

調音特徴抽出に基づく発音マップへの英語発音リアルタイム表示

Real-time Visualization of English Pronunciation Based on Articulatory Feature Extraction

入部 百合絵^{*1}, 森 拓郎^{*1}, 桂田 浩一^{*1}, 新田 恒雄^{*1}
 Yurie IRIBE^{*1}, Takurou MORI^{*1}, Kouichi KATSURADA^{*1} and Tsuneo NITTA^{*1}
^{*1}豊橋技術科学大学 大学院工学研究科 情報・知能工学専攻
^{*1} Graduate School of Technology, Toyohashi University of Technology
 Email: iribe@imc.tut.ac.jp

あらまし：近年、外国語を学ぶ日本人学生を対象に CALL 教材の開発が盛んに行われているが、学習者の発音動作の誤りを正確に指摘できる教材はまだ開発されていない。我々は、学習者が教師との発音の違いを分かり易く理解でき、さらに正しい発音動作への矯正方法を直感的に読み取ることのできる、英語発音訓練システムの開発を進めている。本報告では、母音発音に焦点を当て、学習者音声の発音動作を IPA 母音図上に表示することのできる、英語発音マップシステムとその評価について報告する。提案するシステムは、学習者の音声から発音動作の特徴量（調音特徴）を抽出し、二次元平面上へ座標変換することにより、学習者の英語発音をリアルタイムに発音マップ上へプロットする。英語母語話者に対する評価実験を行った結果、評価対象の全ての英語母音について、良好なプロット精度が得られた。
キーワード：発音学習、調音特徴、IPA チャート

1. はじめに

近年、大学や高校などの教育機関では外国語発音の自学自習用に様々な CALL 教材が導入されている。これまでに学習者の発音と正しい発音の違いを視覚的に評価する機能を備えた CALL 教材が開発されている⁽¹⁾が、音声波形やフォルマント周波数⁽²⁾による比較提示のため、音声学の知識を持つ音声学でない限り、舌の位置や口唇の高低といった調音器官の発話時の動作イメージ（以下、調音動作と呼ぶ）の違いを理解することが難しかった。特に、母語に存在しない調音動作を獲得することに多大な時間を要する。一方、教育機関や語学スクールでは教師が調音動作を学習者にうまく伝えながら、正確な発音ができるように指導している。このように、効果的な発音指導を行うためには、発話の際の口唇の開き具合や舌の位置を学習者に具体的に提示する必要がある。本研究では、舌の位置と口唇の高低の軸から構成される IPA 母音チャート（国際音声記号（International Phonetic Alphabet: IPA）による図表）⁽³⁾⁽⁴⁾を模した母

音発音マップ上に学習者の調音動作をマップし、正しい母音の調音位置と比較することで、調音の違いを容易に理解することのできる発音矯正システムを開発する。発音マップの画面例を図 1 に示す。加えて、これらをリアルタイムに表示することで、学習者がチャート図より自分の口唇や舌の位置を確認しながら正しい位置へと近づけていく発音矯正を可能にする。本研究では、学習者の音声から調音動作の特徴量を示す調音特徴を直接抽出することで、発音マップへリアルタイムに表示することを実現する。

2. 調音特徴に基づく発音訓練システム

我々は以前より調音特徴に基づく音声認識技術を用いた英語発音訓練ソフトウェアを開発している。本報告では、上記ソフトウェアにおいて学習者の音素誤りを指摘後にその発音を正しい発音へ導くため、発音の基本である調音動作を効果的に矯正するシステムを提案する。そのため、調音特徴を用いて IPA 母音チャート上に表示を行う機能（以下、英語母音発音マップと呼ぶ）を実装した。以下に調音特徴の抽出方法と IPA 母音チャートへの表示について述べる。

2.1 調音特徴抽出手法

調音特徴（Articulatory Feature; AF）は、単音分類に用いられる調音様式（発声に要する舌、口唇、口蓋などの調音器官の動作・様式、例：摩擦音、破裂音など）と調音位置（舌の位置や口唇の高低）の諸属性を指す特徴量である。従って、調音特徴が得られれば学習者の調音位置を推定することができ、IPA 母音チャートへのプロットが可能となる。今回は IPA から英語に関する部分を抽出して調音特徴を定

