

令和元年5月22日現在

機関番号：12501

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K00223

研究課題名(和文) モダリティが異なる対話型自然言語としての手話と音声の韻律機能の解明

研究課題名(英文) Analysis of prosody of Japanese Sign Language and Japanese speech as interactive natural language with different modalities

研究代表者

堀内 靖雄 (Horiuchi, Yasuo)

千葉大学・大学院工学研究院・准教授

研究者番号：30272347

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：モダリティの異なる二つの対話型自然言語である日本語音声と日本手話を横断的に分析した結果、どちらもプロソディが非常に重要な役割を担っており、実時間での言語理解や円滑な実時間コミュニケーションに有効に利用されていることが明らかとなった。また、両者には聴覚言語と視覚言語というモダリティの違いがあるため、その特徴にはいくつかの違いが見られたが、言語情報に付加されたプロソディの存在が対話型自然言語での会話において必須であることがわかった。

研究成果の学術的意義や社会的意義

音声や手話は実時間対話言語であり、プロソディ(韻律)と呼ばれる文字で表現できない情報(声の高さ、早さ、声色、頭や視線の動き、表情、ジェスチャー等)が言語理解やコミュニケーションにおいて重要な役割を担っているが、本研究ではそれらプロソディの機能を解明した。近年、人と会話する人工知能が登場しているが、実際に人と会話しているような実感は得られない。自然な会話を実現するためにはプロソディなど、人間が会話で用いている情報を利用する必要があり、本研究は未来の人工知能のために必須となる技術である。

研究成果の概要(英文)：Analysis of two interactive natural languages, Japanese speech (JS) and Japanese Sign Language (JSL) with different modalities showed that prosody plays a very important role in real-time understanding and communication. Although there are some differences between JS and JSL due to the difference between auditory and visual languages, it was suggested that adding prosody to linguistic information is essential in dialogue in both interactive natural languages.

研究分野：ヒューマンマシンインタラクション

キーワード：対話型自然言語 プロソディ 韻律 日本語音声 日本手話 聴覚言語 視覚言語

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

音声の韻律に関しては国内外において研究が盛んであり、科研費特定領域研究「韻律と音声処理」の実施後も韻律と感情表現の関連に関する分析や対話における韻律の機能、音声認識・対話システムへの利用等の研究が続けられている。一方、手話に関する韻律研究はほとんど行われていないが、平成 23 年に障害者基本法が改正され、日本手話は言語であると認められたため、今後、ますます基礎研究が必要となる研究分野である。しかしながら、国内の手話研究者は非常に少なく、手話の韻律に関する研究者はほとんどいない。また、音声と手話という異なるモダリティ（聴覚、視覚）による対話型自然言語において、両者の韻律に関して分析・解明できる研究者は他に類を見ない。本研究ではモダリティの違いを考慮しながら日本語音声と日本手話の韻律を横断的に分析することにより、両言語の韻律情報がどのような機能を持っているのかが明らかになることが期待される。

2. 研究の目的

本研究では音声（日本語）/手話（日本手話）の韻律について横断的に分析することにより、韻律の特徴、機能に関して、各モダリティで共通の性質をもっているもの、モダリティの特性により変化するもの、言語の違いにより変化するものなどの観点から各種韻律情報の分析を行い、個別の研究では得られない韻律と言語との関係を明らかにする。本研究の具体的な目的は以下の通りである。

- (1) 音声と手話における韻律の物理的特徴量の解明
- (2) 音声と手話の韻律機能の解明
- (3) 音声と手話の韻律機能の比較分析

3. 研究の方法

本研究ではモダリティが異なる対話型自然言語である日本語音声と日本手話において、どのような韻律情報がどのような物理的特徴量として表出され、また、どのような機能を有しているかを横断的に分析する。

