

推薦論文

文のトピックを考慮した単語置換によるユーモア発話を行う 対話型エージェント

呉 健朗¹ 長岡 大二² 中原 涼太² 宮田 章裕^{2,a)}

受付日 2019年4月2日, 採録日 2019年10月3日

概要: 対話型エージェントの日常生活における存在感が増している。ユーザが親しみを感じやすい対話型エージェントの研究は多角的な視点で進められており、我々はエージェントがユーモアを提示することで親しみやすさを向上させるアプローチに取り組んできた。しかし、この先行研究を含む従来技術では、ユーザが発話中に用いる単語に特定のルールを設ける必要があり、一般的なコミュニケーションシーンに適用できないという問題がある。この問題を解決するために、文のトピックを考慮してボケるエージェントを提案する。これは、先行研究において聞き間違えを行う元の単語と、聞き間違えた結果の単語の選定時に文のトピックを考慮するようにしたものである。この提案により、ユーザが特定のルールに則った発言をしなくても、エージェントはユーモアを提示できるようになる。プロトタイプシステムを用いた検証実験の結果、従来方式よりも提案方式を用いた場合の方が、ユーザはエージェントに対して親しみを感じやすくなることが示唆された。

キーワード: 対話型エージェント, ボケる, トピック, word2vec

A Conversational Agent Replying with a Humor by Topic-word Substitution

KENRO GO¹ DAIJI NAGAOKA² RYOTA NAKAHARA² AKIHIRO MIYATA^{2,a)}

Received: April 2, 2019, Accepted: October 3, 2019

Abstract: Conversational agents are beginning to become popular in various scenes of daily life. There are a diversity of works on conversational agents that are easy for users to feel familiar with; we too have proposed an agent replying with a joke. However, existing works including our preceding work are hard to be applied to common communication situations because they require users to speak under specific rules. To address this issue, we notice mishearing style jokes performed by professional comedians. Based on this, we propose a conversational agent replying with a topic-sensitive joke. This makes it possible for the agent to bring humor into common situations. We developed a proof of concept and conducted evaluation tasks, confirming that our approach can realize a more friendly agent.

Keywords: conversational agent, joke, topic, word2vec

1. はじめに

対話型エージェントの日常生活における存在感が増して

¹ 日本大学大学院総合基礎科学研究科
Graduate School of Integrated Basic Sciences, Nihon University, Setagaya, Tokyo 156–8550, Japan

² 日本大学文理学部
College of Humanities and Sciences, Nihon University, Setagaya, Tokyo 156–8550, Japan

a) miyata.akihiro@acm.org

いる [1]. 家庭内における事例としては、Apple の Siri [2], Google Assistant [3] が有名である。介護などの人とのコミュニケーションが重要な場面での活躍も期待されており、ユーザが親しみを感じやすい対話型エージェントの研究は多角的な視点で進められている。我々も“笑い”を

本稿の内容は 2018 年 7 月のマルチメディア、分散、協調とモバイル (DICOMO2018) シンポジウムで報告され、グループウェアとネットワークサービス研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

エージェントによって引き起こすことで親しみやすさを向上させるアプローチに取り組んできた [4], [5]. “笑い” は、心理学的観点で人同士の親和的な関係の形成に欠かせないとされており [6], たとえば、リビングにコミュニケーションロボットが導入されたシーンを考えたとき、ロボットがつねに堅苦しい話ばかりでなく, “笑い” を引き起こすことができれば家族の一員として受け入れやすいだろう. あるいは、介護施設で高齢者の話し相手になるロボットが、会話中に生真面目なものばかりでなく、冗談を挟むことができればロボットは施設の人気者になれるだろう.

人同士だけでなく、人とエージェント間の関係でも同様の報告がされている [7], [8]. また、ユーザが継続して対話を続けたいと感じる要因としてユーモア表現を含む発話が有効なことが示されており、ユーモアが対話において重要な役割を持つことがうかがえる [9]. しかし、我々の先行研究を含む既存技術では、エージェントがユーモア発話を行うにはユーザが特定のルールに則った発言をする必要がある、日常生活に行うような一般的なコミュニケーションシーンに適用できないという問題がある.

この問題を解決するために、文のトピックを考慮してボケるエージェントを提案する. これは我々の先行研究 [4] の置換元単語の選定時とボケ単語の選定時において、文のトピックを考慮するようにしたものである. この提案により、ユーザが特定のルールに則った発言をしなくても、ユーモア生起に必要な概念の不適合を生成できるようになり、エージェントはユーモアを提示できるようになる.

本稿の貢献は下記のとおりである.

- 対話シーンにおいて、ユーザが特定のルールに則った発言を行わなくてもユーモア生成を行えるモデルを提案したこと.
- 上記提案概念を実現するプロトタイプシステムを構築し、提案概念の有効性を検証したこと.

2. 関連研究

既存の対話型エージェントはユーモア表現を行うかどうかで大別できる. 本章ではそれらの事例について紹介する.

2.1 ユーモア表現を行わないエージェントに関する事例

近年ではエージェントを対象とした研究は多岐にわたり、様々なシーンでの活用方法が検討されている.

武田らは、エージェントとの対話によりコミュニケーション不足を解決しようとする試みを行っている [10]. 1日の決まった時間にエージェントから利用者に対して会話を行うことで、継続的にコミュニケーションを促している. 野中らは、エージェントに対するユーザの反応を分析することで、健康状態・認知状態の把握を試みている [11]. 小倉らは、ユーザが1人の場合でも効果的な対話型鑑賞を行うための鑑賞支援システムを提案している [12]. 藤堂らは

ドライバの退屈や眠気を低減させる対話型 CG エージェントを提案している [13]. エージェントは車速、運転時間などに基づき、ユーザが飽きないような話題を提供する. 近藤らはスケジュール帳を情報源として、ユーザに余暇の過ごし方を提案するエージェントを提案している [14].

