

平成30年6月18日現在

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2015～2017

課題番号：15K12416

研究課題名(和文) 英語映像アーカイブのマルチモーダルコーパス化と分析：言語学と工学の融合

研究課題名(英文) Development and Analysis of Multimodal Corpora of English Public Speaking Performance: Interdisciplinary Project of Linguistics and Engineering

研究代表者

冬野 美晴 (Fuyuno, Miharū)

九州大学・芸術工学研究院・助教

研究者番号：30642681

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：日本人英語学習者が英語によるパブリックスピーキングを行う際に、効果的なパフォーマンスを行うための音声およびジェスチャー上の条件を明らかにし指導に役立てるべく、英語母語話者・英語学習者による英語スピーチアーカイブをデータベース(マルチモーダルコーパス)化し、音声分析と2D・3D動作解析によって分析した。マルチモーダルコーパス研究においては、従来の分析は手作業によるアノテーションデータを用いて行われることが多く、定量的な分析は工学など他の学術分野で独立して行われることが多かった。本研究ではデータをテキスト分析、音声分析、動作解析の3視点から統合的に分析し、定量的な指標を得た。

研究成果の概要(英文)：This research aimed at developing multimodal corpora of English public speaking performance by both native/non-native speakers of English. We collected multichannel input data of Japanese English learners and NSEs. The data mainly included high resolution video data, sound data, script data, and performance evaluation data. The data were analyzed using 2D motion tracking method and sound analysis software. From the analysis, the characteristics of eye-contact movement and phonological elements of highly evaluated speakers were revealed. The result was then applied into developing a teaching manual for effective public speaking.

研究分野：応用言語学

キーワード：マルチモーダルコーパス パブリックスピーキング 英語教育 モーショントラッキング スピーチ
プレゼンテーション 音声解析 アイコンタクト

1. 研究開始当初の背景

近年、経済及び人材の流動性がかつてないほど高まっており、世界で活躍できる人材を養成することが産学官で急務である(cf. 日本経団連, 2011; 寺内 他, 2010)。多様な文化的背景や価値観を持つ人々を相手に自分の考えを効果的に伝え、確実に意思疎通を行い信頼を得ることが出来るスキルが求められており、特に実践的な力として英語スピーキング能力の向上が急務とされる。申請者が過去に実施した科研費研究における、グローバルビジネスパーソンを対象としたニーズ分析の結果、日本人に関して一般的な英語スピーキング能力のみならず英語スピーチやプレゼンテーション等のパブリックスピーキング能力の改善が特に必要とされていることがわかった。

パブリックスピーキングは代表的なソーシャルフォビアの一因であり効果的な指導が求められているにも関わらず、教材作成者の主観的経験に基づく教科書等が多く、音声・映像・ストレスデータを定量的に分析した成果に基づく教材は希少である。そこで本研究では、定量的なデータを収集しアーカイブ化し、またそのデータを効率的な方法で分析する手法を確立する。

2. 研究の目的

本研究の目的は、日本人英語学習者が英語によるパブリックスピーキング(スピーチ・プレゼンテーション等)を行う際に、効果的なパフォーマンスを行うための音声およびジェスチャー上の条件を明らかにし指導に役立てるべく、英語母語話者・英語学習者による英語スピーチアーカイブをデータベース(マルチモーダルコーパス)化し、音声分析と2D・3D動作解析によって分析することである。

マルチモーダルコーパス研究においては、従来の分析は手作業によるアノテーションデータを用いて行われることが多く、定量的な分析は工学など他の学術分野で独立して行われることが多かった。本研究ではデータをテキスト分析、音声分析、動作解析の3視点から統合的に分析し、スピーチ・パフォーマンスについて定量的な指標を得る。