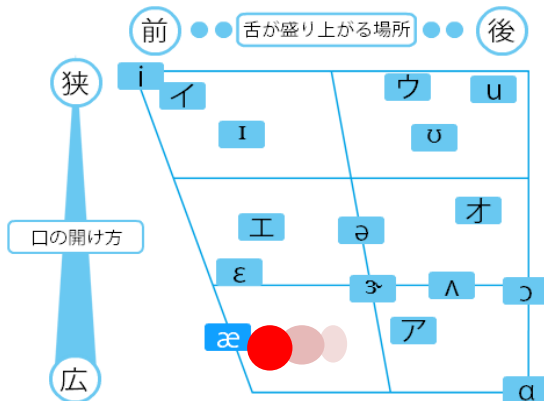


図 1 母音発音マップの画面例

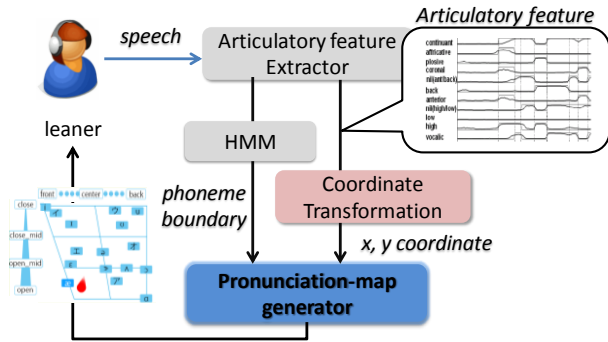


図2 発音マップシステムの基本構成

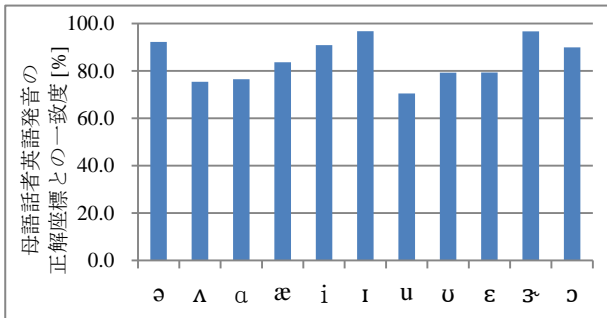


図3 英語母語話者の音声に対するプロット精度

義した。調音特徴は、音声スペクトル系列の時間微分と周波数微分から求めた局所特徴 (Local Feature; LF) を多層ニューラルネットワーク (Multi-Layer Neural network; MLN) に入力して得る。MLNはラベル付き音声データを用いて、時刻 $t-3, t, t+3$ の3フレームを入力し、調音特徴を教師信号として学習する。

2.2 発音マップへのリアルタイム表示

発音マップはIPA母音チャート[4]を模した梯形図に発音記号が配置され、口唇の開き具合を示すスケール(縦軸)、舌の盛り上がる位置を示すスケール(横軸)、そしてユーザの発音位置を示す赤い光点からなる。光点は学習者の発音の調音位置に対応した座標点にプロットされるため、学習者は発音しながら自分の調音位置を確認することが可能である。光点が目標とする発音記号に近いほど正しく発音できていることを示しているため、学習者は口唇の開き具合と舌の盛り上がる位置のスケールを参考に、調音を徐々に修正しながら漸近的に調音位置を矯正することができる。図2に示す発音マップシステムは学習者の発声を検知すると調音特徴抽出器により10ms毎に48次元の調音特徴を抽出する。抽出された調音特徴の母音に関する特徴列が座標変換器により2次元平面上のX,Y座標に変換され、その座標に光点が移動する。特徴列からX座標への変換は式(1)、Y座標への変換は式(2)によって行われる。

$$X = D_{width} * (\alpha + \Delta_x) / 4 \quad (1)$$

Δ_x は調音特徴列の口唇の開き具合を示す「前舌」「中舌」「後舌」の特徴量の差分を示す。 α はマップを水平方向に4分割した際の1セルの幅を示す。これ

らは各特徴量の大小関係ごとに定義される。 D_{width} は発音マップの水平方向の長さである。

$$Y = D_{height} * (\beta + \Delta_y) / 6 \quad (2)$$

Δ_y は音特徴列中の舌の盛り上がる位置を示す「狭」「半狭」「半広」「広」の特徴量の差分を示す。Bはマップを垂直方向に6分割した際の1セルの高さを示している。 D_{height} は発音マップの垂直方向の長さである。発音マップは台形状であるため座標をプロットする際は、適宜台形のスケールに合わせて座標を変換し、マップ上へプロットする。

3. 評価実験

提案する英語母音発音マップは学習者の発音を2次元平面であるIPA母音チャート上にプロットし、マップ上に示す発音記号との相対的な位置から発音動作の違いを視覚的に教示するものである。従って、ネイティブ英語発音に近い発音がなされた場合は、発音記号と同じ座標上にプロットされることが理想である。そこで、英語発音マップの性能を評価するため、今回は英語母語話者の発話から抽出した音声からマップ上へ変換した座標値と各母音の正解座標を比較した。ここで、正解座標とは、国際音声学会が定義したIPAチャート上の各母音の座標値である。MLNの学習には、TIMITの2,600文(男性話者325名)を使用し、評価データはTIMITの896文(男性話者112名)である。プロットの正確さを評価する尺度には、発話から得られた座標と正解座標との正解率を用いた。正解率は発話から抽出された座標と正解座標との距離に反比例し、話者の発音の座標と正解座標の距離が0の場合は100%となる。実験結果を図3に示す。全母音に対するプロット精度の平均は85.5%であった。特に、[ə], [i], [I], [ɜ], [ɔ]の正解率は90%以上である。本実験から、英語母語話者の発音に対して高い精度で発音マップ上にプロットできていることが確認された。

4. まとめ

英語発音訓練ソフトにおける調音動作の教示および矯正機能として、調音特徴に基づき学習者の調音動作をIPA母音図上にリアルタイムにプロットする英語母音発音マップを開発し、評価実験により全ての母音に対して70%以上のプロット精度を得られることが確認できた。今後は更なるプロット精度の向上を図るとともに、子音の調音動作を矯正可能な発音マップの開発を進めていく。

参考文献

- (1) Sonic Print - 株式会社アルカディア
- (2) 菊地歌子, 島崎のぞみ, 境一三: 日本人フランス語学習者のための発音学習教材, 電子情報通信学会技術研究報告 SP, Vol.110(452), pp.25-29(2010)
- (3) 佐伯拓郎他: 3D フォルマント母音図における発声母音のリアルタイム可視化, 電子情報通信学会総合大会講演論文集 2009年_情報・システム(1), pp.169, 2009
- (4) IPA chart: <http://www.arts.gla.ac.uk/ipa/vowels.html>