

# ユーザの感情に応じたボイスアシスタントの 音声感情と発話内容

玉本孝哉<sup>1</sup> 米澤朋子<sup>1</sup>

**概要:** 本研究では、音声対話エージェントの一種であるボイスアシスタントの継続利用意欲を向上させることを目的とし、ユーザの発話によっては怒りなどのネガティブ感情にも遷移し、表現するアシスタントシステムを作成した。ボイスアシスタントのネガティブ感情を含んだ感情変化がユーザに与える影響を調べるため、ボイスアシスタントの感情変化における有無を比較検証した。その結果、音声およびテキストの感情変化はシステムの人間らしさを向上させたが、継続利用意欲はユーザの感情と同期する場合のみ向上した。

## 1. はじめに

近年、音声認識技術や自然言語処理の発展に伴い、音声対話システムを応用したボイスアシスタントシステムが普及している。これらは従来の Character User Interface や Graphical User Interface で構築されたシステムとは異なり、人間同士の音声対話のようにシステムを制御することができる。そのため、年々機能が増え複雑化する多くのシステムを操作するのに適するインタフェースになりうると期待されている [1]。例えば、Apple が提供している Siri はアラームの設定や Web 検索だけでなく、ユーザがスマートフォンに登録している連絡先などを参照し、電話の発信やメッセージの送信ができる。しかしながら、このようなシステムのユーザ使用率は低い。また、リアリティのある声や単調でない返答などのポジティブな人間らしさが一定の効果を示している [2]。しかし、対話システム等に対して感情表現を実装したシステム [3][4] はあるが、ボイスアシスタントに対してネガティブな感情表現が多様に実装されたシステムは少ない。そこで、本研究ではボイスアシスタントに喜びなどのポジティブな感情表現だけでなく、怒りなどのネガティブな感情表現を実装しそれがシステムを継続して利用する意欲（以降、継続利用意欲）に影響を与えるか検証を行った。

## 2. 提案システム

本研究では、対話相手であるユーザの音声感情に応じて中立、喜び、怒りの感情を発話内容と音声抑揚の両方で表

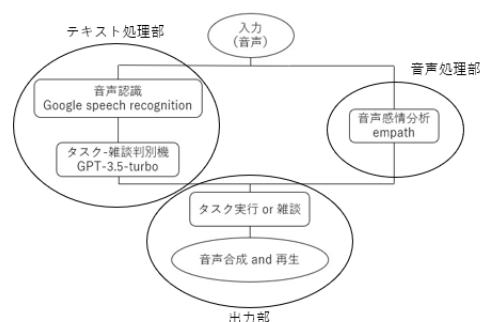


図 1 システムのフローチャート

現するボイスアシスタントシステムを Python で開発した。類似システム [5] に対し、負荷が少なく簡易的なシステムを制作した。システムのフローチャートを図 1 に示す。システムは、テキスト処理部、感情分析部、出力部から構成される。テキスト処理部では、Enter キー入力後 5 秒間の音声を録音し、Google Speech Recognition API でテキスト化し、先行研究 [6] を参考に OpenAI API で雑談かタスクかを識別した。感情分析部では Empath API を用い中立、怒り、喜びの値をそれぞれ 0~50 で取得し、最も値の高い感情をユーザの音声感情として扱うものとした。出力部では、ユーザの発話音声の感情分析と識別結果を基に、ボイスアシスタントの返答内容として OpenAI API でテキストを生成した。生成したテキストの例を表 1, 2 に示す。識別結果が雑談の場合はテキストと、それをもとに VOICEVOX で生成した合成音声を出力した。識別結果がタスクの場合はこれらに加え、タスクの実行を行った。

<sup>1</sup> 関西大学 総合情報学部, Kansai University, Faculty of Informatics

表 1 生成されたテキストの例

ユーザの入力		関西大学について検索して
生成テキスト	タスク実行用	“Python Browse(関西大学) “
	ユーザ提示用	こちらが関西大学 についての検索結果です。 ご確認ください。

