

深層学習による英語発表の評価と練習支援

田上 雄悟¹ 山口 実靖¹ 小島 多香子²

¹工学院大学 情報学部 情報通信工学科

²東京医科大学 国際教育研究センター

➤ 研究背景

- ・英語でのプレゼンテーションを行う機会が増加
- ・時間的制約なく英語の発音練習の積極的な繰り返し練習が可能に
- ・AIの進歩により客観的かつ効率的に学習可能に

➤ 目的

深層学習による英語プレゼンテーション練習システムの構築

→本研究では単語の発音のアクセント(強弱)のつけ方に着目

➤ 関連研究

近年、コンピュータを利用した語学学習環境CALL(Computer-assisted language learning)には機械学習を利用するものが登場している。

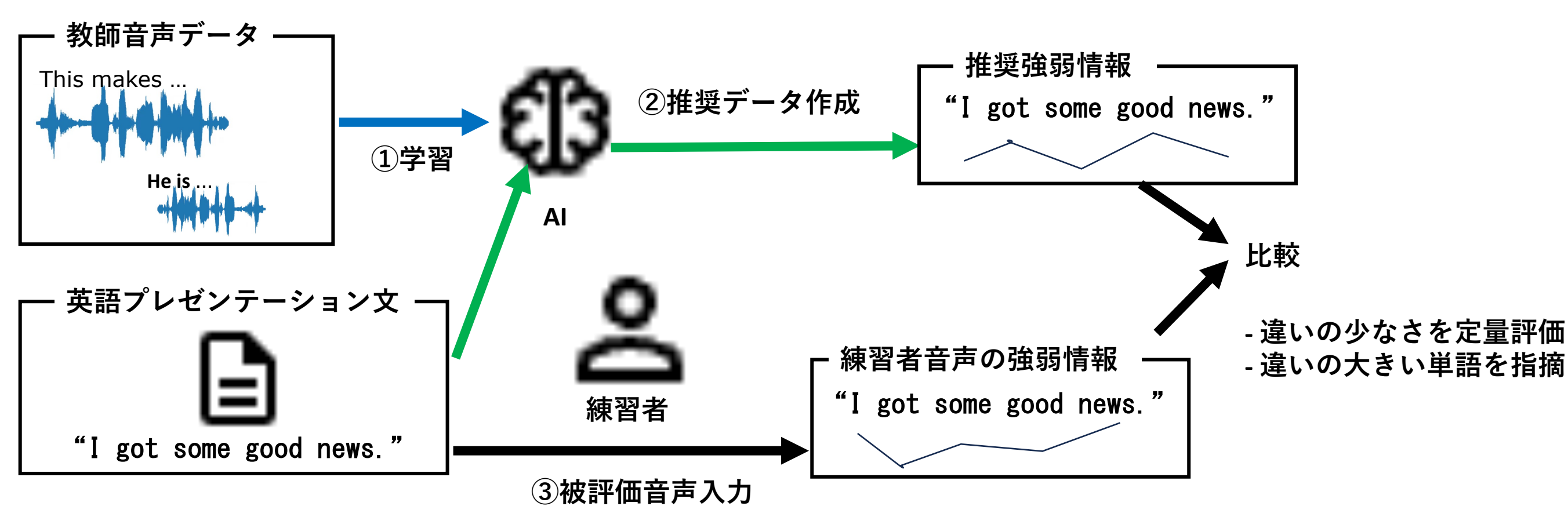
深層学習を用いた英語発音評価システムは従来の音声認識アルゴリズムよりも優れている[1]。

➤ 研究概要

1. 深層学習により英語プレゼンテーション音声データを学習し、任意の文における単語ごとの発音の強弱を推測できるモデルを作成。

2. ユーザが練習システムに対して音声入力を行う。

3. 練習システムが「推奨する単語ごとの強弱」と「練習者の発音の単語ごとの強弱」を比較し、練習者の発音の評価と改善に向けたアドバイスを行う。教師データとしてTED Talksの音声を、音声認識モデルとしてWhisperを使用。



➤ 試作システム

TED Talksの文を学習者が発音し、機械学習による音声認識モデルWhisperを用いた音声認識により単語ごとの音量を計測し、TED Talksの話者の単語ごとの音量と比較する。

比較結果をもとに学習者の音量がTED Talksの話者の音量よりも大きい場合に小さくするように、小さい場合に大きくするようにアドバイスを行う。

TED Talksの話者と学習者の単語ごとの音量の差を文全体で平均をとることで評価値とする。

➤ 性能評価

TED Talksのスピーチからランダムに選んだ(a)から(e)の5つの文をそれぞれ10回ずつ読み、提案手法により評価、発音アドバイスを生成する。アドバイス内容を意識して10回ずつ読み、提案手法により評価を行う。また、アドバイス前後の音声を聞いてよりスピーチとしてよく聞こえる音声を選択する主観評価を行う。

- (a)Because we develop the habits of doing the same things over and over.
(b)I got some good news.
(c)It used to be an easy question to answer.
(d)Now effort moralization makes sense at the individual level.
(e)That's roughly the same as driving 250 miles.

