

# 音素類似関係を用いた聞き間違いによるボケの自動生成

高山 宜之<sup>†</sup> 北村 達也<sup>†</sup> 梅谷 智弘<sup>†</sup> 灘本 明代<sup>†</sup>

<sup>†</sup> 甲南大学知能情報学部 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

E-mail: †s1871055@s.konan-u.ac.jp, ††{t-kitamu,umetani,nadamoto}@konan-u.ac.jp

**あらまし** 現在我々は漫才台本の自動生成の研究開発を行っている。我々の提案する漫才台本自動生成は複数のボケコンポーネントからなる。このボケコンポーネントのひとつに、ある単語を別の単語に言い間違える「言い間違いボケ」を提案してきた。しかしながら、この「言い間違いボケ」は聞き手が聞き間違いを起こしやすい単語を生成できず、ユーザの笑いを誘発することができないという課題が生じていた。そこで、本論文では音素類似関係を用いることにより、聞き間違いを起こしやすく、笑いを誘発する単語の変形手法を提案する。さらに、音素類似関係と単語の親しみやすさと面白さの関係の分析を行う。

**キーワード** 音素類似, 聞き間違い, 単語生成, 漫才, 漫才ロボット

## 1 はじめに

これまで我々は気軽に日々の笑いを提供することを目的として、身近で親しみやすく笑いを生むコンテンツとして漫才に着目し、漫才ロボットを提案している [1]。我々の提案する漫才ロボットの特徴は、漫才台本を自動生成している点である。この自動生成している漫才台本 (以下、自動生成漫才台本と呼ぶ) は複数のボケコンポーネントからなっている [2]。その中で、我々は自動生成漫才台本の中の基本ボケの一つである言葉遊びボケに着目し、音素類似関係に基づく単語変形手法と International Phonetic Alphabet (IPA)<sup>1</sup>に基づく単語変形手法の2つの聞き間違い単語の決定手法及びそれらの手法を用いた『言い間違いボケ』の生成手法を提案し、比較実験を行っている [3]。

比較実験の結果、音素類似関係に基づく単語変形手法の方が言葉遊びボケに適していることがわかった。ここで音素類似関係とは調音を区別するための記号として用いられる音素に対して、音を聞いた際に似通っていると認識される音素の関係性のことである。音素類似度が高く、聞き間違いを起こしやすい単語が、必ずしも面白いと判断されるわけではないことがわかった。そこで本論文では、より詳細に音素類似関係とユーザが面白さを感じやすい単語変形の関係を分析することを目的とし、以下の3つのユーザ評価実験を行う。

- 音素類似度毎の面白さ比較実験

ユーザが面白さを感じやすい音素類似度を決定するために、「どの音素類似度が面白さを感じやすいのか」を評価する必要がある。そこで、各音素類似度で変形した単語を含むボケとツッコミのコンポーネントを作成し、被験者に評価してもらうことで面白さと聞き間違いやすさによる理解度を比較し、面白さを感じやすい音素類似度の評価実験を行う。

- 漫才の音声と文面の比較実験

面白さはユーザの単語の理解度によっても影響するのではない

かと考えた。そこで、音素類似関係を用いた手法と単語中のある1文字を50音順に従い変形する手法を用いて、音声と文面のコンポーネントをそれぞれ作成し、面白さと理解度の観点から比較評価をするための実験を行う。また、音素類似関係を用いた単語変形手法を漫才ロボットの台本自動生成システム [2] に組み込んだ際の面白さを評価するための分析も行う。

- 単語の親しみやすさと面白さの関係実験

面白さは単語の親しみやすさによっても影響するのではないかと考え、実験を行う。具体的には、日本語能力試験の学習用の単語帳<sup>2</sup>に掲載されている単語、4モーラ以上の単語、日本語の語彙特性 第1期<sup>3</sup>に掲載されている文字音声単語親密度が5.0以上の単語である、音素類似度が0.9, 0.8, 0.7の単語である条件をみたす単語を用いた単語変形手法と音素類似関係を用いた単語変形手法を比較する実験により、単語の親しみやすさと面白さを感じやすい単語変形との関係分析を行う。

2章では関連研究について述べ、3章では音素類似関係を用いた単語変形手法について述べる。4章では実験を通して、面白いと感じる音素類似度の分析と単語の親しみやすさの関係を分析する。5章ではまとめと今後の課題について述べる。

## 2 関連研究

単語置換により面白さを感じやすい単語を生成する研究として、呉ら [4] はどのくらい意味が離れているかを示す概念距離、母音の挿入・削除・置換により音がどの程度似通っているかを示す編集距離、Wikipedia コーパス内の出現頻度により単語置換を行うことで面白さのある発話を生成する手法を提案している。また、岩倉ら [5] はユーザの趣味や嗜好を考慮し、word2vec [6] を用いて入力名詞と作成したデータベース内の名詞とのcos類似度により決定した単語を含む面白さのある文を生成する手法を提案している。これらの研究は、特に音素類似関係に着目した変形単語生成手法を用いて音素類似関係、理解度、親しみや

1 : <https://www.internationalphoneticassociation.org/content/full-ipa-chart>

2 : <https://www.ask-books.com/jp/hajimete-jlpt/>

3 : <https://dictionary.sanseido-publ.co.jp/dict/ssd61135>

すさの分析を行う点で本研究とは異なる。

音素の類似性に着目した研究として、于ら [7] は音素に着目した 2 字の漢字間の類似性を計算する手法を提案し、日本語、中国語、韓国語、ベトナム語の類似性からデータベースおよび検索エンジンを作成する手法を提案している。また、馬場ら [8] は音素の持つ意味である音素列に基づくオノマトペ判定モデルとそのモデルを用いた辞書構築手法を提案している。本研究も音韻間や音素間の類似性に着目しているが、検索システムやデータベースの構築を目的としているのではなく、音素類似関係のある単語の面白さの感じやすさを分析する点でこれらの研究と異なる。

