

外国語 speaking 練習に音声認識を取り入れることに関する

予備調査

柏木 治美^{*1}, 康 敏^{*2}, 大月 一弘^{*2}

^{*1} 神戸大学大学教育推進機構 ^{*2} 神戸大学大学院国際文化学研究科

A Preliminary Study on Incorporating Speech Recognition into Foreign Language Speaking Practice

Harumi KASHIWAGI^{*1}, Min KANG^{*2}, Kazuhiro OHTSUKI^{*2}

^{*1} Institute for Promotion of Higher Education, Kobe University

^{*2} Graduate School of Intercultural Studies, Kobe University

本稿では、音声認識を用いた口頭練習のために類似度を取り入れる可能性を検討した。学習者が意欲を失わずに練習できるよう、実際に認識された英文と対象英文を比較し、一定の類似度があれば、対象英文が発話されたと判断することにより、speaking 練習を継続できるか、類似度を取り入れた口頭練習の可能性を検討した。大学生 107 人を対象に 4 英文を上記方法で試した結果、90%以上の参加者の発話した英文が対象英文と判断された。また、音声認識を試した後、speaking 練習や音声認識に関して 5 段階で意識調査を行った結果、学生の 92%が発話練習は必要だと感じているが、78%は発話練習に時間を割いているとは言えない状況で、86%が都合のよい時に練習を行える環境があれば使いたいと考えていた。音声認識については、84%の学生が特に抵抗感はなかったと答えていた。これらの結果から、発話する機会を増やすために類似度を取り入れた方法について、その可能性が示唆された。

キーワード: 音声認識, 類似度, 発話の機会, スピーキング練習, 外国語学習

1. はじめに

外国語教育では、2020 年度から小学 3 年生と 4 年生で「外国語活動」が始まり、小学 5 年生と 6 年生で「外国語」として教科化される⁽¹⁾など、外国語（英語）力強化に向けて体制が整備されてきた。

一方、2023 年度全国学力・学習状況の調査によると、中学校の「英語」においては、特に「話す」ことに関する平均正答率が 12.4%にとどまり、英語で話す力に課題がある結果が報告されている⁽²⁾。英語力向上に向けた取組の支援として、デジタル教科書等 ICT の活用が進められ⁽³⁾、各都道府県において英語教育改善に向けた取組がおこなわれているものの⁽⁴⁾、英語を話す機会が圧倒的に少なく、speaking 練習を行う場を増やすことが重要である。

外国語口頭練習のための環境として、Chatbot を取

り入れた英会話システムの開発^{(5)・(7)}が進められており、その発展が期待される中、EFL (English as a Foreign Language) 学習者にとって、自身の音声認識されずに練習する意欲を失ってしまいかねない課題がある。

本稿では、音声認識を用いた口頭練習のために、類似度を取り入れる可能性を検討する。発音練習等とは別に、音声認識を用いて発話する機会を増やし、学習者が意欲を失わずに練習できるよう、実際に認識された英文と対象英文を比較し、一定の類似度があれば、対象英文が発話されたと判断することにより、speaking 練習を継続できるか、類似度を取り入れた口頭練習の可能性を検討する。但し、本取組では、どの類似度分析を取り入れればよいかといった類似度の比較ではなく、類似度を取り入れることにより、口頭練習の継続が可能かといった点を検討する。ここでは、

主に以下の点を検討する。

1. 類似度を取り入れることにより、練習を継続することは可能か？また、取り入れる際にどのような課題があるか？
2. 外国語 speaking 練習に関して、学習者はどのような状況か、どのような意識を持っているか？
3. 学習者は、外国語 speaking 練習に音声認識を用いることに抵抗感はあるか？

2. 予備調査

2.1 参加者

大学1年生107人の参加者が、類似度を用いた以下の方法で音声認識を試し、その後、アンケートに回答した。

2.2 方法

2種類の類似度を用いた音声認識を含むツールを準備し、対象となる以下の4英文を参加者に発話してもらった。音声認識はGoogle Speech Recognitionを用いた。英文の発話については、それぞれの方法において類似度の値を設定し、その閾値に達しない場合に最大3回まで発話を行い、3回を超えると次の英文に進む。

対象となる英文

- (1) May I see your passport, please?
(→Here you are.)
- (2) How long will you be staying in this country?
(→Six days.)
- (3) Have you ever been to Australia?
(→It's my first time.)
- (4) What's the purpose of your visit?
(→Sightseeing.)

