

語りの傾聴における共感の深さと応答の種類の関係

伊藤 滉一郎^{†1} 村田 匡輝^{†2} 大野 誠寛^{†3} 松原 茂樹^{†1}

^{†1} 名古屋大学 ^{†2} 豊田工業高等専門学校 ^{†3} 東京電機大学

ito.koichiro@a.mbox.nagoya-u.ac.jp

1 はじめに

語ることは人間の基本的な欲求である。語る行為は、聞き手がいて初めて成立する。社会の個人化が進み、聞き手不在の生活シーンが増加している。人が語れる機会を増やすことは現代社会の重要な課題である。

これに対して、コミュニケーションロボットやスマートスピーカーなどの情報機器が語りを聞く役割を担うことが考えられる。これらが聞き手として認められるには、「語りを傾聴していることを話し手に伝達する機能」を備える必要がある。このための明示的な手段は語りに応答することであり、ジェスチャや発話の表出が有力である。以下では、傾聴を示す目的で語りに応答する発話を傾聴応答と呼ぶ。

傾聴応答には語りに共感を示し、話し手の語る意欲を促す効果がある。傾聴応答の代表例は相槌であり、その生成方法はすでに提案されている [1, 2]。相槌以外にも、感心や評価など様々な種類の傾聴応答が存在し、それらの生成方法に関する研究も行われている [3, 4, 5]。傾聴応答が語りに示す共感の深さ (以下、共感度) は、傾聴応答によって異なる。単に共感が深い応答を生成すればよいわけではなく、語りの内容や状況に適した共感を示す必要がある。

そこで本論文では、語りに示す共感に応じた傾聴応答の生成を目的に、傾聴応答の種類と共感度の関係について考察する。本研究では、傾聴応答の種類を共感度の観点から5段階に分類する。語りに対して傾聴応答を付与したデータ (以下、傾聴応答データ) [6] を用いて、この分類の正当性を検証する。

2 共感度と傾聴応答の種類

傾聴応答には語りに共感を示し、話し手の語る意欲を促す効果がある。傾聴応答には相槌を始め、感心、評価など様々な種類が含まれる。図1に語りと傾聴応答の例を示す。()内は傾聴応答の種類を示す。傾聴応答は、その種類によってそれが語りに示す共感度が異なると考えられる。例えば、以下の語りと、それに対する2つの傾聴応答

[語り 1] 書道も好きで総理大臣賞もいただいたりして

語り	傾聴応答
イタリア旅行をしたことが一番楽しかったです	はい(相槌) はい—そうですか—(感心) 素敵ですね—(評価) ふ—ん(感心)
もう二度と行かないかなと 思いながら行ってきましたけど	イタリア旅行(繰り返し応答) いえいえそんな—(不同意) あ—そうですか—(相槌)

図 1: 語りと傾聴応答の例

[応答 1a (相槌)] えー

[応答 1b (評価)] すごいですね

について、応答 1a よりも 1b の方が共感度が高いと言える。語りの内容や状況に適した共感度を示すことは、傾聴の効果をも高めることにつながるため、傾聴応答の種類と共感度の関係性を整理することは重要である。

3 共感度に基づく傾聴応答の分類

本章では、語りに示す共感に応じた傾聴応答の生成を目的に、共感度によって傾聴応答の種類を分類する。本研究では、傾聴応答データ [6] における 16 種類の傾聴応答を分類対象とする。傾聴応答の種類は、役割の観点から文献 [7] を参考に定義されている。表 1 にその種類と役割を示す。

傾聴応答の種類を、表 1 の応答の役割を参考に、共感度の観点から 5 段階に分類した。分類結果を図 2 に示す。Level は共感度の高さに対応し、Level 1 の相槌の共感度が最低で、Level 5 の意見が最高である。以下に、語りに対する相槌と意見の例を示す。

[語り 2] 今 1 番自分で思うことは健康でいること

[応答 2a (相槌)] はい

[応答 2b (意見)] それが 1 番です

4 共感度による分類の正当性の評価

本章では、3 の分類の正当性を傾聴応答データ [6] を用いて定量的に評価する。傾聴応答データでは、1

表 1: 傾聴応答の種類と役割

種類	役割
相槌	聞き取りに成功したことを示す
感心	相手の発話内容に対して、感心、驚き、気づき等の態度を示す
評価	相手の発話内容が示す事態に対する態度を示す
同意	相手の発話内容に対する同意の態度を示す
不同意	相手の発話内容が同意できないものであることを示す
繰り返し応答	相手の発話内容を理解したことを示し、相手に安心感を与える
言い換え	相手の発話内容を理解して共有しようとする態度を示す
納得	相手の発話内容が納得できることを示す
驚き	相手の発話内容に対して、純粋に驚きのみを表す
驚きといぶかり	相手の発話内容に対して、驚くとともに怪しむ気持ちを表す
意見	自分の個人的な体験や意見、感情を表す
補完	相手の発話を熱心に聞いていることを示す
あいさつ	相手の存在の承認と好意的に関わろうとする意志を示す
想起	相手の発話内容から、かつての記憶が呼び起こされたことを示す
考え始め	相手の発話内容について、思考を始めたことを表す
考えている最中	相手の発話内容について、なにかを考えている最中だということを表す

