

# 安心感・信頼感を与える合成音声の特徴抽出と表現

## Feature Extraction and Representation of Synthetic Speech for Providing Comfort and Trust

加藤 愛優美<sup>1)</sup>

指導教員 岩下 志乃<sup>1)</sup>, 櫻 リベカ<sup>1)</sup>, 研究協力者 大竹 正彦<sup>1)</sup>

1) 東京工科大学 コンピュータサイエンス学部 コンピュータサイエンス学科  
岩下・櫻研究室

近年、合成音声は身近なものとなりつつあるが、苦手意識を持つ人も少なくない。その理由として、人声のような安心感や信頼感を得られないことが考えられる。そこで本研究では、安心感・信頼感を得られる合成音声の特徴を分析し、それを表現することを目的とする。

キーワード：合成音声, 安心感, 信頼感

### 1. はじめに

近年、音声合成技術の進歩により、合成音声が使われる場は日々の生活においても増えてきている。一方で合成音声に対して苦手を感じる人も少なくはない。例えば人声と合成音声が脳波に与える影響差に関する研究[1]では、合成音声聞き手に何らかのストレスを生じさせたと推測された。このような状況を鑑みると、より不快感の少ない合成音声について分析する必要があると考えられる。

声の印象に関する研究としては、浅田ら[2]が発話印象モデルにおける声量や抑揚などの要素感覚の条件を変えた音声についてその印象差を調査しているが、用いた要素感覚が声量、抑揚、話速の3要素のみであったこと、条件変更が大小の2段階のみであったことなどが課題として挙げられていた。一方で、声の性格印象に関する研究も行われてきており、これらを活用することで、特定の印象を与える合成音声表現できる可能性がある。

本研究では不快感の少ない声について、聞き手が安心感・信頼感を得る声と定義し、人が安心感・信頼感を得る合成音声の特徴を分析することと、実際にその特徴を表現することを目的とする。

### 2. 研究概要

声の要素の条件を変更した複数の合成音声を用意し、それらを聴いた印象についてアンケートを実施することで、声の各要素と安心感・信頼感の関連性を明らかにする。

本研究で着目する声の要素は、『声量』『速度』

『高さ』『間』『抑揚』『声質』の6項目とする。

各要素の条件変更について、『声量』『速度』『高さ』『間』については5段階に変更する。『抑揚』については合成音声ソフト使用時に自動で調整されたもの(以降、自動調整音声と表記)、抑揚を一切なくしたもの、自動調整音声の語尾の音を上げたものの3種を用意する。『声質』については個性のあるものを10種程度用意することを検討中である。

被験者の負担を減らすために実験は複数回に分けて行う。初めに予備実験として『声量』『速度』『高さ』『間』『抑揚』の5項目について安心感・信頼感との関連性を確認する。次に、予備実験で評価の高かった音声を用いて『声質』を変更する声質比較実験を実施する。予備実験、声質比較実験を踏まえて、聞き手に安心感・信頼感を与える合成音声表現し、その評価をアンケートに回答してもらう。

### 3. 実験詳細

#### 3.1. 実験音声について

予備実験では印象が声質の個性に左右されることを懸念し、一般的なアナウンス等に使用される音声に近いものを使用することにした。男女の声の差については既存研究から大きな影響はないと考えられたため、VOICEPEAKの商用可能6ナレーターセットより女性1を利用した。

声質比較実験にはVOICEVOXを使用する。VOICEVOXで提供されている各音声に対して木戸ら[3]の実験にて示された8組の表現語対で主観評価し、特徴に違いのある10数種の音声を実験に使用

する。発話内容による印象差がないよう、いずれの実験も実験音声の発話内容は全て「こんにちは。よろしくお願ひします。」に統一する。

実施済みの予備実験では条件を変更した 20 種の音声を実験に使用した。近く実施予定の声質比較実験でも実験音声の数は同程度になると思われる。

### 3.2. 実験の流れとアンケート項目

予備実験及び声質比較実験では、各実験音声について、個々の音声の印象評価及び基準とする音声との比較評価を行うアンケートに回答してもらう。アンケートにはそれぞれの音声を聴き終わったタイミングで都度回答してもらうものとする。

