

# 話題の対象に対する親密度に応じて応答する 音声対話システムの検討

加藤 大地<sup>1,a)</sup> 原 直<sup>1,b)</sup> 阿部 匡伸<sup>1,c)</sup>

**概要：**本稿では、傾聴対話システムのとる戦略として話題の対象に対する親密度に応じた応答をすることを検討する。親密度の高い応答とは、話題の対象に対する知識を有していることが感じられる応答であり、親密度の低い応答とは、話題の対象に対する知識を有していないことが感じられる応答である。これにより、最も単純な相槌のみ行うシステムに比べて、システムが話を理解しているという印象をユーザに与えることができると考えられる。本稿では、親密度の高い応答、低い応答をあらかじめ人手で用意しておくことで、音声対話システムを実装した。主観評価実験では、親密度の低い応答をするシステムは、聞き役として良い印象を与えられること、比較的良好な満足度をユーザに与えられることを確認した。親密度の高い応答をするシステムは、ユーザの話に対して関心を持っている、システムとまた話してみたいとユーザに感じさせられることを確認した。

## Investigation of spoken dialog system which respond according to degree of familiarity for topic

### 1. はじめに

近年、音声対話システムの普及が進んでいる。有名な実用化例として、Apple の Siri<sup>\*1</sup> や NTT ドコモの my daiz<sup>\*2</sup> などのスマートフォン向け音声アシスタント、Google Home<sup>\*3</sup> や Amazon Echo<sup>\*4</sup> などのスマートスピーカー、ソフトバンクの Pepper<sup>\*5</sup> のようなロボットなどがある。これらの多くは、ユーザの一発話に対してシステムが答えたり、あらかじめ用意されたタスクを実行する一問一答型であり、天気の確認など情報検索や家電製品の操作など単純なタスクの代行を得意とする。一方で、雑談のような対話そのものを目的とする非タスク指向型の対話システムの普及は進んでいない。

非タスク指向型の一分野として傾聴対話がある。これはユーザが主体となって話し、システムはそれを聞くことに

徹するものである [1], [2]。システムは適切な応答を示すことで、話への理解・共感を示しつつ、話し手がより多くのことを話せるように支援することが求められる。傾聴対話を行うことで、ユーザの話を聴いてほしいという欲求を満たしたり、思考の整理や思い出の振り返り、高齢者の認知機能の維持などの効果が期待される。実際に、カウンセリングにおいて積極的傾聴という方法が実践されたり [3]、認知症へのアプローチとして回想法 [4] や共想法 [5] といった傾聴の方法が開発され実践されたりしているなど、傾聴の需要は高まっている。したがって、傾聴を行う聞き手を音声対話システムで代替することへの期待は高い。

傾聴対話システムを実現するにあたり、聞き役となるシステムがとることができる最も単純な応答戦略は、相槌のみを返すことである。この戦略には、話し手に対して聞いているという意思表示ができ、なおかつ、相槌の多くは文脈を考慮せず使用できるため、対話が破綻しにくいという長所がある。しかし、相槌だけでは対話が単調になりがちという短所もある。

実際の人と人との対話を考えると、自分が知らない話題について話をするときには、相手に対して知らないことを尋ねたり、相手から得た情報から感じた自分の感想を述べ

<sup>1</sup> 岡山大学 大学院ヘルスシステム統合科学研究科

<sup>a)</sup> d.kato@a.cs.okayama-u.ac.jp

<sup>b)</sup> hara@okayama-u.ac.jp

<sup>c)</sup> abe-m@okayama-u.ac.jp

<sup>\*1</sup> <https://www.apple.com/jp/siri/>

<sup>\*2</sup> <https://www.nttdocomo.co.jp/service/mydaiz/>

<sup>\*3</sup> [https://store.google.com/jp/product/google\\_home](https://store.google.com/jp/product/google_home)

<sup>\*4</sup> <https://www.amazon.co.jp/b?node=5364343051>

<sup>\*5</sup> <https://www.softbank.jp/robot/pepper/>

たりすることができる。それに加えて、自分が知っている話題について話をするときには、自らの知識に基づく応答をすることができる。本稿では、この対話の際の知識の有無に注目し、次の2点について検討する。1点目は、対話システムが知識を有する場合の応答戦略の有効性、2点目は、対話システムの知識量に差がある応答戦略について、ユーザに与える印象にどのような違いが出るか、についてである。知識を有する応答戦略により、ユーザに対して、この聞き役は話を聞いている、理解しているという印象を与えられることが期待できる。

