

日本語母語話者の英語弱化母音の発音： 自動音声認識を用いた分析

Pronunciation of English Weak Vowels by Native Speakers of Japanese: Analyses Using Automatic Speech Recognition

中西のりこ¹ 高口 朋也²

Noriko NAKANISHI, Tomoya TAKAGUCHI

(要旨)

本研究では、日本語を母語とする363名の英語学習者による22種類の英文パッセージ朗読2,836音源を対象に、音声がどのように自動認識されたかを調べた。その結果、弱母音 /i/, /u/, /ə/ が認識されにくいこと、二重母音 /ai/ と /ou/ が過剰に認識される傾向があることが示された。この傾向を音素レベルで分析したところ、弱母音 /ə/ が低舌母音 /æ/ や /ɑ/ と置き換わって認識される傾向が見られた。さらに、語レベルの分析では、1語が複数の語として認識されたり、複数の語が1語として認識されたりするケースが見られた。これらのことから、本研究対象者の弱母音の発音に問題があり、母音を弱化させずにすべての母音に等分の力を入れて発音するため語強勢のコントロールがうまくいかず、単語境界が認識されにくいという結果に至ったと考えられる。教育現場では、日本語の音韻体系に存在しない音素に注目して発音指導が行われることが多いが、弱母音もそのうちに含まれるということを念頭に置いた指導が望まれる。

キーワード：自動音声認識、日本語母語話者英語発音、母音の弱化、語強勢

Key Words : Automatic speech recognition, English pronunciation of Japanese native speakers, weak vowels, word stress

-
1. 神戸学院大学 グローバル・コミュニケーション学部 教授
 2. 神戸学院大学 グローバル・コミュニケーション学部 学部生

1. はじめに

1. 1 自動音声認識機能 (ASR) の日常利用および問題点

現代の ICT の発達により、英語学習者が会話をする相手が人間に限定されず、自動音声認識 (Automatic Speech Recognition, ASR) を用いた機器を相手に発話をする時代は近いと考えられる。例えば、旅行先でのタクシー予約サイト (e.g. Taxi Butler, n. d.)、航空券やホテル予約サイト (e.g. Rosen, 1997) で ASR が用いられている場合、目的地の地名が音声認識されなければスムーズな予約は困難となる。学術的なやり取りにおいては、Hartley, Sotto, and Pennebaker (2003) が、28名による英文ライティングをキーボード入力の場合と Dragon Naturally Speaking の自動音声認識機能を用いた場合とで比較し、英文中の文字数などの量の面でも、読みやすさやタイピングエラー、文法エラーなどの質の面でも両者の間に有意差がないことを明らかにした。さらに、英文入力のスピードに関しては、キーボード入力よりも ASR 入力の方が3倍速いという調査結果も報告されている (Ruan, Wobbrock, Liou, Ng, & Landay, 2016)。つまり、従来の World Englishes の議論では英語コミュニケーションの相手が「誰なのか」に焦点が当てられていたが (Kachru, 1992)、これに加えて、相手が「人なのか機械なのか」という視点が必要になると考えられる。

しかし現状の問題点は、特定の英語変種によってチューニングされた ASR では、別の英語変種による発話の音声認識率が下がることである。例えば Minematsu, Hakoda, Zhu, Nakanishi, Nishimura, and Saito (2021) では、12名の日本語母語話者および2名の英語母語話者 (米語1名、英語1名) による30秒程度の発話を、アメリカ英語母語話者2名、日本語母語話者2名、非英語母語話者3名 (合計7名) にシャドーイングさせ、その音声を手動で書き起こした場合と ASR (Amazon) で認識させた場合との認識率の相関を比較した。その結果、7名の音声のうち6名の音声に関しては .85-.91 の相関が得られたが、日本語アクセントの影響をやや強く受けた話者1名の音声のみ、書き起こしと ASR 結果との相関が .69 となった。つまり、非英語母語話者による英語発話に対して、現状の ASR は人間の耳ほど柔軟に対応できないことが示唆された。

1. 2 自動音声認識機能 (ASR) の教育的利用

ASR が人間の耳ほど柔軟な音声認識をしないということは、日常生活や旅行先などで非英語母語話者が英語発話を自動音声認識させなければならない時には不自由するが、英語教育現場で用いれば、学習者音声の伝わりやすさに一定の基準を設けて示すことができる。

