

# 用例に基づく疑似的な聞き間違いによりユーモア表現を生成する対話システム

## A Dialogue System Making Humor by Pseudo Mishearing of a Part of Users' Queries Based on Examples

関 陽介  
Yosuke Seki

徳島大学 高等教育研究センター  
Research Center for Higher Education, Tokushima University  
seki@tokushima-u.ac.jp

**Keywords:** dialogue system, making humor, mishearing

### Summary

Dialogue systems, which give users quick and easy access to required information interactively, have been widely used in various fields. Dialogue systems equipped with interfaces (e.g., humanoid robots and anthropomorphic agents) have been developed in order to enhance familiarity and dialogue continuity. Related studies, in which interactive agents generate humor expressions, have also been reported. Humor is indispensable for the formation of friendly relationships between people and systems, and humor expressions can be applied in situations that generate familiar responses and provide fun to users. In this study, in order to evoke humor through dialogue, a method to generate humorous expression by asking again due to pseudo mishearing of a part of users' queries based on examples is proposed. Specifically, a conversion candidate dictionary for humor expressions, based on Wikipedia of Japanese edition and a classification vocabulary table in which words are classified semantically, is created by word completion using distributed representation. In addition, a word conversion method is designed by approximately 1,000 mishearing survey from Twitter, and the function based on the proposed method is implemented in a dialogue system introduced into a university as a model case. In the results of the comparative evaluation with other methods quantitatively, the proposed methods gave users the most humor by converting singular and multiple words. Thus, the effectiveness of the proposed method was clarified.

## 1. はじめに

ユーザの質問に機械的に回答する対話システムは、様々な分野に導入されており注目を集めている。例えば、Apple の Siri や Google Assistant, 学術機関においては AI-campus<sup>\*1</sup> やとくぼん talk<sup>\*2</sup> など、オープンドメインの雑談や対話的指示、タスク指向型の情報提供などに利用されている。

自然言語処理や人工知能等の分野では、システム技術としての対話システムの導入、実践報告やデザイン研究が増加している[Chen 17]。例えば、親近感や対話持続性を高めるために、ヒューマノイドロボットや擬人化エージェントなどのインターフェースを備える対話システムが開発[西村 04]されている。また、ユーモアは心理学の観点で人とシステム間の親和的な関係形成に必要とされており[Khooshabeh 11, Luger 16], 対話エージェントがユーモア表現を生成する研究(2章で述べる)が報告されている。

ユーモア表現を生成する方法は多様に存在するが、[織田 86]は音韻の類似性が笑いを喚起すると述べている。例えば“ゴディバのチョコをください”という質問に対して“え?ゴリラのチョコをください?”と聞き間違えた場合、“ゴディバ”と“ゴリラ”の音韻の近さ、さ

らには語意の違いによりユーモアが生まれる。このようなユーモア表現を雑談ロボットなどに導入することで、親しみのある応答生成やユーザに楽しみを与える場面に活用できる。

そこで、本研究では先行研究[呉 18, 呉 20]を発展させて、対話を通してユーモアを喚起するために、ユーザ発話文(以下、発話文)の一部をシステムが疑似的に聞き間違えることで、ユーモア表現を含んだ聞き返し文を生成する方法を提案する。具体的には、Wikipedia と分類語彙表、さらには分散表現を用いた語彙検索によりユーモア表現の変換候補辞書を作成し、聞き間違いの用例調査に基づく単語変換方法を設計する。そして、本学の在学生向け対話システムをモデルケースとして、提案手法に基づく機能を実装する。

本研究の貢献は以下になる。

- 聞き間違いの用例を Twitter から約 1,000 件収集して、変換位置や変換内容、変換対象、品詞、変換単語数、文字長、意味的分類、音韻の類似度、出現頻度などを調査したこと。
- Wikipedia と分類語彙表、さらには分散表現を用いた語彙検索により、ユーモア表現の変換候補辞書を作成したこと。
- 用例の調査結果に基づき単数と複数の単語を対象にした変換方法を設計したこと。
- 先行研究等との比較評価により、提案手法の有効

\*1 AI-campus <http://www.saga-u.ac.jp/admissions/>

\*2 とくぼん talk <https://taiwa.honbu.tokushima-u.ac.jp/taiwa/ui/toku>

性を示したこと。

- 本学の在学生向け対話システムをモデルケースとして、提案手法に基づく機能を実装したこと。

## 2. 先行研究

ユーモア表現の生成に関する研究は、発話文を利用しない方法と利用する方法に大別できる。

発話文を利用しない研究としては、[ビンステッド 98]は日本語駄洒落なぞなぞ生成システムを提案しており、辞書から取得した複合語の一部、または複合語を構成する一語を類義語に置き換えることで駄洒落を生成している。[Hayashi 08]はユーザの反応に応じて発話内容や発言区間を変化させるロボット漫才システムを提案しており、[Petrovic 13]は単語の共起頻度や曖昧性等に着目して、Google N-gram データセット<sup>\*3</sup>から駄洒落を生成している。ただし、これらの研究はユーザとの対話を通してユーモア表現を生成しないため、本研究に適用することは困難である。

発話文を利用する研究として、[He 19]は WordNet を用いた単語間の類似度や応答文の例文検索などにより、与えられた同音異義語とその関連語を含んだユーモアな文を生成している。また、Wikipedia 英語記事を使用して、[Luo 19]は敵対的生成ネットワーク、[Yu 18]は LSTM を用いた駄洒落の生成手法を提案している。ただし、これらの研究は音韻を対象としてなく、本研究が目指す聞き間違いによるユーモア表現の生成は困難である。

特定の単語の音韻に着目して、類似する別の単語に置き換えることで笑いを生み出す研究が報告されている。[吉田 12]はボケとツッコミにより笑いを生み出す漫才形式の対話文生成システムを提案している。4 種類の変換語の選択方法を提案しており、その中の一つは DP マッチングや音韻の相違度に関するペナルティを用いてボケに用いる単語を取得している。[Valitutti 13]は音韻辞書等を用いて、ある単語を音韻やアクセントの位置が類似する別の単語に変換することで、ユーモア表現を生成している。[山根 12]は音の長さや音韻類似度、意味類似度、イメージのし易さ、具象度を考慮することで、ユーモアを含む返答文を生成している。発話文の一部に対してユーモアのある聞き間違いをして聞き返すボケをする研究として、[呉 18]は編集距離、概念距離、出現回数を用いて単語変換しており、[呉 20]は対象文のトピックとの意味が離れており音韻が異なる単語を変換語として用いている。以上の研究を本研究に適用することで、対話を通してユーモア表現を生成でき、親和的な関係構築や対話持続性の向上が期待できる。