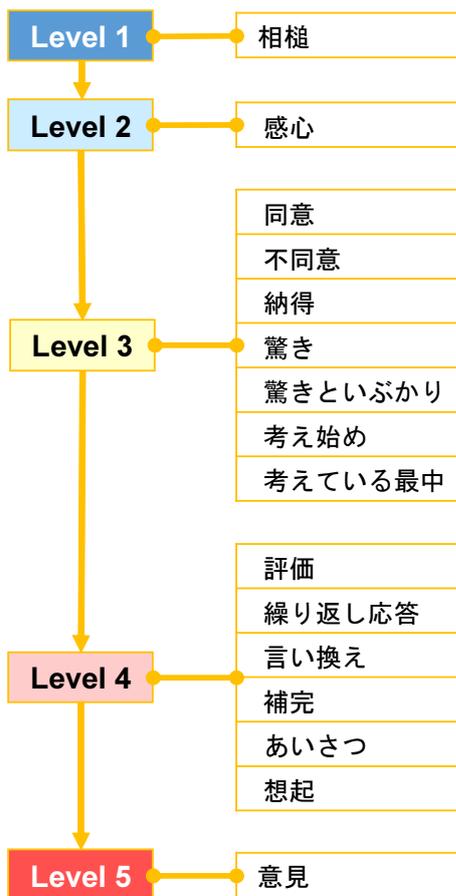


図 2: 傾聴応答の種類と共感度

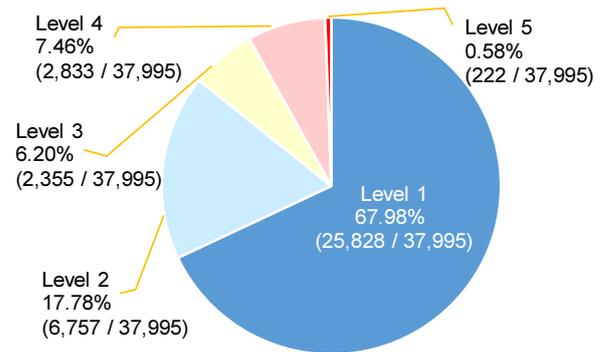


図 3: 評価データにおける傾聴応答の内訳

つの語りに対して 5 人の聞き手がそれぞれ語りを聞きながら、任意のタイミングで発話された傾聴応答を収集している。語りには高齢者のナラティブコーパス JELiCo¹を用いている。

評価に用いた傾聴応答データ (以下、評価データ) には、16 種類の傾聴応答が合計で 37,995 個存在する。図 3 に評価データでの傾聴応答の Level ごとの内訳を示す。Level 1 の応答で全体の 65% 以上を占めている。

4.1 応答の具体性による評価

本節では、応答の具体性を用いて、分類の正当性を評価する。応答の具体性はその種類によって異なる。例えば、以下の語りとそれに対する 2 つの傾聴応答

¹<https://www.gsk.or.jp/catalog/gsk2018-a>

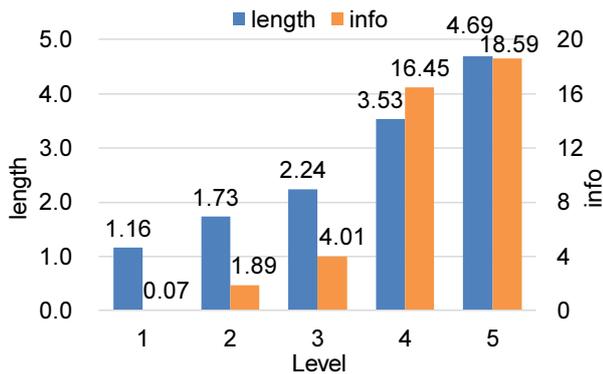


図 4: 傾聴応答の共感度と具体性の関係

[語り 3] イタリア旅行をしたことが1番楽しかった
 [応答 3a (感心)] あー
 [応答 3b (繰り返し応答)] イタリア旅行

について、応答 3b の具体性は 3a よりも高く、共感度も 3b の方が高いと言える。具体性の高い応答を発話するには、語りの内容をより深く考慮する必要があり、語りへの共感度も高くなる。一方、応答 3a のような具体性の低い応答は、語りの内容を深く考慮せずとも発話可能であり、共感度は低くなる。

