

傾聴対話における相槌形態と 先行発話の統語構造の関心の分析

Analysis on relationship of patterns of backchannels and syntactic structure of preceding utterances in attentive listening

山口 貴史^{1*} 井上 昂治¹ 吉野 幸一郎² 高梨 克也² 河原 達也^{1,2}
Takashi Yamaguchi¹ Koji Inoue¹ Koichiro Yoshino² Katsuya Takanashi² Tatsuya Kawahara^{1,2}

¹ 京都大学 情報学研究科

¹ Graduate School of Informatics, Kyoto University

² 京都大学 学術情報メディアセンター

² Academic Center for Computing and Media Studies, Kyoto University

Abstract: We investigate the relationship between backchannels and the syntactic structure in the delimited preceding utterances in attentive listening such as counseling. First, we find out the relationship between particular patterns of backchannels and the category of the clause boundary. Next, we analyze the syntactic structure by using the depth of the syntax tree and the number of cases related to the end of utterance. It is shown that there is a relationship between particular patterns of backchannels and the complexity of the preceding utterances. The results suggest that we can choose different kinds of backchannels depending on the preceding utterance.

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。傾聴を行う際の対話行為として、話し手の発話に対して、「相槌をうつ」「質問をする」「共感を示す」などが挙げられる。質問や共感をする際には相手の発話を十分に認識・理解する必要があるが、技術的に容易ではない。一方、相槌は先行発話の韻律や節末のパターンに基づいて生成できる可能性がある。そこで本研究では、対話の文脈に応じて多様な相槌を生成できることをめざして、相槌の形態と先行発話の特徴との関係の分析を行う。

2 関連研究

2.1 傾聴対話システム

傾聴を行う音声対話システムの研究として、下岡ら [5] は、傾聴時に重要とされる「繰り返し応答」、「問い返し応答」、「共感応答」、「相槌応答」の4つに着目し、これらの応答行為を行うことを検討している。ユーザへの応答行為の選択は、ユーザの発話に対する音声認識結果の信頼度を用いて行っている。信頼度が高いと判定されれば、「繰り返し応答」、「問い返し応答」、「共感応答」を行い、低い場合には「相槌応答」を行う。相槌応答としては、「なるほど」「そうですか」のようなあらかじめ定義された相槌を生成している。横山ら [6, 7] は、ユーザの関心度により、話題提示モードと傾聴モードを切り替える傾聴システムを提案している。話題提示モードではシステムが話し手となって話題を提示し、傾聴モードではシステムはユーザの話の聞き手となる。傾聴時には「うんうん」などの相槌を生成している。Ozkan ら [8] はユーザの発話に応じた質問、相槌、ジェスチャーなどの行為を行うエージェントによる傾聴システムを構築している。

これらの傾聴システムでは相槌をうつ基本的な行為が実装されているが、人間同士の対話と比較して、その形態や韻律のバリエーションは乏しい。

*連絡先：京都大学学術情報メディアセンター 河原研究室
京都市左京区吉田本町
E-mail: takashi@ar.media.kyoto-u.ac.jp

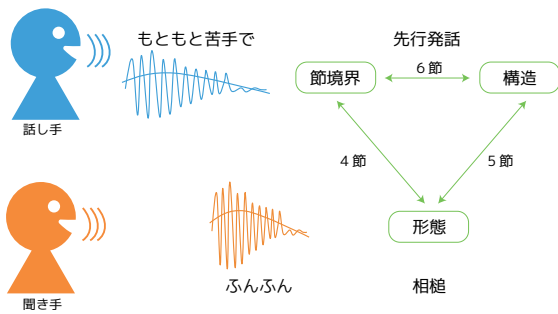


図 1: 本研究における分析の概要

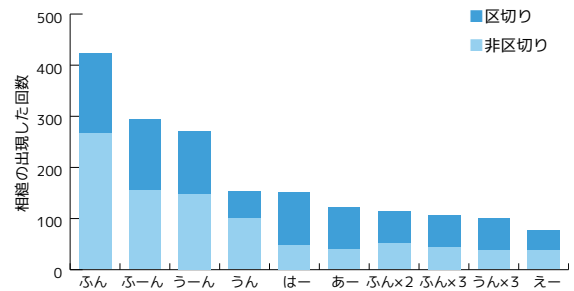


図 2: 傾聴対話で出現する相槌（上位 10 種）

2.2 相槌生成の予測に関する研究

相槌の生成に関する研究としては、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。韻律情報を用いた研究では、先行発話の韻律パターン [9] や、先行発話の F0 が下がった区間の継続時間 [10] を用いて相槌生成タイミングを決定している。言語情報を用いた研究では、発話末の品詞情報や発話長を手がかりとしたルール [11, 12, 13, 14] や、文境界、節境界といった統語的特徴による統計的手法を用いたタイミング予測 [15, 16] がなされている。

