

# 傾聴を示す応答で繰り返される語りの語句の検出

伊藤 滉一郎<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

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。機器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、具体的には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的応答と呼ぶ。傾聴的応答の代表は相槌であり、その生成法が提案されているものの [1, 2], 傾聴を示す応答は相槌以外にも存在する [3, 4]。その1つが繰り返し応答 (echoic response) である。繰り返し応答とは、応答を構成する語句が直前の語りに含まれる発話の形式である。

本論文では、語りに対する傾聴的応答を付与したデータ (以下、応答データ) に基づき、繰り返しの対象となる語句の検出手法について述べる。

## 2 傾聴のための繰り返し応答

繰り返し応答は、対話における機能の観点から、「発話権が移動する繰り返し応答」と「発話権が移動しない繰り返し応答」に分類できる。このうち、本研究では、発話権が移動しない応答を対象とする。

福富 [5] によると、繰り返し応答の機能には9種類のものがあるが、傾聴的応答においては、これらのうち、発話権の移動を伴わない、受信応答や感情表出、共感・一体感表出の機能が出現する。

繰り返し応答の例を以下に下線で示す。【】で囲んだ部分が聞き手の発話である。

- えーとそうですね【えー】あの一外では【はい】あの一ゴム飛びっていう【はい】んで【あっはいはい】【ゴム飛び】ゴムが一本【はい】引いてもらって【えー】

において、「ゴム飛び」は、受信応答の機能を持つ繰り返し応答である。また、以下の例

- 石蹴り【あっはい】っていうんですか【石蹴りやりましたね】あの道路に【はい】【えーえー】丸を描いて【はい】

における応答「石蹴りやりましたね」は、共感・一体感表出の機能を持つ繰り返し応答である。

## 3 繰り返し応答データの作成

語りと応答データを用いて、繰り返し応答データを作成した。

### 3.1 語りに対する傾聴的応答データ

著者らはこれまでに、高齢者の語りに対して傾聴的応答を付与した応答データの収集を行ってきた [6]。応答データは、語りのデータとして、高齢者のナラティブコーパス JELiCo<sup>1</sup> を使用し、その音声に対して高度な接客スキルを要する業務経験を有する作業員1名が応答を実施することで作成した。形態素解析には MeCab<sup>2</sup> を用いた。また、応答データには、文献 [7] を参考に、役割を表すラベル (17種類) を付与している。本データにおける応答 16,684 個のうち、繰り返し応答は 753 個であり、全体の 4.46% を占める。

### 3.2 繰り返し対象語句のタグ付け

繰り返し応答は、直前の語りの形態素列から構成されるため、語りに対して、繰り返しの対象となった語句にタグ付けを行った。作業は以下の手順で行った。

手順 1 語りに対して、繰り返し応答を構成する形態素の基本形の系列と最長一致する系列を、繰り返し対象として機械的にタグ付けする。

手順 2 手順 1 におけるタグ付け結果を確認し、必要に応じて人手で修正する。

本研究では、以下の形態素のみを手順 1 の対象とした。

- 繰り返し応答の発話開始時間からさかのぼって 10 秒以内に発話が終了している形態素、あるいは、
- 繰り返し応答の発話開始までに発話が終了している直近 30 個の形態素

また、機能語のみからなる系列は、最長一致を求めるときに候補から除外した。

以下の発話におけるタグ付けの例を図 1 に示す。【】が応答、それ以外は語りであり、下線で示した「八年」が繰り返し応答である。

- ここ八【うーん】年ぐらいは【はい】歌舞伎は観ており【あー】ません【八年】

語りデータは形態素と sp に、応答データは応答のラベルごとに、それぞれ ID を付与した。sp は無音区間を表す。応答 ID: 589 の繰り返し応答「八年」は、語りに対する最長一致の系列が「八年」なので、語りの形態素 ID 区間:3883-3884 にタグが付与される。