応答の具体性として、以下の 2 つの特徴量を用いた。

1. *length*: 応答を構成する形態素数
2. *info*: 応答の情報量

info は、応答を構成する内容語の情報量の総和とする。応答 r の情報量 $info(r)$ の定義を以下に示す。

$$info(r) = \sum_{w \in (r \cap W_C)} \log_2 \frac{\sum_{w' \in (CSJ \cap W_C)} F(w')}{F(w) + 1} \quad (1)$$

W_C は内容語の集合、 w は応答 r に含まれる内容語、 $F(w)$ は『日本語話し言葉コーパス』短単位語彙表 ver.2018.3.16²における形態素 w の出現頻度、 $\sum_{w' \in (CSJ \cap W_C)} F(w')$ は同様の語彙表における内容語の延べ数 (=3,325,907) である。

評価データの全応答について上記の 2 つの特徴量を算出し、傾聴応答の種類ごとの平均値を用いて、Level ごとの特徴量の平均値を算出した。図 4 にその結果を示す。Level が大きくなると、いずれの特徴量の平均値も高くなる。これは、3 章で行った傾聴応答の分類の正当性を支持していると言える。また、Level 3 から 4 にかけて *info* の値が大きく増加している。このことは、Level 3 と 4 の間で共感度が大きく異なることを示唆している。

²https://pj.ninjal.ac.jp/corpus_center/csaj/chunagon.html

4.2 応答の汎用性による評価

本節では、応答の汎用性の観点から、分類の正当性を評価する。例えば、繰り返し応答は繰り返しの値する語句が語りになければ発話できず、評価応答は評価に値するイベントが語りになければ発話できない。そのような語句やイベントは、語りにおいて頻繁に出現するものではなく、繰り返し応答や評価応答が発話可能なタイミングは限られるため、いずれも汎用性が低い応答と言える。これらの応答を適切なタイミングで発話するには語りの内容を深く理解する必要があり、共感度が高くなる傾向にあると考えられる。

一方、感心の応答の場合、感心に値するイベントは繰り返しや評価に値する語句やイベントよりも語りに出現しやすい。さらに、相槌に関しては語りの内容によらず発話することもできる。したがって、感心や相槌が発話可能なタイミングは多数存在し、これらは汎用性が高い応答と言える。これらの応答の発話に必ずしも語りの内容を深く理解する必要はなく、共感度が低い傾向にあると考えられる。以上から、応答の汎用性と共感度はトレードオフの関係にあると言える。

本節では、3 章で行った傾聴応答の分類を応答の汎用性によって評価する。本研究では、応答の共起で応答の汎用性を定義する。ここで応答が共起するとは、同じタイミングで応答が発話されることを意味する。汎用性の高い応答は多くの場面で発話されているため、多くの応答と共起する。一方で、汎用性の低い応答は発話される場面が限られるため、共起する応答も限られる。つまり、応答 r に関して、 r と共起する応答の多様性を r の汎用性として考えることができる。そこで本研究では、応答 r の汎用性を r と共起する応答のエントロピー $ent(r)$ により定義した。 $ent(r)$ の具体的な定義を以下に示す。

$$ent(r) = - \sum_u \frac{F_r(u)}{\sum_{u'} F_r(u')} \log_2 \frac{F_r(u)}{\sum_{u'} F_r(u')} \quad (2)$$

$F_r(u)$ は r の他の応答 u ($u \neq r$) との共起頻度、つまり r と u が同じタイミングで発話された回数である。

本節では 3 章で行われた傾聴応答の分類を、 ent を用いて評価する。 ent の計算には応答の共起情報が必要である。しかし、傾聴応答データの応答は、聞き手 5 人がそれぞれ任意のタイミングで応答することで収集されており、応答の共起を考えることが難しい。本研究では、応答が語りの言語的境界あるいは音韻的境界の直後で発話されやすいことを踏まえて、応答の発話開始時刻をこれらの境界のうち、最も近い境界に対応付けた。言語的境界、音韻的境界として、それぞれ以下を用いた。

1. 節境界 (言語的境界)

語り	傾聴応答					「はい」と共起した応答
	listener 1	listener 2	listener 3	listener 4	listener 5	
趣味といっても	はい	はい	うん	はい		うん
今続けている		うん	ええ			
ことではないけれど	ええ	はい	はい	ええ	はい	ええ
着付けをしたり	はい	あ / はいはい	はい	着付け	あつ	あ / はいはい / 着付け / あつ
編み物をしたり	へー	はい	うーん / はい	ええ	へえ	へー / うーん / ええ / へえ