- (1) 音声と手話の対話データの収集・書き起こし・タグ付け
音声会話データは既存のデータベースを利用する。日本手話の会話データに対しては、ネイティブ手話者の協力のもと、手の動き、頭の動き、目・眉の形、口の形のタグ付けを行う。
- (2) 音声と手話における韻律の物理的特徴量の解明
音声と手話の韻律の物理的な特徴量として、時間的な特徴量、言語情報に重畳される特徴量、マルチモーダルな特徴量を明らかにする。
- (3) 音声と手話の韻律機能の解明
音声と手話の韻律の機能を下記の観点に基づいて解明する。
 - ・実時間での言語理解を助ける（統語理解、意味理解、談話理解等）
 - ・実時間コミュニケーションを制御する（話者交替等）
- (4) 音声と手話の韻律機能の比較分析
音声と手話の韻律機能に関して、横断的な比較分析を行い、類似の特徴量は類似の機能を有するかどうか、また、異なる特徴量で類似の機能を有する韻律が存在するかどうか、などを分析し、それらについて、モダリティの違いや言語体系の違いからの考察を試みる。

4. 研究成果

(1) 日本語音声の F0 モデル（藤崎モデル）パラメータの推定手法の検討

日本語音声の韻律を表現する情報として、声の高さに相当する基本周波数 (F0) の情報は非常に重要であるが、物理的な特徴量として表現することは困難である。本研究ではその優れた表現方法である周波数パターン生成過程モデル（藤崎モデル）を韻律分析に利用することを検討した。しかしながら、藤崎モデルのパラメータを決定するためには、従来、専門家による手作業で行う必要があった。本研究では藤崎モデルパラメータ推定において、基底周波数の自動決定手法を開発した。また、マイクロプロソディ（微細変動）の除去手法に関する検討を行い、パラメータ推定精度の向上を達成した。

(2) 日本語音声の韻律特徴量と感情表現の関係の分析

音声の韻律として、声の高さ、話速、声色（スペクトル/音色）が挙げられるが、これら韻律特徴量と感情表現の関係を聴取実験により分析し、各韻律特徴量の変化と表現される感情の種類と強度の関係を明らかにした。全体的な傾向としては音声の高さと長さが感情表現の再現性に寄与しているのに加え、音色に相当するスペクトルも感情表現の再現性に寄与することが分かった。また、それらを組み合わせることにより、さらにその影響が大きくなった。

(3) 日本手話における手動作の認識手法の検討

日本手話における手の動きを測定する方法として、これまで磁気センサーをろう者の体に装

着して測定してきたが、このような接触型のセンサーによる測定方法では自然な手話を表出することが困難である。そこで、本研究では Microsoft Kinect(R) 距離カメラを利用し、非接触で手の動きを計測する手法を提案し、有効性を確認した。また、手話の自動認識手法に関する検討を行った。

(4) 日本手話の頭部動作の分析

日本手話では、マルチモーダルな特徴量のうち、頭の動きが非常に重要であるため、日本手話のうなずき等の頭部動作について対話データを分析した結果、日本語音声ではパラ言語または非言語的な情報であるのに対し、日本手話では言語としての文法的な機能を有するうなずきが存在し、意味理解や会話制御に用いられていることが明らかとなった。

(5) 手話の停止動作の分析

日本手話において、音声言語の無音に相当する現象として、手話言語において手の動きが停止する状態を3種類(基準位置、保持、ゆるみ)に分類し、その出現した文脈を実時間コミュニケーションにおける会話制御の観点から分析した結果、保持、ゆるみでの停止は聞き手の発話開始を抑制し、話し手が発話を継続する傾向が見られたことから、音声言語にはない手話特有の会話制御の手法であることを明らかにした。

(6) 日本語音声における母語話者と非母語話者のプロソディの違いの分析

母語話者と非母語話者の日本語音声による話者交替の比較実験の結果、円滑な話者交替を実現するためには、適切なプロソディを付与する必要がある、まだ十分にプロソディを付与できない非母語話者(日本語学習者)の場合、適切な話者交替ができないことが明らかとなった。

