

多様な形態の相槌をうつ音声対話システムのための傾聴対話の分析

山口 貴史[†]

吉野 幸一郎[‡]

高梨 克也[‡]

河原 達也^{†‡}

[†] 京都大学 情報学研究科

[‡] 京都大学 学術情報メディアセンター

1. はじめに

近年、タスク指向型対話システムに加えて、雑談型対話システムも検討されるようになってきている [1]。雑談型対話システムの機能の一つにユーザの話を聞く傾聴がある。傾聴とは話し手の話に共感を示しつつ、話し手がより多く話せるように手助けをして話を聴くことである [2]。音声対話システムが傾聴を行うことにより入院患者や高齢者の話し相手となること [3] や、ユーザの話をしたい、話を聞いてもらいたいといった欲求を満たすこと [4] が期待されている。

傾聴の行為として、相手の発話内容を繰り返す、質問を行う、相槌をうつなどの行為が挙げられる。先行研究における傾聴対話システム [5, 6, 7] はこれらの行為の実現をめざしていた。しかしこれらのシステムでは、実際の傾聴において重要となる対話の文脈に応じた相槌の形態、韻律が考慮されていない。

相槌の生成に関する研究では、相槌の生成タイミングの予測を行うものが数多くされている。これらの研究では、先行発話の韻律情報 [8, 9] や、言語情報を用いた特徴 [10, 11, 12] によってタイミングの予測を行う。しかし、あらかじめ決められた形態の相槌を用いており、多様な形態の相槌をうつような枠組みは検討されていない。これに対して本研究では、対話の文脈に応じて適切な相槌をうつことができる傾聴対話システムの実現を目標として、相槌の形態パターンと先行発話の関係について分析を行う。

2. 傾聴対話における相槌の分析

本研究では、上里ら [13] の研究で収録された傾聴対話を分析する。対話のテーマは「日常の簡単な悩みや困りごと」であり、各セッションの対話は、話し役1名、聞き役1名で行っている。聞き役はスクールカウンセラー2名、話し役は大学生8名で、合計8対話が収録されている。対話時間は20～30分である。

相槌については様々な定義がなされているが、本研究ではメイナード [14] の「話し手が発話権を行使している間に聞き手が送る短い表現」を採用し、相槌の形態ごとの分類には吉田ら [15] による、「うん」や「ふんふん」といった促しや受容を表す応答系感動詞と、「あー」や「はー」といった興味や関心・共感を表す感情表出系感動詞の2つを用いた。また分析対象とする相槌には、先行発話末 500ms 以内でオーバーラップするものも含めた。

文脈に応じた相槌は、区切りに多く出現する。そこで発話の区切りに出現する相槌を対象として、そこで得られる情報と相槌の関係を分析する。ここでは、表層的な区切りである節境界とパーザによる構造的区切りの2種類について検討した。なお、節境界の認定には高梨らの基準 [16] を用いた。節境界には切れ目の強さの順に絶対

