

チャットボットとの雑談対話に対する印象評価と 評価者の性格特性・コミュニケーションスキルとの関連

山本 茉莉¹ ト 秋予¹ 伊藤 紀子²

¹同志社大学大学院 文化情報学研究科 ²同志社大学 文化情報学部
{ctmh0001,ctmh0011}@mail4.doshisha.ac.jp nito@mail.doshisha.ac.jp

概要

従来の雑談対話システムには、いくつかの問題が存在し、ユーザが円滑に対話を継続することができない、すなわち対話が破綻してしまうことがある。しかし、ユーザによって、システムの破綻発話に対する許容度は異なると考えられる。そこで、本調査では、雑談対話におけるシステムの破綻発話に対する不快感の度合いが、ユーザの性格特性・コミュニケーションスキルと関連すると仮説を立て、調査を実施した。その結果、ユーザの「神経症傾向」の特性が高いほど、システムの破綻発話に対するユーザの許容度が低くなり、「自己統制性」の特性が高いほど、システムの破綻発話に対するユーザの許容度は高くなることが分かった。

1 はじめに

近年、Microsoft 社の Cortana, Apple 社の Siri, Google 社の Google Assistant など、各社からさまざまな対話システムが提供されている。これらは主にユーザからの指示や課題を実行・解決することを目的としているが、rinna 株式会社の「りんな」のように雑談ができる対話システムも存在する。従来の雑談対話システムは、いくつかの問題を内包しており、対話の破綻を完全に回避することは困難な段階である。このような対話の破綻は、ユーザに不快感を与え、システムの使用率の低下に繋がる恐れがある。

雑談対話システム発話への評価がユーザの個人特性の違いにより異なりうることが先行研究で示唆されている。東中ら [1]や叶内ら [2]では破綻の判定がアノテーター間で一致しないことが少なくないことが報告されている。また宮本ら [3]は、「ユーザのパーソナリティ特性が心理的近接発話を行う雑談対話システムの受容性にどのような影響を及ぼすか」と「心理的近接発話の戦略の違いが雑談対話システム

の受容性にどのような影響を及ぼすか」という点に着目し、ユーザの雑談対話システムに対する受容性について検討した。その結果、システムが攻撃的冗談を発した際の好感度について、開放性や外向性といった性格特性については正の相関を示したが、神経症傾向については負の相関を示すなど、ユーザの性格特性と関連があることが報告されている。この研究では、雑談対話の終盤のシステム発話を変化させることで受容度にどのような違いが見られるかを調査したものであり、対話中のシステム発話一つ一つに対するユーザの印象評価については調査されていない。

そこで、本調査では、評価者の個人特性の内、性格特性とコミュニケーションスキルに焦点を当て、雑談対話の破綻応答文に対する不快感の度合いの分析を実施する。本調査によって性格特性・コミュニケーションスキルと破綻応答文に対する許容度に一定の規則性を見出すことができれば、それを雑談対話システムに組み込むことにより、ユーザの特性に合わせた対話の展開が可能となり、対話継続欲求の低下を防ぐ効果が期待できる。

2 調査内容

2.1 雑談対話コーパス

システム発話に対する破綻アノテーションを行ってもらうために、岩下 [5]のチャットボットとの対話コーパスから 12 件を抽出した。この対話コーパスは、LINE を通じてインターメディアプランニング株式会社のチャットボット「かたらい」と、以下の 3 つのテーマに則し雑談した際の対話データである。

- 「新しいクラスで初めて知り合った隣の席の学生と始めて会話する感じで」
- 「好きな〇〇について」
- 「小学校の時の思い出」

1人が3つのテーマを合わせて約10分間チャットしたもので、12人分の対話データにシステム発話は2156件含まれている。

2.2 破綻アノテーション実験

実験参加者は20代の日本語母語話者の男女10名である。実験課題は以下の通りである。

1. 直感に従ってシステム発話を○, △, ×の3種類で評価
2. △もしくは×と評価した発話については、5件法のリッカート尺度を用いて、どの程度不快に感じるかについて評価

○, △, ×の基準は、東中ら [1]を参考に、システムの発話を破綻の度合いに応じて、以下の3種類に分類するよう指示した。

- ○: 「破綻ではない」…当該システム発話のあと対話を問題無く継続できる。
- △: 「破綻と言いきれないが、違和感を感じる発話」…当該システム発話のあと対話をスムーズに継続することが困難。
- ×: 「あきらかにおかしいと思う発話(破綻)」…当該システム発話のあと対話を継続することが困難。

不快感には5件リッカート尺度(1.全く不快ではない~5.非常に不快である)で回答してもらい、△, ×のシステム発話数で除することで標準化を行った。

$$\text{不快感の得点} = \frac{[1 \text{ の個数}] * 1 + [2 \text{ の個数}] * 2 + [3 \text{ の個数}] * 3 + [4 \text{ の個数}] * 4 + [5 \text{ の個数}] * 5}{[\Delta, \times \text{ の個数}]}$$

2.3 事後アンケート

実験参加者自身の性格特性・コミュニケーションスキルに関する質問に回答してもらった。性格特性については、小塩ら [6]の日本語版TIPI-Jを使って、全10項目を7件法のリッカート尺度で回答してもらい、「外向性」、「協調性」、「勤勉性」、「神経症傾向」、「開放性」の5因子の得点を計測した。コミュニケーションスキルについては、藤本ら [7]のENDCOREを使って、全6項目を7件法のリッカート尺度で回答してもらい、「自己統制」、「表現力」、「解読力」、「自己主張」、「他者受容」、「関係調整」の6因子の得点を計測した。

2.4 分析方法

性格特性・コミュニケーションスキルと破綻発話への不快感の関係を明らかにするために、TIPI-JとENDCOREの各得点を予測変数、標準化済みの不快感得点を目的変数として重回帰分析を行った。