Nakanishi, Tam, and Ebihara (2020) では、出身地や教育歴などの背景が異なる3名の英語教員が134名の学習者の発話音声をそれぞれ聴き、発音エラーの指摘が必要であると感じた箇所が ASR での誤認識とどれほど重なっているかを調べた。その結果、各教員が指摘した内容と ASR 結果に乖離があるだけでなく、教員間で注目する点に大きな違いがあることが明らかとなった。このように、口頭コミュニケーションにおける発話の「伝わりやすさ」は、話し手がどのような英語変種を用いるかだけでなく、聞き手がどのような変種に慣れ親しんでいるかという要素の影響も受ける。

そこで中西 (2018) では、聞き手を iOS に搭載されている音声認識アプリに固定するこ

とにより、教員の音声学的な興味や指導の得意分野に左右されずに、個々の学習者がどの音素の発音を苦手としているかを提示する手法を検討した。英語学習者に対して音声認識アプリを用いた明示的指導を行った結果、認識されにくい音素を学習者が自ら把握し、発音練習に取り組む様子が観察された。

また、ASR の認識精度が上がってきていることを受け、中西 (2019) では音声認識アプリを利用したライティング活動を実践し、ASR が英語リスニングやスピーキングだけでなくライティングのためにも利用できる学習ツールとなることを示した。学習者が英作文の音声入力を試みても「個別音の発音や連結・脱落などの音声変化、文全体のリズムが整っていない音声はなかなか思うように認識されない」という傾向があることから、「ライティングのためにも英語の発音練習が必要となる時代」において、ASR は「時間や場所を選ばず発音練習のお供をしてくれる自習ツール」となることを予測した (p.67)。

日本語母語話者の英語発音は「machine gun effect (Guerin, 1978)」と呼ばれるように、すべての音節に等分に強勢を置いて発音する傾向があることが知られている。波多野・北村 (2014) は、英語母語話者 ($n = 16$)、日本語母語話者 ($n = 9$) の発話内の dormitory /dɔ:rmətɔ:ri/, understand /ʌndərstænd/, conversation /kənvərséɪʃən/ に含まれる弱化母音 /ə/, /ər/ を音響・調音の両面から観測し、日本語母語話者による発話では /ə/ の調音的弱化が不完全な形でしか実現されないことを指摘した。このような背景から、中西 (2021) では、大学 1、2 年生 ($n = 166$) が音読した音源の ASR (Google API) 結果を、認識率が高い群と低い群で比較し、認識率が低い群では /i/, /u/, /ə/, /ər/ のような弱化母音の認識率が低いことを示した。さらにその後の分析により、音の弱化が起こる部分の音声は、より聞こえ度の大きい母音 (図 1 参照) と置き換わって認識される傾向があることが示唆された。

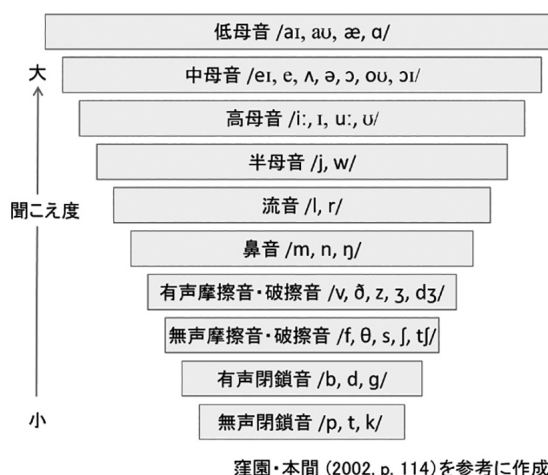


図 1 主要な母音・子音の相対的聞こえ度 (中西・山本, 2022, p.87より抜粋)

1. 3 本研究の目的

本研究では、中西 (2021) で調査対象となった大学 1、2 年生が音読したパッセージの音声はどのように自動音声認識されたかをさらに詳しく分析することにより、日本語母語話

者の英語発音の特徴を明らかにし、英語教育現場での指導に役立てることを目的とする。特に母音は発話のリズムに強く関係するため、本研究では母音に焦点を絞り、以下3点の研究課題（Research Question, RQ）を挙げる。

RQ1：学習者による英語音読音声の認識結果の母音数は音読原稿中の母音数とどう異なるか

RQ2：誤認識された単語列から、学習者の英語発音にどのような特徴が見いだされるか

RQ3：どの母音がどの母音に置き換わって認識される傾向があるか

2. 研究の方法

2. 1 音読音声収録の方法と分析対象とするデータセット

本研究において分析対象とする英語音読音声は、関西の大学でコミュニケーション学を専攻する2020–2021年度の1、2年次生によって収録された。彼らの英語力の目安は、CEFR A1-B2（Versant Speaking Test Overall $M = 36.6$, $SD = 5.51$ 、TOEIC L&R $M = 454.8$, $SD = 131.37$ ）程度であった。音読音声収録は2020年5月から2021年11月の授業実施期間中に、授業進度に合わせて、2–3週間に1度のペースで行われた。感染症予防のための遠隔授業期間中に各学習者の自宅で収録されたため、教室内で録音するよりも雑音が少なく、周囲への影響に対する心理的負担が小さい状態で収録された。