(7) 日本語音声と日本手話の話者交替の比較分析

対話型自然言語におけるコミュニケーションにおいて、話者交替は円滑な会話を実現する上で、非常に重要な問題である。日本語音声では統語情報に加え、声の高さ、話速等の韻律により予測を可能としていることが示された。一方、日本手話では統語情報に加え、手の動きや顔の動きの大きさ、速さ等が話者交替の予測に利用されていると示唆された。また、オーバーラップが生じる話者交替においては、日本語音声と日本手話で顕著な差が表われた。日本手話では言換えや繰返し表現が表出することが多いが、繰返し表現により発話権の譲渡が明示される現象が見出された。これらの分析結果から、視覚言語と聴覚言語のモダリティの違いにおけるプロソディの違いが話者交替に与える影響が明らかになった。

これらの結果から、モダリティの異なる二つの自然言語である日本語音声と日本手話において、どちらもプロソディが非常に重要な役割を担っており、プロソディが実時間での言語理解や円滑な実時間コミュニケーションに有効であることが明らかとなった。また、両者には聴覚言語と視覚言語というモダリティの違いがあるため、その特徴にはいくつかの違いが見られたが、言語情報に付加されたプロソディの存在が対話型自然言語での会話において必須であることがわかった。

5. 主な発表論文等

[雑誌論文](計14件)

- (1) 長谷川愛, 堀内靖雄, 黒岩眞吾, 日本手話対話におけるポーズと話者交替に関する分析, 人工知能学会研究会資料, 査読無, SIG-SLUD-B803, 2019, pp.13-18
- (2) Yoshiko Arimoto, Yasuo Horiuchi, Sumio Ohno, Consistency of base frequency labelling for the F0 contour generation model using expressive emotional speech corpora, Proc. 9th International Conference on Speech Prosody 2018, 査読有, 2018, pp.398-402
- (3) 市川薫, 堀内靖雄, 長嶋祐二, 心的負担が軽い「対話のことば」の構造, ヒューマンインタフェース学会論文誌, 査読有, 20(2), 2018, pp.45-58
- (4) 風間康介, 堀内靖雄, 下元正義, 黒岩眞吾, 言語情報を用いた Kinect による連続指文字認識手法の検討, 情報処理学会研究報告, 査読無, 2018-AAC-6(16), 2018, pp.1-6
- (5) 工藤正景, 堀内靖雄, 黒岩眞吾, 日本手話の呈示速度変化における腕動作変化の分析, 情報処理学会研究報告, 査読無, 2018-AAC-6(17), 2018, pp.1-6
- (6) 高橋遼平, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用い

た HMM-SVM での連続指文字認識, ヒューマンインタフェース学会研究報告集, 査読無, 19(1), 2017, pp.23-28

- (7) 佐藤貴弘, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用いた手話の数字認識手法の検討, ヒューマンインタフェース学会研究報告集, 査読無, 19(1), 2017, pp.17-22
- (8) 市川薫, 大橋浩輝, 仲真紀子, 菊池英明, 堀内靖雄, 黒岩眞吾, 母音対話の話者交換タイミング, 科学・技術研究, 査読有, 5(1), 2016, pp.113-122
- (9) 高橋遼平, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用いた HMM による連続指文字認識の検討, 情報処理学会研究報告, 査読無, 2016-AAC-1(9), 2016, pp.1-6
- (10) 佐藤貴弘, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用いた手話の数字認識における識別手法の検討, 情報処理学会研究報告, 査読無, 2016-AAC-1(10), 2016, pp.1-6
- (11) 古谷佳大, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, 手話認識における位置・動き特徴量の検討, 電子情報通信学会論文誌, 査読有, J99-D(1), 2016, pp.90-92
- (12) 金子紗弓, 堀内靖雄, 黒岩眞吾, 日本手話における手指動作とうなずきのタイミングの分析, 電子情報通信学会技術報告, 査読無, 115(491), 2016, pp.7-12
- (13) 佐賀俊介, 堀内靖雄, 黒岩眞吾, 音韻情報を考慮した話速変換による早口合成音声の明瞭度改善, 電子情報通信学会技術報告, 査読無, 115(491), 2016, pp.31-36
- (14) 市川薫, 川端良子, 大橋浩輝, 仲真紀子, 菊池英明, 堀内靖雄, 黒岩眞吾, 母語話者の交替タイミング機能獲得時期の検討, 情報処理学会研究報告, 査読無, 2015-MUS-107(67), 2015, pp.1-6