<sup>1</sup><http://sociocom.jp/software.html>

<sup>2</sup><http://taku910.github.io/mecab/>

高齢者の語り			傾聴的応答			
発話時間	形態素 ID	基本形	発話時間	応答 ID	基本形	繰り返し対象
77.33 - 77.75	3881	ここ	79.71 - 80.07	586	うーん	
77.75 - 78.16	3882	sp				
79.44 - 79.87	3883	八				
79.87 - 80.19	3884	年				
80.19 - 80.64	3885	ぐらい				
80.64 - 80.81	3886	は	80.73 - 81.08	587	はい	3883 - 3884
80.81 - 80.85	3887	sp	82.55 - 83.42	588	あー	
				589	八年	

図 1: 繰り返し対象語句のタグ付けの例

表 1: 品詞の出現割合の比較

品詞	全形態素 (%)	タグ付けされた形態素 (%)
名詞	21.60	61.12
助詞	32.69	10.01
動詞	13.20	9.80
助動詞	13.33	6.43
接尾辞	1.85	6.01

表 2: 品詞細分類の出現割合の比較

品詞	全形態素 (%)	タグ付けされた形態素 (%)
普通名詞	18.97	45.10
数詞	1.43	9.06
固有名詞	1.19	6.95
助動詞	13.33	6.43
格助詞	14.91	6.01

作業の結果、753 個の繰り返し応答に対して、手順 1 により 802 個の形態素列にタグが付与され、手順 2 により、最終的に 830 個の形態素列にタグが付与された。タグ付け数が繰り返し応答の数よりも多いのは、1 つの繰り返し応答に対して、隣接しない複数の形態素列にタグが付与される場合があるためである。

## 4 繰り返し対象の語句の分析

本データの一部を分析データとした。分析データには、42,153 個の形態素が存在し、繰り返し応答は 466 個含まれていた。また、そのタグ付けの結果として、523 個の語りの形態素列 (949 個の形態素) が存在した。

### 4.1 繰り返し対象の品詞

分析対象の全形態素とタグ付けされた形態素の品詞の出現頻度を表 1 に示す (上位 5 個)。品詞体系は、UniDic<sup>3</sup> に従った。タグ付けされた形態素の 61.12% が名詞であり、名詞の出現割合が 21.60% であることを考えると、名詞のタグ付け割合が特に高いことが確認できる。同様に、細分類 (UniDic における品詞中分類) についても、分析対象の全形態素とタグ付けされた形態素における出現頻度を調べた。その結果を表 2 に示す (タグ付けされた形態素の出現頻度で上位 5 個)。全形態素における数詞と固有名詞の出現割合は、それぞ

<sup>3</sup>[https://www.gavo.t.u-tokyo.ac.jp/~mine/japanese/nlp+slp/UNIDIC\\_manual.pdf](https://www.gavo.t.u-tokyo.ac.jp/~mine/japanese/nlp+slp/UNIDIC_manual.pdf)

表 3: 抽出した固有表現の出現頻度とタグ付け割合

分類	表現例	出現頻度	タグ付け割合 (%)
LOCATION	北海道	301	13.62
DATE	明治 42 年	95	4.21
PERSON	リリアン	77	15.58
ORGANIZATION	三井物産	42	16.67
TIME	30 分前	4	25.00
ARTIFACT	小田急線	2	0.00
合計		521	12.48

れ 1.43% と 1.19% であった。一方で、タグ付けされた形態素における出現割合は、それぞれ 9.06% と 6.95% であり、数詞と固有名詞がタグ付けされやすいことが確認できる。以下の語りの文の点線部は、

1. 神奈川県 の 茅ヶ崎 という 所 に 【茅ヶ崎】
2. 母が えー 106 歳 で まだ 生きて います 【106 歳】

1 が固有名詞を、2 が数詞を含むタグ付けされた形態素列の例である。

### 4.2 固有表現の繰り返しされやすさ

情報量やメッセージ性の強い語句は、聞き手の注意を引き、記憶にも残りやすいので、繰り返しの対象になりやすい可能性がある。固有表現がそのような特徴を持っていると考え、その繰り返しされやすさを分析した。CaboCha<sup>4</sup> を用いて、語りの発話から、固有表現抽出をした。複数の文節で構成される固有表現を許すよう設定し、他のオプションはデフォルトのままとした。CaboCha は、抽出した固有表現に対して、IREX (Information Retrieval and Extraction Exercis) の定義<sup>5</sup> に基づく固有表現タグを出力する。

