

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 12 日現在

機関番号：12601

研究種目：基盤研究(B)

研究期間：2011～2014

課題番号：23300067

研究課題名(和文) 音声の構造的表象と機械学習に基づく頑健・高精度な発音分析と外国語教育への応用

研究課題名(英文) Robust and accurate analysis of pronunciation based on structural representation of speech and machine learning and its application to foreign language learning

研究代表者

峯松 信明 (Minematsu, Nobuaki)

東京大学・工学(系)研究科(研究院)・教授

研究者番号：90273333

交付決定額(研究期間全体)：(直接経費) 15,700,000円

研究成果の概要(和文)：本研究では、音声の構造的表象に基づく発音構造分析を外国語教育、特に発音教育支援に応用することを検討した。その結果、1) 構造的表象と機械学習に基づく、より少数の発音を用いた頑健な発音分析技術の構築、2) 学習者コーパスの IPA ラベリングによる拡充、3) 母語話者・非母語話者による発音に対する高精度分類技術の構築、4) 分類結果の教育用インタフェースの予備的検討を行うことができた。対象言語としては、1) は日本人による中国語発音を対象にし、2)～4) では、母国語を問わず、英語発音を対象にして研究を進めた。これらの成果は対象言語への依存性は少ないため、任意の言語に対して応用可能である。

研究成果の概要(英文)：In this study, pronunciation analysis based on structural representation of speech was investigated and applied to foreign language learning, especially to pronunciation training. The following four results were obtained as contribution. 1) development of a technique to perform structural analysis of pronunciation even for a small number of utterances, 2) extension of a non-native speech corpus by labeling the utterances with IPA (International Phonetic Alphabets), 3) individual-based and precise clustering of native and non-native pronunciations of English, and 4) preliminary investigation of educationally-effective visualization of pronunciation diversity of English. In study 1), Chinese was adopted as target language and in studies 2), 3), and 4), English was used. The technical contributions of these studies are, however, language-independent and they can be applied to any language.

研究分野：音声工学

キーワード：音声の構造的表象 機械学習 外国語学習 発音教育 外国語訛り 世界諸英語

1. 研究開始当初の背景

音声認識技術に基づく発音評価システムの構築は、国内外の研究機関により広く行われている（国外では、蘭国ナイメーヘン大学、瑞国王立工科大、米国カーネギーメロン大、英国ロンドン大、中国科学技術大など）。代表者の峯松は、従来の発音評価技術が有する根本的な欠陥を指摘し、それを解決する理論及び技術を提唱している。音声認識における音響的照合は、ある話者の発声と他者の発声とを音響的に比較し、そのずれを定量化する。これを発音学習に応用すると、発音評価ではなく、声帯模写評価を行うことになる。そのため従来のシステム開発では、教師の発声（モデル発声）を学習者の「声色」に変形・適応し、それを学習者発声と比較している。外国語の音声学習は、教師の声を音響的に模倣することは要求しない。学習者は教師の性別、体格、年齢までを再現しようとはしない。しかし従来の技術には、学習者が真似るべき（教師音声の中の）音響的对象物だけを抽出、モデル化する技術が無かったために、発音学習を声帯模写学習と見立ててシステム開発せざるを得なかった。

代表者の峯松は、音声から、話者の体格、年齢、性別などの非言語的情報を音響的に分離し、学習者が真似るべき音響的对象物のみを抽出、モデル化する理論及び技術を提唱した。峯松による問題点の指摘、及びその解決方法は工学者のみならず、語学教育者からも支持され、近年、招待（依頼）講演・論文の機会が急増した。

本申請は、基盤研究 B（平成 20～22）「音声の分節的・韻律的特徴を包含する発音の構造的表象に基づく外国語教育・学習支援」を先行研究とし、これを発展させる形式をとっている。先行研究では、

- 1) 体格・年齢・性別などの情報が分離・除去された音声表象（音声の構造的表象）に基づいて英語母音、子音発音を表象し、各学習者の発音を個別にモデル化する技術の構築
- 2) 1)に基づいて、学習者単位で発音習熟度で推定したり、音素単位で習熟度を推定する技術の構築
- 3) 体格の違いを無視し、発音の違いのみに基づいて学習者群を分類する技術の構築
- 4) 英語ライミングやリズム学習において重要な二重母音の構造的モデル化に関する初期検討