音素列に着目した言い間違えの研究として、上村ら [9] は文字の挿入・置換・脱落による編集距離に基づく、発音の誤りやすさを表す発音スコアを求める手法を提案している。本研究では、あえて誤りを起こす単語の生成および聞き間違えや面白さの感じ方の分析を行う点で本研究と異なる。

### 3 音素類似関係に基づく単語変形手法

これまで我々の提案している音素類似関係に基づく単語変形手法 [3] を用いて実験を行う。この音素類似関係に基づく単語変形手法を用いた単語変形手法の流れを以下に示す。

- (1) ある単語を形態素解析器 (MeCab<sup>4</sup>) によりカタカナに変換する。この単語を変形元単語と呼ぶ。
- (2) カタカナに変換した単語を音素列に変換する。
- (3) 表 1 に示す音素類似表 [10] に従って音素列に変換した単語を変形する。
- (4) 変形した単語の音素列をカタカナに変換する。
- (5) カタカナに変換した単語が辞書内に存在するか確認する。ここで辞書は IPADIC<sup>5</sup>を用いる。
- (6) 存在した単語を出力する。この単語を変形後単語と呼ぶ。

音素列に変換した変形元単語を表 1 に示す音素類似表 [10] に従って変形する。音素類似度は前田ら [10] の提案を参考に次のように定義する。全ての音素が同一である変形後単語の場合、類似度は 1.0 とする。また、2 つ以上の音素が異なっている変形後単語の場合、類似度は 0.0 とする。そして、1 つの音素のみ異なる場合、類似度は表 1 の比較対象音素に対応した音素類似度とする。さらに、呉ら [11] は語頭の母音と語末の音韻が一致する単語は聞き間違いやすいと述べていることより、この条件を満たす場合には音素類似度に対して 0.3 の加点をする。ただし、加点後の類似度が 1.0 以上となる場合は、上記の定義より全ての音素が同一である変形後単語の類似度を 1.0 としており、類似度 0.9 よりも聞き間違いやすいと考え、類似度を 0.95 に調整する。

表 1: 音素類似度表 [10]

比較対象音素	音素類似度
b, d, g	0.9
p, t, k	0.9
m, n, N	0.9
h, s	0.8
a, o, u	0.7
e, i	0.7
s, z	0.6
b, p	0.6
d, t	0.6
g, k	0.6
上記以外の母音間	0.5
上記以外の子音間	0.1
1 音素の脱落・挿入	0.1

## 4 実験

漫才台本を自動生成するにあたり、より面白さを感じやすい言葉遊びボケを生成するために、音素類似関係とおかしみのある単語変形の関係を分析することを目的とし、3 つのユーザ評価実験を行う。

### 4.1 音素類似度毎の比較実験

被験者が最も面白いと感じる変形後単語を生成する音素類似度を決定するための、面白さ判定実験を行う。本実験では、被験者が音により面白いと感じる場合と単語の意味により面白いと感じる場合があると考え、単語の音と意味の 2 軸による実験を行う。具体的には、単語の聞き取りによる実験と文面を見ての実験の 2 種類の実験を行う。さらに、単語の聞き取り実験では被験者が正しく聞き取れている単語が面白いと感じるのか、もしくは、聞き間違いを起こした単語が面白いと感じるのかを評価するために、生成した文が面白いかを評価する実験 (音声による面白さ判定実験) と変形後単語を聞き取れたかの実験 (単語聞き取り実験) を行う。

以下に本実験の Research Question を示す。

RQ1 面白いと感じる変形後単語を生成する音素類似度の値は何か?

RQ2 変形後単語が正しく聞き取れた方が面白いと判断されやすいのか?

RQ3 変形後単語の意味を理解した方が面白いと判断されやすいのか?

#### 実験条件

提案手法で生成される可能性のある音素類似度は、0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.1 の合計 8 つである。これらの音素類似度を次のように 2 つずつ組み合わせる。

{0.95, 0.9}, {0.95, 0.8}, ..., {0.4, 0.1}

組み合わせの合計は 28 組となり、この 28 組を 1 セットとする。これらの音素類似度の単語変形を行い、変形元単語の箇所を変形後単語に置き換えた文の組を次のボケとツッコミのフ

4 : <https://taku910.github.io/mecab/>

5 : <https://taku910.github.io/mecab/>

レームワークに従い作成する。また、このフレームワークで作成したボケとツッコミを『言い間違えボケ』と呼ぶ。以下に使用したフレームワークを示す。

ボケ: ○○○〈変形後単語〉○○○。

ツッコミ: 〈変形後単語〉じゃなくて〈変形元単語〉やろ。

ここで、○○○は変形元単語が挿入されているニュース記事である。例えば、『センニュウ』を変形元単語とした場合、漫才台本を生成するニュース記事の中で「潜入」を含む文が「特殊部隊がアジトに『センニュウ』した」であったとする。『センニュウ』の音素類似度が{0.95, 0.8}の組み合わせは{「特殊部隊がアジトに『ヘンニュウ』した」, 「特殊部隊がアジトに『サンニュウ』した」}となる。実験に用いる文はYahoo!ニュース<sup>6</sup>から手作業で文および変形元単語を選択し、文の組は2セット作成した。

被験者は関西弁話者の大学生10人であり、男性6人、女性4人である。音声には漫才ロボット[1]に用いているAITalk<sup>7</sup>により生成した関西弁風の合成音声を用いる。なお、合成音声ソフトの特性上、読み間違いが生じる可能性があるため、変形後単語と変形元単語は漢字ではなく、すべてカタカナに変換し音声合成を行う。音声の聞く順番による影響をできる限り排除するため、被験者を音声の前後の順番を入れ替えた2つの群に分ける。また、音声を長時間に渡って聞くことによる被験者の負担を考慮し、単語の聞き取り実験は1セットを2回に分けて行う。なお、各実験の比較検討のために同一の文の組を用いているが、被験者が記憶することによる影響をできる限り排除するために、各実験は3日以上の間隔を空けて行う。

