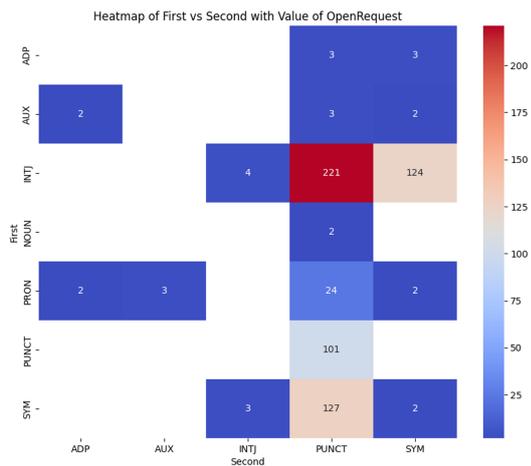


図1 Open Request の品詞タグのシーケンスパターン

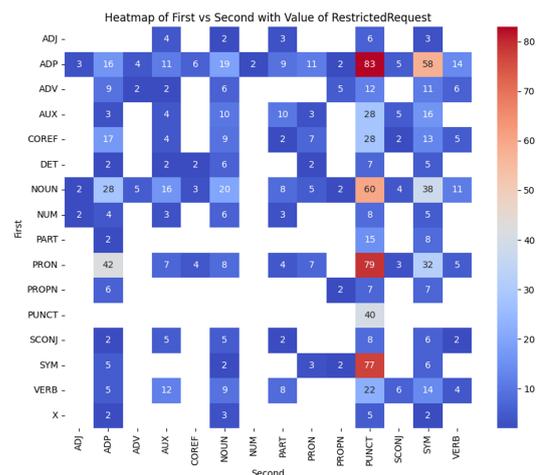


図2 Restricted Request の品詞タグのシーケンスパターン

スク指向対話では OIR 発話タイプの分布が異なることが確認された。OIR 発話タイプ毎の頻出發話を付録 A に示す。

3 他者修復発話の分析

3.1 分析方法

3つの他者開始修復 (OIR) 発話タイプ毎に言語パターンを分析するため、文献 [4] の分析手法を参考にした。まず、日本語自然言語処理オープンソースライブラリ GiNZA [8]⁴⁾を用いて、OIR 発話を形態素解析し、各形態素の品詞タグを得ることで OIR 発話の品詞タグ列を取得する。次に、トラブルの原因となる発話 (Trouble Source 発話) と OIR 発話間の関係进行分析するため、汎用言語モデルに基づく日本語解析器である KWJA [9]⁵⁾を用いて共参照解析を行う。これにより、OIR 発話が Trouble Source 発話を参照しているかを決定する。共参照解析に基づき、OIR 発話から Trouble Source 発話を参照している単語を抽出し、GiNZA により得た品詞タグ列から当該単語の品詞タグを [COREF] に置き換える。以上の手順で、OIR 発話を品詞タグ列に変換したのち、シーケンスパターンを生成するライブラリである Seq2Pat [10]⁶⁾を用いて、OIR 発話の文法的な構造を分析する。なお、発話テキストに含まれる転記タグは手動で削除した後、形態素解析と共参照解析を行った。

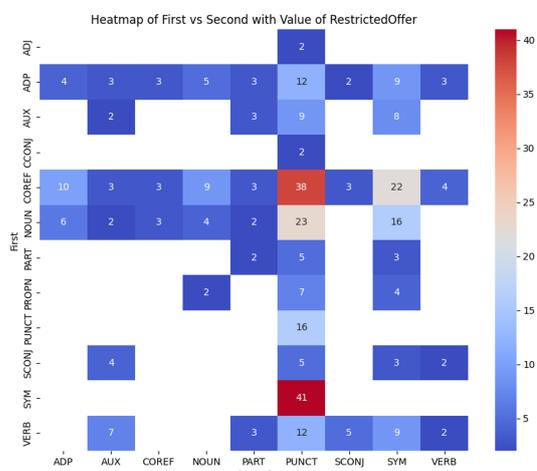


図3 Restricted Offer の品詞タグのシーケンスパターン

3.2 分析結果と考察

まず、3つの OIR 発話タイプ毎の形態素数の比較を行う。OIR 発話の形態素数は GiNZA による形態素解析の結果から算出した。平均形態素数は、Open Request が 2.98, Restricted Request が 4.81, Restricted Offer が 4.16 であった。Kruskal-Wallis 検定により3つの OIR 発話タイプの平均形態素数が異なるか検定を行った結果、有意差が確認されたため ($p < 0.001$), 事後検定として Steel-Dwass 検定を実施した。その結果、すべてのタイプ間で有意差 (1%水準) が確認されたことから、OIR 発話のタイプ毎に形態素数の大きさが異なることがわかった。タイプ毎の形態素数の箱ひげ図を付録 B に示す。

次に、OIR 発話の文法的な構造を分析するため、品詞タグのシーケンスパターンを Seq2Pat により分析した結果を図 1, 図 2, 図 3 に示す。図の縦軸は 1 番目の品詞タグ、横軸は 2 番目の品詞タグを示して