2.2 ユーモア表現を行うエージェントに関する事例

エージェントがユーザにユーモア表現を行う事例は数多く提案されており、これらは3つに大別できる.

1つ目はエージェントが規定のシナリオに沿ってユーモアを提示する方式である. ピンステッドらは複合語と複合語を構成する各語の類音語を見出しとする辞書を用意し、複合語の前半部や後半部を類音語に置換することで駄洒落の生成を行っている [15]. 青木らは、Web ニュース記事内の感情情報を用いて、おかしみの構造図に基づいた形式的なボケによる漫才台本を自動生成する方式を提案している [16]. 竹越らは漫才の自動生成システムを提案している [17]. 入力された文章から単語を選び、その単語をもとにボケを生成している. その後、ボケに対するツッコミ文を生成することで、漫才を生成している. Hayashi らは2体のロボットを用いて漫才を行うシステムを提案している [18]. ユーザの笑い声の大きさなどに反応し発話内容や発言区間を変化させている. 伊勢崎らはエージェントがユーザに笑い感情を誘起させる手段として大喜利を用いている [19]. お笑い構成作家と筆者らによって得られたデータベースをもとに、機械学習的アプローチに基づいてユーモアの生成を行っている. Tinholt らは文中の代名詞関係に着目してユーモア文を生成する試みを行っている [20]. 代名詞が指す正しい対象語とは異なる対象語を発見することで、実現しようとしている.

2つ目は、ユーザの特定の発言に対応するユーモアを提示する方式である. 仕様が開示されていないが、各企業の対話型エージェントはこの挙動を示すことが知られている [2], [3], [21]. たとえば、ユーザが“結婚して”というとき、エージェントは“友達の間までいましょう”と返答する. これらは、ユーザからの特定の入力に対して、あらかじめ設定された、ユーモアを含む応答文を返していると思われる.

3つ目は、ユーザの発言中の特定位置の単語を利用してユーモアを提示する方式である. 著者らの過去の研究では、エージェントが不適合解決モデル [22], [23], [24] に基づいてボケることでユーモアの生成を行っている [4]. 具体的には、エージェントがユーザの発言中の特定位置の単語を、わざと間違えて聞き返すことでユーモアを生成している.

3. 研究課題

今後、対話型エージェントは介護などの人間とのコミュニケーションが重要な場面での活躍も期待されており、様々

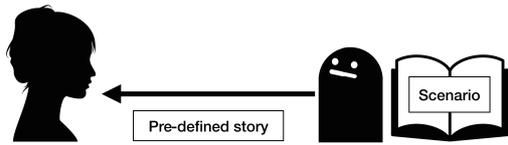


図 1 規定のシナリオに沿ってユーモア表現を行う方式

Fig. 1 A method for telling a story with humor based on a scenario.

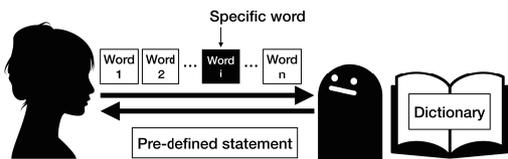


図 2 特定の発言に対応するユーモア表現を行う方式

Fig. 2 A method for replying with humor based on a specific word in user's statement.

な分野での活用方法が研究されている。笑いによってユーザと良好なコミュニケーションを築こうとするエージェントが研究されているが、一般的なコミュニケーションシーンに既存研究を適用するにはいくつかの問題がある。以降、2.2 節で 3 つに大別したユーモア表現方式をそれぞれモデル化し、問題点を詳細に分析する。

1 つ目のシナリオに沿ってユーモア表現を行う方式 [15], [16], [17], [18], [19], [20] は、図 1 のようにモデル化できる。エージェントはユーモア表現を含む規定のシナリオに沿って静的なストーリーをユーザに提示する。この方式は、事前に人手でシナリオを用意しなければならないという問題と、ユーザとエージェントの間で双方向コミュニケーションを行う対話シーンには適用できないという問題がある。

2 つ目のユーザの特定の発言に対応するユーモアを提示する方式 [2], [3], [21] は、図 2 のようにモデル化できる。エージェントは事前に単語（群）とユーモア表現を含む返答を対応付けた辞書を用意しておき、ユーザが特定の単語（群）を発言したら、対応する返答をユーザに提示する。この方式は、事前に人手で辞書を用意しなければならないという問題と、エージェントにユーモア表現を行わせるためにはユーザは特定の単語（群）を用いないとならないという問題がある。

3 つ目のユーザの発言の特定位置の単語を利用してユーモアを提示する方式 [4] は、図 3 のようにモデル化できる。まず、エージェントはユーザの発言の特定位置（例：先頭）の単語 w と、音が近く、意味が遠い単語 w' を単語コーパス内から選定する。次に、エージェントはユーザの発言中の w を w' に置換してユーザに聞き返すことで、ユーザの発言を面白く聞き間違えるスタイルのユーモア提示を行う。この方式は、既存の単語コーパス（Wikipedia コーパ

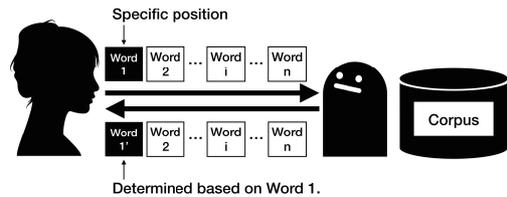


図 3 発言中の特定位置の単語を利用してユーモア表現を行う方式
Fig. 3 A method for replying with humor based on a word at a specific position in user's statement.