一方で、ユーモア表現の生成には言語処理や辞書作成など多様な方法が用いられているが、聞き間違いの用例を対象にした統計的分析やこの調査結果に基づく手法

設計は報告されていない。例えば実際の会話で“集中講義の日程を教えて”を“牛乳講義の日程を教えて?”と一部の単語を聞き間違えた場合、何がどのように変換するかの変換過程を実例に基づき調査することで、システム上でよりユーモア表現に適した聞き返し文を生成できる。そのため、本研究では聞き間違いの用例を収集して変換過程を調査することで、ユーモア表現に適する変換方法を設計する。

## 3. 用例調査に基づく手法設計

### 3.1 聞き間違いの用例調査

ソーシャルセンサーに関する研究では、Twitter の分析による道路交通情報や震源地の推定方法が報告されており、信頼性が高い情報を収集できることが示されている[Sakaki 10, 榊 15]。そこで本研究の用例調査においても、Twitter を用いて聞き間違いに関する日本語のツイートを収集する。

収集期間は 2011 年 9 月 18 日から 2020 年 6 月 3 日までとする。ハッシュタグ (#) を用いて“#聞き間違い”で検索して、聞き間違いの元となる単語と聞き間違えた単語が示されたツイートを手作業で全て収集した。重複を除いて収集した件数は約 1,000 件になる。本研究では用例の説明において、聞き間違いを変換、聞き間違いの元となる単語を変換元、聞き間違えた単語を変換語と表現する。表 1 は収集した用例の変換例を示す。変換元と変換語のローマ字表記を促音は **k**、撥音は **n**、長音は直前の母音を重ねて表現する。例えば“ハーゲンダッツ”は“ha a ge n da k tsu”になる。

単語の変換位置は先頭、中間、語尾、変換方法は追加、削除、置換、変換対象は母音、子音、母音・子音に分類できる。例えば (3) は、先頭の母音・子音置換により“ダチョウ倶楽部 (da cho u ku ra bu)”が“都庁クラブ? (to cho u ku ra bu)”に変換されている。(1~9) の組合せである (10) は、先頭の子音置換と中間の子音置換・削除により、“シークワサー (shi i ku wa a sa a)”が“キングアサー? (ki n gu a a sa a)”に変換されている。(11) は同一の音韻であるが語意が異なる変換で、“下水道”が“ゲス移動?”に変換されている。

MeCab (mecab-ipadic-NEologd 辞書、以下略) を用いて変換元を形態素解析した結果、品詞は名詞または複合名詞 (“芋洗い”や“平日のり弁当”など) が全体の 88.2% を占めていた。また、変換語においてもこれらが全体の 86.2% を占めていた。NEologd 辞書の名詞には約 30 程度の細分類 (サ変接続や固有名詞/組織など) が定義されている。変換語を対象に名詞の細分類を調査した結果、上位から一般 (ゴリラや映画館など) が 36.1%、固有名詞/地域/一般が 11.1%、代名詞/一般が 9.2% を占めていた。一方で、接尾/地域 (兵庫県神戸市など) や接続詞的 (兼や対など) などは該当しなかった。

<sup>\*3</sup> Google N-gram データセット  
<http://storage.googleapis.com/books/ngrams/books/datasetsv2.html>

表1 収集した用例の変換例

番号	変換位置と変換内容	変換対象	変換元	変換語
1	先頭の追加	子音	アイス a i su	ライス? ra i su
2	先頭の削除	母音・子音	手すり te su ri	スリ? su ri
3	先頭の置換	母音・子音	ダチョウ倶楽部 da cho u ku ra bu	都庁クラブ? to cho u ku ra bu
4	中間の追加	母音・子音	博多 ha ka ta	母方? ha ha ka ta
5	中間の削除	母音	来院 ra i i n	Line? ra i n
6	中間の置換	母音	apply pay a k pu ru pe i	アップルパイ? a k pu ru pa i
7	語尾の追加	母音・子音	ひも hi mo	干物? hi mo no
8	語尾の削除	母音・子音	ニート ni i to	二位? ni i
9	語尾の置換	母音・子音	小娘 ko mu su me	小結? ko mu su bi
10	(1~9の組合せ) 先頭置換, 中間置換・削除	子音	シークワサー shi i ku wa a sa a	キングアーサー? ki n gu a a sa a
11	語意変換	-	下水道 ge su i do u	ゲス移動? ge su i do u

MeCab を用いて単語数を調査した結果、変換元と変換語の単語数が異なる組合せが見られた。例えば“厚生労働省”から“高性能ロボット?”は1単語から2単語、“小型家電”から“紅茶花伝?”は2単語から1単語に変換されている。全体の割合は、名詞を対象にすると1単語から1単語が59.4%、1単語から2単語が8.5%、2単語から1単語が5.5%であった。他には、“愛してください”(動詞助詞動詞の3単語)から“アイスティーください?”(名詞動詞の2単語)など、様々な品詞の組合せがある。ただし、このような組合せが占める割合は全て1%以下であった。以降では1単語から1単語の変換を単数変換、2単語以内において変換元と変換語の単語数が異なる変換を複数変換と呼ぶ。

文字列間の異なり具合を示す編集距離を用いて、変換元と変換語における音韻の類似度を調査した。本距離は1文字の追加・削除・置換により、一方の文字列を他方の文字列に変換するために必要な回数を求めている。本研究では単語のローマ字表記を対象に編集距離を求める。例えば“tako (タコ)”と“takosu (タコス)”は2、“koora (コーラ)”と“koara (コアラ)”は1になる。単語*i, j*の音韻の類似度 $s_{ij}$ は編集距離 $l_{ij}$ を式(1)で正規化して求める。 $s_{ij}$ を調査した結果、全体の平均値(標準偏差, 以下略)は0.35 (0.2)であった。これは平均2文字程度の異なりを意味する。なお、全体の81.6%は $s_{ij}$ が0.2以上であり、大部分の変換において4文字以内のローマ字が異なっていた。

$$s_{ij} = \frac{1}{1 + l_{ij}} \quad (1)$$

先行研究[呉 18, 呉 20, 山根 12]では、変換元と変換

語の「先頭の母音と語尾の文字」や「先頭と語尾の母音」の一致具合を考慮して単語変換している。本研究でも聞き返し文を生成する上で、これらの変換位置や変換対象を考慮した変換が重要になる。そこで、変換位置を先頭と語尾、変換対象を文字と母音として調査した結果、「先頭の文字と語尾の母音」が一致は33.7%、「先頭の母音と語尾の文字」が一致は46.1% (前者との重複含む)、「先頭の文字のみ」が一致は13.1%、「語尾の文字のみ」が一致は16.1%であり、他の組合せの割合は微小であった。