#### 4.1.1 音声による面白さ判定実験

以下の例のように、AとBの各音素類似度の組み合わせで作成したボケとツッコミの音声を流し、被験者は「どちらが面白い」を評価する。

【A】ボケ: 特殊部隊がアジトに『ヘンニュウ』した。

ツッコミ: 『ヘンニュウ』じゃなくて『センニュウ』やろ。

【B】ボケ: 特殊部隊がアジトに『サンニュウ』した。

ツッコミ: 『サンニュウ』じゃなくて『センニュウ』やろ。

#### 4.1.2 単語聞き取り実験

被験者に変形後単語部分が空欄となっている文を提示する。変形後単語を含む文の音声を流し、被験者は空欄箇所にはまる単語をひらがな表記で書き取る。例えば、「特殊部隊がアジトに『』した」のように文を提示し、「特殊部隊がアジトに『ヘンニュウ』した」と音声を流す。

#### 4.1.3 文面による面白さ判定実験

聞き取り実験では、単語を正しく聞き取れずに、勘違いした単語による面白さの判断をしている場合がある。そこで、変形後単語を元の文に挿入した際、そもそも面白いかを判断する必要があると考え、文面による面白さの実験を行う。

文面による面白さの実験では、被験者に音声による面白さ判定実験で用いたボケとツッコミの文を漢字に変換し、書き起こ

して提示する。なお、漢字への変換の際、同音異義語が存在する場合は日本語の語彙特性第1期<sup>8</sup>に掲載されている、文字音声単語親密度が最も高い単語を選択する。以下に、被験者に提示するAとBのボケとツッコミ文の例を示す。

【A】ボケ: 特殊部隊がアジトに『編入』した。

ツッコミ: 『編入』じゃなくて『潜入』やろ。

【B】ボケ: 特殊部隊がアジトに『参入』した。

ツッコミ: 『参入』じゃなくて『潜入』やろ。

## 実験結果と考察

音素類似度毎の比較による、図1に音声による面白さ判定実験の結果を、図2に単語聞き取り実験の結果を、図3に文面による面白さ判定実験の結果を示す。

音素類似度毎に比較し、面白さや正しく聞き取れているかを評価する。半数より多くの人が面白いと回答もしくは聞き取れていた場合、その音素類似度の方が面白いもしくは聞き取れていると判断する。例えば、音素類似度が0.95と0.8の面白さを比較する際に、0.95の方が面白いと答えた人が4名、0.8の方が面白いと答えた人が6名の場合、0.95より0.8の方が面白いと評価する。なお、面白いと判定した人、正しく聞き取れた人が同数の場合は無視する。図1、図2において、縦軸は比較してより面白い(図2の場合は正しく聞き取れた)と評価された個数の比率、横軸は音素類似度を表している。

図1の音声による面白さ判定実験と図3の文面による面白さ判定実験の結果より、どちらも音素類似度0.9, 0.8, 0.7の変形単語は被験者の半数以上が面白いと判断している。RQ1の回答として、ユーザが面白さを感じやすい音素類似度は0.9, 0.8, 0.7であることがわかった。

図1の音声による面白さ判定実験結果、図2の単語聞き取り実験の結果より、どちらも音素類似度0.9, 0.8, 0.7, 0.5, 0.4の単語が他の音素類似度の結果と比較して面白いかつ聞き取れている傾向があることがわかる。これにより、面白いと判定されやすい音素類似度は聞き取りやすい傾向があり、変形後単語が正しく聞き取れ、意味を理解した方が面白いと判断されやすいことがわかった。RQ2の回答として、変形後単語を正しく聞き取れた方が面白いと判断されやすいことがわかった。また、RQ3の回答として、変形後単語の意味を理解した方が面白いと判断されやすいことがわかった。

特に音素類似度0.1の単語が聞き取りにくい傾向が見られた。本実験において、被験者全員が『他所(/yoso/)』を『予想(/yosou/)』(母音/u/の脱落の変形)に聞き間違え、『観光書く(/kaNkoukaku/)』を『観光客(/kaNkoujaku/)』(半母音拗音/j/の脱落の変形)に聞き間違えた。これらはいずれも、1音素の脱落による変形単語である。この結果より、母音や半母音における1音素の脱落による変形単語は聞き間違えやすい可能性がわかった。

## 4.2 漫才の音声と文面の比較実験

言葉遊びボケの言い間違えによる音声面白さにどの程度影

6: <https://news.yahoo.co.jp/>

7: <https://www.ai-j.jp/about/>

8: <https://dictionary.sanseido-publ.co.jp/dict/ssd61135>

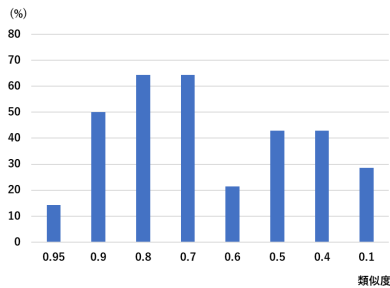


図 1: 音声による面白さ判定実験の結果

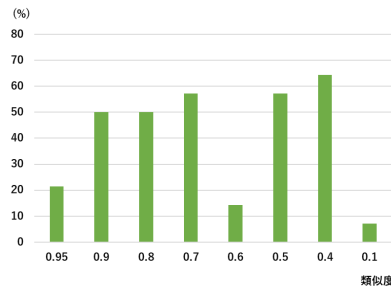


図 2: 単語聞き取り実験の結果

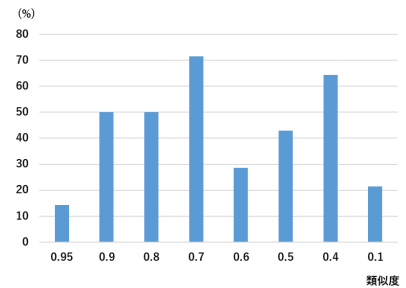


図 3: 文面による面白さ判定実験の結果

響しているか。また、単語の意味理解が面白さにどの程度影響しているかを分析することを目的として、漫才の音声と文面の比較実験を行う。これにより、人は漫才を聞いた場合と漫才を書き起こした文面を見た場合で面白さの感じ方に差が出るかが明らかになると考える。さらに、漫才台本に音素類似関係に基づく単語変形手法が組み込まれた際の面白さと理解度の影響を分析することを目的とし、単語中のある 1 文字を 50 音順に従い変形する手法との比較を行う。