スなど）を利活用できるので人手がかかる事前作業は少ない。しかし、聞き間違えを行う単語の位置・種類を規定しているため、エージェントにユーモア表現を行わせるためにはユーザは特定位置に特定種類の単語を用いなければならないという問題がある。たとえば事例 [4] では、先頭に名詞を用いて発言しなければならない。

先行研究 [4] は、1 つ目と 2 つ目の問題を解決するものであった。すなわち、ユーザが特定シナリオに沿った発言をしたり、特定の単語を用いたりしなくてもユーモアを提示できるものであった。しかし、システムが聞き間違える単語の位置を事前に特定する必要があったため、ユーザは特定位置に特定種類の単語を用いる必要があるという問題があった。そこで本研究では、先行研究が持つこの問題を解決し、システムが聞き間違える単語の位置を事前に特定しなくてもユーモアを提示できる対話型エージェントの実現を研究課題とする。これにより、1 つ目～3 つ目の問題を解決でき、ユーザが特定のルールに則った発言を行わなくても、ユーモアを提示できる対話型エージェントを実現することが期待できる。

4. 提案方式

3 章で定義した研究課題を達成するにあたり、不適合解決モデル [22], [23], [24] に基づく先行研究 [4] の改良を行う。文献 [4] では、音が近く、意味が遠い単語に置換することで大きな不適合を生成できると思われる単語（以降、置換元単語）の位置を事前に指定してユーモアを生成していたため、ユーザが特定のルールに則った発言を行う必要があった。一般的な会話文からユーモアを生成するためには、ユーザに特定のルールを課することは不適切であるため、置換元単語を文中から自動で選定できるようにする必要がある。そこで、単語の位置が特定できない場合、どのようにすればユーモアの生起に不可欠とされる不適合が生成できるかを考える。検討にあたり、国内の人気漫才番組*1を観察したところ、国内の漫才大会で歴代上位 3 組（計 42 組）の漫才師が決勝戦で披露した漫才のうち、20 組の漫才師が文のトピックとなる単語を、その単語から意味がかけ離れた単語に聞き間違えたかのようにいい換えて返答するボケ

*1 M-1 グランプリ、エンタの神様、爆笑レッドカーペットなど。

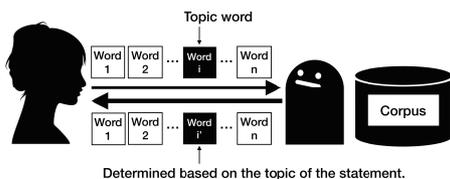


図 4 発言中のトピックを表す単語を利用してユーモア表現を行う方式

Fig. 4 A method for replying with humor using a word determined based on the topic of the statement.

を披露していた。たとえば、ピザを主題とする会話において“トッピングはエビで頼む”という発言に対し、“え、指をピザに入れるんですか？”と文のトピックである“エビ”を意味が離れた“指”に聞き間違えるボケを行っている。

以上より我々は、先行研究 [1] の置換元単語の選定時とボケ単語の選定時において文のトピックを考慮する、文のトピックを考慮してボケる対話型エージェントを提案する [5]。これは、まず、ユーザからの入力文のトピックに最も意味が近い単語を文中から選定し、それを置換元単語とする。次に文のトピックから意味が離れており、置換元単語に音が近い単語を選定し、それをボケ単語とする。最後にユーザからの入力文を、置換元単語をボケ単語に置換した状態で出力する (図 4)。

この提案により、ユーザが特定のルールに則った発言を行わなくても、エージェントはユーモア生起に必要な不可欠な不適合が動的に生成できるようになり、3章で定義した研究課題を達成できると考えられる。

5. プロトタイプシステムの実装

5.1 システムの全体像

提案システムは、置換元単語決定部とボケ単語決定部からなる。置換元単語決定部ではユーザの入力文中から、その文のトピックを最も代表する単語を選定する。1文からその文のトピック語を抽出する方法は、文書からトピックを抽出する方式 [25], [26], [27], [28] の類出語がトピックであるという考えを参考に、類出する類義語群の重心に最も近い語をトピック語と判定する。具体的には、文中の他の全単語との概念距離の平均値が近い単語ほど、文のトピックに意味が近いという考え方を参考に、入力文中の各単語がそれぞれ文のトピックとどれくらい意味が近いかを表す s_t (Topic score) を算出し、置換元単語を決定する。ボケ単語決定部では、文のトピックから意味が離れていて、置換元単語と音が近く、ユーザが理解できる単語を選定する。具体的には、置換元単語に対して、文のトピックからどれくらい意味が離れているかを表す s_s (Semantic score), どれくらい音が近いかを表す s_e (Edit distance score), どれくらいユーザから理解されやすいかを表す s_f (Frequency score) をもとにボケ単語を決定する。これらのスコアを算

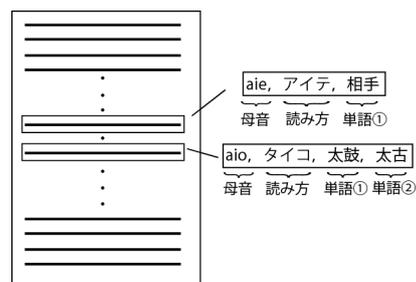


図 5 読み方辞書

Fig. 5 Dictionary with reading pronunciation.