変換元と変換語のひらがな表記を対象に文字長を調査した。変換元の最大文字数は“宮中晩餐会”の12文字、最小文字数は“ハグ”等の2文字、平均値は4.77 (1.82)であった。変換語の最大文字数は“暴力温泉芸者”(アーティスト名)の13文字、最小文字数は“ナス”等の2文字、平均値は4.81 (1.83)であった。なお、2単語から1単語の変換における変換元の最大文字数は“犬 700 円”の9文字、最小文字数は“象さん”の4文字、平均値は5.73 (1.54)、1単語から2単語の変換における変換元の最大文字数は“厚生労働省”の11文字、最小文字数は“論文”等の4文字、平均値は6.37 (1.78)であった。

### 3.2 意味的分類と出現頻度

音韻の類似度のみによる変換は音の近さしか考慮しないため、ユーモア表現に適さない単語が選ばれる可能性が高い。例えば、“教員”が“キャイ〜ン?”と“劉殷?”に変換された場合、音韻の類似度は同じであるが、後者は認知度が低く日常で使用される一般的な単語ではないため、ユーモア表現としては適さない。

そこで、どのような単語が変換語に用いられているかを調査するために、分類語彙表 [国立国語研究所 04] (以下、語彙表) を用いて変換語を分類する。語彙表では単

語が意味的に分類されており、101,070 件の単語が 896 項目に分類されている。例えば“明治維新”の分類項目が時代、“露天ぶろ”が部屋・床・廊下・階段などに分類されている。

語彙表は 2018 年に更新されているが、特に芸能人や歌手、作品名などの新しい固有名詞は十分に反映されていない。そこで、MeCab を併用して固有名詞（一般、人名、組織、地域）を判定する。例えば“板東英二”は語彙表には未登録であるが、MeCab では固有名詞（人名）と判定される。固有名詞（人名）は語彙表の固有人名に相当する。ただし、固有名詞（一般）は他の固有名詞（人名、組織、地域）とは異なり、様々な単語が該当する。例えば、“青菜漬け”、“akb48 劇場公演”、“京阪 1900 系”は全て固有名詞（一般）になるが、それぞれの属性は大きく異なる。そこで、固有名詞（一般）と他品詞において、語彙表に未登録の単語は分散表現を用いて類似語を抽出する。そして、上位の類似語から語彙表を検索して、一致する単語の分類項目を取得する。例えば、語彙表に未登録である“ゴジラ”の類似語は、上位から“ゴジラ対メカゴジラ”、“ガメラ”…“怪獣”、“ゴジラ vs デストロイア”と出力される。この際、上位から語彙表を検索して、登録されている“怪獣”の分類項目（哺乳類）を“ゴジラ”の分類項目として用いる。事前調査では上位 200 件までの類似語が適切な分類項目を取得できる傾向にあった。そのため、上位 200 件までの類似語で分類項目を取得できない場合は終了とする。図 1 に全体の流れを示す。図中の DB は変換語の分類結果を記録する。

分散表現の学習モデルを作成するために、2020 年 4 月時点で最新の Wikipedia 日本語記事全文を用いた。MeCab で単語を分かち書きして、次元数は 300 で fastText (skip-gram) を用いて学習モデルを作成した。本モデルを用いることで、例えば“キンチョール”から“虫よけ”、“スーパードライ”から“生ビール”の類似語を取得で

き、語彙表による分類が可能になった。なお、分散表現により取得した分類項目の精度は 3.3 節で確認する。

表 2 は該当数が多い変換語の分類項目を示す。上位から料理（イタリア料理など）、哺乳類（ゴリラなど）、固有人名（ディズニーなど）の順に該当数が多くなった。一方で、農林水産業や信仰・宗教など該当数が 1 回の分類項目が全体の 48.3% を占めていた。未該当の分類項目は弛緩・粗密・繁簡、身上、叙述などがある。これらには“稠密”、“下賤”、“叙情”などの単語が語彙表に登録されており、日常での使用頻度が低いであろう単語が多くを占めていた。

用例の 99.7% は変換元と変換語の分類項目が異なる結果となった。これは、変換元から意味がかけ離れることで、よりユーモア表現に適する単語に変換されたと考えられる。例えば用例の 1 つに“ソルティライチが飲みたい”が“鈴木大地が飲みたい？”に変換されており、清涼飲料水からプロ野球選手に語意が大きく変わることでユーモアを喚起している。

先行研究[呉 18, 呉 20]では認知度が高い単語に変換するために、Wikipedia 内の出現頻度が 1,000 回未満の単語を除外している。出現頻度が低くても認知度が高い単語が存在する可能性はあるが、大まかな認知度の推測に出現頻度は利用できる。また、分類項目と併用することで、認知度が高く一般的な単語の絞り込みができ、よりユーモア表現に適する単語に変換できると考えられる。そこで、先行研究で用いられる基準値は根拠が明確でないため、本研究では変換語の出現頻度を調査する。Wikipedia を用いて調査した結果、最大は“人間”の 67,141 回、最小は“ペヨンジュン”の 1 回、平均値は 5,800 (17,723) であった。標準偏差は平均値より高く、出現頻度は単語によりばらつきが大きい。

次に、分類項目毎に変換語の出現頻度を調査した結果、分類項目が料理の平均値は 487.6(512.6)、哺乳類は 4,507.6(7,095)、固有人名は 2,705.9(2,607.6)であった。表

