

音声入力テキストチャットにおける 発話強度に基づいた吹き出し形状自動選択システムの提案

青木 俊樹^{†1, a)} 川村 慧^{†1, b)} 中條 麟太郎^{†1, c)}
崔 セミ^{†2} 松井 克文^{†1, d)} ハウタサーリ アリ^{†1, e)}

概要：本稿では、漫画の表現方法として知られている吹き出しの、感情を伝達する機能に着目し、音声入力テキストチャットにおいて、入力された音声の発話強度に基づいた吹き出し形状の自動選択システムを提案する。以前から、感情と吹き出しの形状の対応に関する研究が行われていたが、吹き出しの形状のバリエーションによってユーザーが感じる感情の差異についての研究は行われていない。また、感情を伝達できる従来のテキストチャットシステムにおいては、多くの場合ユーザーによる感情の選択が別途必要であった。本稿では、感情と吹き出しの形状のバリエーションへの対応について質問紙調査を行い、ランキングを行った。また、ランキングに対応した吹き出しを、入力された音声データを元に自動で選択するテキストチャットシステムを実装し、ユーザーの体験から本システムの有用性を確認した。

1. はじめに

漫画は古くから多くの日本人に愛されているメディアであり、現在では日本のポップカルチャーの一つとして、日本国内のみならず海外にもその表現方法は広く認知されている。漫画では、吹き出しがセリフを伝える表現方法として用いられ、その形状や大きさ、文字のフォントによって話者の発話の様子を推測することができる。例えば、ギザギザの吹き出しで表現されているセリフに対して、私たちは「大きな声」で発話されたかと推測することができ、同時に「怒り」などの発話時の感情を推測することもできる。漫画が広く受け入れられている現状において、読者は吹き出しの表現する発話者の感情について、ある程度の共通認識を持っていると推測できる。また、吹き出しはテキストチャットアプリケーションにおけるメッセージの表現方法の一部にも使われているが、その表現の幅は漫画と比べて単調である。本稿では、吹き出しの形状を自動で選択することで、感情をより伝えやすいテキストチャットシステムの開発を目的とする。

漫画における吹き出しの表現方法は筆者によって異なるが、発話が伴う吹き出しについては、「丸いもの」「丸が途中でくびれたもの」「雲状のもの」「多角形」「波型」「破裂型」などのいくつかの形状に分類されており[1]、概してそれぞれの形状に対応する感情があると考えられる。しかし、同じ分類の中での吹き出しの形状のバリエーションによって、ユーザーが感じる感情の差異については明らかになっていない。

そこで本稿では、特に「丸いもの」「破裂型」の吹き出

しの形状に着目し、吹き出しの形状のバリエーションによって読者が感じる感情の差異の調査結果を報告する。また、調査から得られた知見を元に、発話者の発話強度に応じて吹き出しの形状が変化するテキストチャットシステムを開発した。本稿では、測定した音声から算出される RMS の最大値を発話強度と定義する。

2. 関連研究

感情表現を含んだテキストチャットとして実用化されているものには、SEGA によるオンラインゲーム「PHANTASY STAR ONLINE 2」におけるテキストチャットや[2]、Apple による「iMessage」における「吹き出しエフェクト」があげられる[3]。前者では、テキスト入力後に吹き出しの形状、色、効果、出現方法などを指定することにより、感情の表現が可能である。後者では、テキスト入力後に「スラム」「ラウド」「ジェントル」「見えないインク」の4種類からエフェクトを選択することで、吹き出しのアニメーションが変わり、感情が表現される。

しかしながら、これらのテキストチャットシステムでは文字情報の入力に加えて、感情情報の入力が別途必要であり、ユーザーに負担を強いるという問題点がある。これを解消する研究として、ユーザーの生体信号から吹き出しを自動的に生成する手法があげられる[4]。この手法では、筋電位と皮膚伝導を計測することによりユーザーの感情を読み取り、対応する吹き出しを表示しており、ユーザーに操作負担をかけずに感情を伝達できる一方で、通常のテキストチャットでは必要のない装置を装着する必要がある点ではユーザーの負担は依然として存在すると言えるだろう。

そのため、本研究では現在入力方式として実用化されている音声入力に着目し、入力された音声データと感情を関連付けることでユーザーに負担をかけずに感情表現を含んだテキストチャットを実現している。

