

# 人と機械のコミュニケーションを円滑にする VUI の研究

大井一真, 北村壮, 立道大樹, 中森涼太, 細見謙太

大阪ハイテクノロジー専門学校 人工知能学科

**要旨：** 近年, Apple Inc. や Amazon.com, Inc. などの大手 IT 企業から VUI (Voice User Interface) を用いた製品が多数販売されている。一方, 調査会社の実施したアンケートでは, VUI を用いた製品を日常的に利用している人は全体の 1 割という結果となり製品の広がりほど利用が浸透していないことがわかった。そこで本研究では, VUI が持つ要素を変化させるプログラムを作成し, 第三者による評価を考察することで人間にとって心地よい機械との対話の実現を目指し VUI がより多くの方に使われるものに必要なものが何かを探った。

**Keywords :** VUI, ボディ, ペルソナ, アビリティ

## 1. 緒言

VUI とは Voice User Interface の略称であり, コンピュータと人間が音声を通じてやり取りする仕組みである。VUI は, 声質や口調から感じ取れる VUI の性格を表す「ペルソナ」, 視覚的要素である「ボディ」, 能力を示す「アビリティ」の 3 つの要素で構成される[1]。

現在, 音声アシスタントとしてよく知られているものに Siri や Google アシスタント, Alexa などがある。しかし音声アシスタントの利用状況の調査では, 日常的に利用しているのは全体の 1 割という結果がでている。音声アシスタントに話しかけることについては, 全体の 4 割が抵抗を感じると回答した[2]。このように音声アシスタントの利用経験はあるが, 日常的な利用が進んでいないのが現状である。私たちは, VUI とのやりとりにおいて人間が心地よく感じることで, 今以上に VUI の利用率が増加すると考えた。そのため, 人間が VUI とのやり取りで心地よく感じる要素を明らかにしようと考えた。本研究で定義する心地よさとは, VUI を体験する上で次のように感じることにした。(1) 声が気に入る, (2) 口調が気に入る, (3) 機械と会話した体験が楽しく感じる。の 3 つである。

本研究はペルソナとボディを変化させたらどのような効果が現れるか明らかにすることを目的とし, 仮説として, ペルソナとボディが人間に近いイメージの方が会話も楽しく, それによって心地よさを感じ, VUI の利用率の向上に繋がると考えた。

## 2. 方法

### 2-1 検証概要

本研究ではボディ・ペルソナそれぞれ 2 種類の検証を行った。利用シーンとしてネットショッピングでチョコレート菓子を購入する場合とした。これは多くの方が体験しているネットショッピングのシチュエーションで, VUI との会話に何が心地よさを与えることができるのかを検証した。

### 2-2 対象

本検証の対象は本校のバイオ・再生医療学科 2~3 年の学生 20 名とした。

### 2-3 調査項目

2つの項目について検証を行った。1つ目はボディの変化が検証シーンに適切か調査する。2つ目は検証シーンに応じてどのような声質と口調が適切か調査した。

### 2-4 検証1 ボディの変化

1つ目の検証では、ボディの変化がVUI体験の心地よさに影響を与えるのか調査した。この検証では、商品に応じてボディを変化させることで、VUI体験者にどのような印象を与えるかを見た。

表1 検証1におけるボディの変化

		最初のボディ	次のボディ
パターン1	商品に応じて変化する		
パターン2	商品に応じて変化しない		

イは変化させなかった。

### 3. 結果

表1は検証に用いた各パターンでボディをどのように変化させたかを示している。パターン1では、最初のボディを商品に応じて変化させた。今回の検証ではチョコレートの購入となるため、チョコレートを検索したタイミングで、スーパーの店員のようにエプロンを付けたボディに変化させた。また、パターン2ではボデ