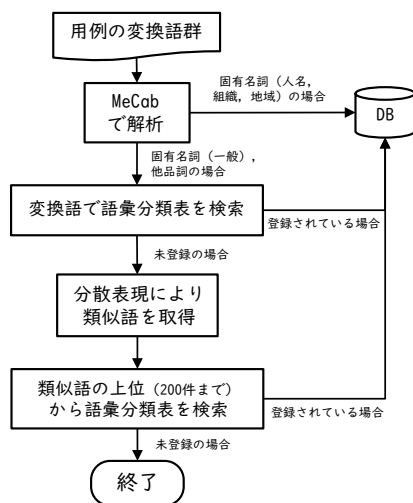


図 1 意味的分類の流れ

表 2 変換語の分類項目（一部）

分類項目	例
料理	イタリア料理 メンチカツ
哺乳類	ゴリラ, なまけもの
固有人名	ディズニー, 森鷗外
固有地名	アフリカ, インド
魚・肉	いくら, ハム
菓子	ソフトクリーム, 水菓子
神仏・精霊	悪魔, 精霊
専門的・技術的職業	坊主, 保険医
スポーツ	ウォーキング, ダンス
飲料・たばこ	生ビール, コーヒー
草本	オクラ, トマト
魚類	鯛, ブリ

表 3 変換候補辞書の例

単語	品詞 (細分類)	ローマ字表記	出現頻度	分散表現による類似語	語彙表の分類項目
桜	名詞, 一般	sa ku ra	7,086	-	木本
ゴジラ	名詞, 固有名詞, 一般	go ji ra	3,860	怪獣	哺乳類
ラーメン	名詞, 一般	ra a me n	3,275	-	料理
木村拓哉	名詞, 固有名詞, 人名, 一般	ki mu ra ta ku ya	73	-	固有人名
魔法使いサリー	名詞, 固有名詞, 一般	ma ho u tsu ka i sa ri i	114	アニメ	演劇・映画

2 の上位 3 件に限定して記載したが、全体的に標準偏差が高く分類内においても出現頻度の差は大きい傾向であった。例えば分類項目が神仏・精霊においては、“神”の出現頻度は 40,772 であり“御主”は 51 であった。

### 3.3 変換候補辞書の作成

3.2 節で述べた分散表現を用いた語彙表の検索方法を用いて、Wikipedia と語彙表からユーモア表現の変換候補辞書を作成する。助詞や助動詞等の語彙的意味を有しない品詞を除いて、重複を除き収集した単語は 2,416,204 件になった。辞書は変換候補の単語、MeCab で判定した品詞 (細分類)、ローマ字表記、Wikipedia 内の出現頻度、分散表現による類似語 (語彙表に未登録の場合のみ)、語彙表の分類項目を記録する。表 3 に変換候補辞書の例を示す。例えば、“ゴジラ”は品詞 (細分類) が“名詞, 固有名詞, 一般”, ローマ字表記が“go ji ra”, 出現頻度が 3,860, 類似語が“怪獣”, 分類項目が哺乳類になる。

約 19 万件の単語が分散表現を用いた語彙検索の対象となった。取得した分類項目の精度を確認するために、無作為抽出した 1,000 件の単語を手作業で調査した結果、87%の単語は正確に分類されていた。一方で、“カメラを止めるな! (映画)”から“クランクアップ”や“七大陸最高峰”から“登頂”など、関連性はあるが属性が異なるものがあつた。分散表現は分布仮説に基づき算出されており、類似する文脈に出現する単語が類似語として選ばれる。本研究では上位類似語から語彙表を検索しており、“映画”や“エベレスト”等は下位に出力されたことで、期待する類似語が得られなかったと考えられる。ただし、関連性が見られることやこのような類似語が占める割合は多くはないため、全ての結果を利用する。

### 3.4 変換元の決定方法

変換元の決定には、文のトピックに最も意味が近い単語を変換元とする先行研究[呉 20]の方法を用いる。これはトピックに近い単語を変換元として選択することで、発話文の中心的な意味を表す語を特定している。具体的には、発話文中のすべての名詞と形容詞を変換元候補群として、郡内の全単語において最も高い $t_i$ を保持する単

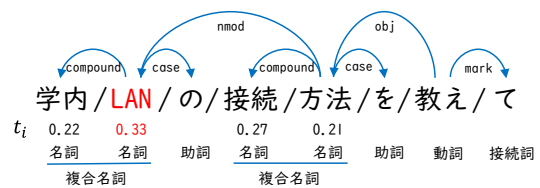


図 2 変換候補の決定方法の例

語を変換元とする。 $t_i$ は自身を除いた全単語との概念距離が近いほど高いスコアになる。これは文中の他の単語との概念距離に近い単語ほど、文のトピックに意味が近いという考え方に基づく。 $t_i$ は式(2)を用いて算出する。 $n$ は文中の名詞と形容詞の総数、 $i$ は $t_i$ を算出する単語の位置、 $d_{ij}(0 \leq d_{ij} \leq 1)$ は $i, j$ 番目の単語間の概念距離を示す。概念距離は 3.2 節で作成した学習モデルを用いた分散表現の余弦類似度とする。

$$t_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^n d_{ij} \quad (2)$$

構文解析には spaCy を用いる。図 2 に変換候補の決定方法の例を示す。図中の compound は複合語を意味する。発話文が“学内 LAN の接続方法を教えて”の場合、 $t_i$ は“学内”が 0.22, “LAN”が 0.33, “接続”が 0.27, “方法”が 0.21 になる。単数変換の場合は、 $t_i$ が最も高い“LAN”が変換元になる。複数変換で変換元が 2 単語の場合は、 $t_i$ が最も高い単語が含まれる複合名詞を変換元とする。例の場合では、“学内 LAN”と“接続方法”が複合名詞と判定されるが、 $t_i$ が最も高い“LAN”が含まれる“学内 LAN”を変換元とする。

### 3.5 単語の変換条件

3.1 節, 3.2 節の調査結果を踏まえて、単語の変換条件は以下とする。なお、各調査対象において用例に該当する割合が低いものは除く。

#### 【単数変換】

- ① 変換元と変換語の品詞は 2 文字以上 (ひらがな) の名詞。

※用例で未使用の細分類 (名詞) は除く。

② 「先頭の文字と語尾の母音」, 「先頭の母音と語尾の文字」, 「先頭の文字のみ」, 「語尾の文字のみ」のいずれかに一致する単語.

③  $s_{ij}(\geq 0.2)$ が最も高い単語.

④ 用例で用いられる分類項目に該当し, 出現頻度が該当する分類の平均値以上の単語. 変換元と同じ分類は除く.

※ $s_{ij}$ が同じ場合は用例の上位分類項目(表 2(一部))から該当する単語を選択. 同分類に該当する場合は出現頻度が高い単語を選択.