^{†1} 東京大学

^{†2} 独立研究者

a) aoki-toshiki1127@g.ecc.u-tokyo.ac.jp

b) wzdmriv@g.ecc.u-tokyo.ac.jp

c) rintaro-chujo@g.ecc.u-tokyo.ac.jp

d) matsui.katsufumi@mail.u-tokyo.ac.jp

e) a.hautasaari@iii.u-tokyo.ac.jp

また、吹き出しと感情の対応に関する先行研究において、いくつかの種類の吹き出しはすでに感情との対応付けが行われており[4][5]、「怒り」の感情が「破裂型」の吹き出しに対応していることや、覚醒度に応じて吹き出しを鋭角から丸みを帯びさせるという対応付けが適していることが明らかになっている。Lang の二次元感情モデル[6]における感情の正負と、吹き出しの形状の違いが対応していないことも先述の研究から明らかになっている。

3. 提案システム

3.1 予備実験

予備実験では、「破裂型」と「丸いもの」に分類される吹き出しの形状のバリエーションに対して、Lang の二次元感情モデル[6]に基づき、覚醒度での順位付けを試みた。質問紙調査を行い、得られたデータに対してアルゴリズムを用いて解析を行った。

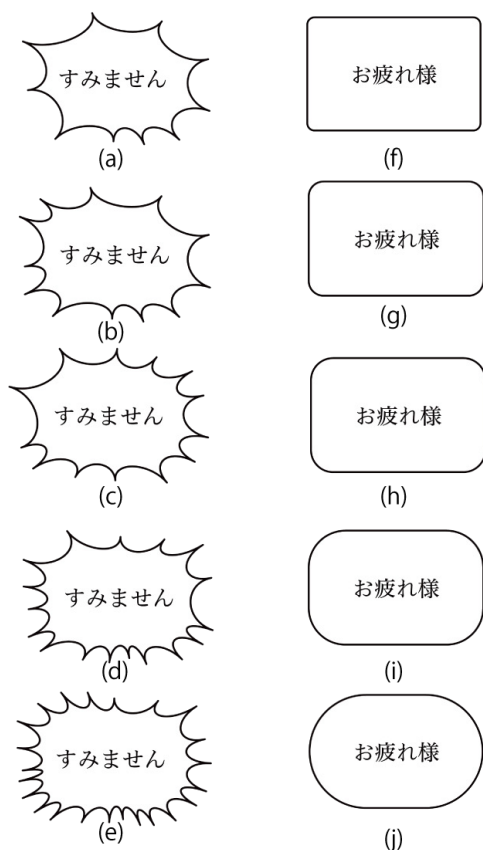


図 1 実験に使用した吹き出し

図 1 の (a) から (j) にある通り、異なる造形の吹き出しを 10 種類用意した。これは、一般的に漫画に使われている吹き出しの中で「破裂型」「丸いもの」の二種類に絞り、それぞれ丸み具合、鋭角の頂点の数を調節したものである。

データを Pairwise-ranking-Aggregation-master アルゴリズム[7]にかけることを前提に、被験者には二種類の吹き出しの画像の比較を提示した。

「破裂型」の吹き出しとして用意した (a) から (e) までの 5 種類から、2 種類の吹き出しを提示し、「どちらの方がより興奮して見えますか？」という質問への回答を求めた。5 種類から作ることのできる、全ての 2 種類の組み合わせ計 10 種類について同じ質問を行った。

同様に、「丸いもの」の吹き出しとして用意した (f) から (j) までの 5 種類においても、2 種類の吹き出しを提示し、「どちらの方がより穏やかに見えますか？」という質問への回答を求めた。5 種類から作ることのできる、全ての 2 種類の組み合わせ計 10 種類について同じ質問を行った。

3.2 結果

「破裂型」の質問紙調査に応じた被験者は、18 歳から 21 歳の 17 名、「丸いもの」の質問紙調査に応じた被験者は、18 歳から 21 歳の 15 名であった。質問紙調査はインターネットを通じて行った。図 2 は、質問紙調査によって得たデータに対して、Pairwise-ranking-Aggregation-master アルゴリズム[7]を用いて解析を行った結果である。

「破裂型」の吹き出しについては、図 1 で示した (b) (c) (e) (d) (a) の順に「興奮して見える」という結果が得られた。また、「丸いもの」の吹き出しについては、図 1 で示した (g) (i) (h) (f) (j) の順に「穏やかに見える」という結果が得られた。