出するための方法は 5.4 節に記す。一部、先行研究 [4] と重複する処理があるが、提案概念の説明に必要な部分についてはあらためて記載を行う。

5.2 事前準備

Wikipedia 日本語記事全文を MeCab [29] を用いて形態素解析し、不要品詞などを除去して分かち書きしたものをコーパスとする (1,004,870 単語)。ここでの不要品詞とは、語彙の意味を担わない助詞、助動詞などである。

また、コーパスをもとに読み方辞書と言語モデルの作成を行う。読み方辞書とは、コーパス内での出現回数が 1,000 回以上かつ日本語である単語の標準形、MeCab で取得した標準形の読み方 (カタカナ)、標準形の読み方をローマ字表記に変換したのち子音を削除して母音のみにしたものを、単語ごとにまとめたリストである (図 5)。本研究における言語モデルとは、単語間の概念距離を計算できるよう自然言語の単語をベクトル表現したものであり、word2vec [30] を用いてコーパスから獲得する。

5.3 トピックを考慮した置換元単語の決定

入力文中の各単語が、文のトピックとどれくらい意味が近いかを表す指標として、言語モデルを用いて文中の単語間の概念距離を測り、 s_t を算出する。これは、既存研究 [25], [26], [27], [28] を参考にし、文中の他の全単語との平均距離が近い単語ほど、文のトピックに意味が近いという考えに基づいている。 s_t を算出する手順としてまず、MeCab を利用して、入力文中の、すべての名詞、形容詞で構成される置換元候補群を作成する。その後、置換元候補群内のすべての単語の s_t を算出し、最も高い s_t を保持する単語を、最も文のトピックを表す単語とし、置換元単語として設定する。 s_t はそれぞれ文中の、自身を除いた名詞、形容詞との概念距離を計算し、その距離が近いほど高いスコアを与える。すなわち、 s_t は下記のように計算される。

$$s_{t(i)} = \frac{1}{n-1} \sum_{k=1, k \neq i}^n \frac{1}{d_{s(k)}} \quad (1)$$

n は文中のすべての名詞、形容詞の数であり、 i は s_t を算出する単語の位置である。 $d_{s(k)}$ は k 番目の単語と s_t を算

表 2 s_e が高い例・低い例

Table 2 Example of high/low score of s_e .

元の単語	母音	s_e が高い単語	母音	d_e	s_e が低い単語	母音	d_e
情報	ouou	乗法	ouou	0	総合警備保障	ououeioou	6

表 1 s_s が高い例・低い例 (入力文: 本を図書館で読む)

Table 1 Example of high/low score of s_s .

トピック	s_s が高い単語	s_s が低い単語
本	路面 (3.543)	書店 (2.282)

表 3 s_f が高い例・低い例

Table 3 Example of high/low score of s_f .

s_f が高い単語	s_f が低い単語
日本 (862,928 回)	詠嘆 (58 回)

出する単語との概念距離である。

5.4 トピックを考慮したボケ単語の決定

本節では、置換候補単語それぞれの各 Score (s_s, s_e, s_f) の計算方法と、それぞれ算出されたスコアをもとに出力文を決定する方法について述べる。ここで、置換候補単語とは、読み方辞書内の単語であり、ここからボケ単語を選定する。その際、先行研究 [4] と同様に、ユーザからの認知度が低い可能性がある人名・地名を置換元単語から外すなどの工夫を行う。なお、頭文字の母音と末尾の文字が置換元単語と一致する単語と、コーパス内での出現回数が 1,000 回以上である単語に絞る理由については、それぞれ 5.4.2 項と 5.4.3 項で述べる。以上より、置換元単語に対する置換候補単語となる条件は下記のようになる。

- MeCab で人名・地域と判定されない。
- 頭文字の母音と末尾の文字が置換元単語と一致する。
- コーパス内での出現回数が 1,000 回以上である。

5.4.1 s_s : 概念距離 Score

入力文のトピックと置換候補単語がどれくらい意味が離れているかを表す指標として、言語モデルを用いて概念距離を測ることで s_s を算出する。置換候補単語と、文中のすべての名詞、動詞、形容詞との概念距離を計算し、その距離が遠いほど高い s_s を与える (表 1)。すなわち、 s_s は下記のように計算される。

$$s_s = \frac{1}{n} \sum_{k=1}^n d_s(k) \quad (2)$$

n は文中の名詞、動詞、形容詞の数の合計値であり、 $d_s(k)$ は k 番目の単語と置換候補単語との概念距離である。

5.4.2 s_e : 編集距離 Score

置換候補単語と置換元単語の音がどれくらい近いかを表す指標として、編集距離^{*2}を測ることで s_e を算出する。置換候補単語と置換元単語の音がどれくらい近いかを表す指標として、編集距離を測ることで s_e を算出する。なお、編集距離測定時には、単語の母音のみを利用する。これは、試行錯誤 [4] の結果、2 つの単語の編集距離をそのまま計

^{*2} 2 つの文字列がどの程度異なっているかを示す距離であり、1 文字の挿入・削除・置換によって一方の文字列をもう一方の文字列に変形するのに必要な手順の最小回数として定義される。

測するよりも、各単語の母音どうしの編集距離を測定する方が、単語間の音の近さを計測しやすいと考えたためである。たとえば、“埋蔵金” という単語に音が近いのは“大貧民” と“映像” のどちらであるか測定する例で考える。通常どおりに“埋蔵金” との編集距離を算出すると“大貧民” が 4、“映像” が 3 となり、“映像” の方が距離が小さくなる。一方、単語の母音どうしの編集距離を測る場合、“埋蔵金 (aiouin)” との編集距離は、“大貧民 (aiinin)” が 2、“映像 (eiou)” が 3 となり、“大貧民” の方が距離が小さくなる。実際、著者ら複数名で確認したところ、母音どうしの編集距離が小さい単語どうしほど音が似通って聞こえる印象を持った。上記より、置換候補単語と置換元単語それぞれの母音どうしの編集距離 (d_e) を測り、その距離に近いほど高い s_e を与える (表 2)。すなわち、 s_e は下記のように計算される。

$$s_e = \frac{1}{1 + d_e} \quad (3)$$

また、試行錯誤の過程で、入力単語との編集距離が同じである単語の中でも聞き間違いに聞こえる単語とそうでない単語があることが分かった。たとえば、“戦闘機 (enoui)” という単語に対し、“倦怠期 (enaii)”, “関東 (anou)” という単語の各母音のみを考慮した編集距離はともに 2 であるが、前者は聞き間違いに聞こえ、後者はそうでないと感じられる人が多かった。多くの試行を行った結果、入力単語と出力単語の頭文字の母音と末尾の文字が一致する場合は、聞き間違いに聞こえやすいということが分かった。これに基づき、 s_e が高い場合でも、入力単語と先頭の文字の母音・末尾の文字が異なる単語は置換候補単語から除く。