※用例で該当件数が低い分類項目は除く.

【複数変換】 ※①~④も適用

⑤ (2 単語から 1 単語変換) 変換元が 2 単語で計 4 文字以上 9 文字以下(ひらがな)の場合は 1 単語に変換.

※文字数により条件③を満たさない場合があるため音韻の類似度の条件(0.2 以上)は除く.

⑥ (1 単語から 2 単語変換) 変換元が 1 単語で 4 文字以上 11 文字以下(ひらがな)の場合は 2 単語に変換. 変換元の区切り位置はランダムとするが, 各単語の最小文字数は 2 文字以上.

### 3.6 システム開発

徳島大学では各種手続きや留学制度, システムの利用方法など, 大学生活に関する質問に回答する在学生向け対話システム(学内版とくぼん talk<sup>\*4</sup>)を 2020 年 4 月から導入している. 本システムは与えられた質問に回答する一般的な対話システムであるが, 親近感の向上を目的に提案手法に基づく機能を実装した<sup>\*5</sup>. 参考画面を図 3



図 3 ユーモア表現を生成する対話システム

に示す. 本図は視認性を高めるために一部の画像を加工している. 本システムでは下から上に対話が続いており, 例えばユーザが“留学について知りたい”と質問すると, とくぼんが“え? 牛肉について!? ...あ, 留学ですね, 留学は~”と回答する. 聞き間違いの頻度は暫定的に 3 回の質問を受ける度に 1 回としており, 単数変換と複数変換はランダムに切り替えている. 導入後のユーザ評価により適切な頻度や変換方式の選択方法を決定する.

## 4. 評価実験

提案手法の有効性を検証するために, 単数変換と複数変換の比較評価を行う.

### 4.1 単数変換の評価

被験者は大学生 20 名で変換方法が異なる以下 6 方式を比較評価で用いる. P1 と P2 間で音韻の類似度(3.5 節の③), P2 と P3 間で分類項目(出現頻度)による絞り込み等(④)を評価して, P3 と P4~P6 間で提案手法と先行研究を比較する. なお, 変換位置と変換対象(②)の評価は, 先行研究[呉 18, 呉 20, 山根 12]で類似手法が用いられており, 本研究においても有効性は明らかと考えられるため対象外とする. 比較対象の先行研究は, 再現する上で必要な情報が記載されているものに限定する. 変換方法を評価するために全方式で変換元(名詞)は同じとして, 辞書は 3.3 節で作成した変換候補辞書を用いる. 変換対象の発話文は学内版とくぼん talk の発話履歴を用いる. 被験者毎に発話文 10 件をランダムに抽出して, 被験者は各方式で生成される変換語を含んだ応答文を評価する. 順序効果と視覚的な影響を避けるために, 応答文の提示順は被験者毎に変更し, 同一画面に全ての発話文と応答文を表示する.

P1: 「先頭の文字と語尾の母音」, 「先頭の母音と語尾の文字」, 「先頭の文字のみ」, 「語尾の文字のみ」のいずれかに一致する単語をランダムに選択.

P2:  $P1 + s_{ij}(\geq 0.2)$ が最も高い単語.  
※ $s_{ij}$ が同じ場合は出現頻度が高い単語.

P3: 【提案手法】P2+用例で用いられる分類項目に該当し, 出現頻度が該当する分類の平均値以上の単語. 変換元と同じ分類は除く.

※ $s_{ij}$ が同じ場合は用例の上位分類項目(表 2(一部))から該当する単語を選択. 同分類に該当する場合は出現頻度が高い単語を選択.  
※用例で該当件数が低い分類項目は除く.

P4: 【[呉 18] の方式】変換元から意味が遠く音が近い単語に変換.

P5: 【[呉 20] の方式】変換対象の文のトピックから意味が遠く音が近い単語に変換.

P6: 【[吉田 12] の方式の一つ】DP マッチングと相違度に関するペナルティにより単語変換.

<sup>\*4</sup> 学内版とくぼん talk [https://www.ait.tokushima-u.ac.jp/service/list\\_out/](https://www.ait.tokushima-u.ac.jp/service/list_out/)

<sup>\*5</sup> 執筆時点では機能実装したシステムは未導入である.

評価項目は先行研究 [呉 18, 呉 20, 山根 12] を参考に以下の通り作成して, 評価尺度は 5 段階 (1 (感じなかった) ~5 (とても感じた)) とした. これらの評価項目はユーモアの面白さを脳科学の立場から分析した研究[Goel 01]に基づいている.

- Q1: ユーモアを感じたか.
- Q2: 意外性を感じたか.
- Q3: 一般的(日常で使用される語, 理解が難しくないと感じたか.

各方式の平均値を表 4 に示す. 検定は有意水準 5% で多重比較法 (Bonferroni 法) を用いた. 表 5 に評価実験で用いた発話文と応答文例 (簡略) を示す.

Q1 は P3 の平均値が最も高く, P3 と他の全方式間で有意差が認められた. そのため, P3 はユーザが最もユーモアと感じる応答文を生成できていた. 一方で, P2 は P1 の平均値より若干高いが有意差は認められなかった. これは, 特定の変換位置と変換対象の一致が, 両方式の変換条件に含まれることが要因として考えられる. ただし, 音韻の類似度を変換条件に加えることで, 若干ではあるがユーモア性の向上に影響を与えている. P3 と P2 間には有意差が認められており, 分類項目 (出現頻度) で単語を絞ることで, P3 ではユーモア表現に適する単語に変換できたと判断できる.

Q2 は P1 の平均値が最も高く, P1 と P4~P6 間, P2 と P5 間, P3 と P4~P5 間で有意差が認められた. P1 は変換条件に該当する単語をランダムに選択しており, 例えば “履修” から “アンディラウ?”, “ウイルス” から “トニー・ブース?”, “インフル” から “フラジール?” 等の変換がみられた. P1 は出現頻度を考慮しないため認知度が低い単語が選ばれ, ユーザが意外性を高く感じた可能性がある. 有意差は認められなかったが, P3 は P2 の平均値より若干高い. 変換語は変換元と同じ分類を除いたことが, 意外性を高めた一つの要因として考えられる.

Q3 は P3 の平均値が最も高く, P1 と P3~P6 間, P2 と P3~P6 間, P3 と P4 間で有意差が認められた. P2 と P3 においては, 例えば変換元が “履修”, “成績”, “大学” の場合, P2 では “李秀?”, “西征?”, “蔡鏢?”, P3 では “英雄?”, “正妻?”, “弾薬?” に変換された. 分類項目 (出現頻度) で単語を絞ることで, P3 ではユーザがより一般的と感じる単語に変換されたと判断できる.