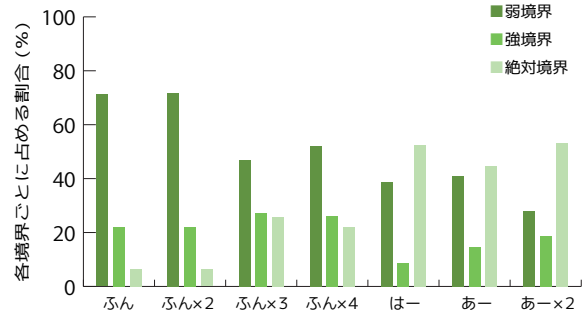


図 1: 節境界の種類と相槌の形態ごとの関係

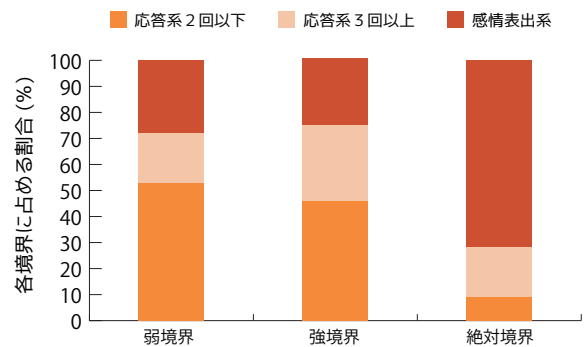


図 2: 節境界の種類と相槌の種類ごとの関係

境界、強境界、弱境界の3種類があり、絶対境界と強境界で節単位（文に相当）を構成する。

3. 表層的区切りと相槌の形態

まず、先行発話の節区切りの種類によって、うたれる相槌の形態が異なるという仮説をたてた。この仮説を検証するために、先行発話の表層的区切りの観点から分析を行う。

表層的区切りとして、節境界の種類ごとにうたれた相槌を分類した。節境界の種類と代表的な相槌の形態ごとの割合を図1に示す。「ふん」「ふん×2」といった2回以下の応答系相槌は、弱境界で多く出現することがわかる。また、感情表出系の相槌は絶対境界で多くうたれることがわかる。またこの図から、形態ごとの各境界でうたれる割合の傾向によって、2回以下の応答系、3回以上の応答系、感情表出系の3つのグループに分けることができる。

次に、この3つにグループ化した場合の境界の種類ごとの出現割合を図2に示す。この結果から、弱境界では2回以下の応答系相槌、絶対境界では感情表出系相槌が多くうたれることが分かる。一方、強境界では、3種の相槌がほぼ同程度の割合でうたれているが、3回以上の応答系相槌に着目すると強境界で多くうたれている。

Analysis of morphological patterns of backchannels toward flexible attentive listening agent.

[†]Takashi YAMAGUCHI, [‡]Koichiro YOSHINO, [‡]Katsuya TAKANASHI, ^{†‡}Tatsuya KAWAHARA

[†]Graduate School of Informatics, Kyoto University

[‡]Academic Center for Computing and Media Studies, Kyoto University

表 1: 構文木の深さ

	節全体		弱境界		強境界		絶対境界	
	1回	2回以上	1回	2回以上	1回	2回以上	1回	2回以上
繰り返し	1回	2回以上	1回	2回以上	1回	2回以上	1回	2回以上
個数	205	458	140	240	48	100	17	118
平均	2.18	2.56	2.22	2.60	2.29	2.62	1.65	2.51
分散	1.41	1.82	1.40	1.56	1.45	2.30	1.24	2.01
有意差	○		○				○	

表層的区切り + 構造的区切り → 相槌の形態

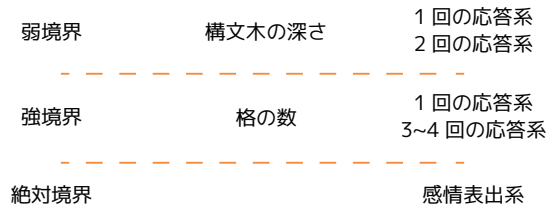


図 3: 分析結果に基づく相槌の予測ルール

表 2: 節終わりにかかる格の数

	節全体		弱境界		強境界		絶対境界	
	1回	2回以上	1回	2回以上	1回	2回以上	1回	2回以上
繰り返し	1回	2回以上	1回	2回以上	1回	2回以上	1回	2回以上
個数	205	458	140	240	48	100	17	118
平均	1.89	1.92	2.01	2.01	1.58	1.96	1.71	1.69
分散	0.92	1.05	1.08	1.24	0.46	0.97	0.60	0.66
有意差					○			

に示す。境界の種類と木の深さ、格の数を組み合わせることにより、次にうつ相槌の形態を予測することができる。

本研究では、多様な形態の相槌をうつ傾聴対話システムの実現を目標として、話し手の発話の区切りにうたれる相槌の形態の種類と先行発話から得られる情報の関係について分析を行った。この結果、境界の種類の違いによる相槌の形態の種類傾向と、先行発話の構造と相槌の繰り返しの有無の関係が明らかになった。今後はこれらの結果に基づいて、相槌の形態予測を行い、傾聴対話システムに応用する予定である。