5.4.3 s_f : 出現頻度 Score

出力候補単語の認知度を表す指標として、Wikipedia コーパス内での各単語の出現回数 (f) を数えることで s_f を算出する。各出力候補単語の出現回数が多いほど高い s_f を与える (表 3)。このとき、単語の出現回数はべき分布に従うため、ごく一部の単語の出現回数が極端に大きい。これらの単語が最終的な総合 Score に与える影響が大きくなりすぎないように、出現回数の対数をとったものを s_f とする。すなわち s_f は下記のようになる。

$$s_f = \log f \quad (4)$$

5.4.4 ボケ単語の決定

算出した3つのScore (s_s, s_e, s_f)を、それぞれ正規化・重み付けを行ったうえで線形和をとったものを、置換候補単語ごとの最終的な s_h (Humor score) とする。すなわち、 s_h は下記のようになる。

$$s_h = w_s s_s + w_e s_e + w_f s_f \tag{5}$$

w_s, w_e, w_f は重み係数である (本稿では重み係数はすべて1.0とする)。算出された s_h のうち、最も高い s_h を保持する置換候補単語を最も文のトピックから意味が離れており、置換元単語と音が近く、聞き慣れたボケ単語とし、文中の置換元単語と置換する。文中の置換元単語をボケ単語に置換した文を、出力文とする。下記に例を示す。

- 入力：水族館でペンギンは人気者だ
出力：水族館で【エンジン】は人気者だ
- 入力：学会の場所が温泉街だから嬉しいな
出力：学会の場所が【商店街】だから嬉しいな

6. 検証実験

提案方式の有効性を確認するには、ユーザが特定のルールに則った発言を行わないような一般的なコミュニケーションシーンでも、エージェントがユーザにユーモアを提示できるか検証する必要がある。そこで本実験では、このようなシーンの一例として、最近、または、近日中に起こる嬉しいことや悲しいことに対する自分の気持ちを我々の提案するエージェントに入力してもらい、その返答にユーモアや親しみを感じられるかどうかについて検証する。

6.1 実験条件

本実験の被験者は20代の学生14名である。実験では、置換元単語・ボケ単語選定時にトピックを考慮するか否かで分けた4つの方式 (BL, P1, P2, P3) を用いる。動作環境はCentOS 7, Intel Core i7 3.30 GHz, 16 GBである。全方式とも、エージェントの見た目・話し方などの外因を排除するため、入力はキーボード、出力はコンソール上でのテキスト表示とした。

BL：入力文中の先頭の単語を、その単語から意味が遠く、音が近い単語に置換する。先行研究 [4] の方式であり、本実験のBaselineである。

P1：入力文中の先頭の単語を、入力文のトピックから意味が遠く、音が近い単語に置換する。ボケ単語をトピックを考慮して選定する方式である。

P2：入力文のトピックを最も表す単語を、その単語から意味が遠く、音が近い単語に置換する。置換元単語をトピックを考慮して選定する方式である。

P3：入力文のトピックを最も表す単語を、入力文のトピックから意味が遠く、音が近い単語に置換する。置換元単語・ボケ単語をトピックを考慮して選定する方式である。

表4 被験者への質問一覧

Table 4 The question list of evaluation.

Q1	出力された返答にユーモアを感じたか
Q2	出力された返答は意外だと感じたか
Q3	エージェントがボケしていると認識できたか
Q4	エージェントに親しみを感じたか

6.2 実験の手順

被験者にはまず“最近または、近日中に起こる嬉しいことや悲しいことに対する自分の気持ちをエージェントに話すシーンを想像してください”と伝え、会話文を10件考えてもらった*3。その後、6.1節に示した4つの方式に対しそれぞれ被験者が考えた10件の会話文を入力してもらい、出力を確認してもらった。4つの方式に対し10回の入出力を行わせた後、表4に示すアンケートに5段階のリッカート尺度で回答してもらった。Q1~3については1つの入出力ごとに回答してもらい、それぞれの平均値をその方式における実験結果とした。Q4については各方式に対して1回ずつ回答してもらった。また、被験者が使用する方式が変わるたびに入力する文の順番を変え、順序効果を相殺した。さらに、被験者ごとに使用する方式の順番もランダムに入れ替えた。システムの仕様を聞いたことによって生じる先入観をなくすため、各方式の出力単語の生成方法は被験者には伝えなかった。被験者の考えた会話文と、その会話文を入力としたときの各方式における出力を一部抜粋したものを記す。大半のケースでは、例1・2のように、被験者の入力に対して各方式は異なる出力を行っていた。

- 例1
入力：先月、中学の友達と海外に行ってとても楽しかったな
BL：【伝説】、中学の友達と海外に行ってとても楽しかったな？
P1：【建設】、中学の友達と海外に行ってとても楽しかったな？
P2：先月、中学の【所在地】と海外に行ってとても楽しかったな？
P3：先月、中学の【装置】と海外に行ってとても楽しかったな？
- 例2
入力：好きなゲームの新作が発表されて嬉しい
BL：え？【月】なゲームの新作が発表されて嬉しい？
P1：え？【武器】なゲームの新作が発表されて嬉しい？
P2：え？好きなゲームの【進学】が発表されて嬉しい？
P3：え？好きなゲームの【神学】が発表されて嬉しい？