音読原稿には、大学1、2年次生「英語会話」授業用にオリジナルで作成し音源付きで配信しているテキスト中のモノローグパッセージを用いた。学習者はモデル音源を聴きながらパッセージ中の語彙や内容理解を問う設問に回答した上で、音読音声を録音した。音声収録にはSpeech Saver（Nakanishi, 2019b）を用いた。このシステムにはGoogle Cloud Speech-to-Text APIがデフォルト（米音）の状態で組み込まれており、音声を録音すると同時に、自動音声認識された文字列が画面上に表示され、サーバーに保存される。音声と認識文字列には学籍番号とともに固有の参照コードが付与され、画面上に表示される。学習者は音声認識結果の文字列を確認しながら繰り返し録音を行い、自分が一番良いと考える音声の参照コードを提出した。1課題のパッセージ音読の中で良い部分だけをつなぎ合わせたり、認識された文字列を学習者が修正したりすることはできない仕様になっている。本研究で用いるデータは、この参照コードから呼び出した、学習者による音読音声の認識文字列である。

合計24種類のパッセージ音読音声および認識文字列が提出された。本研究では、これらのうち、固有名詞や数詞の出現数が比較的少ない22種類のパッセージ（518文、6,481 tokens、1,301 types、Automated Readability Index 6.4、Coleman-Liau Index 8.9¹）を分析対象とする。2種類のパッセージを除外した根拠は、1）固有名詞の認識結果はASRが参照する語リストに含まれているかどうか依存する、2）数詞の認識結果からはASRがどのように聞き取ったかが判断できない（例えば“2020”というASR文字列からは“twenty twenty”と認識したのか“two thousand twenty”と認識したのかが判断できない）の2点である。さらに、上記の基準を満たした3,985音源のうち、マイクの不調などが原因で

パッセージが全て録音されていないものや言い直しをしているものを除外するため、モデル音源と比較して録音時間が1 - 2倍、認識文字列数が0.9 - 1.1倍の範囲に収まるものを抽出した。最終的に、363名の学習者が録音した2,836音源（828,886語、録音時間平均197.6秒 $SD = 36.96$ ）、音声認識文字列数平均95.6%（ $SD = 2.64$ ）が分析対象となった。

2. 2 分析の方法

学習者による英語音読音声の自動認識母音数が音読原稿中の母音数とどう異なるか（RQ1）の分析には、音素カウンター（Nakanishi, 2019a）を用いる。これは、ブラウザ上で入力した英文を音素に分解し、文中に含まれる音素数を算出するシステムである。音読原稿内、および、学習者による音読の認識結果内に含まれる母音（24種類）の出現数をそれぞれパッセージごとに算出し、認識結果の母音出現数を学習者数で割ることにより、原稿内と認識結果内の母音数を比較する。22パッセージ分の原稿と認識結果の間で出現する母音数平均に違いがあるかを調べるため、対応のある t 検定（両側）² を行う。

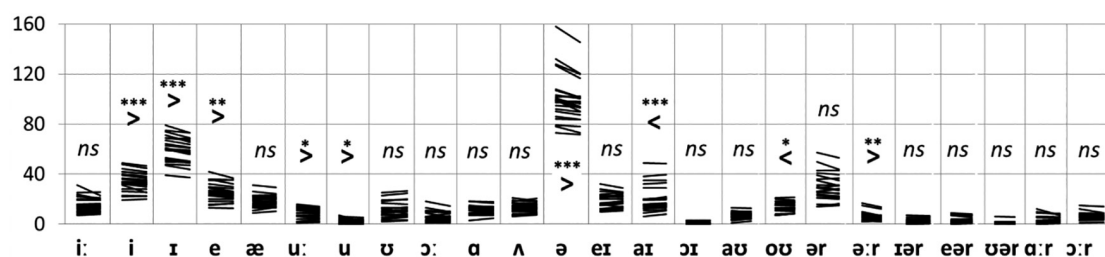
誤認識された単語列から学習者の英語発音にどのような特徴が見いだされるか（RQ2）を調べるため、誤認識が顕著な単語の前後に出現する語も含めた誤認識結果を調べる。音読原稿と認識結果の単語列の対応をとり、どの語がどの語に置き換わって認識されたかを比較する。

最後に、どの母音がどの母音に置き換わって認識される傾向があるか（RQ3）を調べるため、上記の語をさらに音素単位に分け、学習者音声のどこに問題があったのかを検討する。