4) <https://megagonlabs.github.io/ginza/>

5) <https://github.com/ku-nlp/kwja>

6) <https://fidelity.github.io/seq2pat/>

表 1 他者開始修復発話タイプの定義と例

他者開始修復発話タイプ	定義と例
Open Request	問題が正確に何であるか、どこにあるのかを明かさずに、前の会話に何らかの問題があることを示す要求。疑問のイントネーションを持つ感嘆詞や疑問詞を用いるもの。すみませんなどの定型表現もこれに該当する。 例：え？、何？
Restricted Request	問題をより詳細に特定または特徴付けることで、回答を制限する要求。詳細や明確化を求める。いつ？やどこ？などの疑問語によって行われ、多くの場合、部分的な繰り返しと組み合わせて行われる。 例：(B の発話が RestrictedRequest) A：これ持ってる？ B：どれ？
Restricted Offer	問題をより詳細に特定または特徴付けることで、回答を制限する要求。確認を求める。トラブルの原因となる発話の全部または一部の繰り返しまたは言い換えによって行われる。 例：(B の発話が RestrictedOffer) A：タブレット持ってる？ B：タブレット？

おり、数値はその頻度を示している。

3つのヒートマップを比較すると、Open Request は品詞の種類が少なく、次に Restricted Offer, Restricted Request の順で品詞の種類が多くなっており、パターンが複雑になっていることがわかる。図 1 より、Open Request は INTJ (感動詞) の出現頻度が高いことがわかる。また、INTJ (感動詞) から SYM (疑問符を含む記号) への遷移も多く、「ん?。」や「うん?。」が頻出することが確認された。

図 2 を確認すると、Restricted Request は他の OIR 発話タイプより ADP (助詞) が多いことがわかる。上述のように、Restricted Request は3つのタイプの中で形態素数が最も大きく、助詞の使用により意味が付け加えられている可能性がある。Restricted Offer (図 3) は、COREF (共参照) の出現頻度が他のタイプよりも多い。共参照により、Trouble Source 発話の単語を参照することでトラブルの発生を伝えていると考えられる。以上のことから、OIR 発話のタイプ毎に言語パターンが異なることが明らかになった。

3.3 他者開始修復発話タイプの分類

前節で他者開始修復 (OIR) 発話のタイプ毎に文法的な構造が異なることが確認されたため、本節では OIR 発話タイプの自動検出を試みる。まず、各 OIR 発話を BERT (bert-base-uncased) を用いて [CLS] トークンの埋め込みベクトルを取得する。そのベクトルを用いて、Support Vector Machine (線形カーネル) により 5 分割検証を行った。識別の結果、正解率は 79.9%、F1-macro は 0.704 であった。混

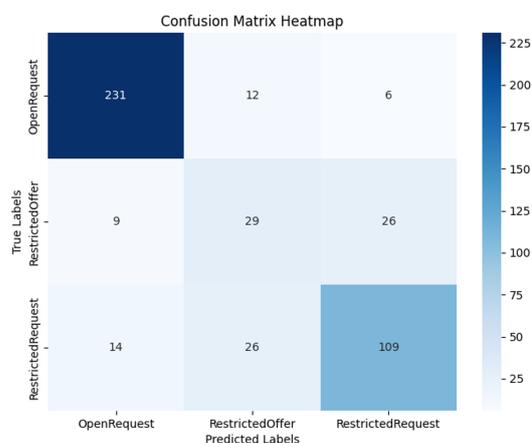


図 4 他者開始修復 (OIR) 発話タイプ識別結果

同行列を図 4 に示す。混同行列から、Open Request と Restricted Request はおおむね識別できているのに対して、Restricted Offer は Restricted Request と誤って識別されるケースが多いことがわかった。

4 おわりに

本研究では、日本語日常会話における人-人コミュニケーションの修復開始発話のパターンを分析するため、日本語日常会話コーパスを用いて分析を行った。日本語日常会話コーパスに対して他者開始修復発話タイプのアノテーションを行い、タイプ毎に品詞の出現パターンを分析した。その結果、日本語日常会話における他者開始修復の品詞の出現パターンをそのタイプ毎に明らかにするとともに、他者開始修復発話タイプの自動分類を行い、79.9%の精度を得た。今後は、修復発話の分析を行い、修復処理の実装を行う。

謝辞

本研究は国立研究開発法人産業技術総合研究所と行っている「高齢者を対象とした音声言語処理に関する研究開発」の一部として実施されたものです。また、愛知県が公益財団法人科学技術交流財団に委託し実施している「知の拠点あいち重点研究プロジェクト第IV期（第4次産業革命をもたらすデジタル・トランスメーション（DX）の加速）」の研究成果です。