以上より, P3 においては変換位置や変換対象, 音韻の類似度や分類項目などの変換条件を用例調査に基づき設計することで, より意外性があり一般的な単語に変換され, 最もユーモアが喚起されたと判断できる.

#### 4.2 変換候補辞書の作成における追加評価

変換候補辞書の作成 (3.3 節) において, 前節の P2 と P3 の比較で分類項目の有無を評価したが, 本節では分散表現を用いた語彙表の検索方法を評価する. 比較方式は P3 と以下を用いる.

- P3': P3 と同一. ただし, 分散表現による検索方法を用いていない変換候補辞書を使用.

被験者は大学生 11 名として, 学内版とくぼん talk の発話履歴からランダムに抽出した発話文 100 件を変換対象とする. この際, 両方式の変換語は 85% が同一であったため, 変換語が異なる応答文のみを被験者が評価する. 応答文の提示順と表示方法は 4.1 節と同一とする.

P3 と P3' の平均値を表 6 に示す. 検定は有意水準 5% で対応のある t 検定を用いた. 全ての評価項目において P3 の平均値が最も高く, Q2 でのみ有意差が認められた. 例として, “ログイン ID” は P3 が “おしん ID?”, P3' が “教員 ID?” に変換された. 分散表現を用いて変換候補辞書が拡張されたことで, 多様な表現が可能になりユーモア性などが高い単語に変換されたと考えられる. また, 3.3 節では正確に分類されなかった単語が確認されたが, 大きな影響はなかったと判断できる.

表 4 単数変換における各方式の平均値 (標準偏差)

	P1	P2	P3	P4	P5	P6
Q1	2.79 (1.51)	2.85 (1.33)	3.33 (1.25)	2.7 (1.38)	2.71 (1.34)	2.72 (1.4)
Q2	3.51 (1.28)	3.28 (1.38)	3.36 (1.19)	2.74 (1.36)	2.73 (1.34)	2.88 (1.5)
Q3	1.84 (1.23)	2.11 (1.4)	3.69 (1.35)	3.34 (1.53)	3.39 (1.55)	3.15 (1.63)

表 6 単数変換における P3 と P3' の平均値 (標準偏差)

	P3	P3'
Q1	2.9 (1.41)	2.48 (1.14)
Q2	3.86 (1.08)	3 (0.92)
Q3	4.48 (0.79)	4.1 (1.41)

表 5 単数変換における発話文と応答文例 (簡略)

発話文	P1	P2	P3	P4	P5	P6
メールの見方を教えて	ニールの見方?	メールの見方?	メイクの見方?	JR の見方?	レベルの見方?	ネールの見方?
留学について知りたい	目目について?	宗学について?	牛肉について?	不足について?	住宅について?	遊学について?
履修の方法について	アンディラウの方法?	李鄒の方法?	英雄の方法?	利用の方法?	利用の方法?	履中の方法?

表 7 複数変換 (2 単語から 1 単語) における各方式の平均値 (標準偏差)

	P3 (単)	P3 (複)	P4	P5	P6
Q1	2.85 (1.23)	2.87 (1.3)	2.46 (1.26)	2.67 (1.23)	2.38 (1.25)
Q2	3.13 (0.94)	3.41 (0.87)	3.31 (1.16)	3.38 (1.12)	3.13 (1.2)
Q3	3.64 (1.25)	3.69 (1.4)	3.87 (1.34)	3.9 (1.32)	3.1 (1.6)

表 9 複数変換 (1 単語から 2 単語) における各方式の平均値 (標準偏差)

	P3 (単)	P3 (複)	P4	P5	P6
Q1	2.53 (1.2)	3.1 (1.08)	2.38 (1.09)	2.36 (1.07)	2.24 (1.03)
Q2	2.67 (1.01)	3.42 (1.13)	3.34 (1.07)	3.34 (1.07)	3.88 (1.05)
Q3	3.67 (1.7)	2.1 (1.22)	2.42 (1.42)	2.52 (1.37)	1.52 (1)

表 8 複数変換 (2 単語から 1 単語) における発話文と応答文例 (簡略)

発話文	P3 (単)	P3 (複)	P4	P5	P6
登録期間はいつ?	盗塁期間はいつ?	トウモロコシはいつ?	6 年間はいつ?	6 年間はいつ?	高速期間はいつ?
学務関係の窓口はどこ?	ガム関係の窓口?	精神状態の窓口?	山岳地帯の窓口?	魔法使いの窓口?	悪夢探偵の窓口?
応用理数のガイダンスはいつ?	相撲理数のガイダンス?	プロ野球のガイダンス?	糖尿病のガイダンス?	女性声優のガイダンス?	ゆうりすうのガイダンス?

表 10 複数変換 (1 単語から 2 単語) における発話文と応答文例 (簡略)

発話文	P3 (単)	P3 (複)	P4	P5	P6
オリエンテーションはいつ?	—	おでんテンションはいつ?	4 年計算はいつ?	4 年計算はいつ?	李媛ケーソンはいつ?
食堂は空いている?	症状は空いている?	ショップ銃は空いている?	属党は空いている?	属党は空いている?	北句網頭は空いている?
大学院の入試はいつ?	映画館の入試?	さい一品の入試?	さい 3 年の入試?	タイ 3 年の入試?	亜維アクインの入試?

