

調音特徴の円唇・緊張・R音を組み込んだ英語母音発音マップ

風張 航[†], 入部百合絵[‡], 桂田浩一[†], 新田恒雄[‡]

豊橋技術科学大学 工学部 知識情報工学課程[†]

早稲田大学 グリーン・コンピューティング・システム研究機構[‡]

1. はじめに

近年、外国語学習者を対象とする CALL 教材の開発が盛んに行われている [1]. 発音学習では、舌や口唇など調音器官の動作 (以下, 調音動作) を示す必要があるが, 多くの CALL 教材では, 具体的にどの調音動作がどのように誤ったか, また, どのように調音動作を矯正すれば良いのか分かりづらい. 正しい発音を身につけるためには, 少人数の会話教室で教師が実施しているように, 学習者の調音誤りをリアルタイムに, 的確に指示して矯正することが重要である.

そこで, 我々は, 「教師と学習者の調音動作の違いを簡単に理解でき, 正しい調音動作への矯正方法を直感的に読み取ることのできる」英語母音発音マップシステム (以下, 発音マップ) を開発している. 発音マップでは, 横軸を舌の盛り上がり位置, 縦軸を口の開け方とする平面上に, 学習者の発音をリアルタイムに表示する [2]. 従来の発音マップでは, 上記の 2 軸に関連する 10 種類 (前舌, 中舌, 後舌, 広, 半広, 半狭, 狭, 母音, 有声音) の調音の特徴を用いていたが, 言語によっては円唇, 緊張, R 音の調音も含まれる. 例えば, 英語母音には円唇, 緊張, R 音の調音が含まれるが, 日本語母音には存在しない. そのため, 日本人が英語発音を学習する際には, これらの調音を訓練する必要がある. そこで, 本研究では, 学習者の音声から円唇, 緊張, R 音を含む 13 種類の調音の特徴を抽出し, 発音マップ上にリアルタイムに表示することを提案する.

2. 円唇, 緊張, R音を組み込んだ発音マップ

2.1 発音マップの概要

発音マップには, 発声時の調音器官の位置 (半広, 前舌など) や様式 (有聲, 母音など) を示す, IPA(International Phonetic Association) に基づく調音特徴を用いている. 我々が開発している音声認識エンジンは, 学習者の音声から多層ニューラルネットワーク (Multi-Layer Neural Network: MLN) を用いて調音特徴を抽出する. 発音マップでは, 音声認識エンジンにより抽出した調音特徴を, IPA 母音チャート図を模した発音マップ上へ座標変換し,

English Vowels Pronunciation Map Indicating Round, Tense and Rhoticity of Articulatory Features.

[†] Department of Knowledge-based Information Engineering, Toyohashi Univ. of Tech.

[‡] Green Computing Systems Research Organization, Waseda Univ.

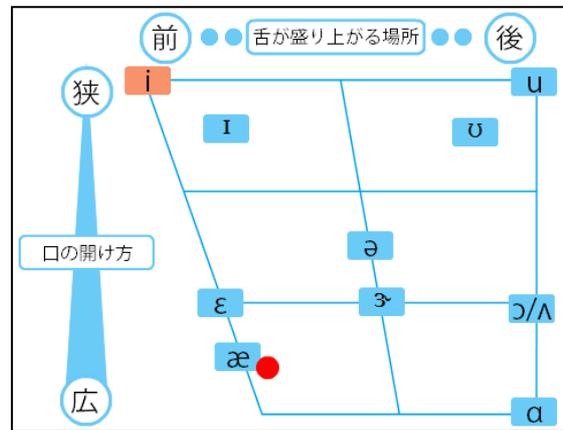


図1 発音マップ

赤い光点 (ポインタ) によってリアルタイムに調音動作を表示する. 具体的には, 調音特徴抽出器より得た調音特徴の「前舌音」「中舌音」「後舌音」を X 座標 (舌の位置), 「狭母音」「半狭母音」「半広母音」「広母音」を Y 座標 (口のの高さ) に変換する. 発音マップの例を図 1 に示す.

2.2 円唇, 緊張, R音の表示

本研究では, 舌の盛り上がり位置と口の開け方の 2 軸に加えて, 円唇, 緊張, R 音を発音マップに表示する. ここで, 円唇とは唇を丸めた状態, 緊張 (張唇) とは唇が左右に広げられる状態を示す. R 音は母音を調音する際に舌先を反らせたり, 舌を盛り上げたりすることで R のような音色を伴う. 本研究では, 円唇, 緊張, R 音を示すための 2 種類のユーザインタフェース (以下, UI) を提案する.

1 つ目は, 図 2 のようにマップに緊張を表す第 3 軸を追加する表示方法 (手法 A) で, 2 つ目は図 3 のように緊張のレベルメータを使用する表示方法 (手法 B) である. 先述の音素 /i/ の場合, 手法 A はマップの左角に第 3 軸を加えて表示する. つまり, 学習者の調音が, 口が狭めで, 舌が前寄りであっても, 緊張がなければ第 3 軸上の始点にポインタが表示され, 音素 /i/ には達していないことが分かる. また, 第 3 軸の調音を教示するため (この例では唇が左右に広がる緊張), 調音を示す吹き出しが表示される. 一方, 手法 B では, マップの左側に緊張の度合いを示したレベルメータが表示されている. このレベルメータは十分に緊張が表れていないときに表示される. 手法 A と同様に, レベルメータが示す調音の説明図が吹き出し内に表示される.