について検討を行った。本申請課題では、これらを踏まえ、基盤技術の高精度化、更には、収集される発音データの効果的な再利用について検討する。

2. 研究の目的

以下の項目を研究の目的とした。

- a) より少数の発声を用いた頑健かつ高精度な発音分析技術の構築
先行研究で構築した発音分析技術は、発音評定に 40 文ほどの音声を要求している。これは評価対象とする音響イベント（音素など）を分布としてモデル化するために、複数のサンプルが必要になるためである。ここでは、Universal Background Model と話者適応技術の融合により、この問題の解決を図る。
- b) 非英語を対象言語とした発音の構造解析
筆者らのこれまでの検討では、日本人による英語音声を対象としてきたが、発音構造解析は、基本的に言語に非依存に適用できる。そこで、日本人による中国語発声に着眼し、これを応用し、その精度を検討する。
- c) 学習者音声の分類とその教育的利用
先行研究において体格、性別、年齢などに影響されない、学習者発音の分類を検討した（但し母音のみを対象とした）。本研究では、これを拡張し、世界諸英語音声データベースである、Speech Accent Archive を用い、子音まで考慮した上で、話者を探知とした発音クラスタリングを実装する。
- d) 上記の作業には、学習者による音声に対する IPA ラベリングが必要となるが、筆者らが従来構築した ERJ (English Read by Japanese) コーパスに対して、IPA ラベリング作業を遂行する。

3. 研究の方法

- a) より少数の発声を用いた頑健かつ高精度な発音分析技術の構築
発音評定を任意の文ではなく、決められた文を対象に行う場合、その文（音声）に対する多数の話者のサンプルをもちいて、UBM をまず構築し、それを当該話者の音声を使うことで、MAP 適応を施し、その話者のその文（章）の音響モデルを得ることができる。ここから構造解析を行うことで、少量データからの発音構造解析は実施できると考えられる。本研究では、この手法の妥当性を実験的に検討する。
と同時に、発音評価の精度向上のため、多段階の回帰分析や、発音矯正度の推定に関して、繰り返し演算の枠組みを導入する。
- b) 日本人の発声した中国語音声に対して構造解析を行い、その精度について検討する。ここでは、構造解析に対してボトムアップ的な機械学習を適用するのか、あるいは、日本語・中国語の音韻体系の違いに基づくトップダウン的な知識を活用

するのか, について両者を比較し, その精度の違いを見る。

- c) a) で検討した少量音声による構造解析が可能となれば, Speech Accent Archive が提供する読み上げ音声を使って, アーカイブ話者を発音に基づいて分類することが可能となる。この場合, SAA にある IPA 書き起しを用い, IPA 書き起し間距離を適切に定義すれば, それに対する回帰処理として, 任意の二話者間の発音差異の定量化が可能となる。そして, 全話者にこれを行うことで, コーパス話者群に対する発音距離行列が推測され, それに基づいて彼らを可視化することが可能となる。
- d) 筆者らが先行研究で拘置した ERJ (English Read by Japanese) コーパスは読み上げた文章, 及び, その文章に対応する音素列はコーパスの一部として提供されている, 個々の発声はどういう音声現象となってしまうのか, についてのアノテーションはない。ここでは, IPA 書き起しを実施する。

4. 研究成果

- a) より少数の発声を用いた頑健かつ高精度な発音分析技術の構築
UBM + MAP に基づく当該話者の発音構造分析の有効性を確認することができ, また, 多段階の回帰分析や, 繰り返し演算に基づく発音矯正度の有効性についても実験的に検証することができた。これらは雑誌論文化することができた。
- b) 日本人の発声した中国語音声に対して構造解析を行い, その効果を検証することができた。これは国際会議論文として発表することができた。
- c) SAA コーパス話者群に対する発音距離行列を IPA 書き起しを基に作成し, DTW により, 書き起し間距離を定義した。これに対して, 音声信号のみからこの距離を推定することを検討した。推定に使う素性として, 構造解析の結果得られる構造的特徴を利用した。良好な相関が得られ, 本成果は二本の国際会議論文として発表した。また, 分類結果の可視化についても予備的検討を行うことができた。
- d) IPA 書き起しを行ない, 日本人英語の様態を客観的に記述することができた。これについても国際会議論文として発表している。

5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文](計 4 件)

河原達也, 峯松信明, "音声情報処理技術を用

いた外国語学習支援", 電子情報通信学会論文誌, vol.J96-D, no.7, pp.1549-1565 (2013-7)

