

# 傾聴対話システムにおける 自分語りを含む多様な聞き手応答の生成

石田 真也<sup>‡</sup> 井上 昂治<sup>‡</sup> 中村 静<sup>‡</sup> 高梨 克也<sup>‡</sup> 河原 達也<sup>‡</sup>

<sup>‡</sup> 京都大学 大学院情報学研究科 知能情報学専攻

## 1. はじめに

近年、人と対話を行うシステムが日常生活においても盛んに利用されるようになってきている。対話の種類として、タスク遂行型の対話だけでなく雑談対話の研究も行われている [1]。雑談対話の一つとして、傾聴対話 [2] がある。システムが傾聴対話を行うことにより、話し相手としてユーザの欲求を満たす効果が期待されている。

傾聴対話を実現するためには、ユーザが発話しやすいように、システムのふるまいを設計する必要がある。従来研究では、システムのふるまいとして相槌に焦点が当てられてきた。システムが相槌をうつ場面において、相槌の韻律をユーザ発話へ同調させたり [3]、相槌の形態をユーザ発話の情報に応じて決定したりすること [4] などが取り組まれてきた。また、相槌以外の聞き手応答として、ユーザの発話内容に応じて、慣用的表現で同意を示す「語彙的応答」、先行発話の一部を繰り返す「繰り返し応答」、先行発話の詳細を問う「掘り下げ質問」、これらの選択、生成 [5] も取り組まれている。さらに、[6] では聞き手応答を「応答系感動詞」、「感情表出系感動詞」、「語彙的応答」、「評価応答」、「繰り返し」、「共同補完」の6つに分類している。本研究では、より自然な傾聴対話を行うために、上記の聞き手応答「掘り下げ質問」、「繰り返し応答」、「語彙的応答」、「評価応答」を生成し、さらに本研究のキーアイデアとして聞き手が自分自身に関する話を行う「自分語り」の生成を目指す。本稿では、自然な対話データに対してアノテーションを行い、その結果を分析することで、実際の対話で登場する聞き手応答の種類や回数の特徴を明らかにする。

## 2. 傾聴対話システム

提案するシステムの概要を図1に示す。入力ユーザ発話の音声である。はじめに、ユーザ発話に対して音声認識を行い、韻律情報とテキスト情報を取得する。次に、得られたユーザ発話のテキスト情報に対して形態素解析・係り受け解析を行い、さらに焦点解析を行う。形態素解析には Juman を、係り受け解析には KNP を、焦点解析には吉野らの手法 [7] を用いる。焦点解析の結果、ユーザ発話の韻律情報などを用いて6種類の各応答を生成する。最後に、韻律情報や過去のシステム自身の応答種類などを考慮して、生成された各応答の中で最も適切なものをシステム応答として出力する。各モジュールの詳細について、次節以降で説明する。

### 2.1 焦点解析

入力ユーザ発話を形態素解析・述語項構造解析したものである。焦点解析として先行研究 [7] の手法を用いる。この手法ではユーザ発話に対して、「各文節の出現位置」、「各単語の品詞」、「各文節が担っている格」、などの特徴量を抽出し、条件付確率場 (以下 CRF) で学習を行い、焦点を解析する。[7] において、対話システムから抽出した計 918 個ユーザ発話にアノテーションを付与

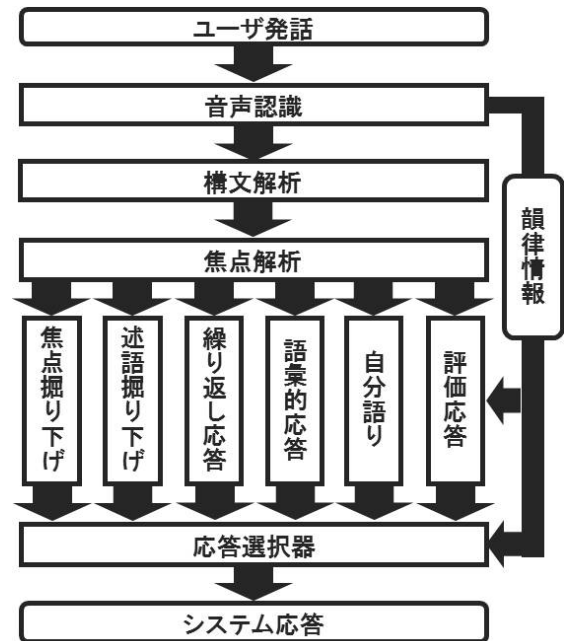


図1: 応答生成システムの処理の流れ

したもの学習データとして検証したところ、ユーザ発話に焦点があるかどうかの判定の精度は 99.9% で、焦点となるフレーズの抽出精度は 78.5% と報告されている。