〔学会発表〕(計 13 件)

- (1) 長谷川愛, 堀内靖雄, 黒岩眞吾, 日本手話対話におけるポーズと話者交替に関する分析, 人工知能学会言語・音声理解と対話処理研究会第 85 回研究会, 京都大学吉田本部キャンパス, 2019
- (2) Yoshiko Arimoto, Yasuo Horiuchi, Sumio Ohno, Consistency of base frequency labelling for the F0 contour generation model using expressive emotional speech corpora, 9th International Conference on Speech Prosody, Adam Mickiewicz University in Poznań, Poland, 2018
- (3) 風間康介, 堀内靖雄, 下元正義, 黒岩眞吾, 言語情報を用いた Kinect による連続指文字認識手法の検討, 電子情報通信学会福祉情報工学研究会, 筑波技術大学, 2018
- (4) 工藤正景, 堀内靖雄, 黒岩眞吾, 日本手話の提示速度変化における腕動作変化の分析, 電子情報通信学会福祉情報工学研究会, 筑波技術大学, 2018
- (5) 佐久間結, 堀内靖雄, 有本泰子, 大野澄雄, 黒岩眞吾, 藤崎モデルのパラメータ推定におけるマイクロプロソディ除去による効果の検討, 日本音響学会 2017 年秋季研究発表会, 愛媛大学, 2017
- (6) 堀内靖雄, Kinect を用いた HMM-SVM での連続指文字認識, 第 141 回ヒューマンインタフェース学会研究会, 京都工芸繊維大学松ヶ崎キャンパス, 2017
- (7) 堀内靖雄, Kinect を用いた手話の数字認識手法の検討, 第 141 回ヒューマンインタフェー

ス学会研究会, 京都工芸繊維大学松ヶ崎キャンパス, 2017

- (8) 天野翔太, 堀内靖雄, 黒岩眞吾, STRAIGHT によるモーフィングを用いた演技音声の感情表現と韻律特徴の分析, 日本音響学会 2017 年春季研究発表会, 明治大学生田キャンパス, 2017
- (9) 高橋遼平, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用いた HMM による連続指文字認識の検討, 情報処理学会第 1 回アクセシビリティ研究発表会, 国立情報学研究所, 2016
- (10) 佐藤貴弘, 堀内靖雄, 川本一彦, 下元正義, 眞崎浩一, 黒岩眞吾, 鈴木広一, Kinect を用いた手話の数字認識における識別手法の検討, 情報処理学会第 1 回アクセシビリティ研究発表会, 国立情報学研究所, 2016
- (11) 佐賀俊介, 堀内靖雄, 黒岩眞吾, 音韻情報を考慮した話速変換による早口合成音声の明瞭度改善, 電子情報通信学会第 83 回福祉情報工学研究会, 筑波技術大学春日キャンパス, 2016
- (12) 金子紗弓, 堀内靖雄, 黒岩眞吾, 日本手話における手指動作とうなずきのタイミングの分析, 電子情報通信学会第 83 回福祉情報工学研究会, 筑波技術大学春日キャンパス, 2016
- (13) 有本泰子, 堀内靖雄, 大野澄雄, 発話の基本周波数分布にもとづく藤崎モデル基底周波数の決定手法, 日本音響学会 2015 年秋季研究発表会, 会津大学, 2015

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

6 . 研究組織

(1) 研究分担者 : なし

(2) 研究協力者 : なし

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。