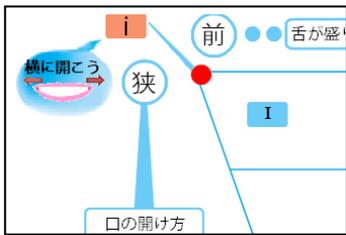


図2 第3軸拡張(手法A)



図3 レベルメータ(手法B)

3. 評価実験

3.1 ユーザインタフェース評価

提案する2つの表示方法の見やすさ分かりやすさを比較するため、各手法に対するUI評価を実施した。チェックリスト評価法を用いて[3]、明瞭性、一貫性、合致性、情報のフィードバックの4基準に対して評価を行った。各4基準について、基準への合致度の評価と満足度の評価を行い、計26設問を双方の手法に対して評価した。被験者は9名である。

4基準の結果を平均した総合評価は、手法B(レベルメータ)が、手法A(第3軸)を2.1ポイント上回り、87.4ポイント(最大値100ポイント)を得た。

各基準の平均評価点と、各基準の満足度の評価を示す平均満足点を図4及び図5に示す。結果は、明瞭性において、平均評価点と平均満足点の逆転が見られた。明瞭性は、画面の表示内容が明瞭で整理されており、読みやすいか、という基準である。手法Bの明瞭性は良い評価を得られたが、満足には至っていない。被験者からは、手法Bはポインタと、ポインタとは少し離れたレベルメータを同時に見ながら調音を調節しなければならないため、学習が難し

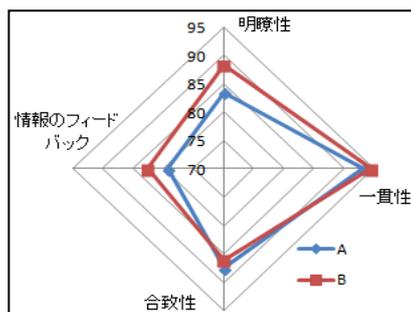


図4 平均評価点

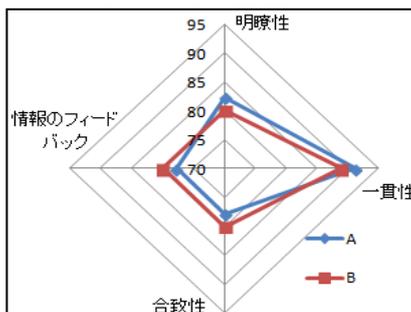


図5 平均満足点

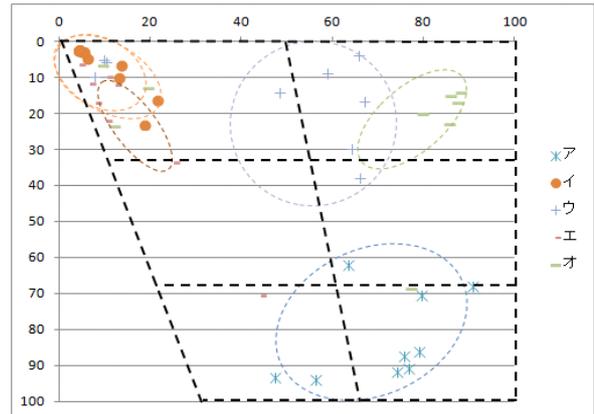


図6 日本語母音のプロット

いという指摘が多くあった。これを軽減させるため、ポインタとレベルメータの片方を見ながら発音した後、もう片方の表示で自らの発音がどのような発音であったかを後から確認できるように、表示を僅かに遅延させる方法が提案された。また、ポインタやレベルメータの更新速度を遅く、滑らかにすることで、学習者の負担を減らすことも考えられる。

3.2 プロット精度評価

発音マップのプロット精度について評価を行った。日本語話者9名が、日本語母音5音素を発話した際のプロット状況を図6に示す。実際にプロットされた座標と正解座標とのユークリッド距離を算出したところ、/ア/や/イ/は距離差が約10であったが、/オ/では35の距離差が生じた。また/ウ/, /エ/, /オ/に関する各ユークリッド距離の標準偏差を求めたところ、約10~30の値であり、話者によるばらつきが生じている。従って、/ウ/, /エ/, /オ/のプロット精度を向上させる必要がある。

4. まとめ

学習者の調音をリアルタイムにプロットする発音マップ上に円唇、緊張、R音を表示する方法を提案した。表示方法として、(A)レベルメータによる方法、(B)第3軸の追加、の2種のUIを開発した。

評価実験の結果、明瞭性が不十分であることが明らかとなった。今後はポインタ及びレベルメータの時間的平滑性及び遅延表示を取り入れることで、更なる進歩を図りたい。また、プロット精度の向上のため、MLN構成の変更や、学習コーパスの変更などが今後の展開として考えられる。

参考文献

- [1] Maxine Eskenazi, "An overview of spoken language technology for education", *Speech Communication*, 51(10), pp. 832-844 (2009).
- [2] 森拓郎 他, "調音特徴抽出に基づく IPA チャートへの英語発音リアルタイム表示" *信学技報 SP2011-169*, pp. 77-82, 2012.
- [3] S. ラブデン, G. ジョンソン: ユーザインタフェースの実践的評価法, 海文堂, 1993