個々の音声の性格印象評価には、Big-Five（外向性、協調性、勤勉性、神経症傾向、開放性）を測定する項目を 10 項目に短縮した TIPI-J[4] を使用する。各音声に対して、各項目の印象を受けた度合を 7 件法で評価してもらう。基準とする音声との比較評価では、以下に示す質問について各音声を基準音声と比較してどちらの方がどの程度当てはまるかを 5 段階で回答してもらう。

1. 安心すると感じる
2. 信頼できる・頼りたいと感じる
3. 任意のサービス利用時に使われていたら嫌だと感じる
4. より自然(人間的)だと感じる
5. 音声が好き

最終評価実験の実験形式及びアンケート内容については検討中である。

### 4. 予備実験の結果

実験では抑揚以外についてはパラメータの値を 5 段階（1～5）に変更しており、全てのパラメータを 3 にしたものを基準音声に設定した。

まず、TIPI-J を用いた性格印象評価の結果、発話速度が極端に遅い 1 音声を除く音声の勤勉性が高かった。これは、説明やアナウンスに用いられるものに近い合成音声を実験に用いたためであると推測される。

続いて基準音声との比較評価結果を示す。声量と間については、パラメータを 2,4 に 1 段階変化させた音声は全ての質問で「どちらでもない」への回答が多かった。このことから、声量と間の多少の変化は印象に大きな影響を及ぼさないと判明した。

速度、高さ、間のパラメータを 2 段階下げて 1 に設定した音声、すなわち速度が極端に遅い音声、高

さが極端に低い音声、間が極端に空いた音声は「音声が好き」の評価が低かった。特に速度、高さについては「使われていたら嫌だと感じる音声」の評価も高く、不快感を得たものと思われる。

一方で最も好印象だったのは、抑揚の変化において自動調整音声の語尾を上げただけのもの（他パラメータの値は基準音声と同一）であった。また、「より自然だと感じる」の評価は「音声が好き」との関連性が薄く、安心・信頼の少なくとも一方は得られる音声を好む傾向にあることが確認できた。このことから人に受け入れられやすい合成音声には、自然さの追求だけでなく、安心・信頼といった感情を聞き手に与えられることも関係していると考えられる。また、安心・信頼の少なくとも一方が得られる音声を好む傾向にあるという結果は、安心感・信頼感を得られる声は不快感を軽減させるといふ仮説を裏付けるものになったと考える。

最後に、性格印象と基準音声との比較評価の結果の相関を見ると、5 因子のうち神経症傾向の得点が高い音声は不快感を与えやすいと分かった。

### 5. おわりに

本研究では、安心感・信頼感を与える合成音声の条件について分析することで、聞き手にとって不快感の少ない合成音声を表現することを目的としている。今後は、実施済みの予備実験、近日実施の声質比較実験の結果をもとに安心感・信頼感を与える合成音声を表現し、最終評価実験を実施する。また、その結果も踏まえて人に安心感・信頼感を与える合成音声の特徴についてまとめる予定である。

### 参考文献

- [1] 大塚敦夫, 富山健, “肉声と合成音声の脳波に与える影響の比較検討”, 情報処理学会研究報告知能システム(ICS), vol.2013-ICS-171, no.10, 2013
- [2] 浅田龍星, 西田悠, 中平勝子, 北島宗雄, “合成音声を利用した好印象発話モデルの構築”, FIT2017, J-012, 2017
- [3] 木戸博, 粕谷英樹, “通常発話の声質に関連した日常表現語の抽出”, 日本音響学会誌, vol.55, no.6, 405-411, 1999
- [4] 小塩真司, 阿部普吾, カトローニピノ, “日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み”, 日本パーソナリティ心理学会 2012, vol.21, no.1, 40-52, 2012