本稿は、以下の通りの構成である。第2章では、提案方式である親密度に応じて受け答える音声対話システムについて述べる。第3章で評価実験とその結果について述べる。最後に第4章で本稿のまとめを述べる。

## 2. 提案方式

### 2.1 音声対話システムの構成

本稿では音声対話システムの構成にMMDAgent (version 1.7) [6] を用いる。MMDAgentとは名古屋工業大学で開発されたオープンソースの音声インタラクション構築ツールキットである。音声認識部にはJulius[7]を使用し、入力音声を単語区切りのテキストに変換する。音響モデルはMMDAgent付属のモデルを使用した。これは、新聞記事読み上げ音声コーパス(JNAS) [8]と、高齢者&子供音声データベース [9] を用いて学習された音響モデルである。言語モデルはMMDAgent付属のモデルに加えて、2.3節で述べるキーワードを認識するための単語辞書を使用した。音声合成部にはOpen JTalk[10]を使用し、応答文テキストを音声に変換する。対話制御部は有限状態トランスデューサで処理し、ユーザによる入力音声に対してシステムが何を応答するか、という形でルールを記述する。

### 2.2 話題の対象に対する親密度

本稿では、ある話題の対象に対して有する知識の度合いのことを「話題の対象に対する親密度」と呼ぶ。例えば、「旅行」という話題に対して「岡山城」という対象があるとき、「岡山城」に対して知識が豊富にある場合は親密度が高い、と表現する。反対に知識が少ない場合は親密度が低い、と表現する。

また、話題の対象に対する親密度の高い応答とは、話題の対象に対して知識を有していることが感じられるような応答のことを指し、反対に話題の対象に対する親密度の低い応答とは、話題の対象に対して知識を有していないことが感じられるような応答のことを指す。例えば、「岡山城に行ってきました」という発話に対する親密度の高い応答は「岡山城って近くに後楽園があるんですね」(岡山城の付近に後楽園があるという知識を有している) のようなもので、親密度の低い応答は「その近くには何かほかにも観

表 1 対象とする観光地

Table 1 Target tourist spots

| 名称     | 住所     | キーワード         |
|--------|--------|---------------|
| 岡山城    | 岡山県岡山市 | 岡山城, 鳥城, 金鳥城  |
| 池田動物園  | 岡山県岡山市 | 池田動物園         |
| 広島城    | 広島県広島市 | 広島城           |
| 安佐動物公園 | 広島県広島市 | 安佐動物公園, 安佐動物園 |

光スポットってありますか」(岡山城の付近に関する知識を有していない) のようなものが例として挙げられる。

話題の対象に対する親密度の高い応答をする、すなわち、ユーザが話している事に関する知識をシステムからユーザに対して提示することにより、この聞き役は話を聞いている、話を理解している、という印象をユーザに与えられることが期待できる。

### 2.3 親密度に応じた応答生成

本稿では対話の話題を「旅行」に限定して対話システムを実装する。話題の対象は岡山県と広島県の観光地から、表 1 に示す4箇所を選定した。

応答文はそれぞれの観光地に関する知識をWeb上から収集し、話題の対象に対する親密度の低い応答(36文)と高い応答(145文)の2種類をあらかじめ人手で作成した。利用した知識の種類は、観光地の住所、周辺の観光スポット、観光地で行われるイベント、観光地で見ることができるもの、観光地で体験できることである。親密度の低い応答は、システムに知識が無いことをユーザに感じさせるために、ユーザに対して質問をするような応答を主に作成した。親密度の高い応答は、システムに知識があるとユーザに感じさせるために、知識を確認するような応答文により、ユーザに対して知識を提示する。システムは、ユーザの発話の中に、表 1 に示すキーワードが含まれるかによってユーザがどの観光地について話しているかを判断し、関連する応答文をランダムに返す。表 2 に親密度の低い応答と親密度の高い応答の応答例を示す。以降、話題の対象に対する親密度の低い応答をするシステムをシステム L、親密度の高い応答をするシステムをシステム H と呼ぶ。