しかし、これらの研究でもあらかじめ決められた形態の相槌が用いられており、多様な形態の相槌をうつ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形態の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統語構造の関係について分析を行う。

3 傾聴対話における相槌の分類

本研究では、上里ら [17] の研究で収録された相談対話を分析する。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 [18] 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」を採用した。また、相槌の形態ごとの分類には吉田ら [19] による、「うん」や「ふんぶん」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分類を参照した。また分析対象とする相槌には、先行発話末に 500 ms 以内でオーバーラップして開始されているものも含めた。

まず、発話の区切りに出現する相槌の形態を調べ、その他で出現する相槌の形態と比較を行った。ここでは発話の区切りの定義として、『日本語話し言葉コーパス』(CSJ) で定義された節境界 [20, 21] を区別せずに用いる。図 2 に傾聴対話でみられた相槌の形態別の出現回数を示す。ここでの「 $\times n$ 」は n 回の繰り返しを意味する。たとえば「ふん $\times 2$ 」は「ふんぶん」である。

傾聴対話で出現する相槌では、「ふん」、「うん」や、それらの引き伸ばし系である「ふーん」、繰り返し系である「ふんぶん」などの応答系相槌が多くみられ、感情表出系では「あー」や「はー」が多くみられた。

このうち、区切りに出現したものの割合をみると、「あー」や「はー」、繰り返し応答系は区切りに多くうたれており、「うん」や「ふん」は区切りではあまりうたれていない。感情表出系や繰り返し応答系は区切りにうたれやすく、「うん」「ふん」は、どこでもうてるのがわかる。

分析結果の傾向から、応答系の「うん」「ふん」を同種として扱い、その繰り返し回数によって、応答系 1 回の相槌、応答系 2 回の相槌、応答系 3 回以上の相槌とカテゴリ化することとした。ここでは、引き伸ばし系である「うーん」と「ふーん」は、機能が異なるものも含まれている可能性があるため除いている。また、「あー」、「はー」、「へー」は応答系と振る舞いが違うため [22]、感情表出系の相槌として一つのカテゴリとした。つまり、応答系 1 回の相槌、応答系 2 回の相槌、応答系 3 回以上の相槌、感情表出系の相槌の 4 つにカテゴリ化した。

文脈に応じた相槌は、先行する発話の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、そこで得られる特徴と相槌の関係を分析する。ここでは、形態素列の表層的な特徴によって得られる節境界 (4 節) と係り受け解析によって得られる構文的な構造的区切り (5 節) の 2 種類について、それぞれの区切りで得られる情報と後続する相槌との関係を分析した。また、これら 2 種類の特徴の関係についても分析を行った (6 節)。分析の概要図を図 1 に示す。

表 1: 節境界ラベルの分類

境界の種類	節境界ラベル
絶対境界	文末・文末候補・と文末 など
強境界	並列節ガ・並列節ケド 並列節ケレド・並列節シ など
弱境界	理由節カラ・テ節 条件節ナラ・トカ節・並列節デ など

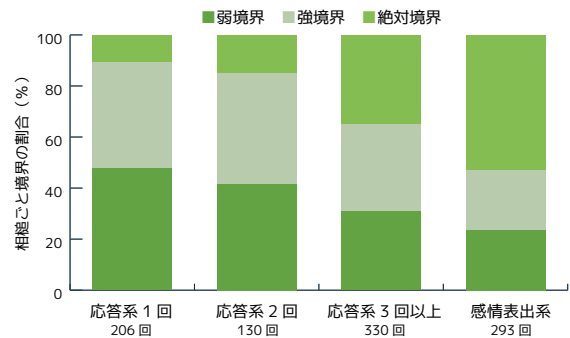


図 3: 相槌の形態の各カテゴリごとの節境界の割合

4 節境界の種類と相槌の形態

まず、先行発話の節境界の種類によって、うたれる相槌の形態が異なるという仮説をたてた。前節で4つにカテゴリ化した相槌の形態と、その相槌がうたれた節境界の種類との関係を調べた。