#### 単語中のある 1 文字を 50 音順に従い変形する手法

単語中のある 1 文字を 50 音順に従い変形する手法による変形手法を次に示す。

(1) 変形元単語中のある 1 文字を 50 音順に従って変形する。

(2) 変形した単語が辞書 (IPADIC<sup>9</sup>) に存在した場合、その単語を変形後単語とする。

例えば、変形元単語『キモチ』の 1 文字目を変形することを考える、変形後単語の候補は『アモチ』、『イモチ』、『ウモチ』、『オモチ』…… となり、『オモチ』が辞書に存在するので、『オモチ』が変形後単語と決定される。

本論文では単語中のある 1 文字を 50 音順に従い変形する手法をベースラインとする。

#### 実験条件

本実験は音声による漫才と文面による漫才の 2 つの形式による漫才を用いて評価する。被験者は関西弁話者の大学生 12 人であり、6 人ずつ漫才の音声を聞いて判定する群と文面を見て判定する群の 2 つの群に分ける。漫才の音声を聞いて判定する群 (音声群) では、漫才ロボット [1] を用いて被験者に面白さと理解度を評価する。一方、文面を見て判定する群 (文面群) では、図 4 に示すチャットアプリケーション [12] を用いて被験者に面白さと理解度を評価する。チャットアプリケーションでは音声と発話毎に表示されるテキストの吹き出しからなっているが、漫才の音声は流さない。本実験で用いる漫才台本は、漫才ロボットの自動台本生成システム [2] で生成される自動生成漫才台本に、4.1 章の音素類似度毎の比較実験で示したボケとツッコミの手法を用いて作成したボケとツッコミを組み込んだ漫才台本を用いる。音素類似度毎の比較実験において面白いと判断されやすい音素類似度 0.9, 0.8, 0.7 で変形した単語を用いた

言い間違えボケを含む漫才を 3 本とベースラインの手法で生成した単語を用いた言い間違えボケを含む漫才を 3 本作成した。被験者は音素類似手法で生成した単語を含む漫才とベースラインの手法で生成した単語を含む漫才を比較評価する。これらの 2 つの漫才を視聴した後に、後述する音素類似関係に基づく単語変形手法とベースラインを比較評価するアンケートに回答する。さらに、各漫才を視聴した後に、後述する漫才毎に 5 段階評価で面白さと理解度をはかるアンケートに回答する。これを 3 回繰り返す。

評価手法としてオムロン株式会社の表情認識カメラ (HVC-P2<sup>10</sup>) による被験者の表情分析と被験者へのアンケートによる面白さと理解度の評価を行う。家根ら [13] の研究を参考に、HVC-P2 の表情推定機能の項目のひとつである「喜び」の数値 (最低: 0 ~ 最高: 100) を用いて、被験者が面白いと感じているかを判定する。図 5 のように、被験者は、新型コロナウィルス感染症拡大防止の対策として、マウスガードを着用している。しかしながら、マウスガードの光の反射で表情が隠れる等の影響により、表情認識カメラでの表情分析をうまくできない場合が多く見られた。そこで、人手で笑いの表情を判定する。アンケートによる面白さと理解度の評価は、漫才毎に 5 段階評価で面白さと理解度をはかるアンケートと、音素類似関係に基づく単語変形手法とベースラインを比較評価するアンケートを行う。漫才毎に 5 段階評価で面白さと理解度をはかるアンケートの質問は以下の 2 問である。

設問 1-1 言い間違えボケ部分について、どの程度面白かったですか？

評価は 5 段階 (1. 全く面白くない ~ 5. 非常に面白い) で行う。

設問 1-2 言い間違えボケ部分について、どの程度理解できましたか？

評価は 5 段階 (1. 全く理解できない ~ 5. 非常に理解できた) で行う。

音素類似関係に基づく単語変形手法とベースラインを比較評価するアンケートは以下の 2 問である。

設問 2-1 漫才 A と漫才 B のどちらがより面白かったですか？

設問 2-2 漫才 A と漫才 B のどちらがより意味を理解できましたか？

評価は漫才 A と漫才 B の 2 択の回答により行う。

9 : <https://taku910.github.io/mecab/>

10 : <https://plus-sensing.omron.co.jp/product/hvc-p2.html>

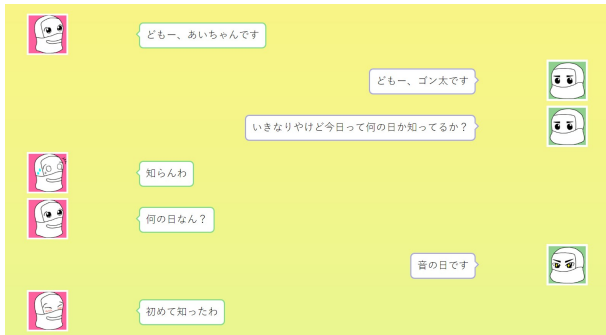


図 4: チャットアプリケーション



図 5: 実験の様子

## 実験結果と考察

カメラでの喜びの表情認識の結果を図 6～図 17 に示す。横軸は時間 (秒), 縦軸は喜びの度合いであり, 図中の青枠ではボケ, 赤枠ではツッコミを示す。また, 図中のひし形のプロットは, 1 人の判定者が人手での表情判定を行い 1 人以上の被験者が笑い判定した時間を示す。ここでは, 被験者の口角が少しでも動いたときを笑いとしている。

