

令和元年6月11日現在

機関番号：56203

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16134

研究課題名(和文) 会話ロボットとの自由な対話に資するコミュニケーション断絶防止技術の研究

研究課題名(英文) Research on communication disconnection prevention technology conducive to dialogue with conversation robot

研究代表者

篠山 学 (SASAYAMA, MANABU)

香川高等専門学校・情報工学科・講師

研究者番号：60508232

交付決定額(研究期間全体)：(直接経費) 800,000円

研究成果の概要(和文)：Webから聞き間違いを収集し、聞き間違いコーパスを構築した。聞き間違いの検出手法は、元発話全てのよみがなを入力とし、音素に変換する。音素列から聞き間違い候補を出力し、聞き間違い発話との類似度を計算することで聞き間違い箇所を検出する。検出実験を行った結果、83%の聞き間違い箇所の検出精度が得られた。書籍やWebから言い間違いを収集し、言い間違いコーパスを構築した。CNNとLSTMを用いた深層学習を行ってモデルを作成し、言い間違いの検出実験を行ったところLSTMの場合に検出精度が約70%となった。

研究成果の学術的意義や社会的意義

聞き間違いを含む対話を収集して分析することで人間が聞き間違いやすい音素や拍の変動などが明らかになった。このことから、コミュニケーションロボットが人間の問い返しに対応できるようになると共にコミュニケーションロボットが発話する単語や表現を聞き取りやすい単語や表現に変更できる。また、言い間違いを含む対話の分析からコミュニケーションロボットが人間の言い間違いを認識できる。これらの成果が現在のコミュニケーションロボットと人間のコミュニケーションの断絶を防止でき、ストレスフリーな対話の実現できる。

研究成果の概要(英文)：We collected mishearing utterances from the web and constructed a mishearing corpus. In the detection method of mishearing, first, the readings of all the original utterances are input and converted to phonemes. Next, we output a candidate for misinterpretation from the phoneme sequence. Finally, the misplacement is detected by calculating the degree of similarity to the mishearing utterance. As a result of conducting the detection experiment, we got a detection accuracy of 83% of misplaced parts. We collected speech errors from books and the web, and built a speech error corpus. We created a model by performing deep learning using CNN and LSTM. We conducted a detection experiment of speech errors, we obtained about 70% detection accuracy in the case of LSTM.

研究分野：自然言語処理

キーワード：問い返し 聞き間違い 言い間違い コミュニケーション

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

コミュニケーションロボットは、介護施設や医療機関がお年寄りや心に傷を負った方のために導入するだけでなく、イベント会場や企業などが来訪者のための情報案内システムとして導入するなど、広く社会に普及しつつある。一般家庭向けにも幼児とコミュニケーション可能なおもちゃが発売されている。しかし、人間と自由な対話ができているとは言い難い。その主な原因となっているのが「コミュニケーションの断絶」である。コミュニケーションの断絶は聞き間違えたり言い間違えたりして会話が成り立たなくなってしまうことである。人間同士では、聞き間違えたときは相手に問い返したり、無意識に言い間違えたときは相手が状況に応じて指摘したり確認するなどによりコミュニケーションを継続している。

現在までコミュニケーションの断絶を防止する技術の研究はほとんどなされていない。これは、これまでの研究が人間の質問を理解し、いかに早く正確に回答するかに主眼を置いていたためである。企業などで情報案内ロボットが実用化されつつあることから、今後はいかにコミュニケーションを維持するかが重要となる。

関連研究では発話者の質問に曖昧性が含まれていることを認識し、曖昧性を除去するための「問い返し文」を生成する研究がある。これらは発話者の質問に意図しない曖昧性が含まれている場合の対処である。本研究では人間がコミュニケーションロボットの発話を聞き間違えたり、無意識のうちに言い間違えたりしたとき、人間の問い返しや言い間違いに対応する。

2. 研究の目的

本研究では、コミュニケーションロボットとの自由な対話を阻害するコミュニケーションの断絶を防止することを目指し、1)人間が聞き間違えたときに発話する問い返しを認識できるようにする、2)人間が無意識に言い間違えたことを認識し、状況に応じて指摘したり確認するための問い返しができるようにする、の2つを目標とする。研究期間内に、1)聞き間違いを含む対話を収集して、音素ごとに書き起こすことで聞き間違いの特徴を明らかにし、人間の問い返しを認識できるようにする、2)言い間違いを含む対話を収集して、人間の言い間違いの特徴を明らかにし、人間の言い間違いを認識できるようにすると共に、言い間違いをどのようなときに指摘してどのようなときに受け流すかを明らかにする、の2つを行う。これにより、人間が言い間違い易い単語や表現の組のデータベースを構築する。

3. 研究の方法

聞き間違えたときに発話する問い返しを認識できるようにするために、聞き間違いを含む対話を収集してコーパスを構築し音素ごとに書き起こして特徴を分析する。また、言い間違いを認識できるようにするために、言い間違いを含む対話を収集してコーパスを構築し人間が言い間違い易い単語や表現の組のデータベースを構築する。最後に、聞き間違えたときの問い返しの実証実験と言い間違いの指摘などの実証実験を行う。

4. 研究成果

(1) 聞き間違いを含む対話を収集して、音素ごとに書き起こすために Web から聞き間違いを収集し、聞き間違いコーパスを構築した。収集したサイトでは不特定多数のユーザから聞き間違いを投稿してもらい、サイト側で編集したあとに公開している。全部で 1147 個収集した。項目は、ID、正しい発話、正しい発話のよみがな、聞き間違い箇所、聞き間違い箇所よみがな、元発話全て、元発話全てよみがな、聞き間違い発話の 8 項目とした。よみがなは発話を音素に変換する処理に必要となる。構築した聞き間違いコーパスの一部を表 1 に示す。