3. 結果と考察

3. 1 音読用原稿中の母音数と学習者音声の母音認識数の比較（RQ1）

図2に、全22パッセージ中の母音出現数を示す。1つの線が1パッセージ内での母音出現数を示すため、母音ごとに22本の線が示されている。線の左側がパッセージ原稿中の出現数、右側が学習者による音読音声の自動認識結果中の出現数である。例えば /ə/ の出現数がもっとも多かった“International affairs”というパッセージでは、原稿中には158回 /ə/ が出現していたが、163人の学習者による音読音声では23,706回認識されていたため、学習者全体の平均出現数は145.4回となる。つまり、右下がりの線は原稿中の出現数より認識数の方が少ないケース、右上がりの線は認識数の方が多いケースである。学習者音声



* $p < .05$, ** $p < .01$, *** $p < .001$.

図2 パッセージ中の母音出現数：原稿（左）と音読認識結果（右）の比較

音素カウンターのデータベースに含まれない語として誤認識されたため音素カウント不可となった語は、認識された総語数828,886語中、9,007語（1.1%）であった。

パッセージごとの原稿と音読認識結果の間で出現する各母音の数に差があるかどうかを調べるため t 検定を行ったところ、0.1%水準で有意差が見られた母音は4つ（ $/i/$ $t(21) = 4.82, r = .72$; $/ɪ/$ $t(21) = 7.73, r = .86$; $/ə/$ $t(21) = 5.35, r = .76$; $/aɪ/$ $t(21) = 5.27, r = .75$ ）、1%水準で2つ（ $/e/$ $t(21) = 3.54, r = .61$; $/ə:r/$ $t(21) = 3.62, r = .62$ ）、5%水準で3つ（ $/u:/$ $t(21) = 2.34, r = .45$; $/u/$ $t(21) = 2.33, r = .45$; $/oʊ/$ $t(21) = 2.38, r = .46$ ）であった。この結果と図2に示された傾向を勘案すると、 $/i/$ 、 $/ɪ/$ 、 $/e/$ 、 $/u:/$ 、 $/u/$ 、 $/ə/$ 、 $/ə:r/$ は学習者が発音したつもりでも認識されない傾向があり、逆に $/aɪ/$ と $/oʊ/$ は発音したつもりはないにもかかわらず誤認識結果の単語列に過剰に含まれていたと解釈することができる。

この結果において特徴的なことは、弱母音 $/i/$ 、 $/u/$ 、 $/ə/$ が全て認識されにくい母音に含まれているということである。さらに「聞こえ度（窪田・本間, 2002）」という観点から分類すると、認識されにくい母音のうち $/i/$ 、 $/ɪ/$ 、 $/u:/$ 、 $/u/$ は聞こえ度の低い母音であるのに対して、認識結果に過剰に含まれる傾向がある $/aɪ/$ と $/oʊ/$ は聞こえ度の高い二重母音である。この傾向は中西（2021）およびその後の研究においても示されており、学習者が、弱化する母音の発音に困難を抱えている（波多野・北村, 2014）ということが改めて支持された。

このことから、弱母音という概念自体を学習者が理解していないためすべての母音に力を入れる machine gun effect（Guerin, 1978）が誤認識の要因となっている可能性が示唆された。英語教育現場における発音指導では $/i/$ 、 $/u/$ 、 $/ə/$ のような弱母音で力を入れすぎず、学習者が強弱リズムを意識することを促す指導が望まれる。

3. 2 誤認識された単語列から分かること（RQ2）

弱母音 $/ə/$ を含む語のうち、誤認識率が高かった内容語は *penalty* /pénəlti/ および *academic* /ækədémik/ であった。パッセージ中に *penalty* は合計7回、*academic* は5回出現していたが、前後の語によって誤認識率が異なるため、これら2語の前後1語ずつを含めた環境ごとに誤認識率をまとめた。表1は、これらをパッセージ内で出現した順に示したものである。

人間の耳による聴解では通常、同じ単語が複数回出現するとき、後になるほど聞き取りが容易になるが、ASR を用いた認識の場合「耳が慣れる」ということが起こらない。その

表1 *Penalty* (94音源)、*academic* (114音源) の出現環境による ASR 誤認識率

penalty				academic			
death	penalty	is	16.0%	your	academic	achievement.	11.4%
death	penalty	would	12.8%	your	academic	achievement.	27.2%
death	penalty	. Our	10.6%	your	academic	performance.	21.1%
death	penalty	should	3.2%	good	academic	standing	9.6%
death	penalty	is	14.9%	to	academic	success.	10.5%
a	penalty	. Moreover,	20.2%				
death	penalty	is	10.6%				
平均			12.6%	平均			16.0%