節境界には、その直後の構造的な切れ目の大きさという観点から切れ目の強さの順に絶対境界、強境界、弱境界の3種類がある [20, 21]。表 1 に各節境界ラベルの例を示す。絶対境界は、形式上明示的な文末表現に相当し、強境界は、文末表現ではないが発話の大きな切れ目に相当する。弱境界は、通常は発話の切れ目になることは少ない節境界である。

図 3 に相槌の形態の各カテゴリごとの節境界の割合を示す。ここでの割合は、相槌が出現した境界のうち各節境界の割合を示している。下段の回数は各相槌カテゴリの生起数である。感情表出系の相槌は絶対境界で多くうたれており、応答系 3 回以上の相槌は強境界もしくは絶対境界で多くうたれている。一方、応答系 2 回の相槌は強境界で、応答系 1 回の相槌は弱境界で比較的多くうたれている。このことから、感情表出系の相槌や繰り返し回数大きい応答系の相槌のような強い相槌は、大きい境界で出現しやすいことがわかる。

次に、先行発話の節境界から得られる情報をもとに後続する相槌を予測可能かを調べるため、各節境界ごとに、出現した相槌カテゴリの割合を求めた。図 4 に各節境界ごとにうたれる相槌のカテゴリの割合を示す。下段の回数は各節境界の総数である。絶対境界では感情表出系の相槌が最も多くうたれ、次に応答系 3 回以上の相槌が多くうたれている。一方、弱境界と強境界では応答系 1 回の相槌と応答系 2 回の相槌が同様の割合が多いが、節境界の種類だけでは区別することができない。

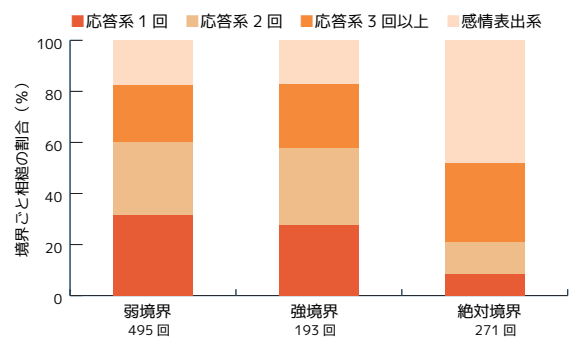


図 4: 節境界ごとの相槌の形態のカテゴリの割合

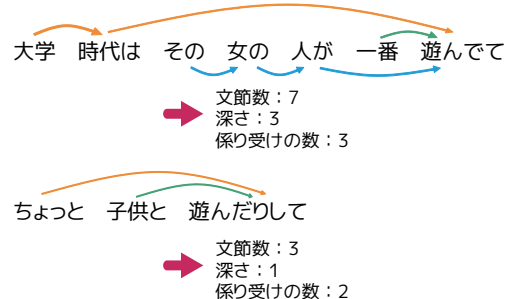


図 5: 先行発話の節末における構文木の深さと係り受けの数の例

5 先行発話の構文構造と相槌の形態

次に、先行発話の構文構造の観点から分析を行う。先行発話の構文構造の複雑さの指標として、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、節末文節にかかる係り受けの数を用いる。この指標の例を図 5 に示す。これらに加えて、直前の節境界には含まれた文節の数と相槌の形態の関係を調べた。

表 2 に文節数との関係、表 3 に構文木の深さとの関係、表 4 に係り受けの数との関係の結果を示す。各カテゴリ間で t 検定 (有意水準 5%) を行った結果、文

表 2: 文節数と相槌の種類との関係

	応答系 1回	有意 差	応答系 2回	有意 差	応答系 3回以上	有意 差	感情 表出系
出現数	206		130		331		209
文節数 の平均	4.73	○	5.52	×	5.42	×	5.15
文節数 の分散	4.94		6.42		6.65		6.89

表 3: 構文木の深さと相槌の種類との関係

	応答系 1回	有意 差	応答系 2回	有意 差	応答系 3回以上	有意 差	感情 表出系
出現数	206		130		331		209
木の深さ の平均	2.18	○	2.57	×	2.56	×	2.54
木の深さ の分散	1.41		1.70		1.87		1.84

表 4: 節末文節にかかる係り受けの数と相槌の種類との関係

	応答系 1回	有意 差	応答系 2回	有意 差	応答系 3回以上	有意 差	感情 表出系
出現数	206		130		331		209
係り受けの 数の平均	1.88	×	2.00	×	1.89	×	1.75
係り受けの 数の分散	0.92		1.12		1.02		1.11

