

# 音声AIを用いた発話者の音声変換が聴覚障がい者の音声聞き取りに与える効果

## Effects of Speech Transformation of Speakers Using Speech AI on Speech Hearing of the Deaf and Hard of Hearing

鈴木祥太<sup>†</sup> 宮川裕之<sup>‡</sup>

Shota Suzuki<sup>†</sup> Hiroyuki miyagawa<sup>‡</sup>

<sup>†</sup> 青山学院大学 社会情報学部

<sup>†</sup> School of Social Informatics, Aoyama Gakuin University.

### 要旨

本研究の目的は、聴覚障がい者が感じる音声の聞き取りづらさを、音声 AI を用いて改善し、音声の聞き取りやすさを向上させる事である。最初に聴覚障がい者が聞き取りづらいつと感じる周波数帯域を調査し、その後、変換前の元音声と AI でアナウンス部所属の男性の声に変換した AI 音声を用いて、語音聴取検査を行った。正答文字率から、周波数帯域によっては、AI での音声変換が音声の聞き取りやすさ向上に効果がある可能性が示唆された。

## 1. はじめに

聴覚障がい者は音声の聞き取りが難しく、日常生活で多くの困難に直面している。これまで、人工内耳装用者の異聴傾向を調査する研究[1]が行われ、補聴器や聴覚トレーニングの精度向上に寄与してきた。しかし、AI を活用した補聴システムに関する研究はまだ進んでいない。そこで、本研究では音声 AI を使って発話者の音声を変換し、聴覚障がい者の音声聞き取りの向上ができるか検証する。

## 2. 聴覚障がい者の課題

聴覚障がい者が感じる課題や音声聞き取りの特徴を把握するため、3つの事を行った。

- ① 2023年6月～現在まで、本大学の障がい学生支援センターで募集をしているUDトーク支援アルバイトを通じて、聴覚障がい者との継続的な繋がりを持った。
- ② 2023年9月から本大学に在籍する聴覚障がい学生1名に継続的にヒアリングを行った。
- ③ 耳鼻咽喉科の専門医へのヒアリング

以上の事から今回把握できた、聴覚障がい者が抱える音声聞き取りに関する課題や特徴を以下に示す。

- ・聞き取りづらい音・音域がある（異聴）
- ・滑舌の悪さ・話すスピード、声質などで聞き取りやすさが大きく違う
- ・急な話題転換についていく事が難しい
- ・異聴傾向や補聴システムの効果は個人差が大きい

## 3. 仮説

音声 AI の音声変換技術を用いて、聴覚障がい者の音声の聞き取りづらさ解消を目指した。音声 AI にアナウンサーのような聞き取りやすい音声を学習させ、発話者の音声を変換する事で、音声の聞き取りがしやすくなるという仮説を立て、検証する。

## 4. 予備実験

### 4.1. 実験目的

実験対象者が聞き取りづらいつと感じる音域を調べる。

## 4.2. 実験対象と期間

本学に在籍する聴覚障がいを持つ 2 名を対象に、2023 年 10 月に実験を行った。

## 4.3. 実験方法

異なる強調周波数を持つ 4 つの音声を使用して語音聴取検査を行った。通常、言葉を理解するために必要な音域は 500Hz から 2000Hz に含まれている[2]。したがって、実験ではこの音域に焦点を当て、4 つの異なる音声を 20 文ずつ作成した。これらはそれぞれ 400Hz、1000Hz、1500Hz、2500Hz の周波数では +6 dB に強調され、その他の周波数では -12 dB に減衰させるためにイコライザーを使用して編集した。

検査文は ChatGPT3.5 を使用して生成し、以下の条件で作成した。

- ・ 2-6 音節の口語的な短文と複文である
- ・ 母音と子音のパターンを網羅し、文章内容と長さがランダムである

実験は防音室内で行われ、ELECOM の MS-76MA スピーカーが対象者の正面約 1 メートルの位置に設置した。スピーカーの音量は、対象者の位置で最大 50 dB に調整した[3]。

被験者には提示された音声を聴いて反芻させ、正答率を計算した。なお、すべての部分が聞こえなかった場合でも、聞こえた部分について反芻させた。



図 1 実験の様子

## 4.4. 実験結果

正答率と完全正答文の数を以下の表に示す。

表 1 被験者 1 の予備実験結果

周波数	正答文字数	正答率	完全正答文 /20
400Hz	130/250	52%	8
1000Hz	138/233	59.20%	7
1500Hz	131/216	60.60%	0
2500Hz	72/193	37.30%	4

表 2 被験者 2 の予備実験結果

周波数	正答文字数	正答率	完全正答文 /20
400Hz	166/250	66.40%	8
1000Hz	205/233	87.90%	13
1500Hz	205/216	94.90%	16
2500Hz	165/193	85.40%	14

両者とも 1000Hz と 1500Hz の正答率が高く、400Hz と 2500Hz の正答率が低かった。すなわち、両者とも 400Hz と 2500Hz の周波数を強調した音声聞き取りづらいという事が分かった。

## 5. 本実験

### 5.1. 目的

音声 AI で音声変換する事で、音声の聞き取りやすさが向上するか調べる。

### 5.2. 実験対象と期間

実験対象者は実験 1 と同様である。2023 年 11 月に実験を行った。

### 5.3. 実験方法

以下の音声を使用して語音聴取検査を行った。

- ① 筆者の録音音声と、その音声を AI で変換した AI 音声
  - ② 予備実験で両者正答率が低かった強調周波数 400Hz の音声と、その音声を AI で変換した AI 音声
- 実験環境・被験者の回答の仕方はいずれも実験 1 と同様である。検査文は、CI-2004(試案)検査のうち、成人用単語リストを使用した[4]。変換前と変換後の正答率を比較し、正答率が向上するか調べた。