表 3 に、IREX の定義に基づく分類ごとに、抽出した固有表現の個数と表現例、タグ付けされた形態素を含む割合 (以下、タグ付け割合) を示す。分析対象のデータから抽出した固有表現の総数は 521 個であり、12.48% にあたる 65 個の固有表現がタグ付けされた形態素を含んでいた。IREX の分類ごとでは、DATE が全体では 2 番目に多く出現するが、タグ付け割合は 5 番目であり、出現頻度に対してタグ付け割合が低いことが確認できる。一方で、PERSON は DATE に近い出現頻度でありながら、タグ付け割合は 15.58% と全体の 12.48% よりも高かった。

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

<sup>5</sup><https://nlp.cs.nyu.edu/irex/NE/df990214.txt>

### 4.3 語句の重要度と繰り返されやすさ

著者らはこれまでに、日本語話し言葉コーパス (CSJ)[8] に収録された模擬講演インタビューの対話データを使用して、重要度が大きい語句ほど繰り返し応答に含まれる確率が高くなることを確認している[9]. 本節では、本研究で扱う語りデータと繰り返し対象の形態素列との間にも同様の傾向が見られるかを調査した. 語句の重要度として、文献 [9] と同様に以下の指標を用いた.

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

$w_i$  は形態素の語彙素,  $F(w_i)$  は『日本語話し言葉コーパス』短単位語彙表 ver. 2018.3.1<sup>6</sup>における  $w_i$  の出現頻度を用いた. 語の異なりは 42,545, 語の頻度の合計  $\sum_j F(w_j)$  は 7,479,773 である.

内容語 (名詞, 動詞, 形容詞, 形状詞, 副詞のいずれか) について, 全形態素とタグ付けされた形態素の重要度  $I$  の中央値は, それぞれ 13.56 と 15.64 であり, タグ付けされた形態素の重要度  $I$  の中央値の方が, 高い値を示していることが確認できる. 以下に, 重要度の高い語句を含む繰り返し応答の例を示す.

- 小さい頃から【はい】女医さんになりました  
 かった【そーなんですわー】【女医さん】

において, 繰り返し対象としてタグ付けされた形態素列「女医さん」における形態素「女医」の重要度は 20.68 である. なお, 内容語に含まれる形態素は, 全形態素において全体の 40.95%にあたる 17,260 個, タグ付けされた形態素において全体の 75.45%にあたる 716 個存在し, タグ付けされた形態素における内容語の割合の高さも同時に確認された.

### 4.4 繰り返し対象の語句とフィラー

ある語句が繰り返しの対象となるかは, その内容だけでなく, 出現するタイミングや位置も重要である. 本節では, 「あー」や「えー」のようなフィラーの出現が, 繰り返し対象の語句の出現に関わっていると考えた. フィラーは, 言いよどみの一種であり, 次に発話すべき言葉を探したり, 言いにくいことを口にすることを示すために用いられる [7]. したがって, フィラーの後の語句は, 比較的重要な意味を持つと言える.

そこで, タグ付けされた形態素列のうち, 直前の形態素がフィラーの割合を調査したところ, 21.99%(115/523) が直前にフィラーが存在した. 以下に, タグ付けされた形態素列のうち, 直前の形態素がフィラーである例を示す. <> で囲んだ部分がフィラーである.

- わたくしの今日までの仕事は【はい】  
 <えー>ライターです【へー】【ライター】

この例において, 繰り返し対象としてタグ付けされた形態素列「ライター」の直前に, フィラー「えー」が出現している. 応答する側として, 語り手が言いよどみながらも, 発した言葉には, 応答したくなるのが推測される.