また、比較対象とするベースラインシステムとして相槌のみを使用するシステムを作成した。使用する相槌の種類は、応答系感動詞である「うん」、「はい」、語彙的応答である「なるほど」、「そうですか」の4種類とした。これらはユーザの発話内容によらず使用できる汎用的な応答であり、対話が破綻するリスクが低いため、無難な対話システムとなることを意図している。システムは、ユーザの発話に対して、4種類の相槌の内どれかをランダムに返す。

### 2.4 対話例

以下に無難な相槌をするベースラインシステムの対話例

表 2 応答の種類と応答例  
Table 2 Response types and response examples

| 応答の種類          | 親密度の低い応答                  | 親密度の高い応答           |
|----------------|---------------------------|--------------------|
| 観光地の住所         | 広島県の動物園のことでしょうか           | 広島市にある動物園ですね       |
| 周辺の観光地         | その近くには何かほかにも観光スポットってありますか | 岡山城って近くに後楽園があるんですね |
| 観光地で行われるイベント   | そこって何かイベントとかありますか         | 夏にはナイトズーがあるんですね    |
| 観光地で観ることができるもの | そこには何があるんですか              | レッサーパンダ見に行きたいなあ    |
| 観光地で体験できること    | 何かそこで体験できることってありますか       | 備前焼体験やってみたいです      |
| 観光地に対する感想      | 広島城行ってみたいなあ               | 広島城はいいですね          |
| その他            | その事は詳しいんですか               | 私、そこによく行きます        |

を示す。

ユーザ 池田動物園に行ってきました  
システム うん  
ユーザ ペンギンのえさやりが楽しかったです  
システム なるほど

以下に親密度の低い応答をするシステム L の対話例を示す。

ユーザ 池田動物園に行ってきました  
システム そこって何かイベントとかありますか  
ユーザ ペンギンのえさやりが楽しかったです  
システム 池田動物園行ってみたいなあ

以下に親密度の高い応答をするシステム H の対話例を示す。

ユーザ 池田動物園に行ってきました  
システム 岡山市のあの動物園ですね  
ユーザ ペンギンのえさやりが楽しかったです  
システム 私もペンギンのえさやりに行きたいです

### 3. 評価実験

#### 3.1 実験条件

##### 3.1.1 実験の概要

学生 11 名を対象として評価実験を行った。ベースラインシステム、システム L、システム H の合計 3 つのシステムを実験参加者に使用してもらい、システム毎に主観評価を行ってもらった。実験の手順は以下のとおりである。

- (1) 実験参加者への実験の説明
- (2) 音声対話システムの使用練習
- (3) 1 つ目のシステムの使用・評価
- (4) 2 つ目のシステムの使用・評価
- (5) 3 つ目のシステムの使用・評価

まず、実験参加者へ実験内容、注意事項等の説明をした後、実験に参加することについての同意書に署名してもらった。その後、実験参加者に、観光地に関する事前アンケートに回答してもらった後、システムと対話するという行為に慣れてもらうために、音声対話システムと対話する練習をした。練習には、評価対象である 3 つのシステムと

は異なる練習用システムとして、「旅行」ではなく「食べ物」という話題の中の、「うどん」というキーワードのみに対して応答するシステムを用意した。練習用システムの応答文には、ベースラインシステムが使用する相槌や、「うどん」に対する親密度の低い応答、高い応答などを用意した。練習の際に、実験参加者には、今回使用するシステムは、システムはキーワードに反応して応答すること、そのキーワードは何か、キーワードを使いながらシステムに対して自由に話すよう説明した。

練習が終了した後、システムの使用順による主観評価への影響を避けるために、実験参加者毎に使用するシステムの順番を入れ替えて実験を行った。システムとの対話はユーザの第 1 発話から開始し、システムとの対話を終了する条件は、ユーザの第 16 発話に対するシステムの応答が終了した後とした。これは、実験参加者がシステムの印象を評価するために、ある程度システムとの発話を繰り返す必要があるためであり、対話システムライブコンペティション [11] を参考に長さを設定した。1 つのシステムの使用が終了するごとに、実験参加者により主観評価を行ってもらった。

##### 3.1.2 評価項目

表 3 に主観評価の評価項目を示す。これらは傾聴対話システムの評価 [2] を参考に、音声対話システムとして適切な振舞いをしているか、聞き役として適切な振舞いをしているか、システム全体として良いものであったかを評価する項目を用意した。

