

話者間の親密性を考慮したマルチモーダル雑談コーパスの構築と分析*
山崎善啓, 千葉祐弥, 能勢隆, 伊藤彰則 (東北大)

1 はじめに

非タスク指向対話システムが日常的に使われるためには、「対話をユーザに楽しんでもらうこと」が重要である。従来の対話システムの多くは、ユーザに対して画一的な応対を行うため、対話をしているうちにユーザが飽きてしまうという問題がある。一方で、人間同士の対話では対話相手との関係性に応じて対話が変化することが知られている [1]。対話システムにおいてもユーザとの親密性を考慮した口調制御の研究がなされており、ユーザ評価の向上に有用であることが報告されている [2]。従来の研究では、口調などの対話の表層的な部分に着目しているが、対話においては対話行為や非言語情報も話者間の親密性によって変化すると考えられる。このような親密性の変化と対話における諸特徴との関係を調べるため、本研究では親密性を考慮したマルチモーダル雑談対話コーパスを構築する。本稿ではまず、関係性による対話行為の違いに関して分析を行った。

2 マルチモーダル雑談コーパスの構築

2.1 対話データの収録

対話システムへの応用を想定して、音声と映像を含んだ人間同士の対一の雑談対話を収録した。Fig. 1 に収録環境を示す。各話者の発話を異なるチャンネルで録音するために対話者を別々の防音室に配置し、個別のダイナミックマイク (AT4055) で収録を行った。また、話者の表情やジェスチャーが映るように調整したビデオカメラ (GoPro Hero7) で話者の上半身を撮影し、撮影した映像と音声をもう一方の防音室のディスプレイとヘッドフォンに互いに出力した。動画像の伝送にはほとんど遅延がなく、対話者は大きな違和感を感じることなく対話を行うことができた。対話者は、「対話相手とより親しくなること」を目的として対話を行った。対話の話題として、Journald によって提案された 60 個の自己開示項目 [3] におけ

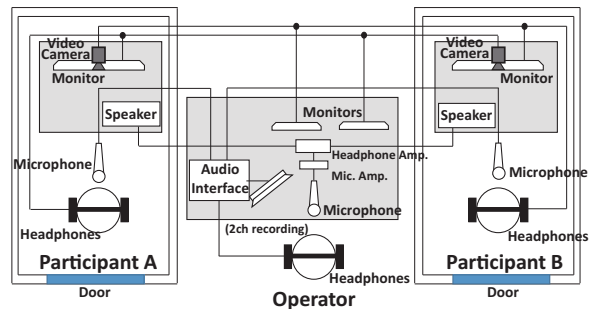


Fig. 1: 収録環境

Table 1: 収録データの概要

関係性	ペア数	対話数	発話数
初対面	8	40	7218
非初対面	11	55	12085

る “Work (or studies)” と “Tastes and interests” の中から、対話の目的を達成するのに適切であると考えられる 10 個の話題 (e.g., 休日の過ごし方) を用意した。それぞれの対話者は収録前に自分が話したいと思う話題を 5 つ選び、最終的な話題は両者の選択を基準にして実験者が決定した。対話収録の参加者は日本人の大学生・大学院生 19 名 (男性 15 名, 女性 4 名) であった。対話データの内訳を Table 1 に示す。なお本稿では、1 つの話題に関する対話を 1 対話とした。結果として、全 19 ペア (初対面: 11, 非初対面: 8) から計 95 対話 (約 10 時間) のデータが収集された。

2.2 親密性に関する事前アンケート

また、事前アンケートとして、対話者は (1) 対話相手とは面識があるかどうか、(2) 対話相手との付き合いの長さ、(3) 相手との仲の良さに関して 5 段階 (1: 全く当てはまらない, 5: 非常に当てはまる) で回答を行った。質問 (2) と (3) については、面識のあるペアのみが回答した。ここで、相手との付き合いの長さの平均値 \pm 標準誤差は 0.877 ± 0.316 [年] であった。また、親密性に関する主観評価のスコアの平均値 \pm 標準誤差は 4.00 ± 0.161 であった。そのため、非初対面のグループは概ね親しい関係であることが示唆される。

2.3 発話内容の書き起こしと対話行為タグの付与

収録した対話は 5 名のクラウドワーカが書き起こし、第一著者が表記ゆれや句点の誤りを修正した。“意味が完結しているまとまり” を基準として発話に分割したところ、19,303 発話が得られた。各発話に対しては 1 つの対話行為タグを付与した。本研究では SWBD-DAMSL タグ [4], JAIST タグ付き自由対話コーパスに用いられたタグ [5], そして聞き役対話の分析に用いられたタグ [6] を参考に Table 2 に示す 12 種類のタグを設定した。また、対話行為タグの付与は第一著者が行った。

3 親密性に基づく対話行為の分析

収録した対話データを初対面と非初対面の対話の 2 グループに分け、対話行為の出現傾向を比較した。本稿では、対話行為タグの出現頻度を分析対象とした。分析は対話単位で行った。

* Construction and analysis of multimodal chatting corpus considering the intimacy in dyads. by YA-MAZAKI, Yoshihiro, CHIBA, Yuya, NOSE, Takashi and ITO, Akinori (Tohoku Univ.)