➤ 結果

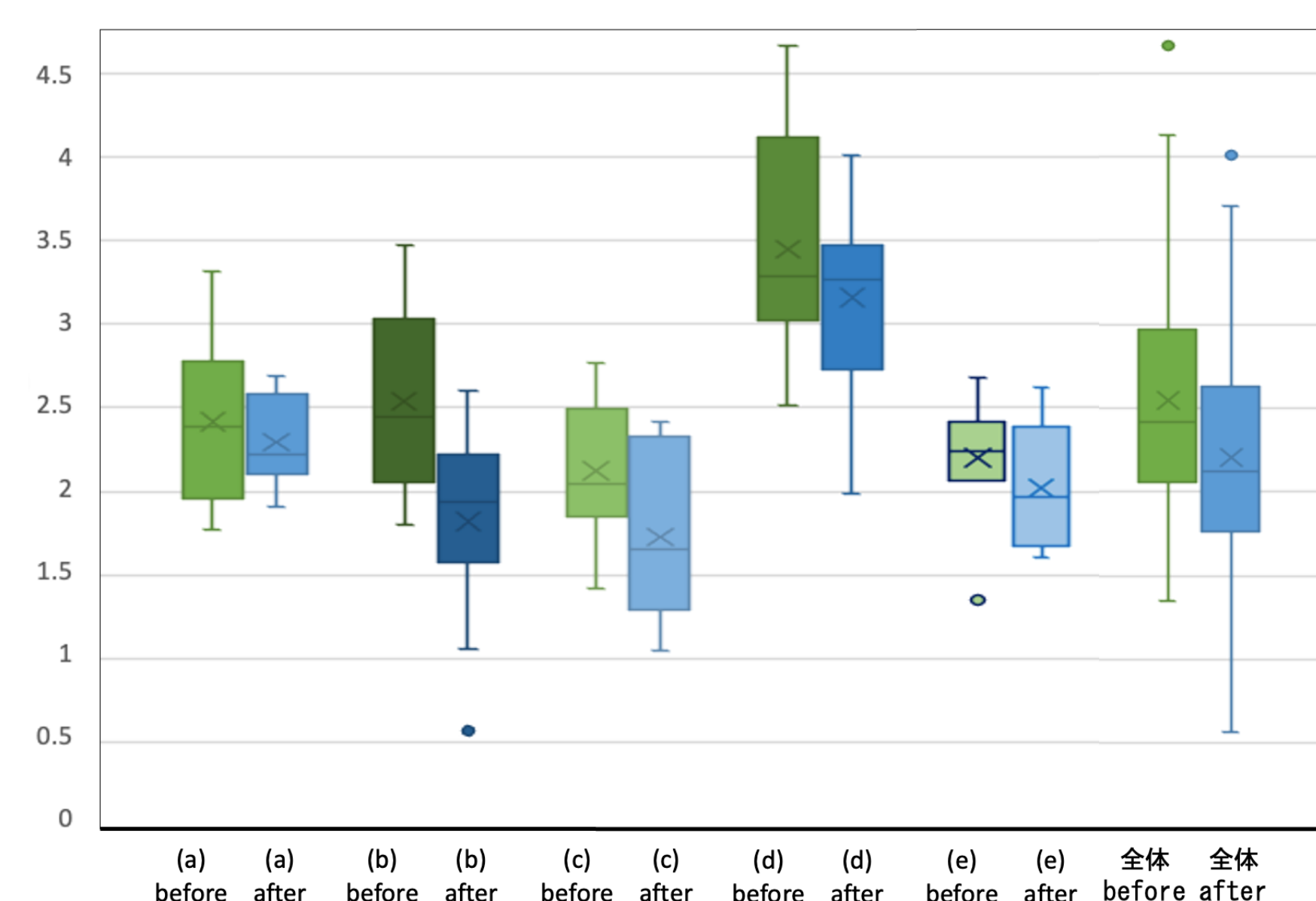


図1 アドバイス前後の評価結果
(縦軸:評価値、小さいほど良い結果を示す)

