

テキスト会話における発話ニュアンス付与のための 吹き出し形状の推薦

Recommendation of Speech Balloon Shapes for
Representing Speech Nuances in Text-based Communication

*1田中 秀樹 *2山西 良典 *2福本 淳一 *2西原 陽子
Hideki Tanaka Ryosuke Yamanishi Junichi Fukumoto Yoko Nishihara

*1立命館大学大学院 情報理工学研究所
Graduate School of Information Science and Engineering, Ritsumeikan University

*2立命館大学 情報理工学部
College of Information Science and Engineering, Ritsumeikan University

The purpose of this study was to communicate acoustic information using speech balloon shapes in text-based communication. Acoustic information of speech disappears in text communication. In Comic, readers can recognize nuance of character voice from speech balloon shape, which expresses characters' nuance, emotion and intention. We analyzed relationships between lines that means speech and speech balloon shapes in comic. We selected candidate for features of line for recommendation of speech balloon shapes based on analysis. We developed a model of relationships between lines that means speech and speech balloon shapes using machine learning and recommended speech balloon shapes from conversation texts. We discussed features of lines related balloon shapes.

1. はじめに

日常生活のあらゆる場面において、テキストを用いたコミュニケーションが頻繁に行われている。例えば、古くからは手紙、近年ではインターネットを介した電子メールや電子掲示板などで、時代を問わずコミュニケーションの主なメディアとしてテキスト情報が用いられている。また、近年では Twitter や LINE などの Social Networking Service (SNS) の普及に伴い、テキスト情報を用いたコミュニケーションはますます活発に行われるようになった。しかしながら、テキスト情報のみを用いたコミュニケーションでは、音声に伴う会話時に話者同士が共有する“声の大きさ”、“声質”といった音響の特徴から把握可能な「発話ニュアンス」が欠落してしまう問題点が存在する。このため、テキストメッセージの受け手は送り手がどのような意図のもとでメッセージを送信したのかを捉えることが難しく、送り手の意図を誤って解釈してしまう可能性がある。

一方で、音響情報を用いることなく、二次元の画像情報とテキスト情報のみによって会話ニュアンスの伝達を実現しているメディアとしてコミックが存在する [長谷 00]。コミックでは、キャラクターのセリフ（発話）を、様々な形状の吹き出し中に記述することでキャラクターのセリフの発話ニュアンスを読者に伝達している。例えば、図 1 中の (a) では、女性の「そんなの困るっ!!」といった発話に風船を破裂させたような形の吹き出しが使われており、この吹き出し形状から読者は「叫び声」を連想する。また、図 1 中の (b) では、男性の「あ…あははそうか」といった発話を波線で囲った吹き出しが使われており、この吹き出し形状から読者は「弱々しい声」を連想する。このように、コミックでは吹き出し形状を用いることで視覚的にテキストのみで発話ニュアンスを表現している [竹内 05]。

本稿では、先行研究 [田中 14] で得られた知見に基づいたセリフ特徴を用いて、セリフ特徴と吹き出し形状の関係性モデルを構築する。そして、構築したモデルを用いて、会話テキストからの吹き出し形状の推薦を行う。



(a) 叫び声の可視化例

(b) 弱々しい声の可視化例

図 1: 吹き出しの使われ方の例。コミックの読者は、吹き出し形状から直感的に (a) は叫び声、(b) は弱々しい声であると認識可能である。

2. 関連研究

松下は、文献 [松下 13] のなかで、コミックコンテンツを対象とした研究を「コミック工学」と位置付けている。松下はコミック工学研究を、コンテンツ情報を (1) どのように計算機利用可能なコードに変換するか、(2) コード化されたコミックコンテンツをどのように利用するか、(3) コミックの表現技法はどのような対象に応用可能か、という 3 つの観点で整理している。本稿は、このうち (1) および (2) に位置付けられる。

吹き出し形状を用いた研究として、紺家ら [紺家 15] や松宮ら [松宮 14] の研究が挙げられる。紺家らは動画の字幕に対し、吹き出しの形状によって話者の音声表現を付加し、音情報の雰囲気や補完する字幕提示方法を提案している。紺家らの提案手法では、吹き出し枠の形状で発話の速度を表現しているが、発話内容には触れていない。松宮らは、音声に内在する感情に合った形状の吹き出しの字幕への付与を目指して、アニメ音声の音響特徴量と言語特徴量の 2 種類の特徴量を用いて吹