### 2.2 応答生成

「掘り下げ質問」、「繰り返し応答」、「語彙的応答」の生成は、[5] で提案した我々の手法を用いる。「掘り下げ質問」は N-gram 確率を用いて焦点単語 (または述語) と相性の良い疑問詞を選択し、それらを組み合わせて生成する。「繰り返し応答」は焦点単語をそのまま用いて繰り返す。「語彙的応答」は「そうですか」という応答内容のみを採用している。

#### 2.2.1 評価応答

「評価応答」は先行するユーザ発話に対してシステムが評価を下す応答である。単語の極性辞書を用いて先行するユーザ発話の内容に対してネガティブかポジティブかを判定する。その後、その判定に即した応答を生成する。例えば、ネガティブなら「それは残念ですね」、ポジティブなら「それはいいですね」などである。

#### 2.2.2 自分語り

「自分語り」は先行するユーザ発話に関連した事象に対して、聞き手自身の考えや情報を提示する応答である。以下が例である。U がユーザ、S がシステムの発話である。

U 「僕の出身地は岐阜だよ」  
 S 「私の出身地も岐阜なんです」 ← 自分語り  
 S 「岐阜のどこが好きですか？」 ← 掘り下げ質問

この「自分語り」は対話をより自然にするためだけでなく、直接「掘り下げ質問」を生成できないような先行発

話に対してクッションとなって「掘り下げ質問」を生成できるようにするという役割果たす。「自分語り」は聞き手自身に関することを矛盾なく述べる必要があるため、システムにバックグラウンドとなる知識を持たせる必要がある。このバックグラウンドが実現できれば、ユーザの先行発話と自分のバックグラウンドを照らし合わせて、矛盾せず「自分語り」を行うことができる。ただし、本研究では単純化のため、「自分語り」の生成方法として、単にユーザの発話に同調するという手法を採用する。具体的には、先行発話の主体をユーザからシステムに入れ替えて、適切な助詞を「も」に変換するというものである。

### 2.3 応答選択器

応答生成モジュールによって生成された6つの応答の妥当性を判定し、その内から一つを選択する。この選択器の学習に用いる特徴量は、先行ユーザ発話のテキストと韻律情報、生成された応答のテキスト、過去のシステム応答の種類などである。実際の自然な傾聴対話においては話し手を質問攻めにしたり、逆に語彙的応答のような無難な応答ばかりを返すことは避けるべきであるので、システムの応答の連続性を考慮し、過去のシステム自身の応答を特徴量とする。

## 3. データセット

### 3.1 作成手順

学習や評価に用いるデータセットを作成するために、前述の6種類の応答タグ(以下、聞き手応答タグ)のアノテーション作業を行った。被験者と自律型アンドロイド ERICA との対話収録実験を行った。当該実験で収録した対話の内容は、序盤でタスク遂行を行ったのちに中盤から後半にかけて雑談を行うというものである。聞き手応答タグを以下の手順でアノテーションした。

1. 各発話を LUU(Long Utterance Unit) 単位 [8] で分割
2. 各 LUU 単位に DA(Dialog Act)[9] のタグ付け
3. 各 LUU 単位に DA とは別に Feedback[9] のタグ付け

その後、各 LUU 単位に対して、DA タグと Feedback タグを考慮して聞き手応答のタグのアノテーションを行った。手順を表1に示す。

DA において“other”かつ“feedback”のタグが付いているものは「評価応答」、「繰り返し応答」、「語彙的応答」のいずれかであるが、その内のどれかは人手で判断する。対話データ内には聞き手応答としてはふさわしくない聞き手の発話(表1では「上記以外」としている)も存在し、それらにも DA のタグ付けは行われるため、それらの発話に対しては「その他」(タグ名は“O”)を付けた。また、隣接ペア [6] の概念における第何部分かに関してもアノテーションを行った。