### 4.3 複数変換の評価

4.1 節と同じ方法で複数変換 (3.5 節の⑤, ⑥) を評価する。被験者は前節と同じ 11 人で、比較方式は P3 の単数変換 (以下, P3 (単)), P3 の複数変換 (以下, P3 (複)), P4, P5, P6 とする。P3 (単) のみ 1 単語から 1 単語の変換になる。P4~6 は複数変換を対象としないが、一部変更<sup>\*6</sup>または変更せずに適用できるため比較方式に用いる。P3 (単) を除く 1 単語から 2 単語の変換において、変換元の区切り位置は同一とする。

2 単語から 1 単語の変換において、各方式の平均値を表 7 に、評価実験で用いた発話文と応答文例 (簡略) を表 8 に示す。Q1 と Q2 は P3 (複) の平均値が最も高くなった。Q3 は P5 の平均値が最も高く、P4~P5 と P6 間で有意差が認められた。Q1 と Q3 の P3 (単) と P3 (複) の平均値は微差であり、被験者により両方式の評価傾向は異なっていた。例として、“登録期間はいつ?” は P3 (単) が“盗塁期間?”, P3 (複) が“トウモロコシ?” に変換された。P3 (単) は変換元の語意をやや保ちつつ、変換語と未変換語の不自然な組合せによりユーモアを喚起している。P3 (複) では変換元の文字数が多いため音韻の類似度は低くなったが、変換元との語意が大きく異なることでユーモア表現を生成している。そのため、P3 (単) と P3 (複) を併用することで、ユーモア表現に

多様性を与えることができると考えられる。

1 単語から 2 単語の変換において、各方式の平均値を表 9 に、評価実験で用いた発話文と応答文例 (簡略) を表 10 に示す。Q1 は P3 (複) の平均値が最も高く、P3 (複) と P4~P6 間で有意差が認められた。4.1 節の評価では P3 の変換語が最もユーモアを喚起しており、このような単語の不自然な組合せにより、P3 (複) はユーモア性が高い単語変換ができたと考えられる。Q2 は P6 の平均値が高く、P3 (単), P4~P5 と P6 間で有意差が認められた。P6 は Q3 の平均値が最も低いため、一般性が低い単語の組合せにより意外性が高くなったと考えられる。Q3 は P3 (単) の平均値が最も高く、P3 (単) と他区分間、P3 (複) と P6 間、P4~P5 と P6 間で有意差が認められた。P3 (単) では、他の方式と異なり単語の不自然な組合せが起こらないため、最も一般的と感じられたと考えられる。

P3 (単) の変換において、全体の 33% が変換元の文字数が多いため変換条件を満たさなかった。一方で、P3 (複) は全て変換された。例えば、“オリエンテーション” は P3 (単) では変換されず、P3 (複) では“おでんテンション” に変換された。P3 (複) は変換元の文字数の影響を受けにくく、様々な発話文に対応できると判断できる。ただし、両変換語が同じ分類の場合、ユーモアが低減する可能性がある。例えば、“おでんテンション?”ではなく“おでん定食?”に変換された場合、組合せの意外性が大きく下がる。そのため、変換元と変換語に加えて、変換語間でも分類が異なる単語を選ぶことが望ましい。

<sup>\*6</sup> P4 においては、A, B, D を変換元、C, E, F を変換候補の単語、— を変換とした場合、A, B→C (2 単語から 1 単語) における C の概念距離は A・C と B・C の平均値、D→E, F (1 単語から 2 単語) における E または F の概念距離は D・E または D・F の値とする。P5 は「A, B」または D を 1 つのトピックとして扱い、前述した方法で概念距離を算出する。



以上より、複数変換においても P3 (複) は最もユーモアが喚起されており、また 2 単語から 1 単語変換では P3 (単) と P3 (複) を組み合わせることで、多様なユーモア表現が可能になると判断できる。

#### 4.4 今後の課題

本研究では名詞を対象にしているが、用例では他の品詞も使用されている。例えば“アフリカゾウ”を“炙りカツオ?” (動詞名詞) や“ポムポムプリン”を“パリパリチキン?” (副詞名詞) など、動詞や副詞、感動詞等が用いられている。変換元と変換語の対象に名詞以外の品詞も加えることで、ユーモア表現を広げることができる。例えば学内版とくぼん talk では、初めに“こんにちは”等の挨拶文を入力するユーザがいる。この際、感動詞を対象に単語変換することで最初に笑いを取ることができ、システムに対する興味を高められる可能性がある。

本研究では分類別に出現頻度を設定したが、出現頻度が低い単語にも認知度がありユーモア表現に適する単語は存在する。例えば、用例では“横断歩道”から“織田無道?” に変換されていたが、変換語の出現頻度は 12 であった。これは世代にもよるが、バラエティ番組に出演していた印象や多少の認知度などが要因として、ユーモア表現に繋がったと考えられる。変換候補辞書には固有人名が占める割合は 24.7% であり、低頻度でもユーモア表現に適する単語は多く存在する可能性が高い。従って、出現頻度とは異なる観点でユーモア表現に適する単語判定の仕組みが求められる。これを実現する一つの方法として変換候補辞書の拡充がある。例えば、芸能人辞典<sup>\*7</sup>は TV や CM などの出演履歴が記録されている。この履歴から出演番組や紹介商品の注目度などに着目することで、多角的に認知度を推測できる可能性がある。

本研究では低頻度の単語は変換語に用いないため、変換可能な単語数は絞られる。そのため、出現頻度の設定における別課題として、異なる変換元から同一単語に変換される頻度の増加がある。例えば提案手法では“退学”と“音楽”の変換語は“山岳?” になるが、比較手法では全て異なる単語に変換された。同一の聞き間違いはユーザに飽きを与える可能性がある。そのため、前述した通り出現頻度とは異なる基準設定や、一度応答した変換語は同一ユーザには用いない等の対策が必要になる。

ユーモア性の向上はシステム導入する上で重要になる。ただし、ユーモアと感じる表現は個人差があるため、全てのユーザに同一基準で単語を変換することは適切ではない。例えば評価実験において、“大学入門について知りたい”の応答文である“え? 弾薬入門について知りたい?” は、ユーザにより評価値が大きく異なっていた。よりユーモアを喚起するためには、個々の好みを考慮した単語変換が求められる。ユーザの嗜好を分析して類似

するユーザを選び出し、それらの嗜好パターンに基づいて情報推薦する協調フィルタリング [Herlocker 99] がある。オンラインストアの商品推薦やネットニュースの記事推薦などで用いられており、ユーザの嗜好を考慮した情報推薦ができる。この技術を用いることで、例えば優先的に動物や芸能人に関する単語に変換するなど、ユーザ単位で単語の変換条件を変えることができる。また、特に学内版とくぼん talk のように教育機関にシステムが導入される場合は、不適切な変換を防止する必要がある。例えば、特定の単語や分類項目の犯罪・罪は使用しないなどが有効と考えられる。

#### 5. おわりに

本研究では先行研究[呉 18, 呉 20]を発展させて、対話を通してユーモアを喚起するために、発話文の一部をシステムが疑似的に聞き間違えることで、ユーモア表現を含んだ聞き返し文を生成する方法を提案した。具体的には、Wikipedia と分類語彙表、さらには分散表現を用いた語彙検索によりユーモア表現の変換候補辞書を作成し、聞き間違いの用例調査に基づく単語変換方法を設計した。そして、本学の在学生向け対話システムをモデルケースとして、提案手法に基づく機能を実装した。