これらの各項目について、「全くそうは思わない」、「そうは思わない」、「ややそうは思わない」、「どちらとも言えない」、「ややそう思う」、「そう思う」、「とてもそう思う」という 7 段階のリッカート尺度で評価してもらった。結果の分析の際には、表 4 に示すように「全くそうは思わない」から「とてもそう思う」に評定値としてそれぞれ 1 から 7 の整数を割り当て、間隔尺度として分析する。

#### 3.2 実験結果

実験中のシステムと実験参加者の対話については、実験参加者が発話したキーワードをシステムが概ね正しく認識

表 3 評価項目

Table 3 Evaluation items

| 番号  | 項目                 |
|-----|--------------------|
| Q1  | システムが話した言葉は自然だった   |
| Q2  | システムが話す言葉の長さは適切だった |
| Q3  | システムは真面目に話を聞いてくれた  |
| Q4  | システムは積極的に話を聞いてくれた  |
| Q5  | 親身に聞いてくれていると感じた    |
| Q6  | システムは話を理解していた      |
| Q7  | システムは関心を示していた      |
| Q8  | システムは共感を示していた      |
| Q9  | このシステムは話しやすい       |
| Q10 | このシステムとまた話したい      |
| Q11 | 会話について満足した         |

表 4 リッカート尺度の対応表

Table 4 Correspondence table of Likert scale and rating value

| 評価値 | 回答        |
|-----|-----------|
| 1   | 全くそうは思わない |
| 2   | そうは思わない   |
| 3   | ややそうは思わない |
| 4   | どちらとも言えない |
| 5   | ややそう思う    |
| 6   | そう思う      |
| 7   | とてもそう思う   |

し、概ねスムーズに対話が進行した。しかし、キーワードを発話したのに認識されなかったり、誤ってキーワードと認識される事（「広島市」を「広島城」と認識）もまれにあった。また、実験参加者が複数のキーワードを発話した際（「安佐動物園の他に広島城にも行きたいです」）に、その内の1つのキーワードのみを認識して適切でない応答（「フラミンゴが見れるんですよ」）をシステムがすることもあった。

各システムの主観評価結果の箱ひげ図を図1に示す。ただし、赤い丸印は評価値の平均値、ひげの下端と上端は評価値の最小値と最大値（外れ値を除く）、黒い菱形は外れ値を表している。また、3つのシステム内の2つに対して、対応ありの両側  $t$  検定 ( $n = 11$ ) によって有意差の検定により得られた  $p$  値を表5に示す。ただし、表の  $p$  値の前にある符号は、システム A と B の平均値を比較し、B の方が大きければ (+), B の方が小さければ (-) を記載している。結果より、全ての評価項目において、ベースラインシステムよりシステム L やシステム H の方が高い評価値を得た。さらに、ベースラインシステムとシステム L の比較では、Q1, Q8 を除く 9 項目で有意差（有意水準 5%）が認められた。ベースラインシステムとシステム H の比較では、Q3, Q4, Q8 を除く 8 項目で有意差が認められた。システム L とシステム H の比較では、Q3 においてシステム L の評価が高く、で有意差が認められた。