マウスガード等の影響により, カメラでの表情認識ができない箇所が多かったため, 人手での評価に着目して考察する。図 7, 図 8, 図 13, 図 14 の音声群, 図 10, 図 11, 図 16, 図 17 の文面群ともに, 比較 2, 比較 3 において, 音素類似手法の方がベースラインよりも人手での表情判定で笑い判定された被験者が存在していることがわかる。これにより, 音素類似手法の方が面白さを感じやすい単語を生成していると言える。さらに, 人手での表情判定で笑い判定された被験者はツッコミに多いことがわかる。これにより, ツッコミは言い間違い単語を正しい単語に訂正していることから, ユーザーが面白さを感じやすいときは間違いに気付くときであることが言える。音声群と文面群を比較すると, 音声群と文面群の被験者が異なることから, 喜びの値が高く出やすい被験者と低く出やすい被験者が存在することがわかる。これにより, 笑いの表情の出やすさは個人差があると考えられる。

漫才毎に 5 段階評価で面白さと理解度をはかるアンケート結果を図 20～図 23 に示す。また, 音素類似関係に基づく単語変形手法とベースラインを比較評価するアンケート結果を図 24, 図 25 に示す。

図 20 と図 21 より, 比較 3 の音素類似手法を除き, 文面群の

方が「全く面白くない」と回答した被験者が多かった。これにより, 聞くことによる面白さの感じ方と見ることによる面白さの感じ方には差があり, 音の響きや感じ方によっても面白さを感じる可能性があると言える。

図 20 の比較 3 では, ベースラインより音素類似手法の方が「全く面白くない」と回答した被験者が多かった。逆に, 図 21 の比較 3 では, 音素類似手法よりベースラインの方が「全く面白くない」と回答した被験者が多かった。これにより, 音素類似手法による言い間違い単語を被験者が聞き間違えたことにより, 意味が理解できず, 面白くないと判断された可能性があることがわかる。さらに, 図 22 と図 23 より, 音素類似手法において, 文面群の方が「非常に理解できた」, 「少し理解できた」と回答した被験者が多かった。これらにより, 音素類似手法のほうがベースラインよりも聞き間違いやすい単語を生成していることがわかった。

図 24 より, 音声群, 文面群ともに半数以上の被験者が音素類似手法の方が面白いと回答している。これにより, 音素類似手法の方が漫才台本として組み込まれた際においても面白さを感じやすい単語変形をしていると言える。図 25 より, 比較 1 の結果において, 音声群より文面群の方が「音素類似手法の方が理解できた」と回答した被験者が多かった。これは音素類似手法の方がベースラインよりも聞き間違いやすい単語を生成しており, 被験者が聞き間違いを起こしたため, このような結果になったと考えられる。

## 4.3 単語の親しみやすさと面白さの関係実験

日常的によく使われていて馴染み深い単語がより面白さを感じるのではないかと考え, 親しみやすさを考慮した変形単語を用いて実験を行う。親しみやすさを考慮した変形単語を用いたボケとツッコミと音素類似手法で生成した単語のボケとツッコミの比較により, 単語の親しみやすさがどれほど面白さを感じる要因となるかを分析できると考え, 実験を行う。

### 親しみやすい単語変形

変形元単語は文中の日本語の語彙特性辞書 第 1 期<sup>11</sup>に掲載されている文字音声単語親密度が最も高い単語とする。また, 親しみやすい変形後単語の定義を以下に示す。

- 日本語能力試験の N5, N4, N3 レベル相当の単語帳<sup>12</sup>に掲載されている単語である。日本語能力試験の単語帳を用いる理由は平易な単語がより親しみやすいと考えたためである。
- 4 モーラ以上の単語である。
- 日本語の語彙特性辞書 第 1 期に掲載されている文字音声単語親密度が 5.0 以上の単語である。
- 音素類似度が 0.9, 0.8, 0.7 の単語である。

これらの定義の変形後単語が本当に日常的によく使う単語であるかを判断する必要がある。そこで, クラウドソーシングを用いて 10 人に以下の内容で日常的によく使う単語であるかのアンケートを行う。