<sup>6</sup>[https://pj.ninjal.ac.jp/corpus\\_center/csaj/chunagon.html](https://pj.ninjal.ac.jp/corpus_center/csaj/chunagon.html)

表 4: SVM で用いた素性

(A) 文節の主辞に関する素性	
1.	文節の主辞の品詞
2.	文節の主辞の品詞細分類
3.	文節の主辞の活用形
(B) 文節の語形に関する素性	
4.	文節の語形の品詞
5.	文節の語形の品詞細分類
6.	文節の語形の活用形
(C) 文節の形態素数	
7.	文節の形態素数が以下の 4 分類のいずれであるか (1, 2, 3~6, 7 以上)
(D) 節境界単位に関する素性	
8.	文節が属する節境界単位の種類
9.	節境界単位内で先頭から何番目の文節か
(E) 文節が含む語句に関する素性 [4.1, 4.2]	
10.	文節が固有名詞を含むか否か
11.	文節が数詞を含むか否か
12.	文節が固有表現を構成する形態素を含むか否か
13.	IREX の分類について, 文節がその固有表現を構成する形態素を含むか否か
(F) 文節の重要度に関する素性 [4.3]	
14.	文節内に内容語が存在し, 文節の重要度 $I$ が全形態素の中央値 13.56 より大きいかな
(G) フィラーに関する素性 [4.4]	
15.	直前の文節の最後の形態素がフィラーか否か

## 5 繰り返し対象語句の検出手法

「節 (clause)」は, 文法的・意味的なまとまりを備えた言語的単位である [10]. 節は聞き手にとっての理解の単位と言え, 傾聴的応答は語りの理解を伝達する発話であることから, 節ごとに応答することは, 傾聴的応答を生成するシステムを実現するための 1 つの方略であると考えられる.

そこで本論文では, 節境界単位内の文節集合から繰り返すべき語句を含んだ文節の検出手法を提案する. 節境界単位とは, 節の終端境界で挟まれた区間である.

### 5.1 手法の概要

本手法では, 節境界単位を入力とし, 節境界単位を構成する文節集合から 1 つ以上の繰り返し対象文節を検出する. なお, 入力される節境界単位には, 繰り返し対象文節が必ず含まれていることを前提とする. 以下に, 手法の概要を説明する. はじめに, SVM(Support Vector Machine)により, 上記の文節集合の中から繰り返し対象文節を検出する. つぎに, 節境界単位内の文節が 1 つも繰り返し対象と判定されていない節境界単位については, その文節の中で SVM が出力するスコア (繰り返されやすさ) が最大の文節を繰り返し対象として検出する.

### 5.2 SVM で用いた素性

SVM で用いた素性を表 4 に示す. 素性 (A), (B), (C), (D) は, 文節情報と節境界単位についての基本的な言語情報であり, それ以外は 4. の分析に基づき定

表 5: 対象データの規模

	学習データ	テストデータ
文節	1946	771
繰り返し対象文節	563	200
節境界単位	474	166

表 6: 各素性の効果

取り除いた素性	再現率	適合率	F 値
(E)	63.50%	62.25%	62.87
(F)	63.00%	63.00%	63.00
(G)	65.00%	64.04%	64.52
本手法	66.50%	64.88%	65.68

めた素性である。各素性に該当する 4. の節番号を表中に括弧書きで記す。素性 7 の分類は、学習データを用いたクロズドテストを試行し、いくつか設定した分類のうち、最も高い性能を示したものを採用した。素性 14 の文節の重要度  $I$  は、文節に属する内容語の重要度  $I$  のうち最大のものとした。