ため、表1で示されるように、回数を重ねるに伴って認識されやすくなることはない。つまり、発音練習に ASR を用いると、毎回別の聞き手が初めてその語を聞くことを想定した認識結果が得られる。聞き手の「慣れ」に依存しない形で、学習者の発音がどれほど伝わるかの目安として有効であると考えられる。

Penalty は“Law and Peace”というパッセージで死刑制度について述べたもので、パッセージ内で出現する7か所のうち6か所は death penalty という名詞句として用いられていた。まず、最も顕著な誤認識はこの2語が disparity という1語として認識された例(6か所計15例)であった。これは1語目の death に十分な強勢が置かれず、しかも語末子音 /θ/ が /s/ と発音されたことに起因すると考えられる。次に、penalty が death と共起せず、a penalty が文末にある環境では、Panera tea /pənérə tí:/ もしくは Panera key /pənérə kí:/ (6例) という誤認識が目立った。これは、前述の例とは逆に、1語の penalty が2語として認識された例である。本来は /pénəlti/ のように最初の音節に語強勢が置かれるべきところを、2音節目に強勢を置いたことに加え、語末の弱母音 /i/ に問題があり強勢が置かれる緊張母音 /i:/ として認識されたことが分かる。他にも語強勢位置の問題に起因すると考えられる誤認識は、personality /pə:rsənəlti/ (4例)、analogy /ənəlti:/, banality /bənləti/, desperately /déspərətli/, mentality /mentəlti/, neurology /njuərlədti/, novelty /nəvəlti/, polarity /pəulərti/, reality /riəlti/, technology /teknəlti/ (各1例) など多岐にわたった。

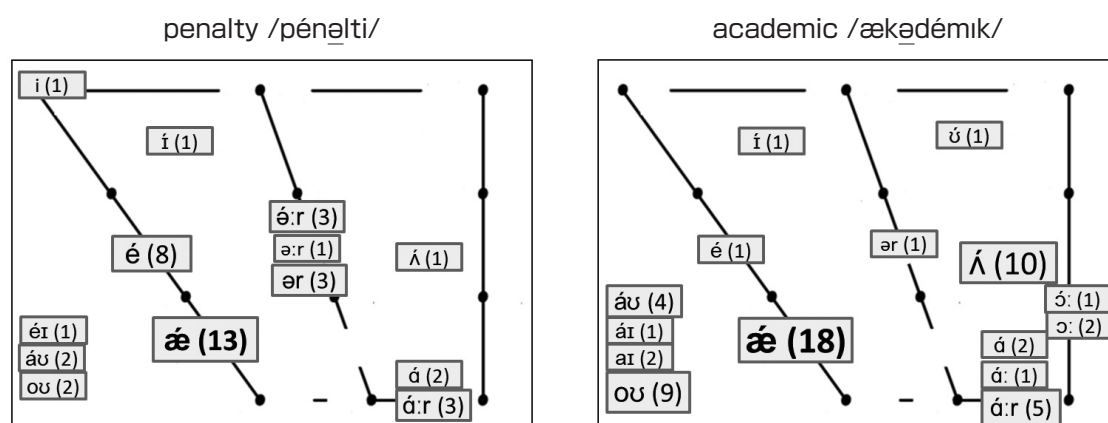
Academic は“College Life”というトピックで、大学での学業成績について述べたパッセージ内で用いられた語である。5か所のうち4か所は、achievement, performance, success という語と共起して文末から2語目に出現していた。特に顕著な誤認識は、academic という4音節の1語が、going to make /góʊɪŋ tə méɪk/ (9例), car to make /ká:r tə méɪk/ (3例), how to make /háʊ tə méɪk/ (3例) のように複数の語として認識されるケースであった。これは、本来なら academic /ækədémɪk/ のように3音節目に第一強勢がおかれる語を、1音節目と最終音節が卓立する形で発音した結果であると考えられる。上記の誤認識例は、最終音節の弛緩母音 /ɪ/ が /éɪ/ という二重母音に置き換わっている点で共通する。/ɪ/ が /éɪ/ として誤認識されたケースは make (上記を合わせて25例)、case、name (各1例) などが見られた。さらに、academic という1語が1語として認識されるケースにおいても、やはり語強勢の位置に起因する例が目立った。例えば、academy /ækədəmi/ (15例)、cousin /kázn/, economy /íkənəmi/ (各1例) などは、本来 academic で第一強勢が置かれる3音節目より前の音節に強勢を置いて発音したことの表れであると考えられる。

3.3 特定の母音がどの母音に置き換わって認識されたか (RQ3)