*3 実装の都合上、Wikipedia コーパス内に存在しない単語のみからなる文ではシステムが動作しないため、被験者がこのような文を考案した場合は再考してもらった。

一方、上述の指示にもかかわらず、一部の被験者が文中の単語数が2つ以下という極端に短い入力を行うこともあり、その場合は例3のように、各方式は同じ出力となることもあった。しかし、このような文中の文字数が10文字未満となる短い入力は全被験者の入力の4.3%であり、実験結果に与える影響は大きくないと考えられる。

● 例3

入力：金欠気味で辛い

BL・P1・P2・P3：え？【新設】気味で辛い？

6.3 実験結果・考察

被験者14名に考えてもらった会話文(計140文)について、置換元単語に対する置換候補単語数は平均312.5単語となり、計算時間は文中に含まれる単語1つにつき平均で0.016秒で処理できていた。続いて、トピックを考慮した置換元単語の決定が適切に行われているか検証する。提案方式では、各入力文中で s_t が最も高い単語が置換元単語となる。つまり、 s_t が最も高い単語とそれ以外の単語群の s_t 間に明確な差があるか否かが、置換元単語選定が適切に行われているか否かの1つの指標となる。これを確認するため、全入力文のうち置換候補単語数が3以上の文*4について、各文中で s_t が最も高い単語の s_t の集合と、それ以外の単語群の s_t の集合の間で、対応のないt検定を行った*5。その結果、集合間には1%水準の有意差が認められ、置換元単語とそれ以外の単語群の間には明確な差があることが確認できた。入力文と選定された置換元単語の例を下記に示す。これらの例のように、直感的にもトピックであると考えられる単語が置換元単語として選定できていることが分かる。

- 入力：車の教習所の授業が楽しかったよ
置換元単語：教習所
- 入力：学会の発表が無事に終わってよかったよ
置換元単語：学会
- 入力文：英語の論文締め切りがもうすぐで怯えているんだ
置換元単語：論文

Q1への回答を図6に示す。各方式における被験者の回答が4以上となった割合は、すべての方式で0%であった。なお、Wilcoxonの符号順位検定を行うと、BL・P3間、P1・P2間、P1・P3間で5%水準の有意差が認められた。また、BL・P2間で10%水準の有意傾向が認められた。ここから、置換元単語をトピックに最も近い単語にすることは、ユーザにユーモアを感じさせるうえで必要である可能性が示唆

*4 式(1)より、1文中の単語数が2である場合は全単語の s_t が等しくなり、1文中の単語数が1以下の場合 s_t が計算できないため、この分析では1文中の置換候補単語数が3以上のものを対象とした。

*5 文間で s_t を比較できるようにするため、各文中の全単語の s_t が平均0、分散1になるよう正規化を行った。

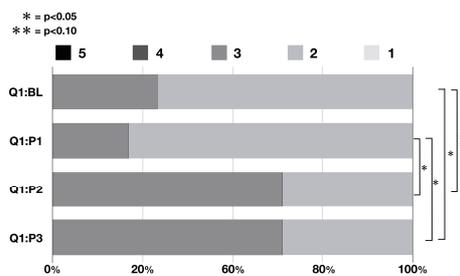


図6 Q1: 出力された返答にユーモアを感じたか (N = 14, 1: まったく思わない~5: とても思う)

Fig. 6 Q1: Did you feel humor in the response? (N = 14, 1: Strongly disagree to 5: Strongly agree).

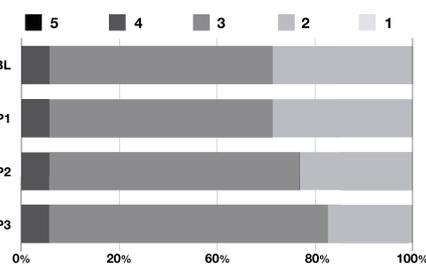


図7 Q2: 出力された返答は意外だと感じたか (N=14, 1: まったく思わない~5: とても思う)

Fig. 7 Q2: Did you feel surprising in the response? (N = 14, 1: Strongly disagree to 5: Strongly agree).

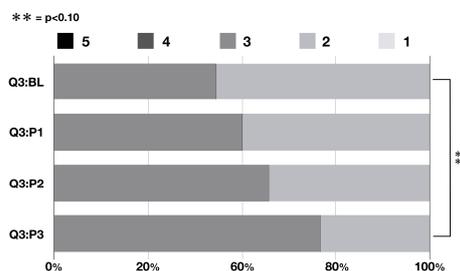


図8 Q3: エージェントがボケていると認識できたか (N = 14, 1: まったく認識できなかった~5: とても認識できた)

Fig. 8 Q3: Did you recognize that the agent was saying a joke? (N = 14, 1: Strongly disagree to 5: Strongly agree).

される。

Q2への回答を図7に示す。各方式における被験者の回答が4以上となった割合は、すべての方式で7.1%であった。また、Wilcoxonの符号順位検定を行ったが、全方式間で有意差は認められなかった。ここから、意外性の感じやすさには置換元単語、ボケ単語ともにトピックを考慮することは影響しないと考えられる。

Q3への回答を図8に示す。各方式における被験者の回答が4以上となった割合は、すべての方式で0%であった。なお、Wilcoxonの符号順位検定を行うと、BL・P3間で10%水準の有意傾向が認められた。ここからボケを認識させやすくするには置換元単語、ボケ単語ともにトピックを考慮する必要がある可能性が示唆される。

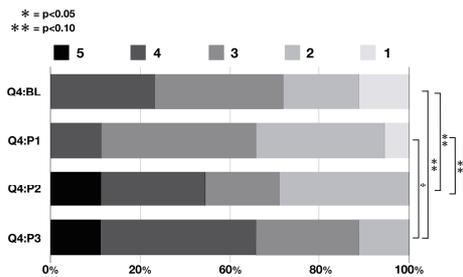


図 9 Q4: エージェントに親しみを感じたか (N = 14, 1: まったく思わない~5: とても思う)

Fig. 9 Q4: Did you feel familiar with the agent? (N = 14, 1: Strongly disagree to 5: Strongly agree).