本学で稼働する在学生向け対話システムの発話履歴を用いた先行研究等との比較評価では、変換位置や変換対象、音韻の類似度や分類項目 (出現頻度) などの変換条件を設計することで、提案手法がユーモアの喚起に有効であることを確認できた。また、単数変換と複数変換の併用がユーモア表現に多様性を与えることや、複数変換は変換元の文字数の影響を受けにくいため様々な発話文に対応できることを確認できた。

提案手法に基づく機能を実装した対話システムを試行評価した結果、ユーザからは「聞き間違いが面白く笑ってしまった。」や「とくぼんが可愛く感じた。」などのコメントがあった。ただし、ユーモアが全く感じられない応答もあり、そのような場面では聞き返し文はユーザに不快感を与えていた。そのため今後の課題としては、さらなるユーモア性の向上が求められる。具体的には、変換元と変換語の対象品詞の拡張、変換候補辞書の拡充、協調フィルタリングなどの推薦アルゴリズムを用いた個々の好みを考慮した変換方法の検討などが挙げられる。また、変換候補辞書はオープンドメインにも対応しているため、雑談ロボット等に提案手法に基づく機能を実装して、対話持続性の向上評価などを行いたい。

#### 謝 辞

本研究は JSPS 科研費 19K14317 の助成を受けたものです。

<sup>\*7</sup> 芸能人辞典 <https://www.oricon.co.jp/prof/>

## ◇ 参考文献 ◇

- [ビンステッド 98] 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, No. 6, pp. 920-927 (1998)
- [Chen 17] Chen, H., Liu, L., Yin, D. and Tang, J.: A survey on dialogue systems: recent advances and new frontiers, *ACM SIGKDD Explorations Newsletter*, Vol. 19, No. 2, pp. 25-35 (2017)
- [呉 18] 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)
- [呉 20] 呉健朗, 長岡大二, 中原涼太, 宮田章裕: 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, 情処学論, Vol. 61, No. 1, pp. 113-122 (2020)
- [Goel 01] Goel, V. and Dolan, J. R.: The functional anatomy of humor: segregating cognitive and affective components, *Nature Neuroscience*, Vol. 4, No. 3, pp. 237-238 (2001)
- [Hayashi 08] Hayashi, K., Kanda, T., Miyashita, T., Ishiguro, H. and Hagita, N.: Robot manzai: robot conversation as a passive-social medium, *International Journal of Humanoid Robotics*, Vol. 5, No. 1, pp. 67-86 (2008)
- [He 19] He, H., Peng, N. and Liang, P.: Pun generation with surprise, *Proc. of NAACL'19*, pp. 1734-1744 (2019)
- [Herlocker 99] Herlocker, J. L., Konstan, J.A., Borchers, A. and Reidl, J.: An algorithmic framework for performing collaborative filtering, *Proc. of SIGIR'99*, pp. 230-237 (1999)
- [Khooshabeh 11] Khooshabeh, P., McCall, C., Gandhe, S., Matthew, Gratch, J. and Blascovich, J.: Does it matter if a computer jokes, *Proc. of CHI'11*, pp. 77-86 (2011)
- [国立国語研究所 04] 国立国語研究所『分類語彙表増補改訂版データベース』(ver.1.0) (2004), [https://pj.ninjal.ac.jp/corpus\\_center/goihyo.html](https://pj.ninjal.ac.jp/corpus_center/goihyo.html)
- [Luger 16] Luger, A E. and Sellen, J A.: "Like having a really bad pa": the gulf between user expectation and experience of conversational agents, *Proc. of CHI'16*, pp. 5286-5297 (2016)
- [Luo 19] Luo, F., Li, S., Yang, P., Li, L., Chang, B., Sui, Z. and Sun, X.: Pun-GAN: generative adversarial network for pun generation, *Proc. of EMNLP-IJCNLP*, pp. 3388-3393 (2019).
- [西村 04] 西村竜一, 西原洋平, 鶴身玲典, 李晃伸, 猿渡洋, 鹿野清宏: 実環境研究プラットフォームとしての音声情報案内システムの運用, 信学論, Vol. 87, No. 3, pp. 789-798 (2004)
- [織田 86] 織田正吉: 笑いとユーモア, 筑摩書房 (1986)
- [Petrovic 13] Petrovic, S. and Matthews, D.: Unsupervised joke generation from big data, *Proc. of ACL'13*, pp. 228-232 (2013)
- [Sakaki 10] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors, *Proc. of WWW'10*, pp. 851-860 (2010)
- [榊 15] 榊剛史, 柳原正, 那和一成, 松尾豊: Twitter を用いた道路交通情報の抽出, 信学論, Vol. J98-D, No. 6, pp. 1019-1032 (2015)
- [Valitutti 13] Valitutti, A., Toivonen, H., Doucet, A., Toivanen, M. J.: "Let everything turn well in your wife": generation of adult humor using lexical constraints, *Proc. of ACL'13* (2013).
- [山根 12] 山根宏彰, 萩原将文: 笑いを生むことわざかしの自動生成システム, 日本知能情報ファジィ学会誌, Vol. 24, No. 2, pp. 671-679 (2012)
- [吉田 12] 吉田裕介, 萩原将文: 漫才形式の対話文自動生成システム, 日本感性工学会論文誌, Vol.11, No.2, pp. 265-272 (2012)
- [Yu 18] Yu, Z., Tan, J. and Wan, X.: A neural approach to pun generation, *Proc. of ACL'18*, pp. 1650-1660 (2018).

(担当委員: 井上 昂治)

2020 年 10 月 14 日 受理

## —— 著 者 紹 介 ——



関 陽介 (正会員)

徳島大学高等教育研究センター特任講師。2008 年甲南大学理工学部情報システム工学科卒業。2014 年徳島大学大学院先端技術科学教育部博士前期課程修了。2017 年同大学院博士後期課程単位修得退学。2008 年から IT 企業での勤務を経て、2011 年から 2016 年まで徳島大学情報センター特任助教。2017 年より同大学総合教育センター特任研究員。2018 年より現職。大学入試、対話システムなどの研究に従事。電子情報通信学会、情報処理学会、日本教育工学会各会員。