表 1: 語尾特徴とその特徴 ID

特徴 ID	特徴	特徴 ID	特徴
f_1^{tail}	“っ” の数	f_{11}^{tail}	“一” の数
f_2^{tail}	“ッ” の数	f_{12}^{tail}	“一” の数
f_3^{tail}	“♡” の数	f_{13}^{tail}	“〜” の数
f_4^{tail}	“☆” の数	f_{14}^{tail}	“です” の数
f_5^{tail}	“♪” の数	f_{15}^{tail}	“ます” の数
f_6^{tail}	“!” の数	f_{16}^{tail}	“ない” の数
f_7^{tail}	“?” の数	f_{17}^{tail}	“かな” の数
f_8^{tail}	“!?” の数	f_{18}^{tail}	“けど” の数
f_9^{tail}	“。” の数	f_{19}^{tail}	“なんて” の数
f_{10}^{tail}	“...” の数		

表 2: 形態的特徴とその特徴 ID

特徴 ID	特徴	特徴 ID	特徴
f_1^{str}	数字の数	f_7^{str}	名詞の数
f_2^{str}	文字数	f_8^{str}	動詞の数
f_3^{str}	人名を含む	f_9^{str}	形容詞の数
f_4^{str}	ひらがなのみ	f_{10}^{str}	一人称代名詞の数
f_5^{str}	カタカナを含む	f_{11}^{str}	二人称代名詞の数
f_6^{str}	漢字を含む		

表 3: 意味的特徴とその特徴 ID

特徴 ID	特徴	特徴 ID	特徴
f_1^{mean}	バッドワードの数	f_8^{mean}	好の数
f_2^{mean}	喜の数	f_9^{mean}	嫌の数
f_3^{mean}	怒の数	f_{10}^{mean}	昂の数
f_4^{mean}	哀の数	f_{11}^{mean}	安の数
f_5^{mean}	驚の数	f_{12}^{mean}	ポジティブ語の数
f_6^{mean}	恥の数	f_{13}^{mean}	ネガティブ語の数
f_7^{mean}	怖の数		

き出し形状を分類している。言語特徴量として形態素単位の uni-gram, bi-gram, tri-gram の頻度を使用しているが、発話量を示す“文字数”といった形態的な特徴は考慮されていない。

3. 吹き出し形状推薦システムの概要

コミック閲読時に、人は吹き出しの形状から、(1) 発話ニュアンスの有無、(2) 発話ニュアンスの種類、を把握している。ここで、発話ニュアンスの有無を把握するとは、通常発話とのニュアンスが異なるかどうかを把握することであり、発話ニュアンスの種類把握とは、叫び声や弱々しい声などの感情や音量などの特性の把握を指す。発話ニュアンスが付与されていない場合には主に“風船型”が用いられる。また、発話ニュアンスが付与されている場合には、大きく7種類の異なる形状(“破裂型”、“波型”、“四角型”、“多角形型”、“雲型”、“角風船型”、“フラッシュ”)によって、それぞれの発話ニュアンスを視覚的に表現している。

提案手法では、Naive-bayes 手法によってコミックのセリフ特徴と吹き出し形状の関係性モデルを構築する。本稿の主な目的は、発話ニュアンスを吹き出し形状によって伝達するために、テキストから発話ニュアンスを推定して吹き出し形状を推薦することである。そのため、上記(2)に関係する7種類の吹き出し形状をセリフ特徴によって推定することを主な目的とする。一方で、(1)の課題である発話ニュアンスの有無もセリフ特徴から推定可能にすることで、発話ニュアンスの有無を判定したのちに、その発話ニュアンスの種類を判定可能な実用性の高い推薦が実現されると考える。本稿では、セリフ特徴を用いて7種類の形状を推定する7クラス分類の関係性モデルと、風船型であるか否かを判定する2クラス分類の関係性モデルをそれぞれ構築し、その推定性能を評価・考察する。

4. コミックにおけるセリフ特徴

先行研究 [田中 14] での分析により、吹き出し形状とセリフの末尾に現れる「語尾特徴」には一定の相関が確認されている。この他に、本稿ではコミックにおけるセリフ特徴として、文字数や品詞情報などの「形態的特徴」、感情語などのセリフが表す意味に関わる「意味的特徴」、キャラクターのセリフを発声した際の「音読時の特徴」を用いる。以下、それぞれの特徴量を列挙し、その抽出方法を示す。

語尾特徴 表 1 に、本稿で用いる語尾特徴の特徴 ID と詳細を示す。語尾特徴の抽出では、セリフの文字列を形態素解析し、セリフの最終端の形態素を起点としてセリフを逆

