

音声認識 AI を活用した英語音読データの分析 —大規模フォニックス学習データに基づく誤読傾向の可視化—

長岡 穂^{*1}

Email: s-nagaoka@bunri-c.ac.jp

*1: 西武文理大学 サービス経営学部

◎Key Words 音声認識 AI, 音読, 第二言語音声習得, 誤読分析, 音素

1. はじめに

教育研究においては、学習活動のデータ化とその分析に基づく指導に関する知見の蓄積の重要性が指摘されている。英語教育における音読活動の評価は、これまで主観的判断に依拠する傾向があり、客観的データに基づく分析の必要性が指摘されている。こうした中、音声認識技術の発展は、学習者の発話を記録・分析する手段を提供し、音読に関する学習過程を定量的に把握することを可能としている⁽¹⁾⁽²⁾。

音読は、語彙認識の自動化や文理解の促進に関与する重要な学習活動であり、そのパフォーマンスは学習者の言語処理能力を反映すると考えられている。しかし、日本の英語教育では語彙や文法に関する指導が中心となる場面も多く、音声と綴りの対応関係に関する体系的な指導は必ずしも十分とは言えない。その結果、学習者は音読において誤読を生じやすく、その傾向を把握することが課題となっている。

音読に関する研究では、流暢性 (fluency) は正確性・速度・表現の統合的指標として位置づけられ、読解能力の発達と密接に関連することが指摘されている⁽³⁾⁽⁴⁾。また、誤読は単なる誤りではなく、読みの処理過程を反映する重要な指標として、誤発音・挿入・省略といった観点から分析されている⁽⁴⁾。

加えて、フォニックス指導に関する研究では、文字と音の対応関係の習得が初期読解能力の発達に寄与することが示されており、音韻認識やデコーディング能力は読解の基盤となるとされる⁽⁵⁾⁽⁶⁾。

一方、音声認識 (ASR) 技術を用いた研究では、音読速度や誤読の自動測定が可能となり、学習ログに基づく大規模分析が進展している⁽⁷⁾。特に教育現場における実践的研究では、Reading Progress を用いた分析が報告されており、音読速度および誤読行動の経時的変化が検討されている⁽⁷⁾。また、音素レベルの分析では、日本人英語学習者の誤りが特定の音素に集中する傾向が示されている⁽⁸⁾。

第二言語音声習得研究では、母語音韻体系が第二言語音の知覚および産出に影響を与えることが知られている。Flege (1995) の Speech Learning Model (SLM) は、第二言語音が母語音韻カテゴリーに同化されることで新たな音韻カテゴリーの形成が困難になることを示している⁽⁹⁾。また、Best & Tyler (2007) の Perceptual Assimilation Model (PAM) は、複数の非母語音が同一カテゴリーに同化される場合、弁別が困難になることを指摘している⁽¹⁰⁾。

この観点から、日本語母語話者は母音体系が比較的単純であり、弱化母音 (/ə/) や二重母音を持たないことか

ら、英語母音の習得に困難を示すとされる。このような音韻体系の差異は、特定の音素における誤読の集中として現れる可能性がある。

以上を踏まえると、音声認識 AI による音読データ分析と第二言語音声習得理論を統合し、音素単位で誤読傾向を検討することには重要な意義がある。しかしながら、既存研究では誤読を単語単位または誤読タイプ単位で扱うものが多く、音素レベルでの大規模分析は十分に行われていない。したがって、本研究は、「どの音が誤られやすいか」という音韻的傾向の解明に焦点を当て、音素単位で誤読傾向を分析する点に特徴がある。

2. 研究目的と仮説

本研究の目的は、音声認識 AI により取得された音読データを用いて、英語学習者の誤読傾向を音素単位で定量的に明らかにし、誤読率の算出および比較を通じて学習者が困難とする音素を可視化することである。加えて、誤読の分布がランダムではなく、音韻体系に依存した体系的傾向を有するかを検討することを目的とする。