3 分析結果

調査によって得られた数値を表1-3にまとめた。

表1 破綻アノテーションに関する基本統計量

	○の 個数	△の 個数	×の 個数	不快感 得点
平均値	406.60	143.40	528.00	3.37
標準偏差	133.54	81.25	86.75	0.63

表2 性格特性の得点の基本統計量

	外向性	協調性	勤勉性	神経症 傾向	開放性
平均値	6.30	10.00	6.20	10.60	7.40
標準偏差	3.92	2.45	2.78	2.72	3.31

表3 コミュニケーションスキルの基本統計量

	自己 統制	表現 力	解読 力	自己 主張	他者 受容	関係 調整
平均値	3.70	3.20	5.00	3.70	5.50	4.70
標準偏差	1.57	1.32	1.33	1.34	1.51	1.89

3.1 性格特性と不快感の関連

性格特性5因子の得点からシステム発話への不快感の得点を予測するために、強制投入法の重回帰分析を行った(表4)。

表4 性格特性を予測変数とした重回帰分析の結果

	非標準化係数		標準偏回帰 係数β
	B	標準誤差	
外向性	-0.067	0.049	-0.417
協調性	-0.169	0.098	-0.658
勤勉性	0.025	0.067	0.111
神経症傾向	0.182	0.071	0.786 ⁺
開放性	0.120	0.061	0.631
切片	2.511	1.438	

調整済み $R^2 = 0.68$, $+0.05 < p < 0.10$

分析の結果、「神経症傾向」の標準偏回帰係数にのみ有意傾向が見られ、不快感に影響を与えていることが分かった。

3.2 コミュニケーションスキルと不快感の

関連

ENDCORE 各因子の得点からシステム発話への不快感の得点を予測するために、強制投入法の重回帰分析を行った。「他者受容」と「関係調整」については VIF が 10 を超えたため分析から除外した (表 5)。

表 5 コミュニケーションスキルを予測変数とした重回帰分析の結果

	非標準化係数		標準偏回帰係数 β
	B	標準誤差	
自己統制	-0.353	0.074	-0.876*
表現力	0.061	0.080	0.128
解読力	-0.067	0.116	-0.142
自己主張	0.201	0.106	0.426
切片	4.074	0.367	

調整済み $R^2 = 0.81$, * $p < 0.05$

分析の結果、「自己統制」の標準偏回帰係数のみ統計的に有意であり、不快感に影響を与えていることが分かった。

4 考察

分析結果から、TIPI-J における「神経症傾向」の得点が高い評価者は破綻発話への不快感の度合いが高くなることが分かった。小塩 [5] によると、神経症傾向の得点が高いとは「心配性で、うろたえやすく」「冷静で、気分が安定し」ない特性を持つ人物である。また、ENDCORE における「自己統制」の得点が高い評価者は破綻発話への不快感の度合いが低くなることがわかった。藤本ら [7] によると、自己統制の得点が高いとは「自分の感情や行動をうまくコントロールする」特性を持つ人物である。

これらを踏まえると、「神経症傾向」が強い評価者は正直に不快感を示したが、「自己統制」が強い評価者は実は不快に思っているけれどもそれをそのまま回答しなかった可能性がある。もしそうだとすれば、アンケート調査のような意識的、主観的な方

法ではなく皮膚電位や心拍などの生体信号データから不快感を計るような工夫が必要かもしれない。

5. おわりに

本研究は、さまざまある個人特性の中から性格特性とコミュニケーションスキルに着目し、破綻に対する許容度が予測可能かどうかを検証した。その結果、TIPI-J の「神経症傾向」と ENDCORE の「自己統制」が破綻発話への不快感の予測に有効であることが示唆された。しかし、現時点で実験参加者が 10 名と少ないため、より多くのデータを収集し、精度の高い分析を目指すことを今後の課題とする。

謝辞

本研究は JSPS 科研費 JP19K12191 「雑談対話システムへの個性の付与とそのコミュニケーションへの影響について」の助成を受けて行われたものである。また、同志社大学と東京工科大学の方々に実験にご協力いただいた。ここに記して感謝いたします。

参考文献

1. 東中竜一郎, 船越孝太郎, 荒木雅弘, 塚原裕史, 小林優佳, 水上雅博. テキストチャットを用いた雑談対話コーパスの構築と対話破綻の分析. 自然言語処理 23(1), pp.59-86, 2016.
2. 叶内晨, 赤崎智, 堀江伸太郎, 小亀俊太郎. Capex 雑談対話コーパスの構築とその分析. 言語処理学会第 28 回年次大会 発表論文集 pp.1090-1094, 2022.
3. 宮本友樹, 磐下大樹, 遠藤水紀, 永井望, 片上大輔. 雑談対話システムにおける心理的近接発話の戦略が対話の評価に及ぼす影響. 人工知能学会論文誌, 36(5), AG21-I_1-14, 2021.
4. 菊池浩史, 楊潔, 菊池英明. 雑談対話システムにおけるユーザの許容範囲を考慮した多様な同調応答の検討. 人工知能学会全国大会論文集 第 35 回, 4E2-OS-11b-02, 2021.
5. 岩下志乃. チャットボットとの対話コーパスと性格特性・社会的スキルのデータ公開. (https://www2.teu.ac.jp/iws/chat_corpus.pdf) 2021.
6. 小塩真司, 阿部晋吾, カトローニピノ. 日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み. パーソナリティ研究, 21(1), pp.40-52, 2012.
7. 藤本学・大坊郁夫. コミュニケーション・スキルに関する諸因子の階層構造への統合の試み. パーソナリティ研究, 15(3), pp.347-361, 2007.