(1) 「はい」と傾聴応答 u ($u \neq$ 「はい」) の共起頻度 $F_{\text{はい}}(u)$ を計算

$$F_{\text{はい}}(u) = 2(u = \text{ええ}), 1(u \in \{\text{うん}, \text{あ}, \text{はいはい}, \text{着付け}, \text{あつ}, \text{へー}, \text{うーん}, \text{へえ}\}), 0(\text{otherwise})$$

(2) $F_{\text{はい}}(u)$ を用いて $ent(\text{はい})$ を計算 $ent(\text{はい}) = -\frac{2}{10} \log_2 \frac{2}{10} - 8(\frac{1}{10} \log_2 \frac{1}{10}) = 3.12$

図 5: 傾聴応答の汎用性の計算例

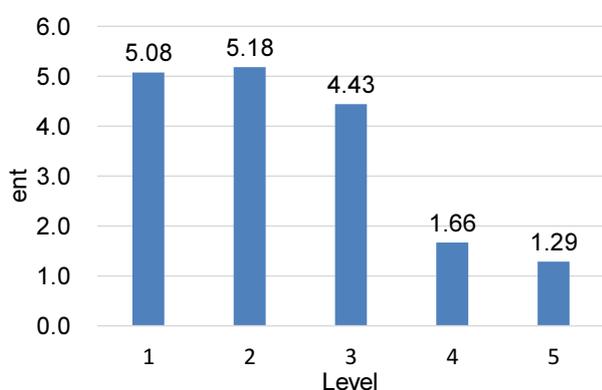


図 6: 傾聴応答の共感度と汎用性の関係

2. 200 ミリ秒以上の無音区間 (音韻的境界)

節境界の検出には、節境界解析ツール CBAP[8] を用いた。以下では、1. または 2. の境界を応答タイミングと呼ぶ。評価データには 10,469 個の応答タイミングが存在した。本研究では、対応先の応答タイミングが同一である応答を共起関係にあるとした。

図 5 に ent の計算例として、傾聴応答「はい」の汎用性 $ent(\text{はい})$ の計算方法を示す。語りの各行には応答タイミングにより区切られた語りが、傾聴応答の各行には同じ行の語りの直後が対応先の応答タイミングである応答がそれぞれ記されている。また、listener 1 ~ 5 は評価データにおける 5 人の聞き手を意味する。

評価データの全応答について ent を算出し、傾聴応答の種類ごとに平均値を用いて、Level ごとの平均値を算出した。図 6 にその結果を示す。Level 1 が Level 2 を僅かに下回っていることを除き、Level が低いほど ent の値が高い。これは 3 章で行った傾聴応答の分類の正当性を支持していると言える。また、Level 3 から 4 にかけて、 ent の値が大きく低下している。このことは 4.1 節と同様、Level 3 と 4 の間で共感度が大きく異なることを示唆している。

5 まとめ

本論文では、語りへの共感度に基づく傾聴応答の分類について述べた。共感度の高さにより傾聴応答を 5 段階に分類したのち、傾聴応答データを用いて分類の正当性を定量的に示した。今後は語りに示す共感に応じた傾聴応答の生成に取り組んでいきたい。

謝辞 高齢者のナラティブデータは、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から提供いただいた。本研究は、一部、科学研究費補助金 挑戦的研究 (萌芽) (No. 18K19811) により実施したものである。

参考文献

- [1] 大野ら. 対話コーパスを用いた相づち生成タイミングの検出. 信学論, Vol. J100-A, No. 1, pp. 53–65, 2017.
- [2] Yamaguchi et al. Analysis and prediction of morphological patterns of backchannels for attentive listening agents. In *Proc. of IWSDS-2016*, pp. 1–12, 2016.
- [3] Kobayashi et al. Design targeting voice interface robot capable of active listening. In *Proc. of HRI-2010*, pp. 161–162, 2010.
- [4] Lala et al. Attentive listening system with backchanneling, response generation and flexible turn-taking. In *Proc. of SIGDIAL-2017*, pp. 127–136, 2017.
- [5] 下岡ら. 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. 4, No. 1, pp. 3–47, 2017.
- [6] 村田ら. 語りの傾聴を話し手に示す応答発話の収集. 電学論, Vol. 138-C, No. 5, pp. 637–638, 2018.
- [7] 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.
- [8] 丸山ら. 日本語節境界検出プログラム CBAP の開発と評価. 自然言語処理, Vol. 11, No. 3, pp. 39–68, 2004.