

口腔内 3D ダイナミックモデルを用いた発音学習法の提案

白鳥 優莉[†] 長谷川 大[†] 佐久田 博司[†]

青山学院大学大学院 理工学研究科 理工学専攻 知能情報コース[†]

1 序論

1.1 背景

近年、生活や仕事のグローバル化により日本でも国際共通語としての英語を必要とする機会が増えつつあるが、従来の日本の中等高等学校における英語教育は、コミュニケーション能力の獲得においては評価が低い。これは、教科書を基にした「読み」「書き」の能力に比べ、「聞く」「話す」の教育機会の差によって能力差が生じやすいことが原因である。これら4つの能力をまとめた英語力の視点から見ると、母国語の長い歴史を持つ国であるフランスやロシア、日本などは一般に能力が低いという報告がある[1]。しかし、日本語を母国語とする人（以下、「日本語話者」と略）の英語力は決して低いわけではなく、EF EPI 英語能力指数は2015年度では53.57である。これは、世界平均と比べて見ると平均的な数値である。それにもかかわらず、日本語話者の英語理解能力の低さが議論されるのは、上で述べた学校教育では学習する機会の少ないリアルタイムのコミュニケーションである「聞く」「話す」の能力不足に原因があると思われる。

近年の研究によれば、乳幼児期の音声認識能力の獲得について、ビデオやテープによる訓練は効果がなく、Face to Face（以下、「FtF」と略）による訓練によって、複数言語の音声認識が可能となることが示されている[2]。これは、実物あるいは立体として認識される対話者の存在の効果が大きいことを示唆している。また、FtF 訓練では、話者の発話の音声を出し、さらに相手の発話を聴覚での聞き取りによる入力として、脳内で繰り返しや試行錯誤を何度も長期的に話者に行わせる。そのため、話者は自身の発音と正確な聞き取りの能力が、同時進行で訓練されていると考えられる。

従って、「聞く」「話す」の能力不足を補うためには、FtF 訓練の機会を増やすことが必要である。しかし、FtF 訓練は、教師に当たる人や、環境を長期的にわたり用意する必

要があり、費用対効果やコスト面そのものから実現が不利であるといえる。そのため、教員の代替手段によって学習者が自分自身の発音と聞き取り能力を向上させることが可能な環境が必要となる。

現在までは、正しい発音を訓練し、改善させる方法として FtF 訓練でも行っているように聞いた発音を真似する方法が主流である。しかし、聴いた音と発音した音の差分を身体知としての発音にフィードバックすることは、学習者自身にとっては難しい。正しく発音できない原因として、音を聴いてもその音の再現方法が分からないことが挙げられる[3]。つまり、日本語には英語側の発音の一部に該当する口腔内および舌の動きが存在しないため、日本人にとって英語を発音する時それらの動きが理解できない。したがって、聴覚だけでは理解できない口腔内を視覚を使って学習者に見せることで、発話のメカニズムを認識させることが必要だと考えられる。

また、聞き取り能力を鍛える方法として熊井ら[4]が聴いた発音をすぐに復唱するシャドーイングを繰り返した結果、リスニング力が上がったという報告がある。そのため、発音を繰り返し行うことで聞き取りが良くなることが期待される。よって、今回は発音改善を主な目標とする。

同時に、以上の手法は、生来の聴覚障がい者の発話トレーニングツールとしても提案することができる。

1.2 関連研究

発音評価を目的とした研究は多くあるが、学習方法を提案して行ったものはあまりない。スウェーデンの Olov ら[5]はバーチャル教師 (ARTUR) を開発し、人間の顔と口腔内を 3D モデルを用いることで学習者と FtF で学習を行った。さらに 3D モデルの断面図で発音時の口腔内の動きを学習者に示した。

この研究では学習者の発音が良くなるという有意傾向はみられた。しかし、このシステムで学習者に提示される画像はモニター上に表示される口腔内のみで、舌の動き等の立体情報は表現できなかった。そこで本研究では、立体視を用いることで発音時の口腔内の提示方法を提案する。立体視は表示オブジェクトの奥行きや空間内での配置と時間変化といった多くの情報量を含むことができるため、発音改善を期待できる。

e-Learning of English Pronunciation by dynamic feedback models of oral cavity

[†] Yuri Shiratori (c5615138@aoyama.jp)

[†] Intelligence and Information Course, Graduate School of Science and Engineering, Aoyama Gakuin University