前項で取り上げた penalty /pénəlti/ および academic /ækədémɪk/ に対する誤認識結果のリストを付録に示す。図3は、これらの語に含まれる弱母音 /ə/ がどの母音に置き換わって誤認識されたかを集計し母音四角形上に示したものである。元の単語がどちらの単語であったかによって若干のばらつきはあるが、/ə/ が /æ/ と置き換わる誤認識が最も多く見

られた。この結果は、日本語母語話者の発話中の /ə/ の第2フォルマントが母語話者と比較して大きくなる、つまり前舌寄りで発音されるという波多野・北村（2014）による記述と一致する。さらに、大半の誤認識において、語強勢が置かれる母音と置き換わっていることも特徴的である。前項で述べた通り、語強勢の置き方や位置に問題があったことが音素レベルの分析によっても支持された。また、/á/ や /á:r/ のような後舌低母音との置き換えも比較的頻繁に起きていたことが分かる。Penalty はペナルティ、academic はアカデミックのように日本語としても使われる語であることから、カタカナ語との混同により「ナ」や「カ」のように口を開いて発音したため、低舌母音として認識されたと考えられる。

通常、英語を指導する教員が日本語母語話者の場合、カタカナの日本語として発音された語も理解できる。しかし実際に世界の英語話者を聞き手とする会話を想定すると、「ペナルティ」「アカデミック」とカタカナで発音された語がスムーズに理解されるとは限らないということを、上記の結果は示している。発音指導の際には、弱母音 /ə/ を発音する際に口を大きく開けすぎないこと、カタカナ語との混同を避ける工夫が必要であることが示唆された。



注. 別の単語に置き換わっているが /ə/ として認識されたケースは penalty 23例、academic 13例であった。元の単語の /ə/ の位置に対応する母音がないケースは penalty 8例、academic 10例であった。

図3 Penalty, academic 弱母音部分の誤認識結果

4. おわりに

本研究では、363名の学習者による22種類の英文パッセージ朗読2,836音源を対象に、音声 Google Cloud Speech-to-Text API によってどのように認識されたかを調べた。その結果、朗読原稿中の母音数と認識された母音数の比較 (RQ1) では、弱母音 /i/, /u/, /ə/ が認識されにくいこと、二重母音 /ai/ と /ou/ が過剰に認識される傾向があることが明らかとなった。誤認識されやすい単語列の傾向の分析 (RQ2) では、語強勢の置き方や位置に問題があるため1語が複数の語として認識されたり、複数の語が1語として認識されたりするケースが見られた。さらに、弱母音 /ə/ の認識に絞った分析 (RQ3) では、/ə/ が低舌母音 /æ/ や /a/ と置き換わって認識される傾向が見られた。これらのことから、日本語母語話者の英語の特徴として、弱母音の発音に問題があることが示唆された。母音を弱化させずにすべての母音に等分の力を入れて発音するため、語強勢のコントロールがうまく

いかず、結果として単語境界が認識されにくいという結果に至ったと考えられる。教育現場では、日本語の音韻体系に存在しない音素に注目して発音指導が行われることが多いが、弱母音もそのうちに含まれるということを念頭に置いた指導が望まれる。

本研究結果から得られる教育的示唆として、まず、ASR を用いることのメリットについて述べる。本研究では学習者から提出された音源を一括して分析をしたが、教育現場では、中西（2018; 2019）の例のように、個々の学習者がそれぞれ自分の音声の ASR 結果を確認することにより、どこに気をつければよいかをその場で知ることができる。特に日本語母語話者にとって発音が難しい単語などは繰り返し練習することが必要であるが、人前で何度も同じフレーズを声に出すことに対して心理的負担を感じる学習者も存在する。ASR 相手ならば気遣いなく繰り返し練習をすることができるため、学習者の心理的負荷を軽減することができる。学習者が自分自身の発音の傾向を把握したうえで教員が手本を見せれば、発音練習のモチベーションの向上につながると考えられる。次に、ASR は日常でも使われているため「何のために練習をするのか」を学習者がイメージしやすいというメリットがある。例えば、タクシーやホテル、航空券の予約の際に自動音声認識が用いられている（Rosen, 1997; Taxi Butler, n.d.）。このように対人だけでなく機械を相手にコミュニケーションをとらなければいけない場面を提示することにより、学習者が現実世界で英語を使用する場面を具体的に想定しやすくなると考えられる。ASR は、様々な聞き手のうちの 1 つでしかないことを念頭におき、機械に認識させることが最終目標ではなく、コミュニケーションを通して何ができるかという学習の目的を顕在化させることも重要である。さらに、ASR を使った英語学習は発話指導のみに限定されない。Ruan, Wobbrock, Liou, Ng, and Landay (2016) や中西 (2019) が示したように、ライティング活動に ASR が用いられる事例もある。従来は英文ライティングをパソコン上で行うにはタイピング練習が必要であったが、その代わりに、ASR による音声入力の練習が英文ライティング活動の一部として取り込まれることも考えられる。最近の英語科の授業では英語 4 技能（スピーキングに関してはやりとりと発表）の統合型が望ましいとされているが、ASR は、発音指導のみならず他の技能においても活用されることが期待できる機能である。

