個人を単位とした世界英語の発音分類地図の構築とその教育利用に関する検討

峯松信明（東大大学院）・牧野武彦（中央大）・原田哲男（早稲田大）
山内豊（東京国際大）・関口幸代（明治学院大）

1. はじめに

英語は母語が異なる者同士が言語コミュニケーションを図る際の共通語である。この場合，話者の母語や言語背景が発音に影響し，その話者特有の訛りを呈することになる。World Englishes（世界英語，Kachru, 2006）はこのような状況を表現した用語である。米語，英語も米国訛り，英国訛りであり，誰の発音であっても訛った発音と見なされ，正しい訛り，誤った訛り，というのは存在しない。発音教育においては，通常，特定の訛った発音をモデルとして採択し，それに近づくよう教育・指導されるが，世界英語の立場に立てば，（十分に了解性が高ければ）その必要は必ずしも無い。

「みんな違って，それでいい」と考えた場合，学習者には「学習者の発音が世界英語の中でどう位置づけられるのか」「他の多様な発音とどう違うのか」を呈示すべきであると考える。訛りを分類する場合，国や地方を単位として分類することが多いが，究極的に考えれば，訛りの最小単位は個人である。本研究は，個人を単位とした世界英語の発音分類地図構築の技術的可能性（Shen, 2013）について述べ，また，構築された地図の教育利用の可能性について検討を行なったのでその様子を報告する。

２．研究方法

N 個の要素から成る集合を分類する場合，二要素間の距離を全て計測し，N×N の距離行列を求めればよい。この距離行列に樹形図などの手法を適用すれば，分類の様子を視覚化できる。世界英語発音地図の場合，二話者間の発音距離を求める必要がある。

Speech Accent Archive（以下 SAA と略す。Weinberger, 2013）は，世界中の英語利用者に，ある特定のパラグラフを読ませ，その音声データと IPA 書き起こし，更には話者の言語背景情報を収集，提供しているコーパスである。現在，1700 名以上の音声データが世界中から集められている。本研究ではまず，1)この SAA の IPA 書き起こしを用いて，任意の二話者間の発音距離を定義する。この IPA 距離を訓練データとして用い，最終的には，2)IPA 書き起こしを使わずに，任意の二話者間の発音距離を推定する技術を構築する。そしてこの技術の教育的価値について検討する。

IPA 距離計算は，二つの IPA 書き起こし，即ち二つの IPA シンボル配列間の距離計算を Dynamic Time Warping (DTW)法を用いて行なった。これは図 1 に示すように，二つの配列間で最適なシンボル対応を求める操作である。異なるシンボルが対応した場合当然「歪み」が生じるため，累積歪みを最小とするシンボル対応を求める。この場合，シンボル間距離を事前に定義する必要がある。本研究では，SAA に出現する約 150 種類の IPA シンボルを音声学者（第二著者）に発声させ，二シンボル間の距離を音響距離（ケプストラム分布のバタチャリヤ距離）として求め，これを用いた。

IPA 書き起こしを使わずに，二話者間の IPA 距離を推定する手法としては，二話者の発音差異に関連する各種の音響特徴を音響分析により求め，これらを説明変数として IPA 距離を回帰により推定する手法を検討した。二話者の音声のスペクトル特徴を

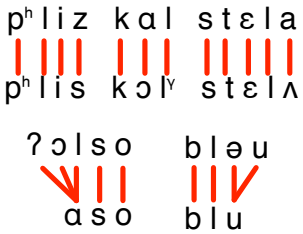

図 1：二つの IPA シンボル配列の対応例

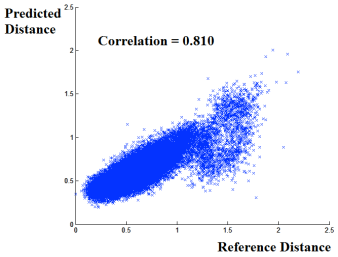

図 2：IPA 距離と推定距離の相関

そのまま説明変数として使うのは得策ではない。スペクトル差異は性別や年齢の差異まで表現するからである。ここでは，話者や年齢の違いを捨象した上で発音の様子を記述する発音構造分析（峯松, 2011）を行ない，分析の結果得られる発音構造パラメータを採択した。即ち，発音構造パラメータの二話者間差異を求め，これを説明変数として IPA 距離を回帰により推定した。回帰モデルとしてはスーパーベクトル回帰を用いた。これは線形回帰に比べ，説明変数増加に伴う過学習を防ぐ効果に優れている。

3. 実験結果

SAA は言い直しもそのまま収録しており，また，極端にノイズレベルが高い音声も散見される。これらを聴取により取り除き，381 名（72,390 話者対）を対象として実験を行なった。全話者対を A/B に二分し，A の IPA 距離を訓練データとして用いて回帰モデルを構築し，B の発音距離を推定し，評価した。更に B を用いて回帰モデルを構築し，A で評価する実験も行なった。IPA 距離と推定距離との相関は平均 0.81 であり（図 2 参照），良好な結果が得られた。ちなみに，SAA 音声サンプルを連続音素認識により音素配列化し，IPA 距離と音素配列間距離との相関を求めたところ，0.31 であった。SAA データの音素認識率は 73.4%であったが，相関は低かった。

4. 考察

SAA の大規模化，回帰モデルの高精度化により，世界人口の 1/4 と言われる英語利用者を対象とした，個人を単位とした発音分類は技術的には可能であると考える。このような発音地図が利用可能となれば，1)自分の発音が他者とどの程度近い・遠いのかを確認する，2)自らの発音が学習によってどう変化したのかを確認する，3)発音的に近い外国人を見つけて会話練習に励む，4)インターネット上に存在する英語コンテンツを発音地図と繋げれば，世界英語ブラウザが構築できる。これを使えば，例えば国際ビジネスの現場で遭遇する様々に訛った英語を事前に聴取し，準備することも可能であろう。また発音分類技術は英語教育のみならず，純粋に世界英語研究にも貢献できると考えられる。各々の応用はそれに特化した技術開発を要求するが，各応用に依存した技術構築も含め，発音地図構築技術の高精度化を行なう予定である。

参考文献

Kachru, B. *et al*. (2006). *The Handbook of World Englishes*. Wiley-Blackwell.

Shen, H.-P. *et al*. (2013). “Speaker-based pronunciation clustering of World Englishes based on pronunciation structure analysis.” *IEICE Technical Report*, SP2012-116, 7-12.

Weinberger, S. (2013). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu

峯松信明他 (2011). 「音声の構造的表象に基づく学習者分類の検証と発音矯正度推定の高精度化」情報処理学会論文誌, 52, 12, 3671-3681.