上り、表 1 に示した形態素以外の形態素が出現するまでの語尾特徴を観測し、その出現回数の特徴量とする。ただし、 f_8^{tail} については“!”、“?” の順で連続して形態素が出現した場合に、これら2つの形態素を1特徴として扱う。例えば、「つけた…っ!?’ というセリフの語尾特徴は $\{f_1^{tail}=1, f_8^{tail}=1, f_{10}^{tail}=1\}$ となる。

形態的特徴 表 2 に、文字数や品詞情報などの形態的特徴の特徴 ID と詳細を示す。字数に関わる特徴量 f_1^{str} , f_2^{str} では、それぞれセリフ中の数字の数と文字数を示す。このとき、セリフの文字数は各セリフで大きく値が異なるため、データベース中全てのセリフを参照した k-means 法を用いて5レベルに離散化し、このセリフの文字数の多さのレベルを特徴量とする。人名に関わる特徴量 f_3^{str} では、形態素解析の結果、“品詞細分類=人名”となる形態素が含まれている場合に $f_3^{str} = 1$ となる。文字種に関わる特徴量としては、 f_4^{str} , f_5^{str} , f_6^{str} を扱う。 f_4^{str} では、セリフがひらがなのみで構成されている場合に $f_4^{str} = 1$ となる。 f_5^{str} , f_6^{str} は、それぞれセリフ中にカタカナ、漢字が含まれている場合に $f_5^{str} = 1$, $f_6^{str} = 1$ となる。品詞に関わる特徴量 f_7^{str} , f_8^{str} , f_9^{str} では、セリフ中に出現した名詞、動詞、形容詞のそれぞれの数をそれぞれ特徴量とする。人称代名詞に関わる特徴量 f_{10}^{str} , f_{11}^{str} では、Wikipedia 中の“一人称単数代名詞の一覧”と“二人称単数代名詞の一覧”を参照し、それぞれの一覧に掲載された名詞が含まれている場合に $f_{10}^{str} = 1$, $f_{11}^{str} = 1$ とする。

意味的特徴 表 3 に、感情語などの意味的特徴の特徴 ID と詳細を示す。セリフの極性を示す特徴量として、 f_1^{mean} , f_{12}^{mean} , f_{13}^{mean} を用いる。これらの特徴量はセリフに含まれるバッドワード、ポジティブ語、ネガティブ語のそれぞれの出現数を値に持つ。このとき、バッドワードとして「ニコニコ生放送: 運営 NG ワード一覧」を参照した。感情に関わる特徴量 $f_2^{mean} \sim f_{11}^{mean}$ では、セリフ中

表 4: 音読時の特徴とその特徴 ID

特徴 ID	特徴
f_1^{voice}	モーラ数
f_2^{voice}	語尾が母音「a」終了
f_3^{voice}	語尾が母音「i」終了
f_4^{voice}	語尾が母音「u」終了
f_5^{voice}	語尾が母音「e」終了
f_6^{voice}	語尾が母音「o」終了

の基本 10 感情 [中村 93] のそれぞれの感情に該当する単語の出現数をそれぞれ値に持つ。

音読時の特徴 表 4 に、キャラクターのセリフを発声した際の音読時の特徴の特徴 ID と詳細を示す。発音数の特徴量 f_1^{voice} では、音韻の単位であるモーラ数を値としてもつ。特徴量 $f_2^{voice} \sim f_6^{voice}$ は、セリフの語尾がどの母音で終了するかを示す排他的な特徴量であり、バイナリで表現される。

5. セリフ特徴と吹き出し形状の関係性モデルの構築

本稿では、(1) モデル構築のための特徴選択、(2) 選択された特徴量を用いた Naive bayes による関係性モデル構築の手順でセリフ特徴と吹き出し形状の対応付けを実現する。モデル構築のための特徴選択では、表 1, 2, 3, 4 に示した全 49 のセリフ特徴について、吹き出し形状の 7 クラスあるいは 2 クラス分類における相互情報量の上位 20 特徴量を導出する。その後、その 20 特徴量について吹き出し形状の推定において最も高い性能を示す特徴量の組み合わせを交差検定によって明らかにし、セリフ特徴と吹き出し形状の関係性モデルを構築する。

