

## 傾聴において表出する評価応答に対応した語りの語句

小杉 駿介<sup>†1</sup> 伊藤 滉一朗<sup>†1</sup> 村田 匡輝<sup>†2</sup> 大野 誠寛<sup>†3</sup> 松原 茂樹<sup>†1</sup>  
<sup>†1</sup>名古屋大学 <sup>†2</sup>豊田工業高等専門学校 <sup>†3</sup>東京電機大学

## 1 はじめに

語ることは人間の基本的な欲求であり、コミュニケーションロボットなどの会話エージェントが人の語りを聞く役割を担うことが期待される。語りの聴き手を担う上で、語りを傾聴していることを語り手に伝える応答（以下、傾聴応答）を表出することは、語り手の語る意欲を促すという点で重要である。傾聴応答の1つに、語りの内容に対し称賛や労いを伝えることによって評価を示す応答（以下、評価応答）がある。評価応答は語りの内容を受けて発話されるため、評価応答を適切に生成する方法を検討する上で、評価応答の対象となりやすい語りの表現を収集し、その特徴を明らかにすることは有用である。

そこで本稿では、傾聴を示す応答で表出される評価応答の対象となりうる表現の獲得を目的として、評価応答に対応した語りの語句（以下、評価対象）の同定について述べる。本手法では、語りの語句および応答の表現の極性、語句の重要度などを用いて同定する。

## 2 評価応答に対応した語りの語句

評価応答は、発話内容に対し称賛や労いを伝えることにより語り手への理解を示す効果があり、語り手への傾聴姿勢を示す上で重要な役割を持つ。語り目と評価応答の例を以下に示す。評価応答を【】で示す。

あの一わたくしはですね歩くのが得意なんです  
 毎日一万歩以上歩いております【すごいですね】  
 そういうことで

評価応答は語りの内容を受けて発話されるため、生成のためには評価対象となりうる表現を適切に検出できる必要がある。上記の例の場合、応答「すごいですね」に対応する評価対象は「毎日一万歩以上歩いております」である。傾聴応答の生成における評価対象の検出を実現するために、どのような表現が評価対象となりうるのかについてデータを観察し、その特徴を明らかにすることは有用である。本研究では、評価応答が対応付けられた語りデータの作成を目的に、語り目から評価対象を同定する方法について検討する。

Phrasal Expressions to be Evaluated in Attentive Listening of Narrative Speech

Shunsuke Kosugi<sup>†1</sup> Koichiro Ito<sup>†1</sup> Masaki Murata<sup>†2</sup>  
 Tomohiro Ohno<sup>†3</sup> Shigeki Matsubara<sup>†1</sup>  
 Nagoya University<sup>†1</sup>  
 National Institute of Technology, Toyota College<sup>†2</sup>  
 Tokyo Denki University<sup>†3</sup>

表 1: 使用した素性

1. 節についての素性
a. 評価応答と節の間の距離 (0~9)
b. 節境界の種類
c. 形態素数
d. 文節数 (1, 2, それ以上)
e. 品詞の個数
2. 節・応答の極性
a. 節の極性 (positive, neutral, negative)
b. 節に含まれる positive な語の個数
c. 節に含まれる negative な語の個数
d. 評価応答の極性 (positive, neutral, negative)
e. 応答の表現
f. 節の極性と応答の極性が一致しているか
3. 節に含まれる単語の重要度
a. 節に含まれる内容語の情報量の平均値

## 3 評価対象の同定手法

## 3.1 提案手法の概要

本手法では、評価応答とその直前の語りの節系列を入力とし、その節系列から評価対象の節を同定する。例えば、2章の例では

- あの一わたくしはですね
- 歩くのが
- 得意なんです
- 毎日一万歩以上歩いております

という直前の語りの4つの節から評価対象となる節を同定する。同定にはSVM(Support Vector Machine)を用い、入力された節の中でスコア（評価対象になりやすさ）が最大であるものを出力する。

## 3.2 使用する素性

本手法で使用する素性を表1に示す。素性1.は節および単語に関する言語情報である。素性2.では、節の極性の他に、応答の極性も使用する。極性解析には日本語評価表現辞書 [1, 2] を使用する。極性は、評価表現辞書に出現する語が、対象となる節または評価応答に現れた場合、語の極性が positive ならば +1, negative ならば -1 とし、その合計が正ならば positive, 負なら

ば negative, 0 ならば neutral とする. 素性 3. の節に含まれる内容語の情報量は以下の式で求める.

$$I(w_i) = -\log_2 \frac{F(w_i) + 1}{\sum_j F(w_j)}$$

$w_i$  は形態素の語彙素,  $F(w_i)$  はテキストデータ上の  $w_i$  の出現頻度である.

