

語りの傾聴において同意が適さない発話の検出

河本 悠太[†] 伊藤 滉一郎[‡] 松原 茂樹[‡]

[†] 名古屋大学情報学部 [‡] 名古屋大学大学院情報学研究科

1 はじめに

社会の個人化が進み、聴き手不在の生活シーンが増加しており、会話エージェントが人の語りを聴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、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点で重要である。

語りの傾聴では、語り手の発話を肯定することが聴き手の基本的な応答方略となる。傾聴応答の代表は、相槌（「はい」「ええ」など）であり、その生成法が提案されている [1]。しかし、語り手の自虐的な発話など、語りに同意した応答を表出することが必ずしも適さないことがある。その場合は、語りに同意しないことを示す応答（以下、**不同意応答**）の表出が求められる。

本稿では、同意を示す応答が適さない発話の検出手法について述べる。語りデータを用いた実験を行い、手法の検出性能を評価した。

2 語りの傾聴と不同意

2.1 傾聴における不同意応答

傾聴応答は、語り手に傾聴態度を伝える効果を持つ。語りの内容を肯定する応答の表出が基本であるものの、場合によっては、不同意応答の表出を必要とすることもある。語りと不同意応答の例を図1に示す。不同意応答は、語り手の自虐的な発話に対して表出されることが多い。そのような語りの発話で、不同意応答を表出しなければ、語り手に不快感を与えることになる。本稿では、不同意応答生成の実現に向け、同意を示す応答が適さない語りの発話の検出可能性について述べる。

まー何しろ色々なものに興味がありまして
年甲斐もないって言われますが【いえいえ】

スチュワードスは憧れだったのですが、
それだけの器量も頭もなかったの
【そんなことないです】

色付けをしていきますけども
そのわたくしはセンスがなくて【いやいや】

図1 語りと不同意応答の例（【 】内は応答を示す）

2.2 傾聴応答コーパスと不同意応答

語りの傾聴応答における不同意応答の出現割合を調査した。調査には、傾聴応答コーパス [2] を使用した。これは、高齢者の語りに傾聴応答を付与することで作成されたコーパスである。

語りデータには、高齢者ナラティブコーパス JELiCo¹ を使用している。JELiCo には、30 名の高齢者による語りの音声収録されている。一方、応答データには、語りデータに対する作業員 11 名の独立した応答を収録している。収録された音声は文字化され、すべての傾聴応答には、その役割を表す応答タイプ（相槌、同意、不同意など計 16 種類）が人手で付与されている。傾聴応答コーパス内の 119,269 個の傾聴応答を調査したところ、不同意はそのうち 353 個で全体の 0.29% を占めていた。

3 同意が適さない発話の検出

応答において不同意応答を生成するには、

1. 同意が適さない発話を検出
2. 不同意応答の表現と表出タイミングを決定が必要となる。本研究では、1. の実現性の検証を目的とし、語りに含まれる節を対象に同意が適さないか否かを分類する手法について述べ

Detection of Narrative Utterances Unsuitable for Showing Agreement in Attentive Listening
Yuta Kawamoto, Koichiro Ito, Shigeki Matsubara
Nagoya University

¹ Japanese Elder's Language Index Corpusv2,
https://figshare.com/articles/dataset/Japanese_Elder_s_Language_Index_Corpus_v2/2082706/1

る。節は、述語を中心とした言語的なまとまりであり、語りの単位として適当である [3]。

本手法では、分類に SVM を用いる。分類のための素性として、節を構成する単語に対応する Bag-Of-Words (BOW) または TF-IDF の特徴量ベクトルを用いる。

4 評価実験

4.1 実験概要

同意が適さない発話の検出可能性の検証を目的として、節の分類実験を行った。

4.1.1 実験設定と実験データ

実験では、不同意応答に対応する節（不同意節）と同意することを示す応答（同意応答）に対応する節（同意節）からなる節集合を対象に、各節を不同意節と同意節に分類する設定とした。同意節は、同意が適する節とみなせるため、不同意節の負例として利用できる。

実験データとして、傾聴応答コーパス [2] の一部を用いた。実験にあたり、コーパス内の不同意応答 353 個と同意応答 2,042 個を、発話時間情報に基づき語りの中の単一の節と対応付けた。節境界の検出には、節境界解析ツール CBAP [4] を用いた。対応付けの結果、不同意節 236 個、同意節 775 個が得られた。不同意節と同意節の合計 1,011 個のうち、812 個を学習データ、199 個をテストデータとして用いた。

4.1.2 実装と評価指標

SVM の実装には、python の scikit-learn¹ の SVC² ライブラリを用いた。SVM は、SVC の class_weight オプションをデフォルト設定と、balanced 設定の 2 通りで学習した。実験データでは不同意節の数が極端に少ないため、balanced とすることで、データの不均衡さが考慮される。その他のオプションは、全てデフォルトの設定とした。

検出性能は、不同意節に対する再現率と適合率、及び、F 値で評価する。また、学習データにおける不同意節と同意節の割合に基づきランダムで分類する手法と比較する。

4.2 実験結果

実験の結果を表 1 に示す。ほとんどの設定でランダム設定よりも F 値が向上し、本手法による検出可能性を確認した。特に、データ不均衡を考慮した場合に F 値が最大となった。

表 1 実験結果

	適合率	再現率	F 値
ランダム	0.150	0.171	0.160
BOW (default)	0.400	0.216	0.280
TF-IDF (default)	0.750	0.081	0.146
BOW (balanced)	0.406	0.351	0.376
TF-IDF (balanced)	0.384	0.405	0.394

4.3 考察

上記の実験では、同意節か不同意節かの決定に関する閾値 τ ($0 \leq \tau \leq 1$) をデフォルトの 0.5 とした。しかし、データ不均衡に対応するため、適当な閾値を設けることが考えられる。そこで、SVC の class_weight オプションをデフォルトとし、検出時の閾値を変化させて実験を行った。

実験の結果、BOW を特徴量とする場合、F 値が最大で 0.460 ($\tau=0.143$)、TF-IDF の場合、F 値が最大で 0.508 ($\tau=0.146$) となった。適切な閾値設定により、検出性能が大幅に向上することを確認した。

5 まとめ

本稿では、同意を示す応答が適さない語りの発話の検出を行った。実験の結果、F 値においてランダム設定よりも高い結果が得られ、不同意節の検出可能性を確認した。今後は、LSTM などの時系列を考慮したモデルや、事前学習済みの BERT などの適用を試みたい。

謝辞 傾聴応答コーパスの利用に関してご教示頂いた豊田高専の村田匡輝先生に感謝します。高齢者ナラティブコーパスは、奈良先端大ソーシャル・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、科研費（挑戦的研究（萌芽））(No. 18K19811) により実施した。

参考文献

- [1] 大野誠寛, 神谷優貴, 松原茂樹. 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, J100-A(1):53–65, 2017.
- [2] 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, 138(5):637–638, 2018.
- [3] 丸山岳彦, 高梨克也, 内元清貴. 日本語話し言葉コーパスの構築法. 国立国語研究所報告, 124:252–322, 2006.
- [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, 11(3):39–68, 2004.

¹ <https://scikit-learn.org/stable/>

² <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>