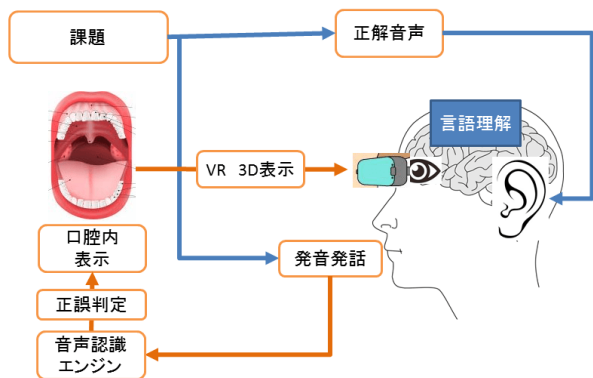


図1 学習システム概念図

2 目的

本研究では、学習者が発音と口腔内の舌などの状態の時間変化を立体視することにより、効率的に発音を改善させるための学習システムを開発し、その有効性を確認することを目的とする。

3 提案手法

本研究で提案する学習システムにおいて学習者は発音音声を聴き、発音し、口腔内の動きを見る。(図1)これにより聴覚、発声、視覚によるフィードバックが発音を効果的に矯正することが期待できる。発音時の口腔内表示は立体視で学習者へ表示する。立体視はHMD(Head Mount Display)を標準オプションとするAndroid端末(Gear VR)によるVirtual Reality(VR)を用いる。

4 実験

4.1 予備実験

図1の機能要素として、(1)音声認識、(2)音声合成、(3)口腔内VR表示(口腔内3Dモデルのアニメーション)の適用性の検証のために、英語話者(米国籍日本在住英語教員)が本システムを試用し、健全性を確認した。

音声認識エンジンはUnityのアセットであるAndroid Speech TTSを用いた。この音声認識エンジンの精度としてrとlだけが異なる単語12組(計24単語)の認識率を英語話者で確認したところ89%であったが、実験で用いる英単語10単語のみの認識率は100%であった。

音声合成はAmazon Pollyを用いて作成した音声を使う。その音声の適用性と口腔内3Dモデルのアニメーションの適用性は英語話者によって確認した。

4.2 本実験

日本人被験者に提案した学習方法で発音学習させ、学習効果を計測する。使用する学習システムの健全性は予備実験で検証した通りである。

実験の学習を行う前に、被験者には事前テストとして発音復唱テストと単語聞き分けテストを行う。事前テスト

後に10分間提案方法の学習法を行う。発音学習後、事後テストとして事前テストでも行った発音復唱テストと単語聞き分けテストを行う。

事前と事後のテストの正誤率や、学習中に単語の発音を繰り返した回数の変化などを分析、比較することで本研究の提案手法の効果を図る。

5 結果

大学生男女8名を被験者として実験を行った。

発音復唱テストの結果、学習後に発音が正しく復唱できた単語数が多かった:4名、変わらなかった:3名、少なくなった:1名であった。単語聞き分けテストの結果、学習前より正解率が上がった:3名、変わらなかった:2名、下がった:3名であった。また、学習過程はそれぞれの単語について正解に至るまでの繰り返し回数にばらつきがあり、被験者ごとの改善プロセスでの個性が現れた。

6 結論

本研究では発音時の口腔内を立体視で見せることで、正しい発音方法を学習者に理解させ改善させるための学習システムを開発し、実験を行い、提案した学習方法の有効性を確認した。その結果、口腔内の動きを見せることで正しい発音を行える傾向がみえたが、統計的に有意性があるとはまでは結論づけることはできなかった。しかし、従来の学習方法と比べ可用性があると思われるため、今後検証実験を行う必要がある。

6.1 謝辞

本研究はJSPS科研費JP26540185の助成を受けたものである。

参考文献

- [1] EF education first. Efi epi (英語能力指数) - 英語力から見る国別総合ランキング. <http://www.efjapan.co.jp/efi/>, 2015.
- [2] PK Kuhl. Linguistic experience alters phonetic perception in infants by 6 months of age. *Foundations of Pediatric Audiology*, Vol. 255, No. 5044, pp. 606-608, 1992.
- [3] 佐伯治中島誠. 日本語・英語の比較研究に基づく英語教育法に関する研究. *教育心理学研究*, Vol. 15, No. 2, pp. 103-118, 6 1967.
- [4] 赤塚麻子熊井信弘. ウェブを利用したシャドーイング練習が学習者のリスニング能力に与える効果について. *言語文化社会*, No. 10, pp. 41-55, 2012.
- [5] Olov Engwall. Analysis of and feedback on phonetic features in pronunciation training with a virtual teacher. *Computer Assisted Language Learning*, Vol. 25, No. 1, pp. 37-64, 2012.