参考文献

- [1] Ryuichiro Higashinaka, Luis F. D'Haro, Bayan Abu Shavar, Rafael E. Banchs, Kotaro Funakoshi, Michimasa Inaba, Yuiko Tsunomori, Tetsuro Takahashi, and João Sedoc. **Overview of the Dialogue Breakdown Detection Challenge 4**, pp. 403–417. Springer Singapore, Singapore, 2021.
- [2] Dennis Benner, Edona Elshan, Sofia Schöbel, and Andreas Janson. What do you mean? a review on recovery strategies to overcome conversational breakdowns of conversational agents. In **International Conference on Information Systems (ICIS)**, Dec 2021.
- [3] Kazuya Tsubokura, Takuya Takeda, Yurie Iribe, and Norihide Kitaoka. Dialog breakdown recovery strategies based on user personality. In **Proceedings of the 14th International Workshop on Spoken Dialogue Systems Technology (IWSDS 2024)**, 2024.
- [4] Anh Ngo, Dirk Heylen, Nicolas Rollet, Catherine Pelachaud, and Chloé Clavel. Exploration of human repair initiation in task-oriented dialogue: A linguistic feature-based approach. In Tatsuya Kawahara, Vera Demberg, Stefan Ultes, Koji Inoue, Shikib Mehri, David Howcroft, and Kazunori Komatani, editors, **Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue**, pp. 603–609, Kyoto, Japan, September 2024. Association for Computational Linguistics.
- [5] Mark Dingemanse and N. J. Enfield. Other-initiated repair across languages: towards a typology of conversational structures. **Open Linguistics**, Vol. 1, No. 1, 2015.
- [6] Hanae Koiso, Haruka Amatani, Yasuharu Den, Yuriko Iseki, Yuichi Ishimoto, Wakako Kashino, Yoshiko Kawabata, Ken'ya Nishikawa, Yayoi Tanaka, Yasuyuki Usuda, and Yuka Watanabe. Design and evaluation of the corpus of everyday Japanese conversation. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odijk, and Stelios Piperidis, editors, **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 5587–5594, Marseille, France, June 2022. European Language Resources Association.
- [7] Yuriko Iseki, Keisuke Kadota, and Yasuharu Den. Characteristics of everyday conversation derived from the analysis of dialog act annotation. In **2019 22nd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)**, pp. 1–6, 2019.
- [8] 松田寛. Ginza - universal dependencies による実用的日本語解析. *自然言語処理*, Vol. 27, No. 3, pp. 695–701, 2020.
- [9] Nobuhiro Ueda, Kazumasa Omura, Takashi Kodama, Hirokazu Kiyomaru, Yugo Murawaki, Daisuke Kawahara, and Sadao Kurohashi. Kwja: A unified japanese analyzer based on foundation models. In **Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics: System Demonstrations**, Toronto, Canada, 2023.
- [10] Serdar Kadioglu, Xin Wang, Amin Hosseininasab, and Willem-Jan van Hove. Seq2pat: Sequence-to-pattern generation to bridge pattern mining with machine learning. **AI Magazine**, Vol. 44, No. 1, pp. 54–66, 2023.

A 他者開始修復のタイプ毎の頻出發話とその度数

表2 他者開始修復のタイプ毎の頻出發話とその度数

OpenRequest		RestrictedRequest		RestrictedOffer	
頻出發話	度数	頻出發話	度数	頻出發話	度数
ん?。	84	これ?。	11	ランタン?。	2
うん?。	54	どれ?。	5	戦争?。	2
え?。	36	何を。	4	(以下は度数が1のため省略)	
えっ?。	16	どっち?。	4		
何?。	8	どうゆうこと?。	3		
何何?。	6	何が?。	3		
はい?。	6	どこ?。	3		
え。	5	誰?。	3		
あ?。	4	何が?。	3		
何。	4	何を?。	2		

B 他者開始修復のタイプ毎の形態素数の比較

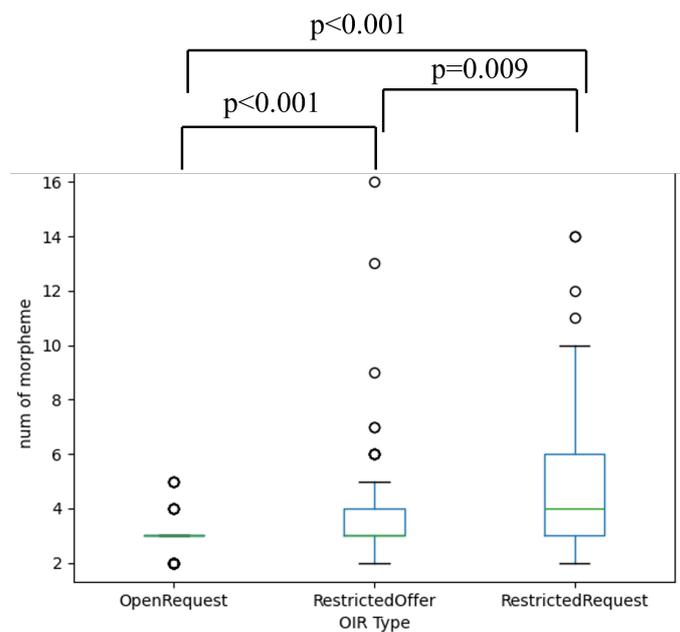


図5 他者開始修復（OIR）発話のタイプ毎の形態素数