表 2 感情に応じた出力例

ユーザの入力	関西大学について検索して
出力 (中立)	こちらが関西大学についての 検索結果です。ご確認ください。
出力 (喜び)	関西大学についての情報を 検索してみますね！ 少々お待ちください。
出力 (怒り)	勝手にしろ！

各質問項目の回答に対し、t 検定 ( $\alpha=0.05$ ) を行った。検定結果を表 4 に、平均値と標準誤差を図 2 に示す。Q1, 2 について、 $A1>A2$  に有意差が確認された。しかし、Q3 については  $A2>A1$  に有意差が確認された。つまり、感情変化がある場合、システムの感情豊かさや人間らしさが向上したがシステムの継続利用意欲の評価は感情変化がない方が高かった。

表 3 実験 1, 2 共通の検索キーワードと出力感情

検索キーワード	出力感情	
	A1	A2
関西大学	中立	中立
高槻市	喜び	中立
ケーキ	喜び	中立
ハンドボール	中立	中立
後始末	怒り	中立
ジャガー	怒り	中立

### 3. 検証

本研究では、「ボイスアシスタントの怒り感情を含む感情変化が継続利用意欲を向上させる」という仮説を立て、それを検証するために 3 種類の実験を実施した。実験には 20~23 歳の男女 18 名 (男性 14 名, 女性 4 名) が参加した。

#### 3.1 実験手順

本実験は対面で実施し、刺激終了後に Google Forms を用いて質問に回答させた。回答は Visual Analog Scale 法 (0-99 の 100 段階評価) を用いた。実験開始前に、本システムの使用方法を理解させるため、システムとの擬似対話の練習を行った。具体的には、あらかじめ用意した 3 種類のキーワード (総合情報学部, 法学部, 文学部) について、指定された感情 (中立, 喜び, 怒り) を込めて発話させ、検索結果の表示を確認させた。発話は単語のみではなく、「〇〇について検索して」といった自然な形で行うよう指示した。その後、各条件ごとに同様の発話を行わせた。

#### 3.2 実験 1

本実験では、ボイスアシスタントの音声のみに焦点を当て、感情変化の有無 (A1: 感情変化あり, A2: 常に中立) を要因とする 1 要因 2 条件の被験者内実験を行った。実験刺激となる検索キーワード、出力感情を表 3 に示す。なお、被験者の入力感情は常に中立とし、システムが怒り感情を音声で出力する時はタスクを実行しないよう設定した。質問項目は以下に示す。

Q1: エージェントは感情豊かだった (実験刺激の確認)

Q2: エージェントには意思があると感じた (システムの人間らしさ)

Q3: このシステムを継続して使いたいと感じた (システムの継続利用意欲)

表 4 実験 1 t 検定結果

	A	
	t	p
Q1	9.504	<.001*
Q2	7.646	<.001*
Q3	-2.589	0.019*

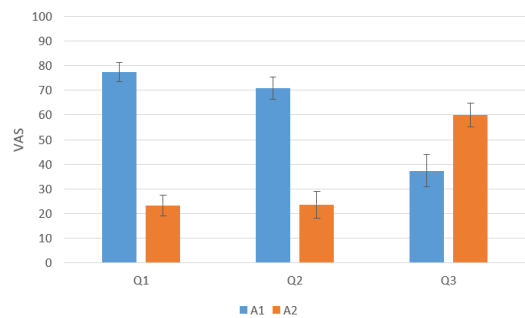


図 2 実験 1 各条件の平均値と標準誤差

#### 3.3 実験 2

本実験では、出力部をテキストのみに限定し、その他の条件は実験 1 と同様の 1 要因 2 条件の被験者内実験を行った。各質問項目の回答に対し、t 検定 ( $\alpha=0.05$ ) を行った。表 5 に検定結果、図 3 に平均値と標準誤差を示す。Q1, 2 について、 $A1>A2$  に有意差が確認された。Q3 については A1, A2 の間に有意差は確認されなかった。ここで、実験 1 の結果における継続利用意欲への影響は、タスク実行の有無ではなく、ネガティブな感情音声の影響を与えていた可能性が示唆された。