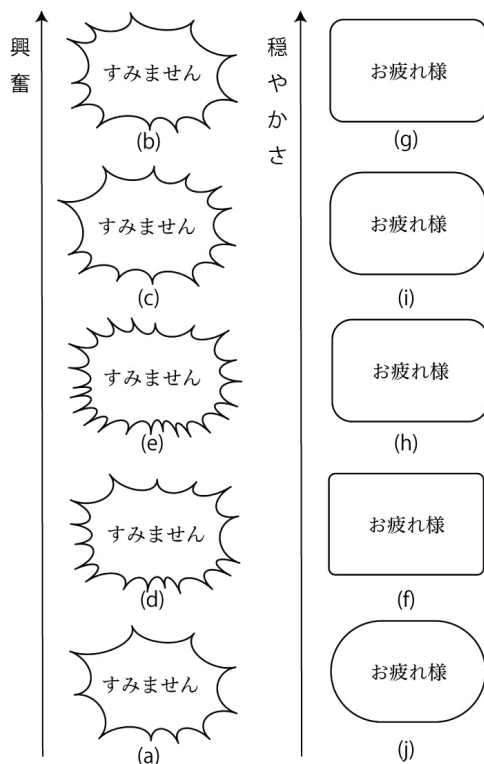


図 2 アルゴリズムを用いて解析を行った結果

3.3 提案システム

実験の結果を元に、発話強度と覚醒度に一定の対応があると仮定して、発話者の発話強度に応じて吹き出しの形状が変化するシステムを実装した。これは、先行研究におい

て、覚醒度と対応する感情である「喜び」について[6]、発話強度と感情強度に関係が大きく見られることが分かっていること[8]を背景としている。今回はタブレットデバイスでの使用を前提とした、テキストチャット型のインタフェースを作成した。以下にその詳細を示す。

図3は、ディスプレイに表示される本システムの様子である。現在のシステムは、評価のためのプロトタイプであるため、デバイス間の通信機能は実装していない。画面中央が吹き出しの表示部分である。これは音声入力テキストチャットを想定しており、新たに音声により入力された文章とそれに伴って選択された吹き出しが画面上に表示される。右下に位置する「Start」ボタンを押下することで録音が始まり、音声認識されると同時に発話強度が記録され、「Start」ボタンの右隣に位置する「Send」ボタンを押下することで、発話強度に対応する形状の吹き出しが文章とともにリストの最下層に出現する。

実験では「丸いもの」「破裂型」のそれぞれの吹き出しについて、覚醒度での順位付けを行ったが、今回の実装では「破裂型」「丸いもの」の順に覚醒度が下がると仮定し、発話強度が小さい場合は「丸いもの」、発話強度が大きい場合は「破裂型」から適切な吹き出しを選択している。また、吹き出しの形状は、音声入力テキストチャット型のインタフェースに合わせて、図4に示した横長のものを利用している。