一方、ASR のデメリットとしては、学習者が、発音が上手になったと勘違いする可能性があることである。自分の発話が認識されると満足し、反復練習を避ける学習者もいるかもしれない。ASR は音素情報とコロケーション情報を元に音声を認識するシステムであり、ピッチ情報は認識に影響をおよぼさない。つまり、ASR に認識された単語列が、実際に発話した単語として人間の聞き手に理解され、感情が伝わるとは限らない。機械による認識結果に一喜一憂するのではなく、人間の聞き手を想定した継続的な発話練習も必要である。もうひとつのデメリットは、ASR を用いた発音練習が学習者の自己肯定感を低下させる原因にもなり得ることである。発話が ASR に認識されないことに対して自信をなくしてしまう学習者もいると考えられるため、教員側の継続的な指導と学習者の練習が必要である。ASR とは音声を自動で認識するシステムであり、個人の発音を評価するものではない。今後さらなる ICT の発達により ASR 自体が進化していくので、音声認識されるかどうかを発音の良しあしと関連付けて論じにくくなる可能性も考えられる。そのた

め、英語教育現場で ASR を使用する際には教員が使い方を提示し、反復練習することの重要性を同時に伝える必要がある。最後に、ASR を教育現場に導入する際の教室環境の整備について述べる。近年、学習者一人一人がタブレット端末を使用できる教育機関も増加しつつあるが、インターネット環境や、パソコンやヘッドセット機器などハード面の整備が不十分であると考えられる。教室で安全なインターネット接続を可能にするためのセキュリティ面の整備も含め、ICT を活用した英語教育の普及が求められる。

これらのことから、本研究で用いた手法は、教育現場で活用できると思われる。だが、上述したように ASR を教育現場に導入する際のメリットデメリットがあることや、導入するための教室環境などの問題点を考慮しながら使用することが望ましい。今後、学習者がより多くの機会でも ASR を使用して自分の発音を矯正し、人や機械に理解される英語を習得してもらいたいと考える。

謝辞

本研究は科研費（17K02914）の助成を受けたものである。

参考文献

- Guerin, T. (1978). Pronunciation problems in English arising from Japanese modes of phonological production. *The Journal of the Faculty of General Education and Women's Junior College* 13, 51 – 72.
- Hartley, J., Sotto, E., & Pennebaker, J. (2003). Speaking versus typing: A case-study of the effects of using voice-recognition software on academic correspondence. *British Journal of Educational Technology*, 34(1), 5 – 16.
- 波多野博顕・北村達也. (2014). 日英母語話者による英語弱化石音の音響・調音特徴：X線マイクロビームデータベースに基づく分析. *日本音響学会誌*, 70(3), 106 – 113.
- Kachru, B. B. (1992). World Englishes: Approaches, issues and resources. *Language Teaching*, 25(1), 1 – 14.
- 窪園晴夫・本間猛. (2002). 『音節とモーラ』（英語学モノグラフシリーズ15）研究社.
- Minematsu, N., Hakoda, R., Zhu, C., Nakanishi, N., Nishimura, T., & Saito, D. (2021, September 6-8). *Objective measurement of instantaneous intelligibility of L2 utterances based on shadowing* [Paper presentation]. ISAPh2021 3rd International Symposium on Applied Phonetics, Tarragona, Spain.
- Mizumoto, A. (2015). *Langtest* (Version 1.0) [Web application]. <http://langtest.jp>
- 中西のりこ. (2018). 「音声認識アプリを用いた自律的発音学習」. 『教育開発センタージャーナル』 第9号, 25 – 37.
- 中西のりこ. (2019). 「グローバル・コミュニケーションのための英語4技能と音声指導」. 『英語教育2019年1月号』（pp.66 – 67）大修館.
- Nakanishi, N. (2019a). *Phoneme Counter* (Ver5.1) [Computer software]. <https://noriko-nakanishi.com/phoneme/>
- Nakanishi, N. (2019b). *Speech Saver* (Ver 1.0) [Computer software]. <https://noriko-nakanishi.com/speech/>
- 中西のりこ. (2021, November 20). 「ASR は英語教育現場で使えるか：学習者の発音得点との関係および誤認識結果の分析」 [口頭発表]. 大学英語教育学会（JACET）関西支部大会.
- Nakanishi, N., Tam, S. Y., & Ebihara, Y. (2020). Spotting English pronunciation errors: Comparison among teachers and with automatic speech recognition. *LET Kansai Chapter Collected Papers*, 18, 125 – 146.
- 中西のりこ・山本晃司. (2022). 「第5章 音節、音連鎖、連続音声」. 長瀬慶来教授 古希記念出版刊行委員会編『英語音声学・音韻論—理論と実践』（pp.84 – 102）大阪教育図書.
- Rosen, C. (1997) *Voicerecognition to ease travel bookings*. Northstar Travel Group. <https://www.businesstravelnews.com/More-News/Voice-Recognition-To-Ease-Travel-Bookings>

