

## 平成30年度 情報工学コース卒業研究報告要旨

松原 研究室	氏 名	伊 籐 滉 一 朗
卒業研究題目	傾聴を示す応答で繰り返される語りの語句の検出	

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の1つは語りに応答することである。傾聴を示す目的で語りに応答する発話(以下、傾聴的応答)の代表は相づちであり、その生成法が提案されているものの、傾聴的応答は相づち以外にも存在する。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である。

本論文では、傾聴的応答の1つである繰り返し応答の自動生成を目的に、語りにおける繰り返しの対象となる語句の検出手法について述べる。本研究では、語りにおける傾聴的応答を付与したデータを用いて、繰り返し応答データを作成した。図1は、応答ID:589の繰り返し応答「八年」の対象を、語りの形態素ID区間:3883-3884としている例である。作成した繰り返し応答データを用いて、繰り返される語りの語句の特徴分析を行った結果、以下の語句が繰り返しの対象になりやすいことを確認した。

1. 固有名詞や数詞などの固有表現
2. 出現頻度が低く、重要度が大きい語句
3. 言いよどみの一種であるフィラーの直後に出現する語句

分析結果に基づき、繰り返される語りの語句の検出手法を開発した。本手法では、文法的・意味的なまとまりを備えた言語単位である節を入力とし、SVMにより、節を構成する文節集合から、繰り返しの対象となる語句を含んだ文節(繰り返し対象文節)を検出する。SVMには、文節と節の基本的な言語情報に加えて、上述の繰り返しの対象となる語句の特徴1~3に基づく素性を導入した。

繰り返し対象文節の検出実験を行い、再現率、適合率、及び、F値により評価し、本手法の有効性を確認した。また、繰り返される語句の特徴に基づく素性の有効性を個別に評価するため、上述の特徴1~3に基づく各素性を取り除いて検出実験を行った結果、いずれの場合もF値が低下しており、本研究で導入した素性の有効性を確認した。

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

図1: 繰り返し応答データの例