



令和2年1月20日

報道機関 各位

東北大学大学院工学研究科

AIにより日本人英語音声の高精度な自動発音評価が可能に ～スマホ等を活用した低コストな英会話練習が可能に～

【発表のポイント】

- 日本語母語話者が発声した英語音声の「発音の上手さ」を高い精度で自動推定することが可能になった。
- 深層学習を用いた音声認識技術により、複数の観点で認識をした結果を比較することで発音の上手さを推定する。
- コンピュータによる語学学習支援システム(CALL)への応用が期待される。

【発表概要】

東北大学大学院工学研究科の伊藤彰則教授、能勢隆准教授、千葉祐弥助教および付江大学院生は、日本人が発声した英語文の発音の上手さ(英語母語話者らしさ)を高い精度で自動的に推定する技術を開発しました。深層学習^{*1}を用いた音声認識システムを使い、英語母語話者としての英語音声認識の結果と、日本語母語話者としての英語音声認識の両方の認識結果を比較することで、実際の英語ネイティブ教師による発音の評価と非常に近い発音評価が可能になりました。

本研究成果は、2019年12月23日にSpeech Communication誌に掲載されました。

【問い合わせ先】

東北大学大学院工学研究科 伊藤 彰則

電話 022-795-7084 E-mail: akinori.ito.a2@tohoku.ac.jp

< 報道に関して >

東北大学工学研究科情報広報室 担当 沼澤 みどり

TEL: 022-795-5898 E-mail: eng-pr@grp.tohoku.ac.jp

【発表内容】

(1) 研究の背景

英語能力は国際的に活動する際に必須の能力です。特に近年は、英語の文章を読んだり書いたりするだけでなく、声で会話をするための「聴く・話す」能力を加えた、いわゆる四技能が重視されています。この中でも、「話す」能力を養成するためには実際に英語での会話練習をする必要があります。コンピュータを用いた語学学習支援(Computer-Assisted Language Learning, CALL)システム^{*2}は、このような練習を手軽に行うための有効な手段として注目されています。

最近の CALL システムでは、英語の発音練習もできるようになってきています。ただし、これまでの技術では、発音練習はシステムがあらかじめ用意した文章の発音に限られており、自由な会話をする中で発音の上手さを自動的に評価することは困難でした。

(2) 研究内容

今回、本研究グループは、複数の音声認識システムによる認識結果を比較することにより、英語ネイティブの教師が採点した「発音の上手さ」を高精度に再現できる自動発音評価技術を開発しました。

この方法では、英語母語話者の英語音声と日本語母語話者の英語音声を認識するための音声認識システムと、それぞれ用意し、その認識結果を比較します。システムに入力した音声ネイティブに近い英語発音であった場合には、両方の音声認識システムで高い認識性能が得られるのに対し、ネイティブとは違う発音であった場合には日本語英語用の音声認識システムでのみ高い認識性能になります。そのため、両方の音声認識システムによる認識結果を比較することにより、発音の良さを推定することができます。音声認識システムの認識結果だけを使って評価をするため、元の文の内容を事前に用意する必要がなく、自由な会話音声に対しても発音の良さが評価できるようになりました。

(3) 社会的意義

初等教育から高等教育まで、英語教育の重要性はますます増していますが、英語四技能、特に話す能力の練習は従来の学校教育では難しく、また英会話学校に通うためには経済的なコストがかかります。本研究を応用した英会話用 CALL システムが実現できれば、スマートフォンなどを使って、いつでも誰でも低コストで英会話練習をすることが可能になります。

本研究は JSPS 科学研究費補助金 (17H00823) の助成により実施しました。

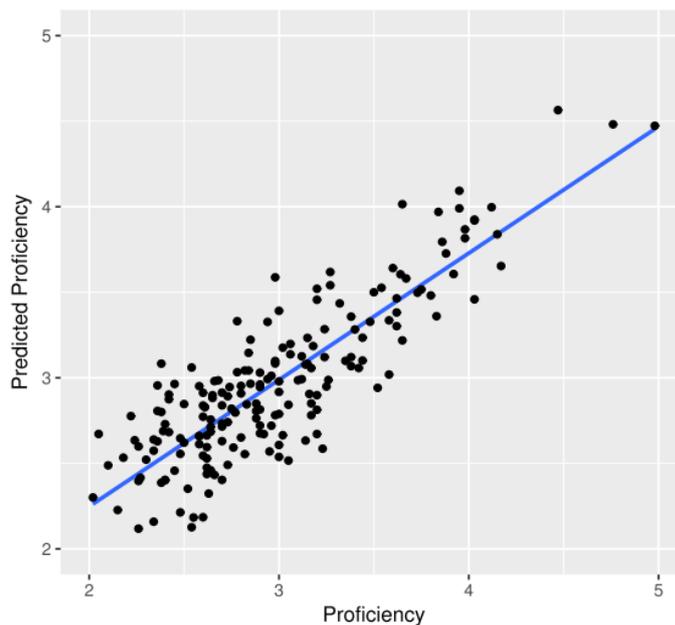


図1 英語ネイティブ教師による評価値と、開発したシステムによる自動評価結果の相関。横軸は人間による評価、縦軸はシステムによる評価。相関係数 0.826 (N=190)。

【発表雑誌】

雑誌名： Speech Communication

論文タイトル： Automatic assessment of English proficiency for Japanese learners without reference sentences based on deep neural network acoustic models

著者： Jiang Fu, Yuya Chiba, Takashi Nose, Akinori Ito

DOI: 10.1016/j.specom.2019.12.002

【用語解説】

※1 深層学習

深層学習(ディープラーニング)とは、生物の神経回路を模擬することにより、さまざまなパターンの認識や生成を行うことができる手法の総称です。近年の人工知能(AI)システムにおける中心的な技術で、音声認識だけでなく、画像認識、自然言語処理など幅広いAIの分野で利用されています。

※2 CALL システム

コンピュータを使って語学の学習をするためのシステムです。古くはコンピュータが選択式の文章読み取り問題や和文英訳問題などを出して自動採点するシステムから始まりましたが、2000年代以降ではビデオなどを使ったリスニング教材による学習もできるものが一般的です。最近のシステムでは、システムが提示した文を学習者が読んで、発音の正しさをシステムが自動的に判定する機能も一般的になってきています。