### 2-5 検証2 声質と口調の組み合わせ

2つ目の検証では、異なるペルソナで印象が変わるかを調べた。表2のように、機械の喋る声質と口調を変え、4つのパターンのペルソナを検証した。

ボディは実験を通じて同じ組み合わせを用い、途中で変化はさせなかった。

表2の機械音声は、Windows標準機械音声である、Microsoft Sayakaを使用した。また、アニメ声は、VOICEVOX ずんだもんを使用した。

表2 検証2におけるペルソナの変化

パターン	声質	口調
パターン1	機械音声	敬語
パターン2		口語
パターン3	アニメ声	敬語
パターン4		口語

### 2-6 検証プログラム

今回の検証を行うにあたりプログラムを作成した。プログラミング言語はPythonを使用した。プログラムに使用したライブラリ一覧を表3に示した。また、本研究で使用したプログラムは次のURLで公開している<<https://github.com/kx47g/2022hightechVUI>>

表3 プログラムに使用したライブラリ

ライブラリ名	活用した機能
speech_recognition	音声データから文章を認識し，文字起こし
win32com	Windows 機械音声による発話
tkinter	ボディの画像を表示
PIL	画像の読み込み・リサイズ
beautifulsoup	検索した Web ページから HTML 情報を取得
selenium	ブラウザの自動操作
pyaudio	音声の録音・保存
simpleaudio	音声ファイルの再生
re	反応するキーワードを設定

## 2-7 検証方法

検証では，人工知能学科の教室に二人ずつ順番に被験者を呼び，音声操作によるチョコレートの購入プログラムを体験してもらった。

検証は私たちが用意した PC を用いて行った。

ボディ・ペルソナの変化による影響を調べるため，機能的に同じ動作をするが，それぞれボディとペルソナが異なるプログラムを体験してもらった。

以下に記すのが検証の流れである。

1. 被験者の誘導
2. 検証の背景を口頭で説明
3. 検証開始
4. チョコレート購入の体験
5. ボディの変化による影響の調査
6. 声質の変化による影響の調査
7. 口調の変化による影響の調査
8. アンケート調査
9. 検証終了

## 3. 結果

### 3-1 検証1 ボディの変化 結果

図1はアプリの途中でボディの見た目は変化したほうが良いか否かについての調査結果である。調査の結果では，ボディの変化に対し肯定的な意見と否定的な意見に分かれた。

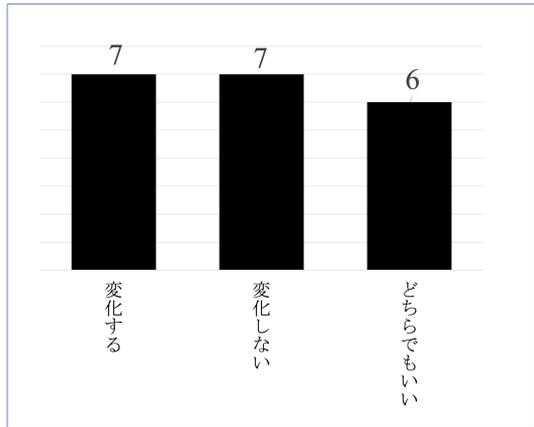


図1 ボディの変化についての調査

肯定的な意見

- 見ていて楽しい
- バリエーションがあった方が面白い
- サイトによって切り替わると雰囲気が変わる

否定的な意見

- いきなり切り替わられると、どこか変なボタンを押してしまったかなと心配になる
- 少し戸惑った

図2 ボディの変化についての意見

3-2 検証2 声質と口調の組み合わせ 結果

2つ目の検証の結果は以下のとおりになった(図3,図4).