本研究では以下の仮説を設定する。

- (1) 日本語母語話者は、母語に存在しない音素 (例: /ʌ/, /ə/ など) において誤読率が高くなる。これは、新たな音韻カテゴリーの形成が困難であるとする第二言語音声習得理論に基づく。
- (2) 弱化母音 (/ə/) など、語中で音価が変化する曖昧な音素は、知覚および産出の両面で不安定となるため、誤読率が高くなる。
- (3) 母語音韻体系において複数の音素が同一カテゴリーに同化される場合 (例: 英語母音が日本語の /a/ に近似される場合)、それらの音素において弁別困難が生じ、誤読率が高くなる。

これらの仮説は、Flege (1995) の Speech Learning Model および Best & Tyler (2007) の Perceptual Assimilation Model に基づくものであり、本研究では大規模音読データを用いて音素別誤読率を算出し、母音・子音の比較や音素間の分布を分析することで、これらの理論的予測との整合性を検討する。

3. 分析方法

3.1 対象者および語彙レベル

本研究の対象は、日本語を母語とする英語学習者 46 名であり、大学 1・2 年生を対象とした 2025 年度の通年必修英語科目の受講者である。当該科目において、週 1 回・

全30週にわたり音読課題を実施し、継続的に音読データを収集した。学習者の英語レベルはCEFRA2～B1相当である。

音読課題には、TOEIC 対策用の語彙を基にした単語フレーズリストを使用した。これらの語彙は短母音・長母音・子音結合など、多様な音素を含むよう設計されている。語彙レベルは概ね CEFRA1～B1 相当であり、学習者のレベルをやや下支えする難易度として設定されている。

3.2 使用データ

本研究では、Microsoft Teams および Learning Accelerator (Reading Progress) を用いた音読課題により収集されたフォニックス学習データ約 10 万件を対象とした。その中から 1000 件を無作為抽出し、分析に用いた。抽出に際しては、特定の音素や語彙に偏りが生じないよう留意し、主要な音素分布を含むことを確認した。

本データは音声認識 AI によりテキスト化されたものであり、誤認識が含まれる可能性がある。しかし、同一条件下で収集された大規模データであることから、個別誤差の影響は平均化されると考えられ、相対的な誤読傾向の分析には一定の妥当性を有すると判断される。

3.3 データ前処理

本研究では、Microsoft Teams の Reading Progress 機能により取得された音読データを用いた。Reading Progress では、誤発音 (mispronunciation)、挿入 (insertion)、削除 (omission)、および音読速度 (WPM: words per minute) といった指標が提供される。本研究ではこのうち誤発音に着目し、フォニックス単位で出力される評価結果を分析対象とした。

次に、フォニックス単位で出力された音読評価データを基に、各単語に含まれる音素情報を整理した。その上で、分析の一貫性を確保するため、フォニックス表記を国際音声記号 (IPA) に対応づけて正規化した (例: broad a → /ɑ:/, uh → /ʌ/, schwa → /ə/)。これにより、表記上異なる項目であっても、同一または類似する音素として比較できるようにした。

また、各音素について誤読回数および出現回数を集計した。ここで出現回数とは、分析対象データ内において当該音素が含まれていた総回数を指す。誤読回数とは、当該音素を含む単語において誤発音が検出された回数を指す。

以上の手順により、音素単位での誤読傾向を定量的に分析した。

3.4 誤読率の算出と可視化

誤読率は、音素ごとに集計した誤読回数を、同一音素の出現回数で除することにより算出した。ここで、誤読回数とは、正解語と音読結果を比較した際に、当該音素を含む項目で誤読が確認された回数を指す。出現回数とは、分析対象データ内において当該音素が出現した総回数を指す。誤読率は以下の式により算出した。

$$\text{誤読率} = \text{誤読回数} \div \text{出現回数}$$

算出した誤読率は、音素ごとに比較できるよう百分率に換算した。同様に、出現回数が極端に少ない音素は誤読率が過大に算出される可能性があるため、本研究では出現回数 30 以上の項目を中心に分析した。また、音素別誤