4. 構造的区切りと相槌の繰り返しの有無

次に、相槌の繰り返しの有無と先行発話の構造的な複雑さに関係があるという仮説をたてた。この仮説を検証するために、先行発話の構造的区切りの観点から分析を行う。

構造的区切りとして、節単位で区切られた先行発話の節末を根とする構文木の深さと、節末にかかる格の数と相槌の繰り返しの有無の関連を調べる。繰り返しの有無に焦点を当てるため、相槌の形態は応答系の「うん」「ふん」を同種として扱い、繰り返しの有るものとして2回以上繰り返されているものをグループ化した。分析は、節境界の種類ごとに分類した場合でも行い、相槌の繰り返しの有無と節境界の種類の間を調べる。

先行発話の構文木の深さに関する結果を表1、節終わりにかかる格の数に関する結果を表2に示す。繰り返しの有る場合とない場合でt検定(有意水準5%)を用いて有意差を確認したところ、構文木の深さでは節全体、弱境界、絶対境界で有意差が示された。一方、格の数では強境界のみ有意差が示される結果となった。このことから構文木の深さと節終わりにかかる格の数を参照することで次にうつべき相槌の繰り返しの有無を予測できると考えられる。

5. まとめ

表層的区切りと構造的区切りの分析結果を組み合わせることで、相槌の形態を予測するための枠組みを検討する。表層的区切りでは、弱境界に2回以下の応答系相槌、絶対境界に感情表出系の相槌がうたれやすいことがわかった。また、構造的区切りでは、繰り返しの有無を決定するために、構文木の深さと、強境界においては節終わりにかかる格の数が有効であることがわかった。これらを組み合わせ、相槌の形態を決定するルールを図3

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, No. 1, pp. 45-51, 2013.
- [2] 楡木満生. 積極的傾聴法. 医学教育, Vol. 20, No. 5, pp. 341-346, 1989.
- [3] 山本大介, 小林優佳, 横山祥恵, 土井美和子. 高齢者対話インタフェース: 『話し相手』となって、お年寄りの生活を豊かに. 電子情報通信学会技術研究報告. HCS, Vol. 109, No. 224, pp. 47-51, 2009.
- [4] 日黒豊美, 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.
- [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. 言語・音声理解と対話処理研究会, Vol. 58, pp. 61-66, 2010.
- [6] 横山祥恵, 山本大介, 小林優佳, 土井美和子. 高齢者向け対話インタフェース-雑談継続を目的とした話題提示・傾聴の切替対話法-. 情報処理学会研究報告. SLP, 音声言語情報処理, Vol. 2010, No. 4, pp. 1-6, 2010.
- [7] D.Ozkan and L.-P.Morency. Modeling wisdom of crowds using latent mixture of discriminative experts. In proc.ACL/HLT, 2011.
- [8] 岡登洋平, 加藤佳司, 山本幹雄, 板橋秀一. 韻律情報を用いた相槌の挿入. 情報処理学会論文誌, Vol. 40, No. 2, pp. 469-478, 1999.
- [9] N.Ward and T.Wataru. Prosodic features which cue backchannel responses in English and Japanese. J.Pragmatics, Vol. 32, No. 8, pp. 1177-1207, 2000.
- [10] 西村良太, 中川聖一. 応答タイミングを考慮した音声対話システムとその評価. 情報処理学会研究報告. SLP, 音声言語情報処理, Vol. 2009, No. 22, pp. 1-6, 2009.
- [11] 竹内真士, 北岡教英, 中川聖一. 韻律・表層的言語情報を発話タイミング制御に用いた雑談対話システム. 情報処理学会研究報告. SLP, 音声言語情報処理, Vol. 2004, No. 15, pp. 87-92, 2004.
- [12] 大野誠寛, 神谷優貴, 松原茂樹. タグ付けの安定性を備えた音声対話コーパスに基づくあいづち生成タイミングの検出. 情報処理学会研究報告, 2010-SLP-84-4, 2010.
- [13] 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.
- [14] 泉子・K・メイナード. 会話分析. ころしお出版, 1993.
- [15] 吉田奈央, 高梨克也, 伝康晴. 対話におけるあいづち表現の認定とその問題点について. 言語処理学会第15回年次大会発表論文集, pp. 430-433, 2009.
- [16] 高梨克也, 内元清貴, 丸山岳彦. 『日本語話し言葉コーパス』における節単位認定. 『日本語話し言葉コーパス』同梱マニュアル, 2004.