

# 五十音の音読に基づく合成音声の印象評価 —授業動画でのアバター音声の効果的な利活用に向けて— Impression Evaluation of Japanese Kana-Based Synthesized Speech: Toward the Effective Use of Avatar Voices in Educational Videos

小島 隆次<sup>†</sup>

Takatsugu Kojima

<sup>†</sup>滋賀医科大学

Shiga University of Medical Science

kojima@kojima-lab.net

## 概要

本研究では、合成音声による五十音の音読と説明文の音読に対する印象評価実験の結果の比較を通じて、授業動画や動画教材における教師役アバターの音声に対する印象評価を五十音の読み上げのみで簡易に実施する方法の有効性を検討した。実験結果から、通常の授業動画が扱う類の内容であれば、音読内容によって合成音声の印象が大きく変化することはなく、五十音の読み上げ音声を利用してアバター音声の印象を簡易的に測定することは有効であると考えられる。

キーワード：合成音声, 印象評価, 五十音

## 1. はじめに

授業動画や動画教材において、教師役やプレゼンター役のアバターのデザインや印象が視聴者に対して重要な役割を担うことは想像に難くない。この点、アバターのデザインや印象が、視聴者の視聴動機づけ、内容理解、アバターとのコミュニケーション頻度やコミュニケーション時の態度などに影響することが示されているが(小林, 2020; 宮地, 2022; 清水・渡邊, 2021)、アバターの外見(視覚情報)のデザインや印象を中心に検討されることが多い。しかし、アバターの音声(聴覚情報)のデザインや印象に対する重要性も忘れてはならない(安久・原田・鷹阪・赤津, 2024)。また、授業動画に限らず、アバターを用いた様々な動画の特徴の一つとして、アバターの外見(視覚情報)は、実在の人物が登場する動画と比較して、相対的に重要性が低くなることが指摘されており(越後・呉・新井・富永・小林, 2023)、動画に登場するアバターの印象形成において、音声をもたらす影響は、実在の人物のケースと比して相対的に高くなることも想定される。

以上のような背景を踏まえて、本研究では、合成音による五十音の音読に対する印象評価実験と説明文の音読に対する印象評価実験との結果の比較を通じて、授業動画や動画教材などの主に何かの説明や解説を行う動画における教師役や解説者役などのアバターの音声

に対する印象評価を五十音の読み上げによって簡易的に実施する方法の有効性や可能性について考える。

## 2. 実験

### 2.1. 実験参加者

実験参加者は大学生・大学院生 100 名であった。参加者間要因である「音読の内容」の 2 条件(「五十音」と「ピワイチ(琵琶湖一周)についての説明)」に対して、各条件に 50 名ずつが参加した。

### 2.2. 実験デザイン

本研究で行った印象評価実験の実験要因は「音読の内容(音読するテキストの種類)」と「合成音声の種類」の二要因であった。音読の内容が参加者間要因であり、合成音声の種類が参加者内要因であった。「音読の内容」は 2 水準で「合成音の種類」は 6 水準であった。従属変数は SD 法による合成音声の印象評価値(7 件法)であった。

### 2.3. 実験刺激・実験課題

印象評価の対象であった音読の内容については、「五十音」と「琵琶湖一周についての解説」の 2 種類を用意した。「五十音」を音読する音声刺激については、「ん」を含めた 47 文字の読み上げとし、読み上げ音声の長さは全体で約 10 秒であった。「ピワイチ(琵琶湖一周)についての説明」については、平仮名で全 394 文字相当の文章で、読み上げ音声の長さは全体で約 50 秒であった。

音読のための合成音声については、フリーの合成音声ソフト VOICEVOX で使用可能な合成音声の中から 6 種類(男性声 3 種類; 女性声 3 種類)を選択して使用した。

実験参加者が取り組む課題は、提示された音声を聞き、その音声の印象を SD 法によって評価するというものであった。音声の印象評価の際に SD 法で用いた形容詞対は全部で 22 対であった (表 1 参照)。これら 22 対の内 20 対については、小島 (2023) が教師アバターの印象評価実験で用いた形容詞対と同じもので、8 対が対人的魅力、4 対が社会的魅力、8 対が教師イメージに関するものであった。対人的魅力と社会的魅力に関する形容詞対の選定では川名 (2012) を参考にし、教師イメージに関する形容詞対の選定では木谷・野仲・太田 (2016) を参考にした。本研究では、以上の 20 対に「聞きやすい - 聞きにくい」と「好き - 嫌い」の 2 対を加えた。

2.4. 手続き

実験は LimeSurvey を用いた web ブラウザによるオンライン実験であった。参加者に対して今回の実験に関する諸注意の提示と参加の同意確認等が行われ、実験時の音声聞き取りのためにブラウザ上での音声再生可否の確認と音量調整を実施した。これら確認作業の後、全部で 6 試行の本試行を行った。本試行では、異なる 6 種類の合成音による音読をそれぞれ聞き、聴取後に各合成音の印象を 22 種類の形容詞対で評価するというものであった。音読内容は「五十音」か「ピロイチ (琵琶湖一周) についての説明」のいずれか一方であった。本試行での 6 種類の合成音声の提示順は、参加者ごとにランダムであった。各音声は、必ず最初から最後まで一通り聞くように教示した。聴取回数に制限はなかった。