読率を棒グラフとして可視化し、学習者がどの音素において誤読しやすいかを視覚的に把握した。

4. 結果

分析の結果、誤読率は特定の音素に偏って分布していることが確認された (図 1)。

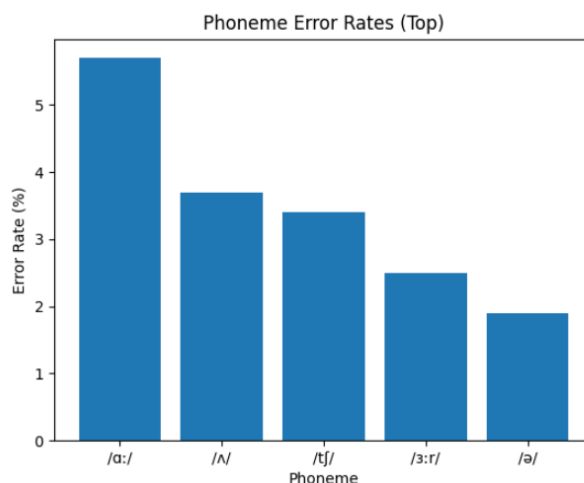


図 1 音素別誤読率 (上位)

図 1 は音素ごとの誤読率を示したものであり、誤読率の高い音素ほど棒の長さが大きくなるように表示している。

特に、/ɑ:/ が 5.7% と最も高く、次いで /ʌ/ (3.7%)、/tʃ/ (3.4%)、/ɜ:r/ (2.5%)、/ə/ (1.9%) の順であった。これらの音素は主に母音または母音を含む複合音であり、子音単独の音素と比較して誤読率が高い傾向が見られた。加えて、英語特有の母音体系に関わる音素に誤読が集中している点特徴的である。

その上で、母音と子音を比較した結果、母音の誤読率は 0.31%、子音は 0.10% であり、母音の誤読率は子音の約 3 倍であった (図 2)。

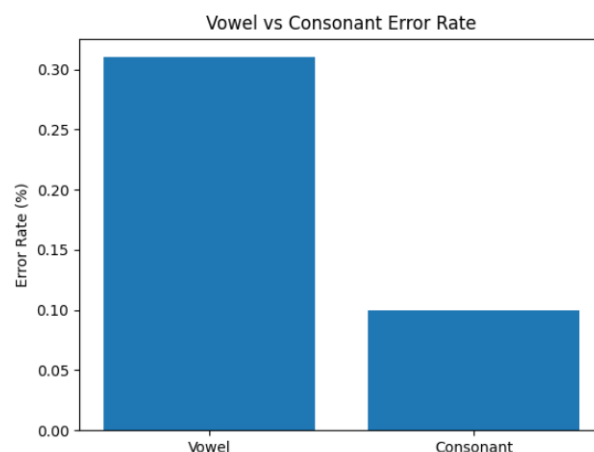


図 2 母音と子音の誤読率比較

図 2 は母音と子音の平均誤読率を比較したものであり、母音の方が一貫して高い値を示していることが分かる。母音の誤読率は子音の約 3 倍に達しており、音読における誤りが母音に顕著に集中していることが示唆される。

母音を短母音・長母音・二重母音に分類して誤読率を比

較した結果、長母音において最も高い誤読率が確認され、次いで短母音、二重母音の順となった（図3）。

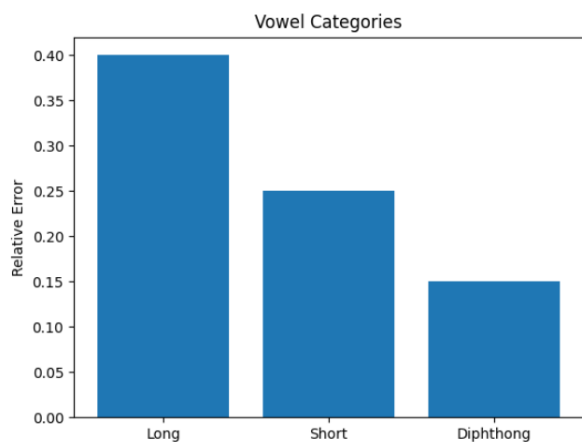


図3 母音カテゴリ別誤読率

特に /ɑ:/ を含む長母音において顕著な誤読が見られた。このことは、日本語において音長の対立が限定的であることに起因し、長母音の知覚および産出が不安定となるためと解釈できる。一方で、二重母音については構成要素が比較的明確であるため、誤読率が相対的に低く抑えられている可能性がある。