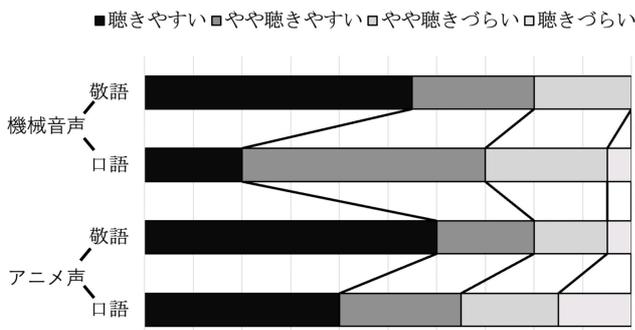


図3 声質と口調の聴きやすさについての調査

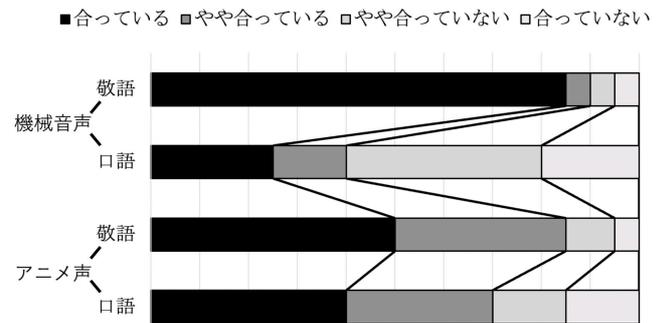


図4 声質と口調のシーンとの整合性についての調査

声質と口調の聴きやすさについての調査結果では、口調が敬語のパターンで少し評価が高い傾向が見られた(図3).

声質と口調のシーンとの整合性についての調査結果では、機械音声と敬語のパターンと、アニメ声と敬語のパターンが比較的高い評価となった(図4). また、声質に関しては敬語と口語による差は見られなかった.

4. 考察

ボディの変化に関して、検証前はパターン1のボディを変化させた場合の方が評価が高くなると予想した。商品に関係のあるボディが出てくることで、専門の店員に対応されているといった体験ができると考えたからである。しかし検証を行ってみると、変化の有無による評価の差はなかった(図1).

一方で具体的な評価理由を見ていくとボディが変化することに対して、肯定的な意見では体験としての面白さが評価され、否定的な意見では、急に切り替わることによって驚いたといった意見が見られた(図2). そのため、切り替わること自体に大きなマイナスは無く、事前の予告なく急に切り替わることで、違和感や不安に繋がったと推察した.

ペルソナの組み合わせに関して、検証前は、パターン4のアニメ声と口語のパターンの評価が高くなる

と予想した。抑揚のあるアニメ声で喋り機械音声よりも人間に近い喋り方ができることから、聴く人に親しみやすさを感じさせられると考えたからである。

しかし、検証の結果では、機械音声と敬語のパターンと、アニメ声と敬語のパターンが比較的高い評価となった。つまり今回のシーンでは、声質による影響は見られず、口調が敬語であることが有効であったと考えられる(図3)(図4)。

これは、買い物と言う利用シーンを選択したことが影響していると考えた。買い物をする時に、店員と客としての立場に対し節度をもった対応を無意識に望んでいることが影響していると推察した。

## 5. 結語

本研究で扱ったシーンにおいて、結論を3つ得た。

1. ボディを対話の状況に合わせて変化させることで使用者に体験としての面白さを与えられるということである。但し、この時に変化することを伝える演出がなければ、使用者に違和感や不安が生まれ低評価に繋がる可能性が示唆された。
2. ペルソナの声質による影響は見られなかった。
3. ペルソナの口調は敬語にした方が良い評価を得られた。

この研究を統括して、今回行ったネットショッピングのシーンにおいては、ボディを対話の状況に合わせて変化させ、ペルソナの口調を敬語にすることで心地よさを与えられることが明らかとなった。

今後の課題として他のシーンでボディとペルソナの変化の影響を調べる。心地よさを与えることで VUI の利用率の向上に繋がるのか調べていく。

## 謝辞

本研究を進めるに当たり製品開発の見地からご意見をいただきました、Fairy Devices 株式会社藤野真人様に厚く御礼申し上げます。

## 参考文献

- [1] "Fairy Devices 音声システムの開発プロセス", <<https://fairydevices.jp/consulting/sds>>.
- [2] "音声アシスタントとスマートスピーカーに関するアンケート調査", <<https://u-site.jp/survey/voice-assistant-2022-4>>.