11 : <https://dictionary.sanseido-publ.co.jp/dict/ssd61135>

12 : <https://www.ask-books.com/jp/hajimete-jlpt/>



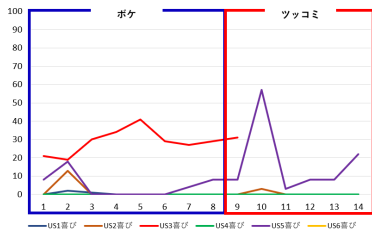


図 6: 音声群 音素類似手法の喜びの表情  
比較 1

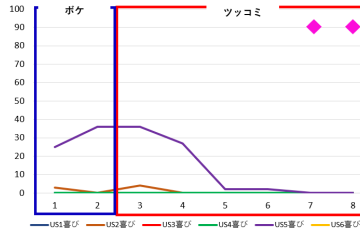


図 7: 音声群 音素類似手法の喜びの表情  
比較 2

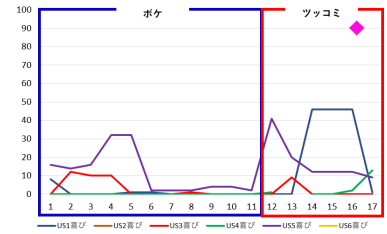


図 8: 音声群 音素類似手法の喜びの表情  
比較 3

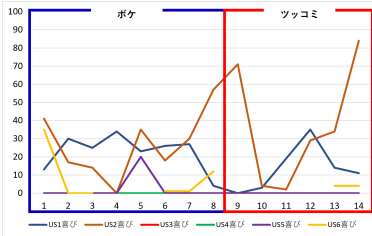


図 9: 文面群 音素類似手法の喜びの表情  
比較 1

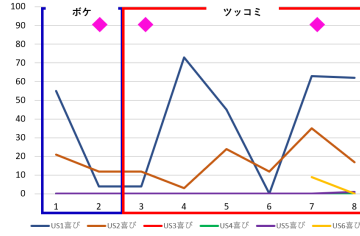


図 10: 文面群 音素類似手法の喜びの表情  
比較 2

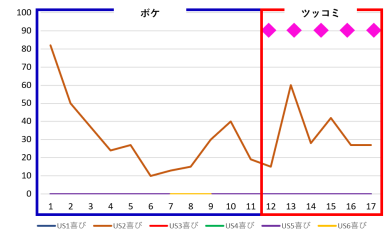


図 11: 文面群 音素類似手法の喜びの表情  
比較 3

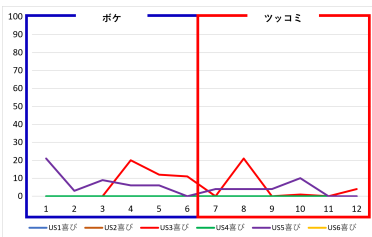


図 12: 音声群 ベースラインの喜びの表情  
比較 1

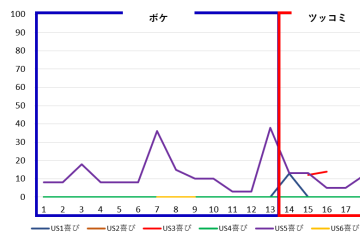


図 13: 音声群 ベースラインの喜びの表情  
比較 2

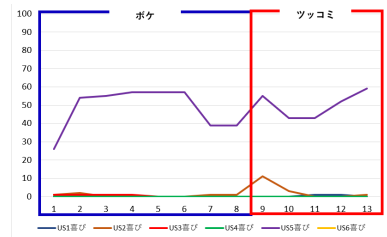


図 14: 音声群 ベースラインの喜びの表情  
比較 3

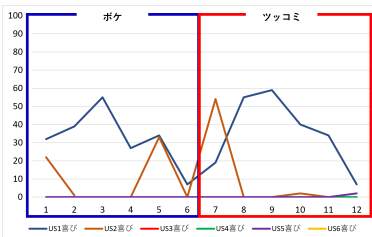


図 15: 文面群 ベースラインの喜びの表情  
比較 1

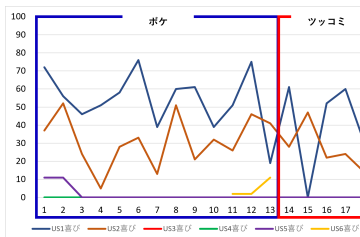


図 16: 文面群 ベースラインの喜びの表情  
比較 2

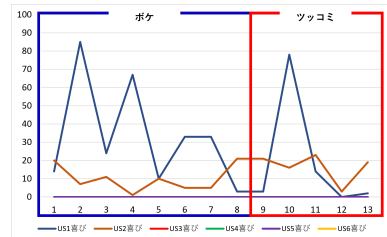


図 17: 文面群 ベースラインの喜びの表情  
比較 3

Q.「あなたが日常的によく目にするもしくは、よく使う単語であるか」について、以下の 5 段階で評価してください。

- 5: 非常によく目にする (非常によく使う) 単語
- 4: 少し目にする (少し使う) 単語
- 3: どちらでもない
- 2: あまり目に見えない (あまり使わない) 単語
- 1: ほとんど目に見えない (ほとんど使わない) 単語

5 人以上が、「5: 非常によく目にする (非常によく使う) 単語」または「4: 少し目にする (少し使う) 単語」であると回答した単語を変形後単語に用いる。

## 実験条件

被験者は関西弁話者の大学生 9 名。4.1 章の音素類似度毎の比較実験で示したボケとツッコミの手法を用いてボケとツッコミを作成する。Yahoo!ニュースから取得した 18 文中、親しみやすい単語生成手法を用いて作成したボケとツッコミを 9 つ、

音素類似手法を用いて作成したボケとツッコミを 9 つ作成する。音声は関西弁風の合成音声を用いる。被験者は親しみやすい変形単語と音素類似手法で生成したボケとツッコミの音声を聞き、各設問に対して「どちらがより面白いですか」と「どちらがより理解できましたか」の 2 つの質問に回答する。

## 実験結果と考察

面白さの比較評価結果を図 18、理解度の比較評価結果を図 19 に示す。横軸は問題番号、縦軸は「より面白い」または「より理解できる」と回答した人数である。

この実験の結論として、親しみやすさを考慮した単語であっても、面白いと感じやすく、理解されやすいとは限らないことが言える。例えば、3 問目の親しみやすい変形単語で作成されたボケの「今年、皆さんのおかげで目標にしていた R-1『友情』ができた。」(変形元単語:『優勝』, 変形後単語:『友情』)のように、文が成り立っていないと考えられるボケは、面白いと感

じず、理解されにくい傾向が見られた。6 問目は全員が「親しみやすい単語生成手法」の方が面白く、理解できたと回答している。この問題の音素類似手法における変形元単語は『氏(シ)』、変形後単語は『灯(ヒ)』であり、親しみやすい単語生成手法における変形元単語は『ドキドキ』変形後単語は『ときどき』である。これより、1 モーラの単語は変形箇所を理解しにくいいため、面白さを感じにくいことがわかった。したがって、あまりにも短い単語の場合は、ユーザが聞き取れず、面白さを感じにくいと考えられることがわかった。