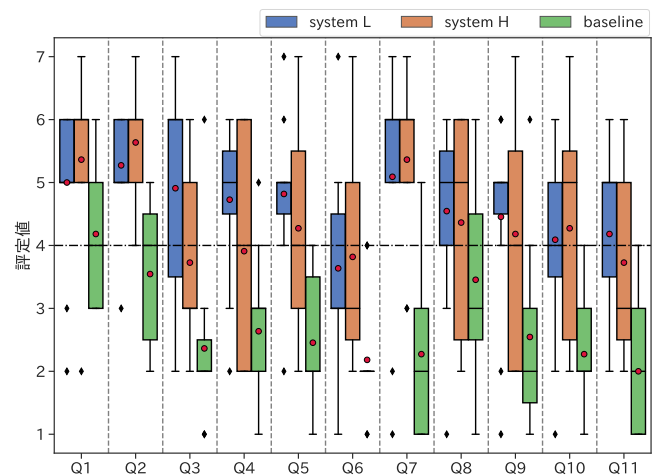


図 1 主観評価の結果

Fig. 1 Results of subjective evaluation

表 5 各システムの主観評価の  $t$  検定の結果

Table 5 Results of  $t$ -test for subjective evaluation of each system

| 項目番号 | ベースライン<br>とシステム L | ベースライン<br>とシステム H | システム L<br>とシステム H |
|------|-------------------|-------------------|-------------------|
| Q1   | (+)0.146          | (+)0.019 *        | (+)0.506          |
| Q2   | (+)0.010 *        | (+)0.001 *        | (+)0.307          |
| Q3   | (+)0.008 *        | (+)0.072          | (-)0.011 *        |
| Q4   | (+)0.005 *        | (+)0.121          | (-)0.082          |
| Q5   | (+)0.000 *        | (+)0.022 *        | (-)0.294          |
| Q6   | (+)0.024 *        | (+)0.031 *        | (+)0.733          |
| Q7   | (+)0.003 *        | (+)0.000 *        | (+)0.676          |
| Q8   | (+)0.074          | (+)0.211          | (-)0.792          |
| Q9   | (+)0.014 *        | (+)0.046 *        | (-)0.557          |
| Q10  | (+)0.008 *        | (+)0.004 *        | (+)0.659          |
| Q11  | (+)0.002 *        | (+)0.017 *        | (-)0.242          |

(\*  $p < .05$ )

### 3.3 考察

#### 3.3.1 各システムの評価の比較

表5において、有意差が認められた項目、および、提案システムの評価値の平均値が、「どちらともいえない」に相当する4未満だった項目について考察する。

「Q1 システムが話した言葉は自然だった」において、ベースラインシステムの平均値4.2に対し、システム H が平均値5.4で有意差が認められた。これは、相槌のみの応答より、文による応答の方がより自然だと評価されたとと言える。

「Q2 システムが話す言葉の長さは適切だった」についての平均値は、ベースラインシステムが3.6、システム L が5.3、システム H が5.6であった。これは、ベースラインシステムの相槌は応答として短すぎると評価されたのに対し、システム L、システム H の応答文の長さは聞き手の応答としてより適切だと評価されたとと言える。

「Q3 システムは真面目に話を聞いてくれた」についての平均値は、ベースラインシステムが2.4、システム L が

4.9, システム H が 3.7 であった。システム L については応答文の多くがユーザに対して質問をするようなものであった事が、ユーザに対して話を聞いているという印象を与えられたことにつながったと考えられる。一方、システム H の平均値は「どちらともいえない」に相当する評定値 4 よりも、やや評価が悪い。これは、システム H の応答文の多くは、システムが話題の対象に対してどう思っているか、システムによる知識の提示で構成されていたため、話を聞いてもらっているとユーザに感じてもらえなかったことが原因だと考えられる。

「Q4 システムは積極的に話を聞いてくれた」についても、ベースラインシステムとシステム L に有意差が見られたが、これは Q3 と同様の理由だと考えられる。

「Q5 親身に聞いてくれていると感じた」についても、ベースラインシステムとシステム L に有意差が見られたが、これは Q3 と同様の理由だと考えられる。ベースラインシステム（平均値 2.5）とシステム H（平均値 4.3）に有意差が見られたことについては、知識を活用した応答が、ユーザの話に対して丁寧に対応しているという印象を与えられたと考える。しかし、回答の幅が広く、今後更に詳細な分析をする必要がある。

「Q6 システムは話を理解していた」については、システム L とシステム H の両方ともベースラインシステムに対して評価が高かったが、平均値はそれぞれ 3.6, 3.8 と低い。これは、本研究で作成した音声対話システムは言語理解部を簡単なキーワードスポッティングのみで実装したため、順当な結果といえる。

「Q7 システムは関心を示していた」については、ベースラインシステムの平均値 2.3 に対して、システム L は 5.1, システム H は 5.4 とどちらも高評価だった。システム L については、Q3 と同様の理由により高評価が得られたと考えられる。システム H については、Q3 では低評価の原因となった知識の提示が、「知識のある話題に食いついてきた」というような印象をユーザに与えられたことにより、高評価に寄与したと考えられる。

「Q9 このシステムは話しやすい」については、ベースラインシステムの平均値 2.6 に対して、システム L は 4.5, システム H は 4.2 でどちらも有意差が見られた。システム L については、Q3 や Q7 で高評価だったことが影響していると考えられる。システム H についても、Q2 や Q7 で高評価だったことが影響していると考えられる。

