

# 話し手の語りに傾聴的な応答の収集

村田 匡輝<sup>†</sup>      大野 誠寛<sup>‡</sup>      松原 茂樹<sup>††</sup>

<sup>†</sup>豊田工業高等専門学校 情報工学科      <sup>‡</sup>名古屋大学 情報基盤センター

<sup>††</sup>名古屋大学 大学院情報科学研究科

murata@toyota-ct.ac.jp

## 1 はじめに

音声でやりとりする会話ロボットの研究開発が盛んである。会話ロボットは、人と会話すること自体に目的があり、コミュニケーションの過程が利用者にとって快適であるかが重視される。ロボットと会話するユーザのストレス要因として、ロボットが自らの発話を聞き、理解しているのかに関する不安がある。ユーザのロボットとのコミュニケーションへの動機付けのためにも、ロボットはむしろ積極的にユーザ発話への傾聴態度を表明することが有効である。

そこで本研究では、会話ロボットにおける、発話により話し手への傾聴を示す機能の実現を目指す。これまでに、高齢者の語りの音声を対象とした応答を集め、その分類を試みている [1]。今回、傾聴的な応答を追加で収集し、データの規模を拡張した。本稿では、収集の概要と収集した応答発話の種類について報告する。

## 2 傾聴的な応答

会話ロボットにおける音声コミュニケーションでは、ユーザとロボットが、あらかじめ役割を定めることなく、ユーザがロボットとの自然なやりとりに、楽しさや心地よさを感じられることが理想である。しかしながら、現状の音声会話技術で、人が快適に感じられるようなレベルのやりとりを日常的に遂行することは容易ではなく、そのような状況下でユーザに音声でのやりとりを強いても、それは不快をもたらすことになる。そこで、ユーザとロボットがある役割に徹することを前提とした会話の遂行が考えられる。ロボットの役割の例として、以下のようなものがある。

1. 話し役 発話コンテンツをあらかじめ備える。ロボット主導で会話を展開し、自らのコンテンツを音声で提示する。

2. 聞き役 ユーザ主導で会話が進められる。ユーザは自らが話したいことを語り、ロボットは傾聴的に振る舞い、ユーザの語る意欲を高める。

本研究では、上記のうち、2. の役割を想定し、傾聴的な会話ロボットの実現を目指す。本章の以下で、人との音声コミュニケーションにおいて会話ロボットが傾聴的であることの利点とそのための方略について論じる。

### 2.1 会話ロボットによる傾聴的な応答

人は話している相手に対して、相づちや頷きなどの行為を敢えて行うことがある。これらの行為は、理解状態を開示する方法の一つであり、相手の話す行為を円滑化する効果がある。会話ロボットが傾聴しているという印象をユーザが持つためには、上述の行為をタイミングよく実行することを意味する。

相づちは、話し手の語りを聞いているとき、その内容を聞き取っているという合図を話し手に知らせる行為である [2]。さらに、相づち以外にも、話し手への同意や共感を示す発話を行って相手と同じ気持ちであることを伝える、自身の感情の動きを示す発話を行うことで相手の話から何らかの思いを感じ取っていることを示す、といった行為により、相手の語る行為が円滑化すると考えられる。本研究では、このような応答を傾聴的応答と呼び、傾聴的応答を適切に生成できるロボットの開発を目標とする。

### 2.2 関連研究

聞き役としての役割を担うことを目的とした対話システムの開発が行われている [3-5]。これらのシステムでは、対話システムがユーザの話の展開に積極的に関わっていくのに対して、本研究が目指す傾聴的応答では、ユーザに質問や要求をするなど、対話の発話権を保持する応答を行うことなく、あくまで一方向的な応答を行いつつも、ユーザに話を聞いているという

表 1: 対象データの規模

話者数	30
合計発話時間	8時間 42分 57秒
発話数	13,381
形態素数	73,513

表 2: 収集データの規模

合計発話時間	2時間 48分 26秒
応答発話数	15,113
時間あたり応答発話数	28.90
形態素数	29,711

ことを伝えることにより、対話を促進することを目指しており、目標とする対話の形態に違いがある。

自らの理解状態の開示方法である相づち表現の生成については、生成タイミングの検出に関する研究 [6] がある一方で、相づち表現の使い分けに言及した研究は少ない [7]。一方、本研究では、傾聴的応答として、相づちに限定せず、多様な形態の応答を適切に使い分けて生成することを目標としている。

### 3 傾聴的応答の収集

傾聴的応答は会話の様々な場面で観察でき、その機能や現れ方は多様である。本研究では、傾聴的応答の特徴を明らかにするために、話し手の語りに対し、聞き役に徹して応答した発話を収集した。

#### 3.1 収集の概要

応答発話を収集するための対象データとして、高齢者のナラティブコーパス JELiCo<sup>1</sup>を利用した。これは、高齢者の語りの研究を目的に構築されたコーパスであり、30名の高齢者による平均20分の音声が発録されている。高齢者は用意された10個の質問に回答することによって、語りを遂行している。