方法1

方法1ではコサイン類似度を用いた。参加者が対象となる英文、例えば「May I see your passport, please?」を発話すると、参加者の発話英文に対して、実際に認識された英文結果「May I see your passport freeze」等が表示される。次に、対象となる英文と、実際に認識された英文に対する類似度解析の結果で

ある「0.929」等の0～1までの類似度計算結果が表示され、閾値以上の場合、参加者が発話したと想定される英文「May I see your passport, please?」が表示される。閾値を下回った場合は、「I didn't understand what you said.」等が応答としてかえてくる。

方法2

方法2では、Chatbot (ChatterBot) を用いた。参加者が対象となる英文、例えば「May I see your passport, please?」を発話すると、参加者の発話英文に対して、実際に認識された英文結果「May I see your passport freeze」等が表示される。方法2では、レーベンシュタイン距離による類似度解析の結果(0～1までの値)が閾値以上であれば、レスポンス文として準備されている「Here you are.」が応答としてかえてくる。閾値を下回った場合は、「I'm sorry, but I don't understand.」等が応答としてかえてくる。

アンケート調査

実施後、表1に示すspeaking練習や音声認識に関する参加者の意識や状況について、5段階で回答してもらった。

表1 speaking練習や音声認識に関する調査

(1)外国語の発話練習は必要だと思う。
1. そう思わない 2. あまりそう思わない 3. どちらとも言えない 4. ある程度そう思う 5. そう思う
(2)外国語の発話練習に時間を割いている。
1. 割いていない 2. あまり割いていない 3. どちらとも言えない 4. ある程度割いている 5. 割いている
(3)自分の都合のよいタイミングで外国語の発話練習ができる環境があれば使ってみたい。
1. そう思わない 2. あまりそう思わない 3. どちらとも言えない 4. ある程度そう思う 5. そう思う
(4)音声認識を使った外国語の発話練習に抵抗感がない。
1. そう思わない 2. あまりそう思わない 3. どちらとも言えない 4. ある程度そう思う 5. そう思う

3. 結果と考察

3.1 類似度を取り入れることによる練習継続の可能性について

上述した2つの方法において、最大3回まで各英文を発話することにより、参加者107人のうち何%が類似度により対象英文を発話したと判断されたか、表2・表3に結果を示す。

表2 方法1において各英文が発話されたと判断された参加者の割合(%)

方法1	英文1	英文2	英文3	英文4
1回目,又は2回目,又は3回目で 閾値以上	100%	100%	99%	92%
1-3回目全てで 閾値未満	0%	0%	1%	8%

表3 方法2において各英文が発話されたと判断された参加者の割合(%)

方法1	英文1	英文2	英文3	英文4
1回目,又は2回目,又は3回目で 閾値以上	100%	100%	100%	96%
1-3回目全てで 閾値未満	0%	0%	0%	4%

全体的な結果を見ると、今回対象とした4英文について、方法1・方法2とも、90%以上の参加者の発話した英文が対象となる英文と判断されていた。英文ごとでは、英文3と英文4において3回とも閾値未満の結果が出ていた。それぞれにおいて、英文のどういった点が認識されていなかったかを見ると、英文3では冒頭のHaveがHowと認識されており、英文4では、purposeがPapa, Papas, poppers, pop world's, pops等と認識されていた。これらの結果から、どちらの類似度を取り入れた方法ともspeaking練習を継続するという点において、その可能性が示唆された。

一方、今回実施した英文は4英文に留まり、また、比較的簡単な英文であった。英文によっては何度発話

しても対象の英文と判断されない場合があり、英文の中に含まれる単語や英文の長さについて、さらに対応方法を検討する必要がある。

今回の類似度を取り入れた方法2はChatbot(ChatterBot)を用いることにより、参加者が英語で問いかけるとレスポンスがかえってくることで、会話練習のツールへと発展でき、発音練習といった目的とは別に、英語を話す機会を増やすために、場面別に想定される英文を口頭練習するといったことが考えられる。speaking練習については、ここからさらに人との練習や指導が必要となり、speakingの練習段階や目的を考えた指導計画の中で、こういった方法を取り入れることを考えていく必要がある。また、Google Speech Recognitionについて音声学的、英語学的な評価や課題も考えられ、それらを考慮して取り組む必要がある。

3.2 外国語 speaking 練習に関する参加者の意識や状況について

表4～表6は、表1(1)～(3)の項目に対する回答結果について、参加者107人中の割合を示している。表4の結果から、「外国語の発話練習が必要だと思う」について、参加者の75%が「そう思う」、17%が「ある程度そう思う」と答え、全体の92%が外国語の発話練習の必要性を感じていた。

表4 アンケート(1)の回答に関する参加者107人中の割合(%)

選択肢の項目	(1)外国語の発話練習は必要 だと思う。
そう思わない	7%
あまりそう思わない	1%
どちらとも言えない	0%
ある程度そう思う	17%
そう思う	75%