5.1 相互情報量を参照した分類性能が高い特徴量の導出

各吹き出し形状を各カテゴリとしたとき、ある特徴 f の出現を表す確率変数 U 、あるカテゴリ c の出現を表す確率変数 C を用いることで相互情報量 $I(U; C)$ は、下式で算出される；

$$I(U; C) = \sum_{e_f} \sum_{e_c} P(U = e_f, C = e_c) \log_2 \frac{P(U = e_f, C = e_c)}{P(U = e_f)P(C = e_c)}, \quad (1)$$

ここで、 e_f および e_c はそれぞれ特徴量の使用の有無と吹き出し形状の種類を示す。

相互情報量は 2 つの確率変数の相互依存の尺度を表す量であるため、高い値を持つほど吹き出し形状に関連する特徴であると考えられる。この相互情報量が高い値を示す上位 20 特徴量を導出する。

5.2 交差検定の分類精度に基づいた特徴量の組み合わせ

交差検定を複数回行うことで、5.1 節で導出された 20 特徴量について 20C1~20C20 までの特徴量の組み合わせについて、最も高い推定性能を示す特徴量の組み合わせを得る。学習データには先行研究 [田中 14] で作成した吹き出しデータベースを用いて、Naive bayes 分類器による 10 分割交差検定を行った。表 5 に、7 クラス分類（発話ニュアンスを示す吹き出し形状の分類）と 2 クラス分類（風船型であるか否か）それぞれの推定において高い推定率を示した特徴量群を示す。

表 5: 特徴選択によって得られた 7 クラス分類（発話ニュアンスを示す吹き出し形状の分類）と 2 クラス分類（風船型であるか否か）それぞれの吹き出し形状推定に用いるセリフ特徴量群

7 クラス分類	2 クラス分類
f_1^{tail}	f_1^{tail}
f_4^{tail}	f_2^{tail}
f_6^{tail}	f_4^{tail}
f_7^{tail}	f_6^{tail}
f_8^{tail}	f_7^{tail}
f_9^{tail}	f_8^{tail}
f_{10}^{tail}	f_9^{tail}
f_{13}^{tail}	f_{10}^{tail}
f_{14}^{tail}	f_{13}^{tail}
f_1^{str}	f_{14}^{tail}
f_4^{str}	f_4^{str}
f_6^{str}	f_6^{str}
f_4^{mean}	f_1^{mean}
f_9^{mean}	f_{13}^{mean}

表 6: 風船型を除いた各吹き出し形状の適合率、再現率、F 値

吹き出し形状	適合率	再現率	F 値
破裂型	44.41%	67.36%	53.35%
波型	47.97%	45.45%	55.47%
多角形型	50.00%	5.50%	9.88%
雲型	50.00%	14.52%	22.50%
四角型	34.59%	57.50%	43.19%
角風船型	0.00%	0.00%	0.00%
フラッシュ	0.00%	0.00%	0.00%

6. 吹き出し形状推定実験

5. 章で構築したモデルを用いて、吹き出し形状の推定実験を行った。吹き出しデータベースには、5.2 節と同様に先行研究で構築した吹き出しデータベースを用いた。各セリフに対する正例は、吹き出しデータベース上の吹き出し形状データとする。構築した吹き出し形状モデルの性能評価には、10 交差検定を行った際の適合率、再現率、F 値を用いる。

6.1 7 クラス分類の推定結果：発話ニュアンスを表現する吹き出し形状の分類

表 6 に、風船型を除いた各吹き出し形状毎の推定結果を示す。破裂型、波型、四角型の吹き出し形状の推定において、比較的高い F 値が確認された。吹き出し形状モデル作成のために使用した表 5 に示した特徴量によって、特に破裂型、波型、四角型の吹き出し形状推定において比較的高い推定率を示すことが確認された。

一方で、角風船型、フラッシュ型については適合率、再現率ともに 0% となり全てのサンプルにおいて、正しい推定が行われなかった。これら角風船型とフラッシュ型の推定結果を見てみると、どちらも約 40% が破裂型であると誤推定されていた。破裂型として約 40% を推定することは、他の吹き出し形状における推定結果と比較したとしても高い値である。ここで、各吹き出し形状毎の使用されるシーンを見てみると、角風船型は少女漫画上で女性がセリフを強調するシーンで扱われやすく、フラッシュ型は心の中の叫び声を表現するときに扱われやすい

表 7: 風船型と風船型以外の吹き出し形状の適合率, 再現率, F 値

吹き出し形状	適合率	再現率	F 値
風船型	68.64%	87.63%	76.98%
風船型以外	67.61%	39.20%	49.64%

傾向にある。破裂型は叫び声や怒りなどを表すことが多い。これらのシーンで用いられるセリフでは、「!」や「っ」などが用いられることが多く、言語的特徴が類似している。このことから、セリフ特徴のみでは「心のなかの叫び」「女性のセリフ」といったメタな状況を判別できず、同様に強調されるセリフに対して用いられる破裂型と誤推定する割合が高くなったものとみられる。