「Q10 このシステムとまた話したい」については、ベースラインシステムの平均値 2.3 に対して、システム L は 4.1 で有意差が見られた。これは、Q3 と同様に、システム L による質問がユーザに好印象を与えられたことによると考えられる。また、システム H も平均値が 4.3 と高評価であった。これは、システムによる知識の提示によって、ユーザは知識の獲得ができるので、またシステム H の話を聞いて

表 6 各システムの合計評定値と最も良かったシステム  
Table 6 Total rating of each system and the best system

| ID | ベースライン | システム L | システム H | BEST   |
|----|--------|--------|--------|--------|
| 1  | 26     | 49     | 55     | システム H |
| 2  | 34     | 63     | 54     | システム L |
| 3  | 32     | 18     | 35     | システム H |
| 4  | 26     | 58     | 61     | システム H |
| 5  | 29     | 57     | 44     | システム L |
| 6  | 36     | 52     | 38     | システム L |
| 7  | 34     | 57     | 67     | システム H |
| 8  | 39     | 47     | 46     | システム L |
| 9  | 21     | 65     | 72     | システム H |
| 10 | 17     | 48     | 27     | システム L |
| 11 | 35     | 44     | 36     | システム L |

みたいと評価されたと考えられる。しかし、この考えにおいては、システムが話し役として機能していたとも解釈できるため、聞き役となるシステムの開発という本稿の提案方式の目的から考えると、必ずしも望ましい結果ではない。

「Q11 会話について満足した」については、ベースラインシステムの平均値 2.0 に対して、システム L は 4.2 と高評価であった。これは、Q3 や Q5 で聞き役として高評価だったことが影響していると考えられる。また、システム H の平均値は 3.7 と低評価であった。これは、Q3（平均値 3.7）や Q4（平均値 3.9）で聞き役として低評価だったことが影響していると考えられる。

### 3.3.2 最も高い評価を得たシステム

表 6 に実験参加者毎の各システムの合計評定値を示す。表には、実験参加者毎に最も合計評定値が高かったシステムも示している。この表によると、最も合計評定値が高かったシステムは、システム L が 6 名、システム H が 5 名と概ね二分される結果となった。これは、3.3.1 項から、システム L は聞き役、システム H は話し役の特徴が現れていることから、システムに話を聞いてもらえて満足した実験参加者と、システムの話が面白かったので満足した実験参加者の 2 種類に分かれたと考えられる。

### 3.3.3 重回帰分析による評価項目の影響の分析

表 3 の評価項目の内、「Q9 このシステムは話しやすい」、「Q10 このシステムとまた話したい」、「Q11 会話について満足した」については、音声対話システムの評価として総合的な項目と言える。これらの項目に対して、Q1 から Q8 の評価項目が及ぼす影響が分かれば、今後の音声対話システムの改善方針の目安になる。そこで、重回帰分析により、回帰係数からその影響を調べる。Q9, Q10, Q11 の評定値を目的変数として重回帰分析により得られた回帰係数を図 2 に示す。

まず、Q9, Q10, Q11 全てにおいて、Q4 の係数が最も大きい。したがって、システムは真面目に話を聞いているとユーザに感じさせることが、総合的な評価を向上させる

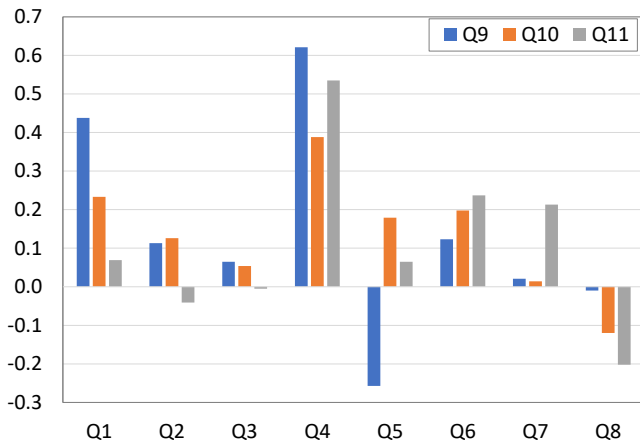


図 2 評価項目の重回帰分析

Fig. 2 Multiple regression analysis of evaluation items

ために最も重要であると考えられる。

また、Q9, Q10については、Q1の係数が2番目に大きい。したがって、システムの応答文をより自然なものにすることもユーザーに話しやすいと感じさせるため、また話したいと感じさせるために有効であると考えられる。