2.5. 結果

SD 法による印象評価で用いた 22 項目の各形容詞対尺度の平均評定値について、有意水準を 5% として二要因 (「音読の内容」と「合成音声の種類」) 混合計画の分散分析を行った。その分析結果をまとめたものが表 1 である。

「音読の内容」の主効果については 5 項目 (「責任感がある - いい加減な」「独創的な - ありきたりな」「明るい - 暗い」「つまらない - おもしろそう」「聞きやすい - 聞きにくい」) で統計的に有意であり、17 項目で有意ではなかった。

「合成音声の種類」の主効果については 18 項目で統計的に有意であり、有意ではなかったのは 4 項目 (「神

経質な - のんきな」「凶暴な - 温厚な」「弱気な - 強気な」「厳しい - 緩い」) であった。

交互作用については 18 項目で統計的に有意であり、有意ではなかったのは 4 項目 (「無計画な - 計画的な」「神経質な - のんきな」「つまらなそう - おもしろそう」「聞きやすい - 聞きにくい」) であった。また、交互作用が有意になった項目について単純主効果の分析を行い、「音読の内容」が「合成音声の種類」の各水準に及ぼす効果について有意水準 5% で有意な効果が見られた結果を表 1 の単純主効果の列にまとめた。単純主効果の列内の数字は「合成音声の種類」要因内の 6 種類の各合成音声に付与した識別番号である。

表 1 印象評価に関する統計的検定の結果

印象評定項目	音読の内容	合成音声の種類	交互作用	単純主効果
1. 親切な - 7. 冷淡な		**	**	1
1. 平凡な - 7. 奇抜な		**	*	6
1. 素直な - 7. 強情な		**	**	1
1. 無計画な - 7. 計画的な		**		
1. 話しかけやすい - 7. 話しかけにくい		**	*	1
1. 勤勉な - 7. 怠惰な		**	*	6
1. 神経質な - 7. のんきな				
1. 凶暴な - 7. 温厚な			**	1
1. 責任感がある - 7. いい加減な	*	**	**	3, 6
1. 協力的な - 7. 敵対的な		**	**	1
1. 弱気な - 7. 強気な			**	1, 4
1. 親しみにくい - 7. 親しみやすい		*	**	1, 4
1. 説得力がある - 7. 説得力がない		**	*	3
1. 独創的な - 7. ありきたりな	*	**	*	6
1. 愚かな - 7. 知的な		**	*	
1. 活動的な - 7. 無気力な		**	**	6
1. 厳しい - 7. 緩い			**	1, 4
1. 内向的な - 7. 社交的な		**	**	6
1. 明るい - 7. 暗い	*	**	*	1, 6
1. つまらなそう - 7. おもしろそう	*	**		
1. 好き - 7. 嫌い		**	*	1
1. 聞きやすい - 7. 聞きにくい	*	**		

\*  $p < .05$ , \*\*  $p < .01$ .

3. まとめ

実験結果から、音読内容が何かを説明したり解説したりするもの場合には、全体的な傾向として音読内容によって個々の合成音声の印象があまり大きく変化することはないだろうと予測できる。そのため、一般的な授業動画や動画教材に登場するアバターの音声もたらず印象については、五十音の読み上げ音声をを用いて簡易的に測定することは有効な方法であると考えられる。但し、表 1 の単純主効果の列のデータからわかるように、一部の音声 (特に 1 と 6 の 2 種類の合成音声) については、音読内容の違いで印象が比較的大きく変化する可能性があることには注意を要する。

## 謝辞

本研究は JSPS 科研費 23K02657 の助成を受けた。

## 文献

- 安久 絵里子・原田 悦子・鷹阪 龍太・赤津 裕子 (2024). 音声に対する“かわいい”感情がアバター評価に与える効果とその加齢変化—仮想的2者対話システムの実現に向けて— 日本感性工学会論文誌, 23, 233-248.  
<https://doi.org/10.5057/jjske.TJSKE-D-23-00068>
- 越後 宏紀・呉 健朗・新井 貴紘・富永 詩音・小林 稔 (2023). カテゴリ別におけるVTuberとYouTuberの配信スタイルによる印象評価 情報処理学会論文誌, 64, 86-95.
- 川名 好裕 (2012). 笑顔および外見による魅力変化 立正大学心理学研究紀要, 10, 31-47.
- 木谷 智子・野仲 真理子・太田 美里 (2016). 教師を目指す青年の教師イメージの質的検討 中国四国心理学会論文集, 49, 65.
- 小林 英恵 (2020). オンライン授業におけるアバターを活用した個別最適化 DHU JOURNAL, 7, 78-81.
- 小島 隆次 (2023). 教師であるという認識がアバターの外見の印象にもたらす効果 日本心理学会第 87 回大会発表論文集, 883.
- 宮地 英生 (2022). アバターの外見による講義での学生の活性化度の違い 日本バーチャルリアリティ学会誌, 27, 12-13.
- 清水 千夏・渡邊 慎二 (2021). アバターの外見と動きとその印象に関する研究 日本デザイン学会第 68 回春季研究発表大会概要集, 214-215.