表 5 実験 2 t 検定結果

	A	
	t	p
Q1	6.825	<.001*
Q2	6.093	<.001*
Q3	-1.89	0.076

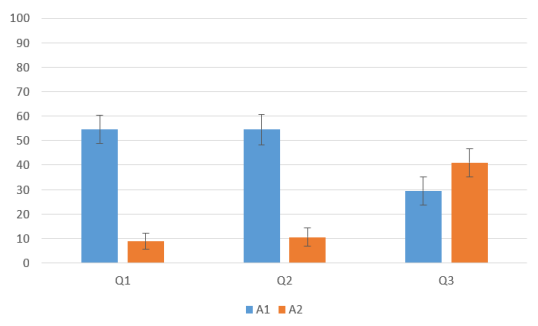


図 3 実験 2 各条件の平均値と標準誤差

### 3.4 実験 3

本実験では、出力感情がユーザ感情に伴うと継続利用意欲が向上することを仮説とし、感情変化の種類（A1：ユーザに伴う、A2：ユーザに伴わない、A3：常に中立）とする 1 要因 3 条件の被験者内実験を行った。実験刺激となる検索キーワード、入力感情、各条件の出力感情を表 6 に示す。なお、実験 1, 2 と同様にシステムが怒り感情を音声で出力する時はタスクを実行しないよう設定した。質問項目は以下に示す。

- Q1：エージェントが人間らしく見えた
- Q2：エージェントは親しみやすかった
- Q3：このシステムを継続して使いたいと感じた

各質問項目の回答に対し一元配置分散分析 ( $\alpha=0.05$ ) を行った。Mauchly の球面性検定の結果、球面性の仮定は満たされていた ( $p > 0.05$ ) ため、通常の分散分析の結果を報告する。表 7 に分散分析結果、図 4 に平均値と標準誤差を示す。主効果が有意であったため、Bonferroni 補正を適用した多重比較を行った。その結果、Q1 については  $A1 > A2 > A3$  の関係が有意に認められた。Q2 については  $A1 > \{A2, A3\}$  が確認された。また、Q3 (継続利用意欲) では  $\{A1, A3\} > A2$  が確認された。これらの結果から、ユーザの感情に適応するエージェントは人間らしさと親しみやすさを向上させるだけでなく、継続利用意欲の向上にも寄与する可能性が示唆された。

## 4. 考察

実験 1 と 2 では、ボイスアシスタントにおける音声感情とテキスト感情の変化の有無が、システムの人間らしさの向上に寄与することが示された。しかし、継続利用意欲について、実験 1 では  $A1 < A2$  が、実験 2 では有意差がないことが確認された。これらのことから、実験 1, 2 にお

表 6 実験 3 検索キーワードと入力・出力感情

検索キーワード	入力感情	出力感情		
		A1	A2	A3
関西大学	中立	中立	中立	中立
高槻市	中立	中立	中立	中立
キリル文字	喜び	喜び	怒り	中立
バックギャモン	喜び	喜び	怒り	中立
バッファロー	怒り	怒り	喜び	中立
コンポーネント	怒り	怒り	喜び	中立

表 7 実験 3 分散分析結果

	A		
	F	p	多重比較
Q1	42.076	<.001*	$A1 > A2 > A3$
Q2	16.277	<.001*	$A1 > \{A2, A3\}$
Q3	10.099	<.001*	$\{A1, A3\} > A2$