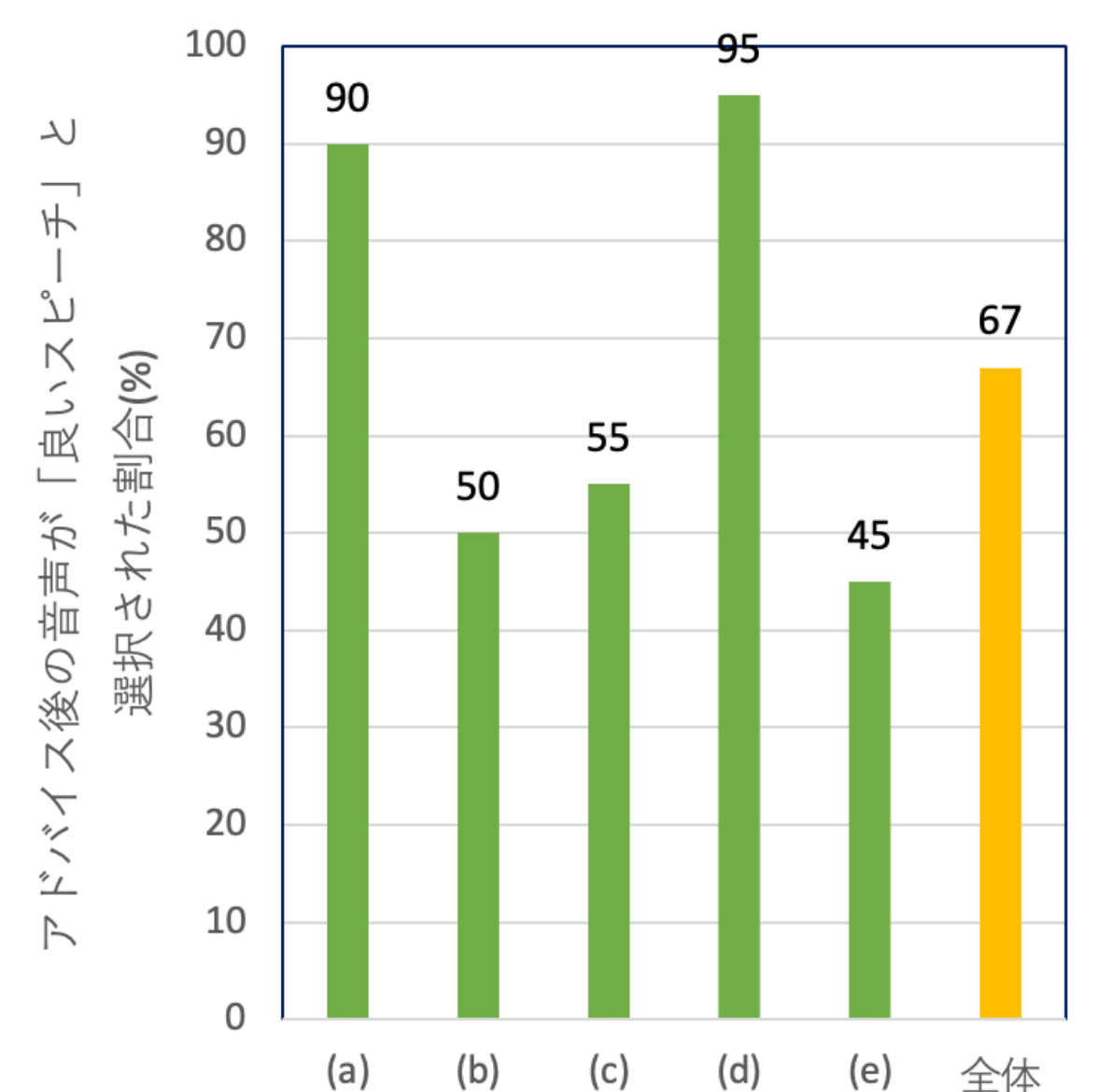


図2 主観評価結果

提案手法によるアドバイス前後の評価値を図1に示す。評価値の平均値は全体で13.7%、中央値は11.9%改善された。

また、2人の主観評価でアドバイス後の音声により良いスピーチに聞こえた割合を図2に示す。(a) (d)においてはアドバイス後の音声により良いと判断されたが、(b) (c) (e)では意見が分かれた。

➤ まとめ

提案手法により英語スピーチの発音は改善された。今後、深層学習による予測モデルの実装を進める。

➤ 参考文献

[1] Teng Haikun, Wang Shiyang, Liu Xinsheng, and Xiao-Guang Yue. 2019. Speech Recognition Model Based on Deep Learning And Application in Pronunciation Quality Evaluation System (ICDMML 2019). <https://doi.org/10.1145/3335656.3335657>