6.2 2クラス分類の推定結果：風船型と風船型以外の判別

表 7 に、風船型と風船型以外の判別についての 2 クラス分類の推定結果を示す。風船型であることの推定においては、特に再現率が高く、比較的高い F 値が確認された。しかしながら、風船型以外であることの推定においては、低い推定性能となった。このことから、表 5 に示した特徴量のみでは、風船型であるか否か分類において十分な推定性能が得られないことが確認された。これらの風船型であるか、他 7 つの形状の吹き出し形状であるかを判定するためには、単一のセリフのみならず会話のコンテキストなどの他の特徴量の検討が必要である。

風船型はコミックにおいて標準的な吹き出し形状である。そのため、コミュニケーションツール内での実利用場面においてはデフォルトの吹き出し形状を風船型としておき、ユーザの任意で吹き出し形状を変化させる場合にのみ、6.1 節での 7 クラス推定器によって吹き出し形状を推薦するといった応用的な適用も考えられる。

7. おわりに

本稿では、テキストコミュニケーションにおける吹き出し形状を用いた発話ニュアンスの伝達を目指して、入力テキストからの吹き出し形状の推薦手法を提案した。特徴選択の結果、発話ニュアンスを表現する吹き出し形状の推定に有用なセリフ特徴は、セリフ語尾に関わる特徴量であった。選択された特徴量を用いた吹き出し形状推定実験では、破裂型、波型、四角型の吹き出し形状の推薦において比較的高い F 値が確認された。また、発話ニュアンスの有無を判定する 2 クラス分類では、風船型の F 値 76.98% の推定性能を示し、発話ニュアンスの有無の判定についても一定の性能を示すことが確認された。

今後は、特徴選択および機械学習手法の比較検討および文脈を意識した吹き出し形状の推薦に取り組んでいく。また、テキスト会話において対話者の発話ニュアンスを吹き出し形状によって視覚的に表現することは、テキストコミュニケーションにおける発話ニュアンスの理解性向上のみならず、神経発生的障害患者の心情の読み取りにおいても高い有用性 [キャロル 05] が期待される。今後は、吹き出し形状を伴うテキスト会話の心的影響についても考察を行っていく。

謝辞

本稿では、例示のためにコミック中のセリフを一部引用させて頂いた。また、コミックのコマ画像の例示のために、Manga109

のデータセット [Matsui 15] を利用させて頂いた。データセット内の赤松健氏の「ラブひな」、新沢基栄氏の「ハイスクール! 奇面組」を利用させて頂いた。記して謝意を表す。

参考文献

- [Matsui 15] Matsui, Y., Ito, K., Aramaki, Y., and Aizawa, T. Y. K.: Sketch-based Manga Retrieval using Manga109 Dataset, *CoRR*, Vol. abs/1510.04389, pp. 1–13 (2015)
- [キャロル 05] キャロル グレイ, 門 眞一郎 (訳): コミック会話 自閉症など発達障害のある子どものためのコミュニケーション支援法, 明石書店 (2005)
- [紺家 15] 紺家 裕子, 椎尾 一郎: 吹き出し枠形状を利用した音声表現の字幕提示方法, 電子情報通信学会論文誌, Vol. J98-A, No. 1, pp. 85–92 (2015)
- [松下 13] 松下 光範: コミック工学の可能性, 第 2 回 Web インテリジェンスとインタラクション研究会, pp. 63–68 (2013)
- [松宮 14] 松宮 翔, Sakti, S., Neubig, G., 戸田 智基, 中村 哲: 音響特徴量を用いた吹き出しテキストの生成, 日本音響学会, pp. 11–12 (2014)
- [竹内 05] 竹内 オサム: マンガ表現学入門, 筑摩書房 (2005)
- [中村 93] 中村 明: 感情表現辞典, 東京堂出版 (1993)
- [長谷 00] 長谷 邦夫: 漫画の構造学!, インデックス出版 (2000)
- [田中 14] 田中 秀樹, 山西 良典, 福本 淳一: コミックにおけるセリフ内容と吹き出し形状の関係性についての基礎検討, 電子情報通信学会研究報告. HCG シンポジウム 2014, pp. 286–289 (2014)