- Ruan, S., Wobbrock, J. O., Liou, K., Ng, A., & Landay, J. (2016). *Speech is 3x faster than typing for English and Mandarin text entry on mobile devices*. arXiv preprint arXiv: 1608. 07323.
- Someya, Y. (2009). *Word Level Checker*. [Computer software]. <http://someya-net.com/wlc/>
- Taxi Butler. (n. d.). *Taxi Butler is a success in over 40 countries*. Taxi Butler.
<https://www.taxibutler.com/clients/>

注

- 1 音読原稿の readability 分析には Someya (2009) を用いた。
- 2 統計分析には Mizumoto (2015) を用いた。

付録 Penalty, academic 弱母音 /ə/ 部分の誤認識結果

pén	ə	lti	n	æk	ə	démik	n
dispær	ə	ti	14	æk	ə	démiks	3
pær	ə	di	2	i:k	ə	námik	2
déspær	ə	tli	1		ə	deməkrætík	1
díd	ə	nt sí:	1		ə	támik	1
kén	ə	di	1	ɑ:rk	ə	démik	1
náv	ə	lti	1	æd	ə	mənt	1
njuərál	ə	ɔʒi	1	ət	ə	límpik	1
pær	ə	ti	1	gét	ə	míkænik	1
w	ə	lti:ɸ	1	hwʌts ð	ə	témp mí:n	1
	ə:r	ði	1	kak	ə	tú:méik	1
n	ə:r	di	2	ʌð	ər	neim	1
w	ə:r	ld	1	k	ʌ	tɪŋ	2
despéin	ər	der	1	k	ʌ	zn	2
pæn	ər	əti	1	ðə	ʌ	ðərmiks	1
tæn	ər	art	1	ək	ʌ	znkeis	1
ən	ʌ	ðər	1	k	ʌ	mɪŋmi	1
pæn	é	rəti:	5	k	ʌ	tɪŋméik	1
péin	é	vriðer	1	k	ʌ	tɪŋmi	1
pæn	é	ləpi	1	k	ʌ	tɪŋmí:t	1
pæn	é	rəki:	1	kr	é	dətméik	1
déspærət	ér	bl tə	1	ək	æ	dəmi	15
pə:rsən	æ	ləti	4	ək	æ	dəmikəd	1
	æ	ldi	1	ətək	æ	dəmi	1
bən	æ	ləti	1	g	æ	vən	1
despən	æ	lti:	1	d	ai	næmík	1
ən	æ	ləɔʒi	1	t	ai	tæník	1
kən	æ	l	1	tr	ái	ɪŋtméik	1
ment	æ	ləti	1	h	áo	təméik	3
p	æ	rədi	1	h	áo	ðeiméik	1
poul	æ	rəti	1	g	óo	ɪŋtəméik	9
ri	æ	ləti	1	k	ɔ:	lɪŋmi	1
d	áv	ti	1	k	ɔ:	lɪŋmikát	1
téstənd h	áv	ðer	1	k	ɔ:	tɪn	1
déspær	ou	tí:	1	g	á	ttəméik	1
n	ou	təs	1	ɪk	á	nəmi	1
m	ó:r		1	ævouk	á:	douméik	1
bət n	á	t	1	k	á:r	təméik	3
tekn	á	ləɔʒi	1	k	á:r	ðætme'iks	1
bərn	á:r	dɪt	1	k	á:r	mík	1
dástbɪn	á:r	tí:	1	ðək	í	ɪdəm	1
spɪn	á:r	di	1	k	ó	dəntméik	1
évr	i	der	1	æ		démik	4
fəs	í	ləti	1			déməkræt	2
æ		ldɔʒi:	1			ðəmi:k	1
léi		di	1			déniz	1
páp			1			rú:m	1
pé		ni	1	á:rk		ðætmeík	1
sí		ti	1				
trém			1				
θó:r		ti	1				