分析結果を実際の指導に取り入れて試行し、効果を検証することで、今後の英語教育における実践的なスピーチ指導へと反映させる。

3. 研究の方法

英語スピーチのマルチモーダルコーパス(音声、映像、発話内容などを統合したデータベース)を作成した。日本語を母語とする英語学習者(以下、日本人英語学習者)、英語母語話者それぞれについてスピーチデータの収集を行った。

収集したデータを整理・選別し、ジャンルごとに映像・音声・テキスト、またデータの

タイプによってはストレス計測値を合わせてコーパスに組み込んだ。

次に、作成したマルチモーダルコーパスの分析を行った。英語母語話者、日本人英語学習者それぞれの群の無音区間の特徴、特に意味チャンク上どのような位置でどのような長さのポーズが挿入されているのか、どのような位置でアイコンタクトが行われているのか等を分析した。その際、パフォーマンス評価情報を合わせて検討し、高評価の話者と低評価の話者にどのような違いが見られるかを導き出した。

音声分析においては音声分析ソフトウェアPraatを用いて無音区間を自動抽出しながら分析した。動作解析では特に頭部の動きに着目し、スピーチ中に話者がどのようにオーディエンスとアイコンタクトを取っているかを定量的に分析した。これまでの我々の研究では、トラッキング・ポイントを鼻のみにつけて左右の顔向きを測定していたが、本研究では3DモーションキャプチャシステムとComputer Vision技術を応用した独自のプログラムを用いて、トラッキング・ポイントを40箇所以上につけ、トラッキング精度を高めた。

ストレス値については話者の心拍変動を用いて計測した。日本人の英語学習者に日本語スピーチと英語スピーチをそれぞれ行ってもらい、スピーチ前後とスピーチ中の心拍変動を計測した。

分析結果から、効果的な英語パブリックスピーキングのための、音声的特徴と動作的特徴を定量的に指標化した。以上の研究成果を取り入れ、パフォーマンス向上とストレス軽減とを目的とした指導マニュアルを開発した。それを用いて、日本人英語学習者に英語スピーチ指導を行い、従来の指導法との指導効果を比較した。

4. 研究成果

まず、非英語母語話者(英語学習者)および英語母語話者それぞれのスピーチ・パフォーマンスの収集を行い、動画・音声データ・スクリプトデータを集めマルチモーダルデータベースを作成した。

非英語母語話者データについては既存のデータに加えて日本人英語学習者を中心とした英語学習者のデータを新たに録画・録音することに成功し、英語暗唱スピーチデータ、英語自由スピーチデータの双方でデータの拡充を行った。オーセンティックなパブリックスピーキング環境で録音・録画し、中学生・高校生・大学生それぞれのグループのデータを得ることができた。データベースには高画質ビデオデータ、音声データ、パフォーマンス評価データ、発話内容スクリプトを含んでいる。

英語母語話者データについてはアメリカ大学卒業式祝辞スピーチ、国外の大学で収集した母語話者の大学生・大学院生約30名の

スピーチ等新たなデータを収集した。特に、話者がスピーチ内容を自由に考えるタイプのスピーチ(規定の課題などの読み上げではないスピーチ)で話者がスクリプトを参照しているシーンのデータを得られたことから、スクリプトを参照する頻度や参照している時間長などを分析することで、日本人向けの教育に応用しやすい分析成果を得ることができた(山下・冬野, 2015)。

非英語母語話者のデータについて、応用言語学と画像工学の観点から分析し、特に話者の顔向きを角度推定に成功した。これにより半自動的に動作の詳細なアノテーションを行えるようになった。更に、動作分析については当初計画していた顔向きの動作推定に加え、黒目の動きを分析することに成功し、目そのものの動作推定を実施した(Komiya et al., 2016; 2017)。