Table 2: 初対面と非初対面のグループにおける各対話行為の出現確率 [%] (平均値 ± 標準誤差)

対話行為	初対面	非初対面	p 値
自己開示	21.07±0.96	17.88±0.67	0.008**
情報提供	19.42±1.11	22.13±0.99	0.072†
提案	0.27±0.08	0.38±0.08	0.315
要求	0.06±0.03	0.19±0.05	0.024*
確認	7.28±0.52	7.47±0.42	0.776
質問	8.78±0.54	6.88±0.44	0.008**
同意・共感	11.53±0.78	13.27±0.59	0.079†
不同意・不共感	0.45±0.10	1.04±0.28	0.050†
相槌	17.22±1.04	17.93±1.24	0.663
感嘆	6.55±0.56	5.33±0.37	0.073†
フィラー	6.20±0.60	5.93±0.37	0.709
その他	1.16±0.16	1.54±0.20	0.131

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

3.1 対話行為の出現確率

各対話行為の出現確率をそれぞれのグループに対して計量した結果を Table 2 に示す。Table 2 には各対話行為の出現確率の平均値 ± 標準誤差, 2 群間での対応のない t 検定の結果を示した。結果として, 「自己開示」, 「要求」, 「質問」において有意差が見られた。

非初対面のグループにおいては, 初対面のグループに比べ, あまり自己開示が起こらないことが示された。社会浸透理論 [1] では, 二者間の親密性が向上するにつれて, より個人の内面的な自己開示が増えたとされているが, 本稿で設定した話題には “好き, もしくは嫌いな食べ物・飲み物” などの表層的な内容が多かったため, より親密なグループでは自己開示が少なかったと考えられる。

また, 非初対面のグループでは, 初対面のグループに比べて対話相手への要求や不同意がより多く出現した。これは, 関係が構築されると, 比較的抵抗なく対話相手に要求したり, 反対意見を述べやすくなるためだと考えられる。一方で, 初対面のグループにおいては, 対話相手に対する質問が増加した。これは, 初対面のグループの方が, 積極的に対話相手の情報を得ようとするからであると考えられる。

3.2 対話行為の遷移

続いて, 対話行為の遷移について分析を行った。式 (1) によって表される遷移確率によって, 2 グループ間の各対話行為の遷移を比較した。

$$P(a_{i+1}|a_i) = \frac{C(a_i, a_{i+1})}{C(a_i)} \quad (1)$$

ここで, a_i は i 番目の発話の対話行為を表しており, $C(a_i)$ と $C(a_i, a_{i+1})$ はそれぞれ対話において出現する対話行為 a_i と, a_i から a_{i+1} への対話行為の遷移の出現回数を表している。分析では, 各グループにおいて起こりやすい遷移を比較するため, 遷移確率に関して上位 10 項目を選出した。Table 3 に選出した遷移確率の平均値 ± 標準誤差, 2 群間での対応のない t 検定の結果を示す。なお, 選出された遷移は全て一方の話者から相手話者への遷移であった。

結果より, “質問から自己開示への遷移” と “感嘆から自己開示への遷移” において有意差が見られた。

Table 3: 上位 10 項目の対話行為遷移確率 (平均値 ± 標準誤差)。

$a_i \rightarrow a_{i+1}$	初対面	非初対面	p 値
要求 → 自己開示	0.50±0.29	0.18±0.09	0.361
質問 → 自己開示	0.39±0.03	0.29±0.02	0.009**
確認 → 同意・共感	0.32±0.03	0.31±0.02	0.809
相槌 → 自己開示	0.33±0.03	0.30±0.02	0.265
相槌 → 情報提供	0.26±0.03	0.29±0.02	0.652
その他 → 情報提供	0.23±0.06	0.20±0.04	0.674
質問 → 情報提供	0.22±0.03	0.21±0.02	0.652
感嘆 → 自己開示	0.24±0.03	0.16±0.01	0.046*
提案 → 同意・共感	0.24±0.11	0.15±0.07	0.505
自己開示 → 相槌	0.19±0.01	0.18±0.01	0.543

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

初対面においては質問に対してそのまま自己開示をするといったような単純な構造の対話が多いことが示された。

3.3 対話システムへの適用

分析の結果, 対話行為の出現頻度, 遷移確率において, 親密性の段階によって傾向が異なることが示唆された。これによって, 例えば, 初対面のユーザに対しては標準的な質問応答を行うが, ユーザがシステムに慣れるにしたがって質問応答の間に意見を述べるような逸脱のある対話制御を行うことが, ユーザの親しみの向上に有用であることが示唆された。

4 おわりに

本研究では, ユーザとの親密性を考慮した対話システムの実現を目標として, マルチモーダル雑談コーパスの設計と構築を行った。分析によって, 話者間の親密性によって対話行為の出現傾向が異なることが明らかになった。今後は, コーパスの信頼性を高めるためさらなるデータの拡充を行うとともに, 非言語情報に関しても分析を行う予定である。また, 分析から示唆された対話制御手法を対話システムに導入し, ユーザ評価によりその有効性を評価する。

謝辞 本研究は, JST COI JPMJCE1303 及び科学研究費補助金 (課題番号 JP18K18136) の助成を得た。

参考文献

- [1] Altman and Taylor, “Social penetration: The development of interpersonal relationships,” New York: Holt, Rinehart & Winston, 1973.
- [2] Kageyama *et al.*, “Improving user impression in spoken dialog system with gradual speech form control,” in Proc. SIGDIAL, 235–240, 2018.
- [3] Jourard and Lasakow, “Some factors in self-disclosure,” The Journal of Abnormal and Social Psychology, 91–98, 1958.
- [4] Jurafsky *et al.*, “Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual,” draft 13, Technical Report 97-01, University of Colorado Institute of Cognitive Science.
- [5] Shirai and Fukuoka, “JAIST annotated corpus of free conversation,” in Proc. LREC, 741–748, 2018.
- [6] Meguro *et al.*, “Analysis of listening-oriented dialogue for building listening agents,” in Proc. SIGDIAL, 124–127, 2009.