次に、子音を破裂音・摩擦音・破擦音に分類し、誤読率を比較した結果、破擦音および摩擦音において相対的に誤読率が高く、破裂音では低い傾向が確認された（図4）。

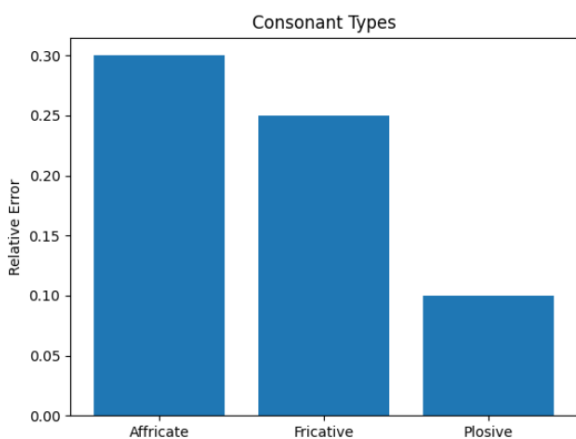


図4 子音タイプ別誤読率

特に /tʃ/ に代表される破擦音では、複数の調音動作を伴うことから、調音の複雑性が誤読の要因となっている可能性がある。また、摩擦音についても持続的な気流制御を必要とするため、音声産出の安定性が低下しやすいと考えられる。

併せて、音素の出現位置に着目し、語頭・語中・語末の各位置における誤読率を比較した結果、語中において最も誤読率が高く、次いで語末、語頭の順となった（図5）。

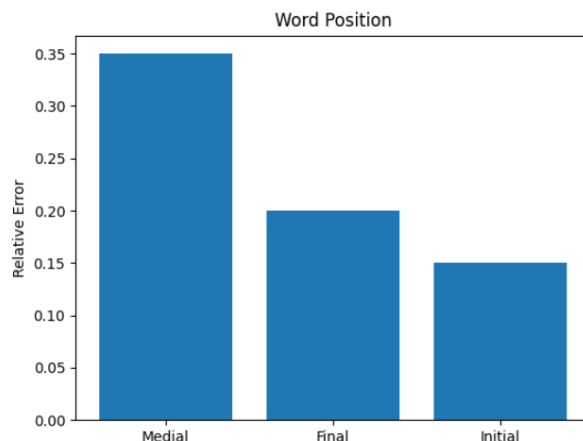


図5 語中位置別誤読率

語中で誤読率が高い要因としては、弱化母音 /ə/ の出現頻度が高いことや、強勢の影響により音価が変化しやすいことが挙げられる。一方で語頭は比較的明瞭に発音されるため、誤読率が低く抑えられていると示唆される。

以上の結果から、誤読は特定の音素に集中するだけでなく、音素の種類（母音・子音）、音韻的特徴（音長・調音様式）、および語中位置といった複数の要因が相互に影響して生じていることが示唆された。

5. 考察および結論

本研究では、音声認識 AI を用いた大規模音読データの分析により、英語学習者の誤読が特定の音素に集中する傾向が確認された。特に /ɑ:/, /ʌ/, /ɜ:/, /ə/ といった母音において誤読率が高く、誤読はランダムではなく音韻的特性に依存して分布することが明らかとなった。

これらの結果は、第二言語音声習得理論により説明可能である。Flege (1995) の Speech Learning Model (SLM) によれば、第二言語音が母語の音韻カテゴリーに同化される場合、新たな音韻カテゴリーの形成が阻害される。本研究で誤読率が高かった /ʌ/ や /ə/ は、日本語母語話者にとって対応する音が不明確であり、「新規音」または「曖昧に同化される音」として処理される可能性が高い。その結果、音韻カテゴリーが安定せず、誤読として現れると示唆される。また、/ɜ:/ のような母音と /r/ の連続を含む音素は、日本語の音韻体系に類似構造が存在しないため、単一単位としての知覚が困難であると示唆される。このような複合音は部分的に既存カテゴリーに同化され、不完全な音声表象となることで誤読を引き起こす可能性がある。