表 1. 構築した聞き間違いコーパスの一部

ID	元発話全て	聞き間違い箇所	聞き間違い箇所よみがな
	よみがな	聞き間違い発話	よみがな
1	エムアールアイ行 ってきた	エムアールアイ	エムアールアイ
	エムアールアイイ ッテキタ	芋洗い	イモアライ
2	五名様ご案内	五名様	ゴメイサマ
	ゴメイサマゴアン ナイ	お姉様	オネエサマ
3	次は三滝口	三滝口	ミタキグチ
	ツギハミタキグチ	豚キムチ	ブタキムチ

聞き間違いの検出手法は、元発話全てのよみがなを入力とし、音素に変換する。音素列から聞き間違い候補を出力し、聞き間違い発話との類似度を計算することで聞き間違い箇所を検出する。類似度は音素の弁別素性表を基にした重み付きレーベンシュタイン距離を用いた。検出実験を行った結果、83%の聞き間違い箇所の検出精度が得られた。これにより、聞き間違いに対する問い返し箇所の検出が可能となりロボットが人間の問い返しに対応できるようになった。

聞き間違いの特徴を考察するために、聞き間違いによる音素の置き換わりを調査した。音素の類似によらない置き換わりとして聞き逃しやその補完、母音もしくは子音のみの置き換わり、音素の入れ替わりについての知見が得られた。

(2) 言い間違いを含む対話を収集して、人間の言い間違いの特徴を明らかにし、人間の言い間違いを認識できるようにするために、Web から言い間違いを収集し、言い間違いコーパスを構築した。また、静岡県立大学寺尾康教授のご厚意で寺尾先生が収集した言い間違いコーパスをご提供いただいた。全部で 3420 個となった。項目は、ID、元発話全て、言い間違い箇所、正しい発話、言い間違い分類の 5 項目とした。構築した言い間違いコーパスの一部を表 2 に示す。

表 2. 構築した言い間違いコーパスの一部

ID	元発話全て		
	言い間違い箇所	正しい発話	言い間違い分類
1	おねーちゃん寝返ったんだって		
	寝返った	寝違えた	\$語,\$代用
2	じゃけとさこひとつずつ		
	じゃけとさこ	じゃことさけ	\$音韻,\$交換
3	皆既日食肉まんで見えたんだよ		
	肉まん	肉眼	\$語,\$代用

言い間違いの検出は類似文字列検索を用いた検出手法と深層学習を用いた手法の 2 通りを試した。類似文字列検索を用いた手法では、言い間違った箇所は形態素解析の結果が未知語となった箇所とし、Wikipedia の単語との類似度から元の単語を推定した。実験の結果、F 値 0.7 を得たが、Wikipedia には日常的な単語以外にも多く含まれており、推定精度の向上が難しい。次に深層学習を用いた手法では、学習に必要なコーパスの量が不足しているため、データ拡張を行って学習した後、テストデータを用いて精度を確認した。拡張方法はモーラ単位で行った。構築した言い間違いコーパスをモーラに変換し、言い間違いが発生している個所の bi-gram をカウントし、出現回数が一定回数以上の bi-gram を変換ルールとした。この変換ルールを正しい発話 1 個に対して 1 回だけ適用しコーパスを拡張した。CNN と LSTM を用いた深層学習を行ってモデルを作成し、テストを行ったところ LSTM の場合に検出精度が約 70%となった。

## 5. 主な発表論文等

〔雑誌論文〕(計 件)

〔学会発表〕(計 7 件)

1. 大谷 優果, 篠山 学, 聞き返しに対する誤聴箇所の推定, FIT2016 (第 15 回情報科学技術フォーラム) 第 3 分冊, pp. 305-308, 2016 年 9 月.
2. 小松 聖矢, 篠山 学, 類似文字列検索を用いた言い間違いの検出と訂正, 平成 29 年度電気関係学会四国支部連合大会, pp. 234, 2017 年 9 月.
3. 片岡 麻輝, 篠山 学, 松本 和幸, 歌詞感情コーパスの拡張についての検討, 人工知能学会第 82 回 SLUD(言語・音声理解と対話処理)研究会予稿集, pp. 1-6, 2018 年 3 月.
4. 小松 聖矢, 篠山 学, 言い間違いの検出のためのデータ拡張, 平成 30 年度電気関係学会四国支部連合大会, pp. 246, 2018 年 9 月.
5. Kazuyuki Matsumoto, Manabu Sasayama, Lyric Emotion Estimation Using Word Embedding Learned from Lyric Corpus, IEEE 4th International Conference on Computer and Communications, Peer-reviewed, pp. 2295--2301, 2018.
6. Asaki Kataoka, Manabu Sasayama and Kazuyuki Matsumoto, Data Augmentation for Lyric Emotion Estimation, 2018 International Conference on Creativity, Innovation, and Invention in Digital Technology (CIIDT2018), Peer-reviewed, Dec. 2018.
7. 小松 聖矢, 篠山 学, 言語学的な単位に応じた言い間違いの検出, 情報処理学会第 81 回全国大会, pp. 353-354, 2019 年 3 月.

〔図書〕（計 件）

〔産業財産権〕

○出願状況（計 件）

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年：  
国内外の別：

○取得状況（計 件）

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年：  
国内外の別：

〔その他〕

ホームページ等

## 6. 研究組織

### (1) 研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号（8桁）：

### (2) 研究協力者

研究協力者氏名：

ローマ字氏名：

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。