節数と構文木の深さについて、応答系 1 回の相槌と 2 回の相槌の間で有意差がみられた。前節の節境界の種類に関する分析では、これらの区別はできなかったが、直前の発話の構文構造によって違いがみられることがわかった。

6 節境界と構文構造

前節までの分析の結果で、節境界の種類によって応答系 1、2 回と応答系 3 回以上および感情表出系の間で、先行発話の構文構造の複雑さによって応答系 1 回と応答系 2 回の間で、それぞれ出現傾向の違いがみられることがわかった。そこで、節境界の種類と構文構造の複雑さの関係を調べた。一般に、境界の大きさが

表 5: 節境界の種類と文節数の関係

	弱境界	強境界	絶対境界
出現数	406	172	236
文節数 の平均	5.31	5.07	5.00
文節数 の分散	5.94	6.52	7.34

表 6: 節境界の種類と構文木の深さの関係

	弱境界	強境界	絶対境界
出現数	406	172	236
木の深さ の平均	2.50	2.52	2.40
木の深さ の分散	1.52	1.96	2.09

表 7: 節境界の種類と係り受けの数の関係

	弱境界	強境界	絶対境界
出現数	406	172	236
係り受けの 数の平均	2.00	1.85	1.63
係り受けの 数の分散	1.22	0.89	0.71

大きくなると、係り受けの係り受け先である述語項が完成され、構文構造が複雑になる。つまり、境界の大きさが大きくなるにつれ、構文木の深さと係り受けの数は大きくなると考えられる。

節境界の種類と文節数の関係を表 5 に、構文木の深さの関係を表 6 に、係り受けの数との関係を表 7 に示す。結果は予想に反し、節境界の種類によって、構文木の深さについては差がほとんどみられず、文節数と係り受けの数は、大きい境界ほど小さくなる傾向がみられた。

しかし、節境界の種類と先行発話の構文構造の関係

に相関がないという結果は、4節で節境界の種類によって感情表出系と応答系3回以上が、5節では応答系1回と応答系2回を、文節数と構文木の深さによって区別できることが独立に寄与することと整合する。

7 まとめ

本研究では、多様な形態の相槌をうつことのできる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言語的情報との関係について分析した。その結果、感情表出系の相槌や応答系3回以上の相槌には節境界の種類の大きさと、また、応答系1回の相槌と応答系2回の相槌は先行発話の文節数および構文木の深さとの関係があることがわかった。このような先行発話の言語情報の違いに応じて相槌の使い分けができる可能性が示された。

今後は、先行発話末の韻律情報や当該の先行発話単位が話し手の発話の中や、対話構造のなかで占めている位置も考慮することによって、今回得られた知見に基づいて傾聴対話システムを構築する予定である。

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] 榎木満生. 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] 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.
- [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・傾聴の切替式対話法-. 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] Y.Kobayashi, D.Yamamoto, T.Koga, S.Yokoyama, and M.Doi. Design targeting voice interface robot capable of active listening. In *proc.HRI*, pp. 161–162, 2010.
- [8] D.Ozkan and L.-P.Morency. Modeling wisdom of crowds using latent mixture of discriminative experts. In *proc.ACL/HLT*, 2011.
- [9] 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.
- [10] 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.
- [11] H.Koiso, Y.Horiuchi, S.Tutiya, A.Ichikawa, and Y.Den. An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and Speech*, Vol. 41, No. 3, pp. 295–321, 1998.
- [12] 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.
- [13] 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.
- [14] N.Kitaoka, M.Takeuchi, R.Nishimura, and S.Nakagawa. Response timing detection using prosodic and linguistic information for human-friendly spoken dialog systems. 人工知能学会論文誌, Vol. 20, pp. 220–228, 2005.
- [15] 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.
- [16] Y.Kamiya, T.Ohno, and S.Matsubara. Coherent back-channel feedback tagging of in-car spoken dialogue corpus. In *proc.SIGDIAL*, pp. 205–208, 2010.
- [17] 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.
- [18] 泉子・K・メイナード. 会話分析. くろしお出版, 1993.

- [19] 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.
- [20] 日本語話し言葉コーパスの構築法. 独立行政法人国立国語研究所, 2006.
- [21] 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.
- [22] 常志強, 高梨克也, 河原達也. ポスター会話におけるあいづちの韻律的特徴に関する印象評定. 言語・音声理解と対話処理研究会, Vol. 56, pp. 31–36, 2009.