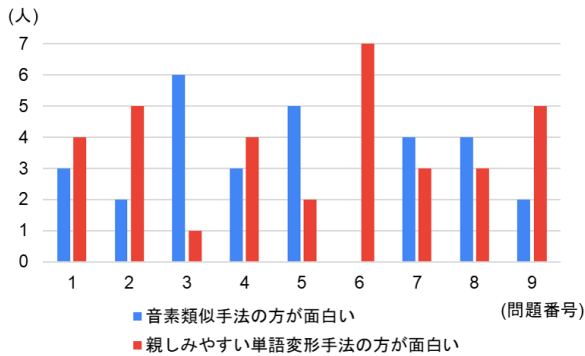


図 18: 面白さの比較評価結果

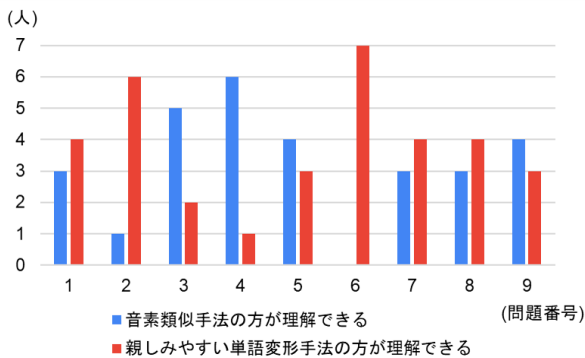


図 19: 理解度の比較評価結果

## 5 まとめと今後の課題

音素類似関係と単語の親しみやすさの観点から、面白さとの関係性を評価するために、音素類似度毎の面白さ比較実験、漫才の音声と文面の比較実験、単語の親しみやすさと面白さの関係実験の3つの実験を行った。実験の結果、音素類似度 0.9、0.8、0.7 の変形後単語が面白いと感じやすく、聞き間違えが起こりにくく、被験者が単語の意味を理解できていると面白さを感じやすいことがわかった。また、親しみやすさを考慮した単語は面白さの感じ方や理解度との関係があるとは言えないことがわかった。文が成り立っていないと考えられるボケは、面白いと感じず、理解されにくい傾向が見られた。また、変形後単語があまりにも短い場合には、被験者が変形箇所に気付かないため、面白さを感じにくいことがわかった。

今後の課題として、変形後単語を含むボケが文として成り

立っていることが面白さの感じ方に影響があるのかを評価する実験を行う。また、合成音声による関西弁特有のイントネーションや抑揚がおかしみや聞き間違えによる影響の可能性が考えられる。そこで、合成音声とヒトの発話によるおかしみの関係を分析することが今後の課題である。さらに、変形後単語が現実には存在しない場合の面白さの感じ方の評価、また、本論文では1音素のみの変形に着目したが、2音素以上の変形後単語についても面白さの感じ方の評価を行うことを今後の課題とする。

## 謝 辞

論文の一部は JSPS 科研費 19H04218, 19H04221, 20K12085, 及び私学助成金 (大学間連携研究助成金) の助成によるものである。ここに記して謹んで感謝の意を表する。

## 文 献

- [1] Tomohiro Umetani, Ryo Mashimo, Akiyo Nadamoto, Tatsuya Kitamura and Hirotaka Nakayama, "Manzai Robots: Entertainment Robots based on Auto-created Manzai Scripts from Web News Articles", Journal of Robotics and Mechatronics, vol.26, no.5, pp.662-664, 2014.
- [2] 真下遼, 梅谷智弘, 北村達也, 灘本明代, "つかみ・本ネタ・オチから構成される漫才ロボット台本自動生成手法の提案", ARG W12 no.4, 2014.
- [3] 高山宜之, 原口和貴, 北村達也, 灘本明代, "音素類似関係を用いた言い間違えボケの自動生成", 第13回データ工学と情報マネジメントに関するフォーラム (DEIM2021), H25-3, 8pages, 2021.
- [4] 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", 情報処理学会論文誌, Vol.61, No.1, pp.113-122, 2020.
- [5] 岩倉 亮介, 吉川 大弘, ジメネス フェリックス, 古橋 武, "Twitter データを用いたユーモア語句自動生成手法に関する一検討", 人工知能学会全国大会論文集, 2017, JSAI2017 巻, 第31回, 2017.
- [6] Mikolov, Tomas, et al, "Efficient estimation of word representations in vector space.", arXiv preprint arXiv:1301.3781, 2013.
- [7] 于劭賢, 金志宣, 玉岡賀津雄, Hoang Thi Lan Phuong, Jingyi Zhang, "2 字漢字語の音韻類似性・音韻距離に関する日韓中越データベースのオンライン検索エンジンの構築", 2018 年度日本語教育学会春季大会予稿集, pp.129-134, 2018.
- [8] 馬場睦也, 楠和馬, 波多野賢, "音素単位の分散表現に基づくオノマトペ辞書構築法の提案" 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020), G4-3, 7pages, 2020.
- [9] 上村航平, 鷹野孝典, "音素列を対象とした発ミス検知による正解語提示手法の検討", 人工知能学会研究会資料 知識ベースシステム研究会, vol 117, pp.61-64, 2019.
- [10] 前田実香, 鬼沢武久, "単語の関連性とおもしろさを取り入れたなぞかけ生成", エンタテインメント感性特集, Vol.5, no.3, pp.17-22, 2005.
- [11] 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.
- [12] 家根和希, 原口和貴, 佐藤亮, 荒牧英治, 宮代勲, 灘本明代, "チャット型漫才アプリケーションの提案", 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020), A3-3, 8pages, 2020.
- [13] 家根和希, 原口和貴, 佐藤亮, 荒牧英治, 宮代勲, 灘本明代, "がん患者を対象とした漫才ロボットによる笑いの実証実験", 信学技報, vol.119, no.201, DE2019-20, pp.29-34, 2019.