主な結果の一部として、日本の高校で実施されたオフィシャルな英語スピーチ大会のパフォーマンスデータを用いて音声・動作・発話内容・審査員のパフォーマンス評価をクロス分析したところ、以下のことが示唆された。まず、評価の高い話者は音声ポーズの分布に共通した特徴が見られ、コマ、ピリオドにおけるポーズ長の割合をコントロールしている。次に、アイコンタクトにおける顔向き動作のリズムと方向が安定しており、平均しておよそ何秒に1回の頻度で向きを変えるかが明らかになった。更に、アイコンタクト動作における顔向き動作と黒目の動きの間には高い相関は見られず、評価には顔向きの影響が強いことが示唆された(Fuyuno, Komiya & Saitoh, 2018)。

また、発話内容と音声データのクロス分析を行い、スピーチ中の適切なポーズ位置について新たな知見を得た。特に、パフォーマンス評価の高い話者について、意味チャンク上の切れ目とポーズが同期する割合が高く、反対に評価の低い話者では意味チャンク上切れ目でない部分にポーズが置かれやすく、内容に応じたポーズのコントロールが評価に重要な役割を果たすことが示唆された(Fuyuno et al., 2016.01)。

更に、以上の結果から得られた定量的な指標を基に、日本人向けの英語パブリックスピーキング学習マニュアルを作成し、従来の指導法と比較して話者の緊張や不安を軽減する効果があるか検証実験を行った。実験結果より、話者の主観評価において、開発したマニュアルを用いて練習した場合に緊張度が下がる傾向が見られた(Fuyuno, Yamada, Yamashita & Nakajima, 2016)。

マルチモーダルコーパス研究においては、動画・音声データを分析する際に手作業でのアノテーションを行う場合の膨大な作業コストの問題があったが、本研究ではモーショントラックや音声解析を応用して分析の半自動化に成功したこと、およびデータをテキスト分析、音声分析、動作解析の3視点から

統合的に分析し、スピーチ・パフォーマンスについて定量的な指標を得た点が新規的と言える。

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8件)

Fuyuno, M., Komiya, R., and Saitoh, T. (2018). Multimodal analysis of public speaking performance by EFL learners: Applying deep learning to understanding how successful speakers use facial movement, *The Asian Journal of Applied Linguistics*, 4, 3, 117-129.

冬野美晴, 山下友子. (2017). 英語パブリックスピーキングにおける音声ポーズパターンとアイコンタクト動作パターンがパフォーマンスの総合評価へ及ぼす影響に関する統計的考察, マルチモーダルコーパスデータに基づくパブリックスピーチの統計的解析, 1-15.

小宮凜子, ワラポン・チンサティ, 齊藤剛史, 冬野美晴. (2017). スピーチ指導のための映像情報を用いた話者の顔解析, *DIA2017*, 195-200.

Komiya, R., Saitoh, T., Fuyuno, M., Yamashita, Y., Nakajima, Y. (2017). Head pose estimation and motion analysis of public speaking videos, *International Journal of Software Innovation*, 5, 1, 57-71.

Komiya, R., Saitoh, T., Fuyuno, M., Yamashita, Y., Nakajima, Y. (2016). Head pose estimation and movement analysis for speech scene, *Proc. of 15th IEEE/ACIS International Conference on Computer and Information Science*, 409-413.

Yamashita, Y. & Fuyuno, M. (2015). Three Aspects of Public Speeches with Competent Speakers: Speech Pauses, Head Movements, and Lexical Choices. *STEM Journal*, 5, 215-231.

山下友子・冬野美晴. (2015). アメリカの大学の卒業式における著名人による祝辞スピーチの録画映像の分析: 音声の無音区間と顔上げ動作. *ATEM Journal*, 20, 33-44.

冬野美晴. (2015). アジア圏グローバルビジネスパーソンを対象とした英語スキルのニーズ分析: グローバル人材育成教育に向けた考察. *東アジア英語教育研究*, 5, 13-27.

[学会発表](計11件)

1) 冬野美晴 (2017). 没入型バーチャル・リアリティによる英語パブリックスピーキング練習用システムの開発: マルチモーダルコーパス分析に基づく教材の展開, 外国語教育メディア学会.