Q11については、Q6の係数が2番目に大きい。したがって、システムが話を理解しているとユーザーに感じさせることが、ユーザーの満足度を向上させるために有効であると考えられる。

#### 4. おわりに

本稿では、傾聴対話のための音声対話システムの応答戦略として、話題の対象に対する親密度の高い応答をする戦略を提案した。この戦略は、知識がある話題に対して知識に基づく応答をすることで、ユーザーに対して、対話システムが話を理解しているという印象を与えられることを期待したものである。

提案した音声対話システムを評価するために、11名の実験参加者による主観評価実験を行った。実験の結果から、親密度の低い応答をするシステム、親密度の高い応答をするシステム、両方の有効性が示された。親密度の低い応答をするシステムは、聞き役として良い印象をユーザーに与えられること、比較的良好な満足度をユーザーに与えられることを確認した。親密度の高い応答をするシステムは、ユーザーの話に対して関心を持っていると感じさせられること、システムとまた話してみたいとユーザーに感じさせられることを確認した。ただし、親密度の高い応答をすることで、聞き役ではなく、話し役としての役割が強くなる可能性も示唆された。

今後の課題として、親密度の低い応答をするシステムと親密度の高い応答をするシステムで評価の高い項目が分かれたため、両者の長所を併せ持つ応答戦略を検討することが挙げられる。解決するためには、本稿で実装した2つの提案システムの中間の親密度をもつシステムや、話題の対

象によって親密度が異なるシステムなどが考えられる。また、重回帰分析による評価項目の分析結果を基に、実際に音声対話システムの総合的な評価を向上させる方式の検討や、ユーザーに対して、システムが話を理解していると感じさせられるような言語理解部、応答生成部の検討も課題として挙げられる。

#### 参考文献

- [1] 下岡和也, 徳久良子, 吉村貴克, 星野博之, 渡部生聖: 音声対話ロボットのための傾聴システムの開発, 自然言語処理, Vol. 24, No. 1, pp. 3-47 (オンライン), DOI: 10.5715/jnlpp.24.3 (2017).
- [2] 井上昂治, Lala, D., 山本賢太, 中村静, 高梨克也, 河原達也: 自律型アンドロイド ERICA による傾聴対話システムの評価, 人工知能学会研究会資料 SIG-SLUD, Vol. B5, No. 02, pp. 19-24 (2019).
- [3] 池見陽: 傾聴とフォーカシングの臨床心理学, 聴能言語学研究, Vol. 13, No. 3, pp. 213-220 (オンライン), DOI: 10.11219/jjcomdis1983.13.213 (1996).
- [4] Butler, R. N.: The Life Review: An Interpretation of Reminiscence in the Aged, *Psychiatry*, Vol. 26, No. 1, pp. 65-76 (online), DOI: 10.1080/00332747.1963.11023339 (1963).
- [5] 大武美保子: 認知症の予防と支援に役立つ人工知能と高齢者とともに作る認知症予防支援サービスの開発, 人工知能, Vol. 31, No. 3, pp. 363-370 (2016).
- [6] : MMDAgent - Toolkit for building voice interaction systems -, <http://www.mmdagent.jp/>. accessed on Jan. 28, 2020.
- [7] : 大語彙連続音声認識エンジン Julius, <https://julius.osdn.jp/>. accessed on Jan. 28, 2020.
- [8] Itou, K., Yamamoto, M., Takeda, K., Takezawa, T., Matsuoka, T., Kobayashi, T., Shikano, K. and Itahashi, S.: JNAS: Japanese speech corpus for large vocabulary continuous speech recognition research, *Journal of the Acoustical Society of Japan (E)*, Vol. 20, No. 3, pp. 199-206 (online), DOI: 10.1250/ast.20.199 (1999).
- [9] : 高齢者&子供音声データベース, [https://www.ntt-at.co.jp/product/onsei\\_db/](https://www.ntt-at.co.jp/product/onsei_db/). accessed on Feb. 5, 2020.
- [10] : Open JTalk, <http://open-jtalk.sourceforge.net/>. accessed on Jan. 28, 2020.
- [11] 東中竜一郎, 船越孝太郎, 稲葉通将, 角森唯子, 高橋哲朗, 赤間怜奈, 宇佐美まゆみ, 川端良子, 水上雅博: 対話システムライブコンペティション2, 人工知能学会研究会資料 SIG-SLUD, Vol. B5, No. 02, pp. 42-49 (2019).