表 5 s_s , s_e , s_f とユーモアの感じやすさの相関関係

Table 5 Correlation between s_s , s_e and s_f and humor.

スコア	相関係数
s_s	0.6360273856
s_e	0.3331779753
s_f	-0.3200686156

Q4 への回答を図 9 に示す. 各方式における被験者の回答が 4 以上となった割合は, BL で 28.6%, P1 で 14.3%, P2 で 42.9%, P3 で 57.2% であった. Wilcoxon の符号順位検定を行うと, P1・P3 間で 5% 水準の有意差が認められた. また, BL・P2 間, P1・P2 間, BL・P3 間で 10% 水準の有意傾向が認められた. ここから, エージェントに親しみを感じさせやすくする方式として, 置換元単語をトピックに最も近い単語にすることはある程度有効であると考えられる.

以上より, 一般的なコミュニケーションシーンにおいてエージェントがユーモア発言を行う際に, 置換元単語選定時に文のトピックを考慮することで, ユーザはエージェントに親しみを感じやすくなると思われる. この傾向が普遍的なものかどうか今後も検証を行う必要がある.

最後に, ボケ単語選定時に s_s , s_e , s_f のすべてのスコアが, ユーモア・親しみを感じやすくするために有効であったかどうかを検証する. P3 で選出されたボケ単語の s_s , s_e , s_f とユーモアの感じやすさとの相関関係を表 5 に示す. 表 5 に示すように, ユーモアの感じやすさについて, s_s は正の相関, s_f は弱い正の相関が確認された. なお, s_f は弱い負の相関が確認された. これは, 文のトピックとなる単語を別の単語に置換するとき, あまりにも聞き慣れている単語に置換された場合は, ユーモア生起に不可欠な不適合をうまく生成できなかつたためだと考えられる. この問題は, 出現回数が一定以上となる単語を候補から外すことで解決する. P3 で選出されたボケ単語の s_s , s_e , s_f と親しみの感じやすさとの相関関係を表 6 に示す. 表 6 に示すように, 親しみの感じやすさについて, s_s は弱い正の相関, s_e は正の相関があることを確認できた. ここから,

表 6 s_s , s_e , s_f と親しみの感じやすさの相関関係

Table 6 Correlation between s_s , s_e and s_f and friendliness.

スコア	相関係数
s_s	0.3090030438
s_e	0.4124568514
s_f	0.0431604857

親しみを感じやすくするために, ボケ単語選出時に s_s , s_e を考慮することは有効であったと考えられる. なお, s_f と親しみの感じやすさについては相関が見られなかった. 今後は, s_f への重み付けの見直しや, ボケ単語選定時には s_f を考慮しないようにするなどの改良を行っていく.

7. おわりに

本稿では対話シーンにおいて, ユーザが特定のルールに則った発言を行わなくてもユーモアを提示できる対話型エージェントを実現するために, 先行研究 [4] を改良し, 文のトピックを考慮してボケるモデルを提案した. 具体的には, 置換元単語をユーザから入力された会話文のトピックに最も意味が近い単語にし, ボケ単語を文のトピックから意味が離れた単語にするというものである. 検証実験では一般的なコミュニケーションシーンにおいて, 従来方式よりも提案方式を用いた場合の方が, ユーザはエージェントに対し, ユーモア・親しみをそれぞれ感じやすくなることを確認できた. 今後の課題はいくつかある. ユーモア生成のアルゴリズムについては, ユーザの直前の発話のトピックのみを考慮しているため会話のコンテキストをとらえていない可能性がある. このことから現在の 1 文のトピックを考慮してボケている方式を改良し, 過去の複数文のコンテキストを考慮してボケるエージェントの実現を行っていく必要がある. システムの入出力については, 現在の入出力はテキストのみであり, 多様なモダリティを考慮できていない. すなわち, ユーザの見た目・表情に基づいてユーモアを生成したり, 生成したユーモアの効果を高めるために表情・ジェスチャを駆使したりすることはできていない. このことから, 生成したユーモアの効果を高められるような入出力の方法について検討を続ける必要がある.

参考文献

- [1] McDuff, D. and Czerwinski, M.: Designing emotionally sentient agents, *Comm. ACM*, Vol.61, pp.74-83 (2018).
- [2] Apple inc.: Siri, available from (<https://www.apple.com/ios/siri/>) (accessed 2019-04-02).
- [3] Google inc.: Google Assistant, available from (<https://assistant.google.com/>) (accessed 2019-04-02).
- [4] 呉 健朗, 中原涼太, 長岡大二, 中辻 真, 宮田章裕: ボケて返す対話型エージェント, 日本バーチャルリアリティ学会論文誌, Vol.23, No.4, pp.231-238 (2018).
- [5] 長岡大二, 中原涼太, 呉 健朗, 鈴木 奨, 宮田章裕: 文脈を考慮してボケる対話型エージェントの実装と評価, マルチメディア, 分散, 協調とモバイルシンポジウム論

文集 2018, Vol.2018, pp.627-632 (2018).

[6] 井上 宏:「笑い学」研究について, 笑い学研究, No.9, pp.3-15 (2002).

[7] Khooshabeh, P., McCall, C., Gandhe, S., et al.: Does it matter computer jokes, *Proc. CHI '11*, pp.77-86 (2011).

[8] Lugar, E. and Sellen, A.: Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents, *Proc. CHI '16*, pp.5286-5297 (2016).