- 2) Fuyuno, M., Komiya, R., & Saitoh, T. (2017). *Analyzing Public Speaking for EAP Pedagogy: Factors of Better Performance*, Faces of English 2: Teaching and Researching Academic and Professional English.
- 3) Liu, S., Nakajima, Y., Yamashita, Y., Fuyuno, M. (2017). *Quantitative analysis of English public speaking presented by Japanese EFL learners*, the 6th Conference of the Asia-Pacific Society for the Cognitive Sciences of Music.
- 4) Fuyuno, M., Yamashita, Y., Saitoh, T., & Nakajima, Y. (2016). *Semantic Structure, Speech Units and Facial Movements: Multimodal Corpus Analysis of English Public Speaking*, 8th International Conference on Corpus Linguistics.
- 5) 小宮凜子, 齊藤剛史, 冬野美晴, 山下友子, 中島祥好. (2016). *スピーチ指導のための顔特徴点を用いた話者頭部の動き解析*, 電子情報通信学会.
- 6) Fuyuno, M., Yamashita, Y., Saitoh, T., Nakajima, Y. (2016). *Multimodal Corpus Approach to Speech Psychology: The Effect of Public Speakers' Facial Movements on Audience Reception*, 31st International Congress of Psychology 2016 (ICP2016).
- 7) Fuyuno, M., Yamada, Y., Yamashita, Y., & Nakajima, Y. (2016). *Developing effective instructions to decrease Japanese speaker's nervousness during English and Japanese public speeches: Evidence from psychological and physiological measurements*, 31st International Congress of Psychology 2016 (ICP2016).
- 8) 山下友子, 冬野美晴, 横森大輔. (2016). *日本人大学生の英語プレゼンテーションにおける効果的な間の取り方*, 日本認知科学会研究分科会「間合い—時空間インタラクション」.
- 9) 冬野美晴, 山下友子. (2015). *マルチモーダルコーパスを用いた英語母語話者によるパブリックスピーキング・パフォーマンスの分析: 発話速度・ジェスチャー・語彙特性の観点から*, 英語コーパス学会第41回全国研究大会, 2015.10.
- 10) Fuyuno, M. (2015). *Analysis and application of multimodal corpora of English public speaking by EFL learners: Toward evidence-based teaching*, International Conference on Corpus Linguistics and Technology Advancement 2015.
- 11) Yamashita, Y. & Fuyuno, M. (2015). *An analysis of speech pauses, head movements, and lexical choices in graduation speeches at the universities in the United States*, The 19th STEM

International Conference.

〔図書〕(計1件)

Fuyuno, M., Yamashita, Yuko., & Nakajima, Y. (2016). *Input a Word, Analyse the World.-Selected Approaches to Corpus Linguistics*, Cambridge Scholars Publishing, Multimodal Corpora of English Public Speaking by Asian Learners: Analyses on Speech Rate, Pause and Head Gesture.

〔産業財産権〕

○出願状況(計1件)

名称: パブリックスピーキング支援装置、及びプログラム

発明者: 冬野美晴, 山下友子, 齊藤剛史, 中島祥好

権利者: 九州大学、九州工業大学

種類: 特許

番号: 2017-077706

出願年月日: 2017年9月26日

国内外の別: 国内優先

6. 研究組織

(1)研究代表者

冬野 美晴 (FUYUNO, Miharu)
九州大学・芸術工学研究院・助教
研究者番号: 30642681

(2)研究分担者

山下 友子 (YAMASHITA, Yuko)
芝浦工業大学・工学部・助教
研究者番号: 10726334

中島 祥好 (NAKAJIMA, Yoshitaka)
九州大学・芸術工学研究院・教授
研究者番号: 90127267

齊藤 剛史 (SAITOH, Takeshi)
九州工業大学・大学院情報工学研究院・准教授
研究者番号: 10379654