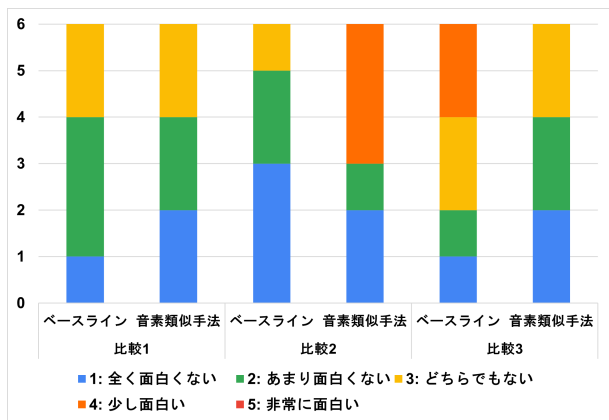


図 20: 音声群 面白さ評価の結果

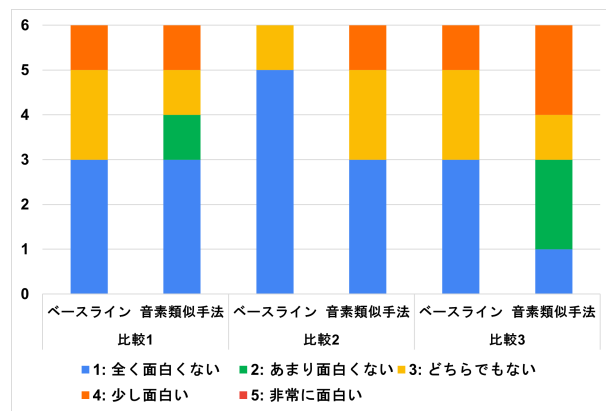


図 21: 文面群 面白さ評価の結果

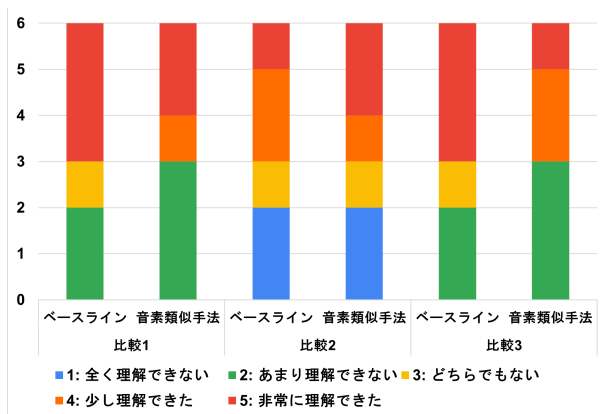


図 22: 音声群 理解度評価の結果

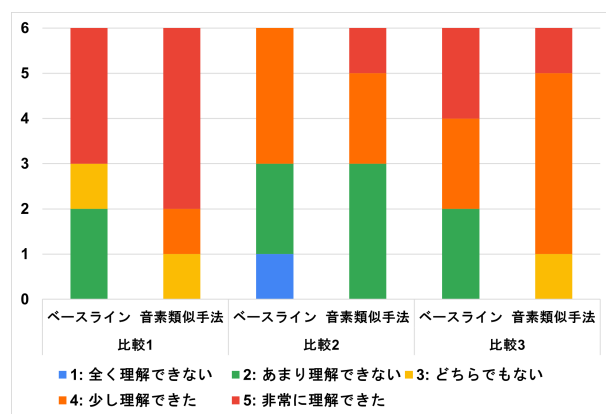
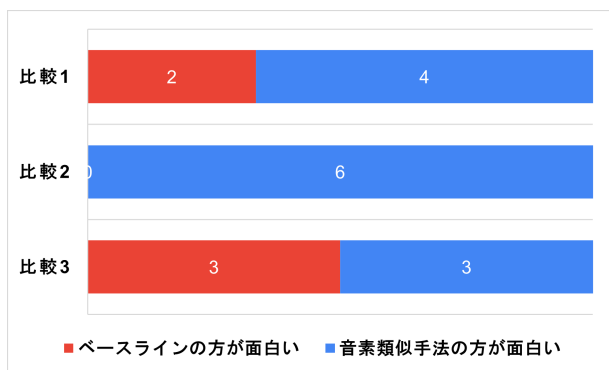
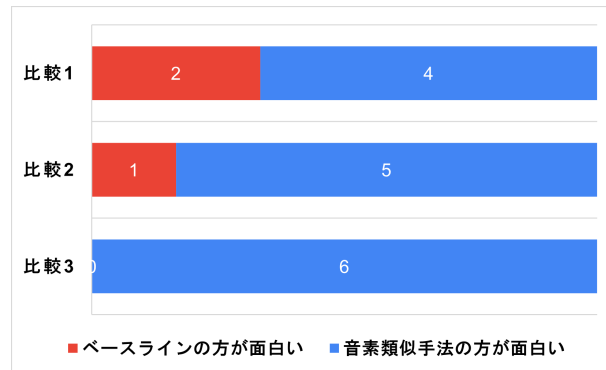


図 23: 文面群 理解度評価の結果

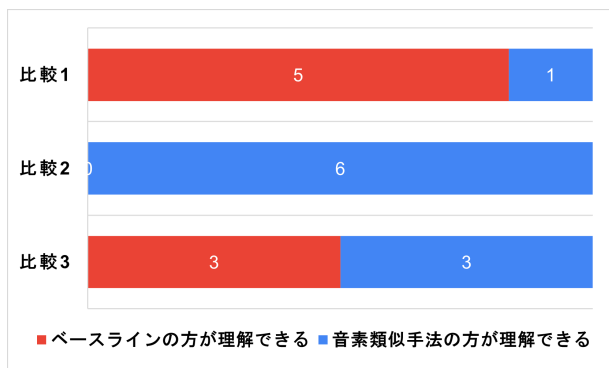


(a) 音声群

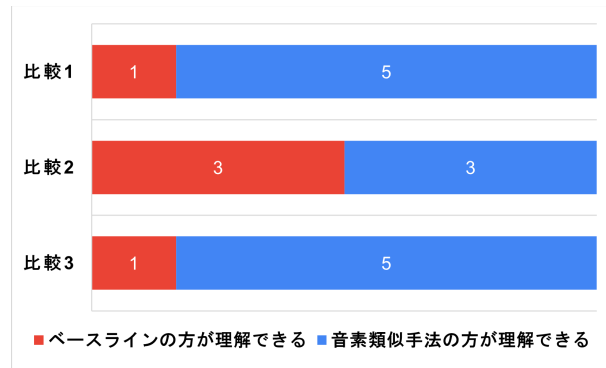


(b) 文面群

図 24: 面白さの比較評価の結果



(a) 音声群



(b) 文面群

図 25: 理解度の比較評価の結果