[9] 宮澤幸希, 常世 徹, 榊井祐介ほか: 音声対話システムにおける継続欲求の高いインタラクションの要因, 電子情報通信学会論文誌, Vol.J95-A, No.1, pp.27-36 (2012).

[10] 武田風太, 鈴木慎太郎, 矢島敬士: 対話型キャラクターエージェントを用いたコミュニケーションによる高齢者の見守りシステム支援, 情報処理学会第 78 回全国大会 2016, Vol.78, pp.401-402 (2016).

[11] 野中裕子, 酒井洋一, 安田 清ほか: 語りかけエージェントとの会話における韻律情報の分析, 情報処理学会第 75 回全国大会 2013, Vol.75, pp.179-180 (2013).

[12] 小倉拓人, 徳永隼人, 松村冬子ほか: 進行役と鑑賞者のエージェントを用いた対話型鑑賞システム, じんもんこん 2015 論文集, Vol.2015, pp.43-50 (2015).

[13] 藤堂祐樹, 西村良太, 山本一公ほか: ドライバの退屈・眠気回避を目的とした会話型 CG エージェントに関する検討, 情報処理学会第 65 回全国大会 2003, Vol.2003, pp.411-412 (2003).

[14] 近藤亮太, 畑中聖二, 加藤誠巳: 余暇の過ごし方を提案する会話型エージェントシステムに関する検討, 情報処理学会第 69 回全国大会 2007, Vol.2007, pp.543-544 (2007).

[15] ビンステッド キム, 滝澤 修: 日本語駄洒落なぞなぞ生成システム “BOKE”, 人工知能学会誌, Vol.13, pp.920-927 (1998).

[16] 青木 亮, 義 尚晃, 原口和貴ほか: 理解しやすい対話を用いた漫才台本の自動生成, データ工学と情報マネジメントに関するフォーラム (DEIM Forum), C1-5 (2018).

[17] 竹越智也, 萩原将文: ロボット漫才自動生成システム—動作が漫才に与える影響の考察, 日本感性工学会論文誌, Vol.15, No.1, pp.47-54 (2016).

[18] Hayashi, K., Kanda, T., Miyashita, T., et al.: Robot conversation as a passive-social medium, *International Journal of Humanoid Robotics*, Vol.5, No.1, pp.67-86 (2008).

[19] 伊勢崎隆司, 小林明美, 有賀玲子ほか: お題に対してユーモアを生起する回答文選択の検討, マルチメディア, 分散, 協調とモバイルシンポジウム論文集 2018, pp.643-648 (2018).

[20] Tinholt, H.W. and Nijholt, A.: Computational humour: Utilizing cross-reference ambiguity for conversational jokes, *International Workshop on Fuzzy Logic and Applications*, Vol.4578 of LNCS, pp.477-483 (2007).

[21] Softbank: Papper, available from <https://www.softbank.jp/robot/> (accessed 2019-04-02).

[22] Coulson, S. and Williamsab, R.F.: Hemispheric Asymmetries and Joke Comprehension, *Neuropsychologia*, Vol.43, No.1, pp.128-141 (2005).

[23] Samson, C.A., Hempelmann, C.F., Hubera, O., et al.: Neural Substrates of Incongruity Resolution and Nonsense Humor, *Neuropsychologia*, Vol.47, No.4, pp.1023-1033 (2009).

[24] 伊藤大幸: ユーモアの生起過程における論理的不適合及び構造的不適合の役割, 認知科学, Vol.17, No.2, pp.297-312 (2010).

[25] 松尾 豊, 石塚 満: 語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム, 人工知能学会論文誌, Vol.17, pp.217-223 (2002).

[26] 松山 学, 平岡佑介, 渡邊 倫ほか: 収集論文を利用したキーワード抽出に基づくユーザプロファイルの生成について, 情報処理学会第 66 回全国大会, Vol.66, pp.125-126 (2004).

[27] 武田善行, 梅村恭司: キーワード抽出を実現する文書頻度分析, 情報処理学会研究報告自然言語処理, Vol.2001, pp.27-32 (2001).

[28] 福田雅志, 延澤志保, 太原育夫: 話し言葉における出現位置情報を用いたキーワード抽出, 情報処理学会研究報告自然言語処理, Vol.2005, pp.1-6 (2005).

[29] Mecab: Yet another part-of-speech and morphological analyzer, available from <http://taku910.github.io/mecab/> (accessed 2019-04-02).

[30] Mikolov, T., Chen, K., Corred, G., et al: Efficient estimation of word representations in vector space, *Proc. ICLR '13* (2013).

推薦文

本稿は, 対話エージェントの返答に対して文脈を考慮する方式に関する検討に関するものであり, 新規性の面から特筆すべきものであるため推薦する。

(グループウェアとネットワークサービス研究会主査
齊藤 典明)



吳 健朗 (学生会員)

2018 年日本大学文理学部情報科学科卒業。同年日本大学大学院総合基礎科学研究科博士前期課程に進学, 現在に至る。2017 年 VR 学会サイバースペース研究賞, 2018 年情報処理学会 GN 研究賞, 山下記念研究賞。ヒューマンコンピュータインタラクションの研究に従事。



長岡 大二

2019 年日本大学文理学部卒業。



中原 涼太 (正会員)

2019 年日本大学文理学部卒業。2019 年情報処理学会山下記念研究賞。



宮田 章裕 (正会員)

日本大学文理学部情報科学科准教授.
2005年日本電信電話株式会社入社.
2008年慶應義塾大学大学院博士課程
修了. 2016年より現職. ヒューマン
コンピュータインタラクションの研
究に従事. 情報処理学会 2017年度・

2018年度論文賞. ACM, 日本バーチャルリアリティ学会,
ヒューマンインタフェース学会, 日本データベース学会各
会員. 博士 (工学). 本会シニア会員.