さらに、Best & Tyler (2007) の Perceptual Assimilation Model (PAM) によれば、複数の非母語音が単一の母語カテゴリーに同化される場合、弁別が困難になる。本研究で観察された /ɑ:/, /ʌ/, /ə/ は日本語の /a/ に近似的に知覚される可能性があり、複数の母音が同一カテゴリーに統合されることで弁別困難が生じ、誤読の増加につながったと解釈できる。

以上より、本研究の結果は、SLM および PAM が示す「非母語音の知覚・産出困難」が実際の音読における誤読として現れることを示しており、理論と実証データの整

合性を支持するものである。

教育的観点からは、フォニックス指導において単なる文字一音対応にとどまらず、音韻カテゴリーの再編成を促す指導が重要である。特に /ə/ や /ʌ/ のような弱化母音については、強勢や語中位置との関連を明示する指導が有効と考えられる。また、/æ/ - /ʌ/ - /ɑ:/ のようなミニマルペア練習は、単一カテゴリー同化の回避に寄与する可能性がある。加えて注目すべきは、音声認識 AI による即時フィードバックは、学習者の誤読傾向の自覚と修正を促す点で有効である。

本研究の意義は、大規模音読データに基づき誤読の体系的傾向を音素単位で明らかにした点にある。特に、誤読が母音に偏在することを定量的に示した点は、音韻指導の重点を明確化する上で有用である。また、第二言語音声習得理論と学習データを接続した点にも意義がある。

一方で、本研究には限界もある。第一に、音声認識 AI による誤発音検出には誤差が含まれる可能性がある。第二に、本分析はテキストベースの評価結果に依存しており、知覚過程や調音運動を直接測定したものではない。第三に、対象者が特定の教育環境に限定されているため、一般化には慎重な検討が必要である。

今後は、音声波形や知覚実験を組み合わせた分析により誤読の発生メカニズムを詳細に検討する必要がある。また、縦断的データを用いた分析により、音韻カテゴリーの形成過程や指導効果の検証が求められる。

以上より、本研究は音声認識 AI と第二言語音声習得理論を統合することで、英語学習者の音韻的困難を実証的に明らかにし、音読指導に対する新たな分析枠組みを提示した。

参考文献

- (1) Vančová, H.: “AI and AI-powered tools for pronunciation training”, *Journal of Language and Cultural Education*, 11(3), pp.12-30 (2023) .
- (2) Alonzo, D., Abril, J. M. V., & Oo, C. Z.: “The use of artificial intelligence in English language assessment: Empirical evidence and future directions”, *Cristal: Critical Studies in Teaching and Learning*, 13(2) (2025) .
- (3) Kuhn, M. R., & Stahl, S. A.: “Fluency: A review of developmental and remedial practices”, *Journal of Educational Psychology*, 95(1), pp.3-21 (2003) .
- (4) Rasinski, T. V.: “Speed does matter in reading”, *The Reading Teacher*, 54(2), pp.146-151 (2000) .
- (5) Ehri, L. C.: “Phases of development in learning to read words”, *Journal of Research in Reading*, 18(2), pp.116-125 (1995) .
- (6) Adams, M. J.: *Beginning to Read: Thinking and Learning about Print*, MIT Press (1990) .
- (7) 長岡 穂: “AI 音読評価ツール Reading Progress による音読速度と誤読の経時的変化”, *CIEC 春季 PC カンファレンス論文集 Vol.17*, pp.16-23 (2026) .
- (8) 長岡 穂: “日本人英語学習者における音素別発音誤りの体系的分析”, *JACET 言語エキスポ 2026 予稿集* pp.129-130 (2026) .
- (9) Flege, J. E.: “Second language speech learning: Theory, findings, and problems”, in Strange, W. (Ed.), *Speech Perception and Linguistic Experience*, York Press, pp.233-277 (1995) .
- (10) Best, C. T., & Tyler, M. D.: “Nonnative and second-language speech perception: Commonalities and complementarities”, in Munro, M. J. & Bohn, O.-S. (Eds.), *Second Language Speech Learning*, John Benjamins, pp.13-34 (2007) .