## 6 実験

### 6.1 実験概要

本手法の有効性を評価するため、繰り返し対象文節の検出実験を行った。学習データには、分析データと同じ語りデータから、繰り返し対象文節を含む節境界単位を抽出したものを使用した。テストデータには、分析データとは別の語りデータから同様にして抽出した節境界単位を使用した。学習データとテストデータの規模を表 5 に示す。文節への分割には CaboCha を 4.2 節と同じ設定で、節の終端境界の検出には節境界解析ツール CBAP[11] を用いた。CBAP には 147 種類の節境界ラベルが存在している。また、SVM のツールとして SVM-light<sup>7</sup>[12] をデフォルトのオプションのまま使用した。繰り返し対象文節を正解とし、正解に対する再現率、及び、適合率により評価した。再現率、適合率はそれぞれ、

$$\text{再現率} = \frac{\text{正しく検出できた繰り返し対象文節数}}{\text{正解の繰り返し対象文節数}}$$

$$\text{適合率} = \frac{\text{正しく検出できた繰り返し対象文節数}}{\text{繰り返し対象として検出された文節数}}$$

を測定した。さらに、4. の分析に基づいて定めた素性 (E), (F), (G) の有効性を個別に評価するため、表 4 に示した素性から、各素性を取り除き検出実験を実施し、本手法と性能を比較した。

### 6.2 実験結果

実験の結果、本手法の再現率、適合率はそれぞれ、66.50%(133/200)、64.88%(133/205) であり、これらの調和平均である F 値は 65.68 であった。表 6 に、各素性を取り除いた場合の検出結果を示す。いずれの場合も、F 値が低下しており、各素性の有効性を確認した。特に、文節が含む語句についての素性 (E) を取り除くことによる F 値の低下が最も大きかった。

<sup>7</sup><http://svmlight.joachims.org/>

### 6.3 節境界単位レベルの精度

テストデータにおける 166 個の節境界単位において、節境界単位内の全ての文節に対して繰り返し対象であるか否かを正しく判定できた節境界単位の割合を調べた。文節を 1 つしか含まない節境界単位 13 個を除いた 153 個の節境界単位を用いて精度を評価した。また、節境界単位からランダムに文節を 1 つ選択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。本手法の精度は全体で 49.02% であり、ベースラインの精度 27.69% を上回っていた。

## 7 まとめ

本論文では、傾聴を示す応答で繰り返される語りの語句の検出手法について述べた。本手法では、繰り返される語句の特徴として、固有名詞、数詞、固有表現との関係、語の重要度、フィラーとの位置関係に注目し、繰り返し対象の文節を検出する。語りデータを用いて、繰り返し対象の文節の検出実験を実施した結果、再現率 66.50%、適合率 64.88% であり、本手法の有効性を確認した。今後は、本手法によって検出した繰り返すべき語句を用いた、応答生成方法を検討したい。

**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] 大野, 神谷, 松原. 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] 山口, 井上, 吉野, 高梨, G.Ward, 河原. 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成. 人工知能学会論文誌, Vol. 31, No. 4, pp. C-G31.1–10, 2016.
- [3] 目黒, 東中, 堂坂, 南. 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. 情処論誌, Vol. 53, No. 12, pp. 2787–2801, 2012.
- [4] 下岡, 徳久, 吉村, 星野, 渡部. 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.
- [5] 福富. 日本語会話における「くり返し」発話について. 言語文化研究, Vol. 5, pp. 105–125, 2010.
- [6] 村田, 大野, 松原. 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.
- [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] 前川, 籠宮, 小磯, 小椋, 菊池. 日本語話し言葉コーパスの設計. 音声研究, Vol. 4(2), pp. 51–61, 2000.
- [9] 村田, 大野, 松原. 会話ロボットにおける繰り返し応答の生成. FIT(情報科学技術フォーラム), Vol. 14, No. 2, pp. 251–252, 2015.
- [10] 丸山, 高梨, 内元. 報告書『日本語話し言葉コーパスの構築法』, 第 5 章. No. 124. 国立国語研究所, 2006.
- [11] 丸山, 柏岡, 熊野, 田中. 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.
- [12] Joachims T. Learning to Classify Text Using Support Vector Machines: Methods, Theory, and Algorithms. *Computational Linguistics*, Vol. 29, No. 4, pp. 655–664, 2003.