S. Shimizu, M. Suzuki, N. Minematsu, and K. Hirose, "An experimental study on dynamic features of speech structure," Journal of Research Institute of Signal Processing, vol.16, no.4, pp.319-322 (2012-7)

峯松信明, 鎌田圭, 朝川哲, 牧野武彦, 西村多寿子, 広瀬啓吉, "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, vol.52, no.12, pp.3671-3681 (2011-12)

鈴木雅之, 峯松信明, 広瀬啓吉, "音声の構造的表象と多段階の重回帰を用いた外国語発音評価", 情報処理学会論文誌, vol.52, no.5, pp.1899-1909 (2011-5)

[学会発表](計 13 件)
国際会議発表のみを示す。

H.-P. Shen, N. Minematsu, T. Makino, S. H. Weinberger, T. Pongkittiphan, C.-H. Wu, "Automatic pronunciation clustering using a world English archive and pronunciation structure analysis," Proc. ASRU, pp.222-227 (2013-12)

N. Minematsu and T. Makino, "Automatic clustering of world English pronunciations on an individual basis," Annual meeting of International Association of World Englishes (IAWE), (2013-11)

T. Pongkittiphan, N. Minematsu, T. Makino, K. Hirose, "Automatic detection of the words that will become unintelligible through Japanese accented pronunciation of English," Proc. SLaTE, pp.109-111 (2013-8)

H.-P. Shen, N. Minematsu, T. Makino, S. H. Weinberger, T. Pongkittiphan, C.-H. Wu, "Speaker-based accented English clustering using a world English archive," Proc. SLaTE, pp.184-188 (2013-8)

T. Zhao, A. Hoshino, M. Suzuki, N. Minematsu, K. Hirose, "Automatic Chinese pronunciation error detection using SVM trained with structural features," Proc. Spoken Language Technology (SLT), pp.473-476 (2012-12)

Y. Luan, M. Suzuki, Y. Yamauchi, N. Minematsu, K. Hirose, "Performance

improvement of automatic pronunciation assessment in a noisy classroom," Proc. Spoken Language Technology (SLT), pp.428-431 (2012-12)

S. Kato, G. Short, N. Minematsu, and K. Hirose, "Effects of learners' language transfer on native listeners' evaluation of the prosodic naturalness of Japanese words," Proc. Speech Prosody, pp.198-201 (2012-5)

Y. Qiao, M. Suzuki, N. Minematsu and K. Hirose, "Structure constrained distributions matching using quadratic programming and its application to pronunciation evaluation", Proc. Asian Conference on Pattern Recognition (ACPR), CD-ROM (2011-11)

N. Minematsu, K. Okabe, K. Ogaki, K. Hirose, "Measurement of objective intelligibility of Japanese accented English using ERJ (English Read by Japanese) database," Proc. INTERSPEECH, pp.1481-1484 (2011-8)

S. Kato, G. Short, N. Minematsu, C. Tsurutani, K. Hirose, "Comparison of native and non-native evaluations of the naturalness of Japanese words with prosody modified through voice morphing," Proc. Int. Workshop on Speech and Language Technology in Education (SLaTE), CD-ROM (2011-8)

T. Makino, "Are both American and British pronunciations necessary in learner's English dictionaries in Japan?" Proc. ASLALLEX2011, pp.343-347 (2011)

Y. Yamauchi "How to assess shadowing performance: manual or automatic?" Proc. 16th World Congress of Applied Linguistics, p. 24 (2011)

Y. Yamauchi "Investigation on the possibility of shadowing as an overall proficiency test," Proc. Combined Conference of the Applied Linguistics Association of Australian and Applied Linguistics Association of New Zealand, p. 134 (2011)

〔図書〕(計 1 件)

中川聖一, 小林聡, 峯松信明, 宇津呂武仁, 秋葉友良, 北岡教英, 山本幹雄, 甲斐充彦, 山本一公, 土屋雅稔, "音声言語処理と自然言語処理"(第二章担当), コロナ社 (2013-3)

〔産業財産権〕

出願状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
取得年月日:
国内外の別:

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

峯松信明(東京大学工学系研究科・教授)
研究者番号: 90273333

(2) 研究分担者

牧野武彦(中央大学・経済学部・教授)
研究者番号: 00269482

(3) 連携研究者

山内豊(東京国際大学・商学部・教授)
研究者番号: 30306245