本研究では、上記コーパス30名分の高齢者の音声データを利用して、応答発話の収集を実施した。収集のために、コーパスに収録されている音声から、各質問に対する高齢者の回答音声を個別に抽出し用いた。対象データの規模を表1に示す。なお、発話数とは、回答音声を、人が知覚できるポーズで区切った単位の総数である。

<sup>1</sup>奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室, <http://sociocom.jp/software.html>

イタリア旅行をしたこと【はい】が一番楽しかった【はーそうですかー】です【素敵ですねー】【ふーん】【イタリア旅行】もう二度と行けないかなと【いえいえそんなー】思いながら行ってきましたけど【あーそうですかー】未だに【はい】頭に残っていること【はい】は【えー】イタリア旅行です【はい】

図 1: 語りと応答の例

表 3: 応答発話と出現頻度

応答発話	出現頻度	出現割合
はい	5,870	38.84%
えー	1,091	7.22%
ええ	836	5.53%
えーえー	648	4.29%
あー	496	3.28%
うーん	327	2.16%
うん	293	1.94%
あっ	225	1.49%
んー	175	1.16%
はー	171	1.13%

抽出した305個の音声に対して、高度な接客スキルを要する業務経験を有する作業員1名が聞き役を遂行した。作業員は、接話マイクを使用し、高齢者音声を聞きながら、傾聴を示す応答を行った。収集した応答に対して、書き起こし、形態素情報、形態素単位の発話時間情報を付与した。形態素解析には MeCab [8] を、発話時間情報の付与には Julius [9] の音素セグメンテーションキット<sup>2</sup>を、それぞれ用いた。

#### 3.2 収集した傾聴的応答

収集した応答の例を図1に示す。【】で囲んだ部分が作業員による応答である。収集データの規模を表2に示す。収集した応答を、応答を行った作業員とは別の作業員1名が知覚できるポーズで区切った単位を応答発話とした。30名分の高齢者の語りに対して、合計15,113個の応答発話が行われ、その異なり数は2,051であった。出現頻度の高い応答発話を表3に示す。上位10個で全体の67.04%を占めている。

## 4 おわりに

本論文では、傾聴機能を備えた会話ロボットの実現を目的に、聞き役として適切な応答を明らかにするた

<sup>2</sup><http://julius.osdn.jp/index.php?q=ouyoukit.html>

め、傾聴的応答の収集を行った。

今後は、新たに収集した傾聴的応答の分析を進めるとともに、発話時間情報や音響情報等を用い、応答のタイミングや高齢者の発話内容との関連について分析する。さらに、その分析に基づき、聞き役としての適切な応答発話の生成手法を開発することを目指す。

謝辞 本研究は、一部、科学研究費補助金（挑戦的萌芽研究）(No. 15K12095) により実施したものである。

## 参考文献

- [1] 村田匡輝, 大野誠寛, 松原茂樹. 話し手への傾聴を示す応答発話の収集と分析. 情報処理学会研究報告. 知能システム, Vol. 2016-ICS-183, No. 1, pp. 1-7, 2016.
- [2] 日本語記述文法研究会. 現代日本語文法 7. くろしお出版, 2009.
- [3] 目黒豊美, 東中竜一郎, 堂坂浩二, 南泰浩. 聞き役対話の分析および分析に基づいた対話制御部の構築. 情報処理学会論文誌, Vol. 53, No. 12, pp. 2787-2801, 2012.
- [4] 小林優佳, 山本大介, 土井美和子. 高齢者対話インタフェース — 発話間の共起性を利用した傾聴対話の基礎検討—. FIT2011 講演論文集, Vol. 10, No. 2, pp. 253-256, 2011.
- [5] 下岡和也, 徳久良子, 吉村貴克. 音声対話ロボットのための傾聴システムの開発. 人工知能学会研究会資料, 言語・音声理解と対話処理研究会, Vol. 58, pp. 61-66, 2010.
- [6] 大野誠寛, 神谷優貴, 松原茂樹. 対話コーパスを用いた相づち生成タイミングの検出. 電子情報通信学会論文誌, Vol. J100-A, No. 1, pp. 53-65, 2017.
- [7] 山口貴史, 吉野幸一郎, 高梨克也, 河原達也. 多様な形態の相づちをうつ音声対話システムのための傾聴対話の分析. 情報処理学会第 77 回全国大会講演論文集, pp. 145-146, 2015.
- [8] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In *Proc. of EMNLP-2004*, pp. 230-237, 2004.
- [9] Akinobu Lee, Tatsuya Kawahara, and Kiyohiro Shikano. Julius — an open source real-time large vocabulary recognition engine. In *Proc. of EUROSPEECH-2001*, pp. 1691-1694, 2001.