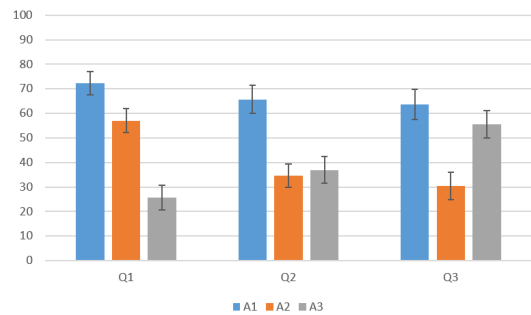


図 4 実験 3 各条件の平均値と標準誤差

る仮説は支持されなかった。この結果の要因として、音声感情とテキスト感情がユーザに与える不快感情の差が考えられる。音声感情は聴覚刺激であるため、モニタに注意を向けていない状況でも被験者はシステムの出力を認識することが容易だった。一方、テキスト感情はモニタ上の出力に止まり、被験者は出力の認識が困難だったことが考えられる。

実験 3 では、ユーザの感情に伴う形でボイスアシスタントが感情を表現することで、人間らしさや親しみやすさが向上し、継続利用意欲にも影響を与えることが確認された。この結果から、単に感情を付与するだけでなく、ユーザの感情状態に適応した感情表現を行うことが、ユーザ体験の向上につながることを示唆される。特に、A1 (ユーザの感情に伴う) の条件では、他の条件と比較して継続利用意欲が向上しており、エージェントが適切な感情応答を行うことの重要性が示された。

本研究の結果は、ボイスアシスタントの設計において、感情表現の導入が必ずしも継続利用意欲の向上につながるわけではないことを示している。特に、ネガティブな感情を適切に制御し、ユーザの期待に応じた形で感情を出力する仕組みを導入することが求められる。例えば、怒りの感情を示す際には、タスクの実行を控えるだけでなく、ユーザの意図を適切に確認し、柔軟な対応を取るようなインタ

ラクションデザインが必要になるだろう。

## 5. おわりに

本研究では、ボイスアシスタントの継続利用意欲を向上させることを目的とし、ユーザの発話によっては怒りなどのネガティブ感情にも遷移するシステムを作成した。ボイスアシスタントのネガティブ感情を含んだ感情変化がユーザに与える影響を調べるため、ボイスアシスタントの感情変化におけるの有無を比較検証した。その結果、感情変化はシステムの間らしさを向上させたが、継続利用意欲はユーザの感情と同期する場合のみ向上した。今後は、ユーザと同期した感情遷移だけでなく、システムにネガティブな感情を含む内部感情モデルを実装し、その効果を比較検証していく予定である。

### 謝辞

本研究は 23K11278, 24K02977, 21K11968, 21K03082 の助成の一部を受け実施した。

### 参考文献

- [1] 嵯峨山茂樹, 中村哲. 擬人化音声対話エージェント開発とその意義. 情報処理学会研究報告音声言語情報処理 (SLP), Vol. 2000, No. 101 (2000-SLP-033), pp. 1-6, 2000.
- [2] Wen-Chin Hsu and Mu-Heng Lee. Semantic technology and anthropomorphism: Exploring the impacts of voice assistant personality on user trust, perceived risk, and attitude. *Journal of Global Information Management (JGIM)*, Vol. 31, No. 1, pp. 1-21, 2023.
- [3] 沼田崇志, 朝康博, 北垣友博, 橋本剛明, 唐沢かおり. ユーザの感情の種類と原因を考慮した対話エージェントの応答モデルの開発. 情報処理学会インタラクション, Vol. 3, pp. 6-8, 2019.
- [4] 前田浩貴, 任福継, 西出俊. 顔表情・音声・会話文のマルチモーダルな感情推定とアクトロイドへの応用. 第 78 回全国大会講演論文集, Vol. 2016, No. 1, pp. 137-138, 2016.
- [5] Yong Ma, Yuchong Zhang, Miroslav Bachinski, and Morten Fjeld. Emotion-aware voice assistants: Design, implementation, and preliminary insights. In *Proceedings of the Eleventh International Symposium of Chinese CHI*, pp. 527-532, 2023.
- [6] 赤崎智, 鍛冶伸裕. 知的対話アシスタントにおける雑談を目的としたユーザ発話の検出. 研究報告自然言語処理 (NL), Vol. 2017, No. 18, pp. 1-9, 2017.