一方、表5の結果から、参加者の78%（「割いていない」23%、「あまり割いていない」54%）は外国語の発話練習に時間を割いているとは言えない状況であり、表6の結果から、参加者の86%（「そう思う」62%、「ある程度そう思う」24%）が都合のよい時に練習を

行える環境があれば使いたいと考えていた。

これらの結果から、参加者は speaking 練習の必要性は感じているものの、実際に英語に触れる機会、特に話す機会が少ない状況で、発話練習を行える環境に対するニーズがあることがうかがえる。

表 5 アンケート(2)の回答に関する参加者 107 人中の割合(%)

選択肢の項目	(2)外国語の発話練習に時間を割いている。
割いていない	23%
あまり割いていない	54%
どちらとも言えない	6%
ある程度割いている	15%
割いている	2%

表 6 アンケート(3)の回答に関する参加者 107 人中の割合(%)

選択肢の項目	(3)自分の都合のよいタイミングで外国語の発話練習ができる環境があれば使ってみたい。
そう思わない	4%
あまりそう思わない	1%
どちらとも言えない	9%
ある程度そう思う	24%
そう思う	62%

3.3 外国語 speaking 練習に音声認識を用いることに対する抵抗感について

表 7 は、表 1(4)の項目に対する回答結果について、参加者 107 人中の割合を示している。表 7 の結果から、参加者の 84% (「そう思う」48%, 「ある程度そう思う」36%) が音声認識を使った外国語の発話練習に、特に抵抗感はなかったと答えていた。今回は、指定された英文を発話するという比較的易しい課題であり、また、実際に認識された英文と対象英文が完全に一致することがなくとも類似度による対応が準備されていたため、音声認識に対する否定的な反応は少なかった可能性が考えられる。

表 7 アンケート(4)の回答に関する参加者 107 人中の割合(%)

選択肢の項目	(4)音声認識を使った外国語の発話練習に抵抗感がない。
そう思わない	0%
あまりそう思わない	6%
どちらとも言えない	10%
ある程度そう思う	36%
そう思う	48%

4. おわりに

本稿では、発音練習等とは別に、音声認識を用いて発話する機会を増やし、学習者が意欲を失わずに練習できるよう、類似度を取り入れる可能性を検討した。その結果、90%以上の参加者の発話した英文が対象英文と判断され、継続して speaking 練習を行うという点において、類似度を取り入れた方法の可能性が示唆された。音声認識については、84%の学生が特に抵抗感は無かったと答えており、発話する機会を増やすために、類似度を取り入れた方法を用いて練習する可能性が示唆された。

今後は、次の段階として、Chatbot および類似度を取り入れた練習ツールを準備し、具体的な場面で使用される表現について実際に口頭練習を行い、その運用状況や課題点を探っていきたいと考える。また、ChatGPT のような生成 AI の発展がめざましい。ChatGPT はテキストデータの広範なコーパスで訓練されることにより、人間と対話しているかのような自然なレスポンスを生成できる特徴があり、口頭練習への活用可能性が考えられる。一方、ChatGPT は基本的に、説明的でまとまったレスポンスを返す傾向があるため、口頭による対話的な結果(出力)を生むために、プロンプトと呼ばれる指示を工夫することが重要である。こういった条件や関連情報、役割を与えれば、口頭練習に役立てることができるかについても検討していきたいと考える。

参 考 文 献

- (1) 文部科学省: “小学校学習指導要領（平成 29 年告示）解説 外国語活動・外国語編”, (2017)
- (2) 国立教育政策研究所: “令和 5 年度 全国学力・学習状況調査の結果（概要）”, (2023)
<https://www.nier.go.jp/23chousakekkahoukoku/report/data/23summary.pdf> (2023 年 8 月 30 日確認)
- (3) 文部科学省: “外国語の指導における ICT の活用”, (2020)
https://www.mext.go.jp/content/20200911-mxt_jogai01-000009772_13.pdf (2023 年 8 月 30 日確認)
- (4) 文部科学省: “令和 4 年度英語教育改善プラン推進事業”, (2022)
https://www.mext.go.jp/a_menu/kokusai/gaikokugo/index_00009.htm (2023 年 8 月 30 日確認)
- (5) 與那覇啓人, 小渡悟: “自然対話ボットによる英会話体験システムの検討”, 教育システム情報学会学生研究発表会, 241-242 (2019)
- (6) Muhammad, A.F., Susanto, D., Alimudin, A., et al.: “Developing English conversation chatbot using Dialogflow”, 2020 International Electronics Symposium (IES), 468–475 (2020)
- (7) Kim, H., Yang, H., Shin, D., et al.: “Design principles and architecture of a second language learning chatbot”, Language Learning & Technology, 26(1), 1–18 (2022)