### 5.4. 学習音源と AI 設定

今回は、音声合成などの研究目的に制作された ITA コーパスの emotion 朗読者用 100 文[5]を、アナウンス部所属の高校生（男性）に録音してもらい、学習音源とした。

本実験では音声変換の際に、RVC WebUI (Retrieval-based Voice Changer) という音声 AI ソフトウェアを Google Colaboratory で起動して使用した。RVCWebUI の設定は以下の通り行った[6]。下記以外の項目は全てデフォルトのままである。

<学習設定>

- ・ピッチ抽出アルゴリズムの設定 : harvest を選択 (品質高く処理が出来る)
- ・学習回数 : 200 回 (学習回数の許容範囲は 200 回)

<推論設定>

- ・変調 : +2
- ・検索特徴率 : 1 (1 に近づけるほど学習音源の特徴に近づく)

### 5.5. 実験結果

正答率と完全正答文の数を以下の表に示す。

表 3 被験者 1 の実験結果

試験条件	音声	正答文字数 /78	正答率	完全正答文 /25
①	筆者	57	73.00%	15
	AI 音声	39	50%	9
②	400Hz 強調	51	65.30%	9
	AI 音声	55	66.60%	15

表 4 被験者 2 の実験結果

試験条件	音声	正答文字数 /78	正答率	完全正答文 /25
①	筆者	70	89.70%	20
	AI 音声	54	69%	14
②	400Hz 強調	33	42.00%	6
	AI 音声	55	70.50%	12

①では、筆者の音声と AI 音声の正答率を比較すると、AI 音声の正答率は低くなった。対照的に、②では 400Hz の音声を変換した AI 音声の方が、正答率が高くなった。したがって、周波数帯域によっては AI 音声変換が聞き取りやすさ向上に効果がある可能性が示唆されるが、全ての音声で効果があるわけではない事が分かる。

## 6. 考察

耳鼻咽喉科の専門医のヒアリングによると、聴覚障がい者の聞こえ方には大きな個人差があるとされている。今回の実験では、AI による音声変換が特定の周波数帯域において音声の聞き取りやすさを向上させる可能性が示唆されたが、その効果には個人差があると考えられる。また、試験条件①と②の AI 音声の正答率を比較すると、②の 400Hz 強調音声を変換した AI 音声は、両者とも正答率が高かった結果が得られた。変換前の元音声を変換後の聞き取りやすさに影響を与えると考えられるため、どのような元音声に適しているか、またどの学習音源が個々の人に合ったものかを検討する必要がある。

## 7. 今後の予定

今回の実験では、アナウンス部所属の男子高校生の音声を学習音源として使用した。今後は、学習音源の種類を検討し、より聞き取りやすさ向上に効果のある音声を目指す。

## 8. 謝辞

本研究を行うにあたり、多大なご指導・ご助言を頂いた、国際医療福祉大学 三田病院 耳鼻咽喉科 高橋優宏先生、青山学院大学 社会情報学部 居駒幹夫教授に深謝いたします。

※論文執筆に当たり、文章の校正を目的として ChatGPT を使用しました。

### 参考文献

- [1] 射場 恵, 熊谷文愛, 熊川孝三, 鈴木久美子, 武田英彦, 語音聴取評価検査「CI—2004 (試案)」を用いた人工内耳装用者の聴取能, 虎の門病院 耳鼻咽喉科・聴覚センター, 2011
- [2] 中山 禎子, 純音聴力型と難聴度評価の関連について, 東邦大学耳鼻咽喉科学教室, 1977
- [3] Nishio SY, Tono T, Iwaki T, Moteki H, Suzuki K, Tsushima Y, Kashio A, Akamatsu Y, Sato H, Yaegashi K, Takeda H, Kumagai F, Nakashima T, Matsuda Y, Hato N, Dairoku T, Shiroma M, Kawai R, Usami SI. Development and validation of an iPad-based Japanese language monosyllable speech perception test (iCI2004 monosyllable). *Acta Otolaryngol.* 2021;141(3):267-272.
- [4] iCI-2004 導入・検査マニュアル iPad 版 日本語語音弁別検査 (単音節・単語検査), 一般社団法人日本耳科学会, 2023
- [5] 小田恭央, 金井郁也, 小口純矢, 細田計, 齊藤剛史, 森勢将雅, ITA コーパス emotion 朗読者用, [https://github.com/mmorise/itacorporus/blob/main/emotion\\_%E6%9C%97%E8%AA%AD%E8%80%85%E7%94%A8.pdf](https://github.com/mmorise/itacorporus/blob/main/emotion_%E6%9C%97%E8%AA%AD%E8%80%85%E7%94%A8.pdf), 2022 (参照日 2023 年 11 月 13 日)
- [6] 子供プログラマー 大人のためのプログラミング・コンピュータ学習支援サイト「【RVC v2 モデル対応版】日本人のための RVC WebUI の使い方: AI ボイスチェンジャー」, [https://child-programmer.com/ai-voice-change-tutorial-v2/#RVC\\_v2RVC\\_WebUI\\_-\\_Train](https://child-programmer.com/ai-voice-change-tutorial-v2/#RVC_v2RVC_WebUI_-_Train), 2023 (参照日 2023 年 11 月 11 日)