### 3.2 アノテーション結果

聞き手応答のタグのアノテーションは先述の収録データの内の13対話に含まれる1136発話(ERICA)に対して行った。各応答の種類を分布を表2に示す。“Q”と“F”が多い。ただし、“E”や“R”も生成される回数は少ないとはいえ、一定数は存在することが判明したのでこれらの生成は必要である。また、本研究で特に重要と考えている“S”の発話も“Q”や“F”と同等の量が生成されており、傾聴対話における「自分語り」の重要性を確認することができた。表3に ERICA の各種類の応答の例を示す。表3中のAは被験者を、Bは ERICA を指している。

表 1: 聞き手応答のアノテーション規則

Dialog Act	聞き手応答
“inform”	S (自分語り)
“question”	Q (掘り下げ質問)
“other” / “feedback”	E (評価応答) R (繰り返し応答) F (語彙的応答)
“other” / not “feedback”	O (その他)
上記以外	

表 2: 聞き手応答タグのアノテーション結果

S	Q	E	R	F	O
214	209	15	39	163	496

表 3: 聞き手応答の具体例

タグ	発話例
S	A「就職の方を主に考えてます。」→ B「私も今年の四月から社会人してます。」
Q	A「梅田劇場とか、見に行きますね。」→ B「最近は何か見られたんですか。」
E	A「友達の劇とか見に行ったりします。」→ B「あー、素敵ですね。」
R	A「私、三回生です。」→ B「あ、三回生」

## 4. おわりに

本稿では、自然な傾聴対話システムを実現するために、傾聴に必要なと考えられる6種類の応答の生成・選択の手法を述べ、また、モデルの学習と評価を行うためのデータセットの作成の手順、結果を示した。本研究のキーアイデアは、傾聴対話の聞き手があえて「自分語り」をする点である。対話において聞き手が「自分語り」する場合がかなりあることを確認した。しかし、本手法では単に相手に同調するのみで対話中に矛盾を生じさせてしまう可能性がある。それを解決するために今後はシステム側のバックグラウンドも実現する。さらに、予測モデルの実装に加えて、応答選択器での生成された6種類の応答の適切性を図る。

謝辞 本研究は、JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクトの支援を受けて実施された。

## 参考文献

- [1] 河原達也. 音声対話システムの進化と淘汰: 歴史と最近の技術動向. 人工知能学会誌, Vol. 28, pp. 45-51, 2013.
- [2] 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. 言語・音声理解と対話処理研究会, Vol. 58, pp. 61-66, 2010.
- [3] 上里美樹, 吉野幸一郎, 高梨克也, 河原達也. 傾聴対話における相槌の韻律的特徴の同調傾向の分析. 言語・音声理解と対話処理研究会, Vol. 70, pp. 7-13, 2014.
- [4] 山口貴史, 井上昂治, 吉野幸一郎, 高梨克也, Nigel Ward, 河原達也. 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成. 人工知能学会研究会資料, Vol. 76, pp. 49-54, 02 2016.
- [5] 石田真也, 井上昂治, 中村静, 高梨克也, 河原達也. 傾聴対話システムのための発話を促す聞き手応答の生成. 人工知能学会研究会資料, Vol. 77, No. 01, pp. 1-6, 08 2016.
- [6] 佐藤晴. 対話への情報付与. 小磯花絵(編), 日本語コーパス3: 話し言葉コーパス-設計と構築-, pp. 101-130. 朝倉書店, 東京, 2015.
- [7] 吉野幸一郎, 河原達也. Conversational system for information navigation based on pomdp with user focus tracking. *Computer Speech Language*, Vol. 34, pp. 275-291, 2015.
- [8] Y. Den, H. Koiso, T. Maruyama, K. Maekawa, K. Takanashi, M. Enomoto, and N. Yoshida. Two-level annotation of utterance-units in japanese dialogs: An empirically emerged scheme. *Proceedings of the 7th International conference on Language Resources and Evaluation*, pp. 1483-1486, 05 2010.
- [9] H. Bunt, J. Alexandersson, J. Carletta, J.W. Choe, A. Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary, C. Soria, and D. Traum. Towards an iso standard for dialogue act annotation. In *Proceedings of the 7th International conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, 05 2010.