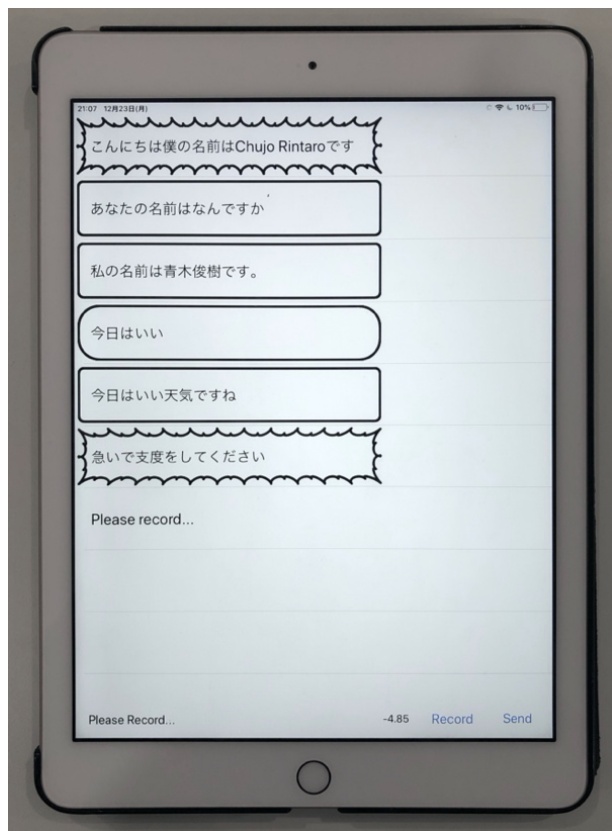


図3 吹き出しの形状が変化するテキストチャットシステムの実装例

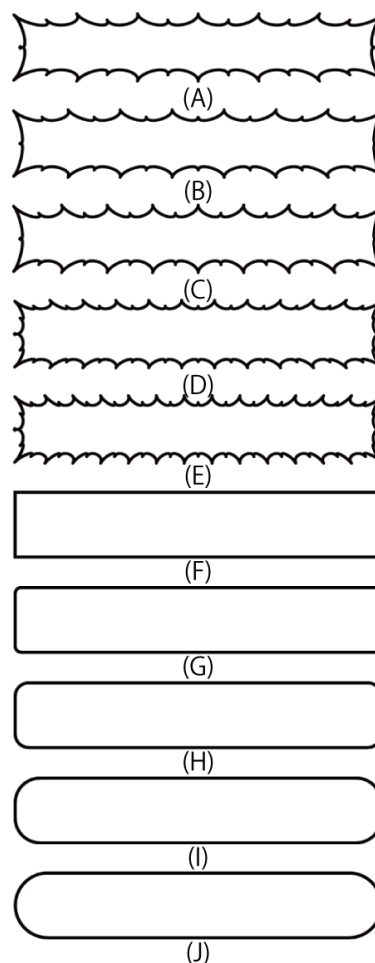


図4 実装に使用した吹き出し

4. ユーザスタディ

実装した音声入力テキストチャットシステムを用いて、フィードバックを得るためにユーザの観察とインタビューを行った。

4.1 方法

18歳と21歳の2名を被験者として、提案システムを用いた音声入力テキストチャットを複数回体験させた。終了後、本システムの有効性について回答を求めた。

4.2 結果

体験中、ユーザが様々な発声方法で入力を試す様子が観察できた。また、何回か入力続けることで、任意の吹き出しの生成するために必要な音声について探るといった様子も見られた。

終了後、発話時の感情に適切な吹き出しが選択されていたかについて確認した。「破裂型」の吹き出しが表示される発話強度での入力については概ね想定通りの表示であるが、「丸いもの」の吹き出しが表示される発話強度での入力については感情と吹き出しが一致していないという意見があった。「破裂型」と「丸いもの」のそれぞれの吹き出しの関係については今後検討したい。

体験の感想では、テキストチャットにおける感情伝達は有用である、また新しいコミュニケーション手段であるといったポジティブな感想が多く見られた一方で、音声入力を常用しておらず、音声入力自体に違和感があるというネガティブな感想も得られた。しかしながら、調査では若年層ほど音声入力の利用経験がある人の割合が高くなることが明らかになっていることから[9]、将来的に音声入力はより広く使われるようになると考えられる。このことは、音声入力テキストチャットシステムにおいて、入力された音声に対応する吹き出しを表示する本システムが社会に受容される可能性を示唆している。

5. まとめ

本稿では、吹き出しの形状の違いから感じられる感情の違いについて実験を行った。また、実験結果に基づいた音声入力テキストチャットの実装を行った。実装では、入力された音声の発話強度から吹き出しの形状を自動で選択するテキストチャットシステムを開発し、ユーザスタディによりユーザからのフィードバックを得た。

吹き出しという表現方法は、多くの人にすでに馴染みがあるだけでなく、感情との関連が強く、コミュニケーションに活用することが有意義であると考えられる。今後は、本研究で実装した音声入力テキストチャットに限らず、ビデオ通話や異言語間コミュニケーションへの展開を目指すと同時に、感情と吹き出しの形状の対応について、対応関係の精査をさらに行っていく。

謝辞 本研究は JSPS 科研費 JP18K18085 の助成を受けたものです。

参考文献

- [1] 夏目房之介: マンガはなぜ面白いのか: その表現と文法, NHK 出版, 1997.
- [2] “SEGA チャット | 『ファンタシースターオンライン 2』プレイヤーズサイト | SEGA”, <http://pso2.jp/players/manual/communication/chat/>(参照 2019-12-22)
- [3] “iPhone, iPad, iPod touch の iMessage でメッセージエフェクトを使う - Apple サポート”, <https://support.apple.com/ja-jp/HT206894> (参照 2019-12-22)
- [4] 小俣昌樹, 内藤雄也, 今宮淳美: 生体信号に連動する吹き出しオンラインチャットの設計と評価, ヒューマンインタフェース学会論文誌, Vol. 10, No. 2, pp. 179-190, 2008.
- [5] R. Yamanishi, H. Tanaka, Y. Nishihara, J. Fukumoto: Speech-balloon Shapes Estimation for Emotional Text Communication. Information Engineering Express, Vol. 3, No.2, pp. 1-10, 2017.
- [6] P. J. LANG, The emotion probe. Studies of motivation and attention, The American psychologist, Vol. 50, No. 5: pp. 372-385, 1995.
- [7] X. Chen, P. N. Bennett, K. Collins-Thompson, E. Horvitz: Pairwise ranking aggregation in a crowdsourced setting, Proceedings of the sixth ACM international conference on Web search and data mining, ACM, pp. 193-202, 2013.
- [8] 伊東久美子: 感情を含む音声に関する研究 II 合成単母音 [え] による音響パラメータ評価, 人間工学, Vol. 21, No. 2, pp. 81-87, 1985.
- [9] “「文字入力による情報検索『面倒』は半数 「人前での音声検索は『恥ずかしい』7割超, 家電などの音声操作, 4割が『自宅に人がいなければ』利用したい」 | 2017年 | KDDI株式会社”, <http://news.kddi.com/kddi/corporate/newsrelease/2017/10/05/besshb2726.html> (参照 2019-12-22)