## 4 評価実験

### 4.1 実験データの作成

実験に使用するデータを, 傾聴応答コーパス [3] を用いて作成した. 傾聴応答コーパスは, 語りデータとして高齢者のナラティブコーパス JELiCo<sup>1</sup> を使用し, その音声に対して作業者 5 名が応答を実施することで作成している. 本データにおける応答 64,543 個のうち, 評価応答は 2,000 個で, 全体の約 3.1% を占める. このコーパスから, 語りと応答に基づき, 評価対象と考えられる節系列 (ただし, 長さを最大 3 に限定) を評価応答と対応付けた<sup>2</sup>. この手順により, 全 2,000 個の評価応答のうち, 1,924 個の評価応答について評価対象と対応付けることができた. 残り 76 個の評価応答については, 評価対象が一意に定まらないため評価対象と対応付けることができなかった. 評価対象と評価応答の対応付けの例を以下に示す.

【評価対象の節系列】 現在その先生僕をあの一手術してくれた/先生は/肝臓の権威っていわれるぐらいの大先生になっているんです/

【評価応答】 いい先生

対応付けた評価応答の節数は平均 1.87 個であった.

### 4.2 実験の概要

本手法の有効性を確認するため, 評価対象の同定実験を行った. 4.1 節で作成した節系列と評価応答の対のうち, 学習データとして 1,240 対, テストデータとして 345 対を用いた. 入力する節の数はそれぞれ 11,500 個, 3,234 個である. 形態素解析には MeCab<sup>3</sup>, 構文解析には CaboCha<sup>4</sup> を使用した. 節境界解析には, 日本語節境界検出プログラム CBAP[4] を用いた. 情報量の計算に用いるテキストデータには『日本語話し言葉コーパス』(CSJ) 短単位語彙表<sup>5</sup>を用いた. 語の異なり数は 42,542, 語の頻度の合計  $\sum_j F(w_j)$  は 7,479,773 である. SVM には, scikit-learn の SVC<sup>6</sup> をデフォルトのまま使用した. 4.1 節で評価応答と対応付けた評価対象を正解とし, 正解に対する精度により評価した. 精度は,

$$\text{精度} = \frac{\text{正しく検出された節数}}{\text{評価対象として検出された節数}}$$

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

<sup>2</sup>節系列は評価応答の直前 10 個の節から選択した.

<sup>3</sup><http://taku910.github.io/mecab/>

<sup>4</sup><https://taku910.github.io/cabocha/>

<sup>5</sup>[https://pj.ninjal.ac.jp/corpus\\_center/csj/chunagon.html](https://pj.ninjal.ac.jp/corpus_center/csj/chunagon.html)

<sup>6</sup><https://scikit-learn.org/stable/index.html>

表 2: 実験結果

	ベースライン	提案手法
精度	68.50%	76.59%

により測定した. 評価応答の直前に出現した節を評価対象として同定する場合をベースラインとして, 本手法の精度と比較した.

### 4.3 実験結果

実験の結果を表 2 に示す. 本手法の精度は 76.59% であった. ベースラインの精度を上回っており, 本手法の有効性を確認した. 同定に成功した例を以下に示す. 正解である節系列を下線で示す.

【語り】 多くの先生方や友達にもえー素晴らしかったという/風に褒められましたで弓道は/高校時代から 60 年間やっていますので/えーま自信がないってことは/ないんですが/まー当たらなかつた時の/ことを考えると/不安だったんですけど/2 本のうち 1 本当たりましたので/えー非常に良かった

【応答】 凄い

【同定された節】 2 本のうち 1 本当たりましたので

## 5 まとめ

本稿では, 傾聴を示す応答で表出する評価応答に対応した語りの語句の同定手法について述べた. 本手法では, 語りや応答の極性, 節の重要度に注目して評価応答の対象となる節を同定した. 傾聴応答コーパスを用いた同定実験の結果, 本手法の有効性を確認した.

謝辞 高齢者のナラティブコーパスは, 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] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出のための評価表現の収集. 自然言語処理, Vol.12, No.3, pp.203-222, 2005.
- [2] 東山昌彦, 乾健太郎, 松本裕治. 述語の選択選好性に着目した名詞評価極性の獲得. 言語処理学会第 14 回年次大会論文集, pp.584-587, 2008.
- [3] 村田匡輝, 大野誠寛, 松原茂樹. 語りの傾聴を話し手に示す応答発話の収集. 電学論 C, Vol.138, No.5, pp.637-638, 2018.
- [4] 丸山岳彦, 柏岡秀紀, 熊野正, 田中英輝. 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.