

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 17 日現在

機関番号：12501

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22500148

研究課題名（和文）異なるモダリティによる対話型自然言語の韻律機能の分析

研究課題名（英文）Analysis of Prosody in Interactive Natural Languages Expressed Using Different Modality

研究代表者

堀内 靖雄（HORIUCHI YASUO）

千葉大学・大学院融合科学研究科・准教授

研究者番号：30272347

研究成果の概要（和文）：対話型自然言語では韻律（プロソディ）が言語理解やコミュニケーションの向上に貢献していると考えられている。本研究では伝達モダリティの異なる日本語音声と日本手話を分析対象とし、プロソディの機能や効果について、モダリティの違いを考慮しながら、横断的に分析し、(1)韻律情報の計測手段の開発・改良、(2)実時間コミュニケーションにおける韻律の必要性の解明、(3)モダリティの異なる言語間における類似点・相違点の解明を行った。

研究成果の概要（英文）：Prosody in interactive languages is important for language understanding or real-time communication. In this study, the function and the effect of prosody were analyzed for Japanese oral language and Japanese sign language with different modalities. The results of research were (1) the development of measurement method of prosody information, (2) showing the importance of prosody in real-time communication, (3) finding of similarities and differences between different languages with different modalities.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,400,000	420,000	1,820,000
2011年度	1,200,000	360,000	1,560,000
2012年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：対話型自然言語・コミュニケーション・プロソディ・韻律・日本語音声（聴覚言語）・日本手話（視覚言語）

1. 研究開始当初の背景

音声の韻律に関して、国内外において研究が盛んであり、国内では特定領域研究「韻律と音声処理」が平成12年～15年に実施され、

当申請者も研究に参加した。主に韻律と感情表現の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が研究されている。一方、手話に関する韻律研究は非常に少ないが、最近、研究を開始したグループも

わずかながら存在し、今後、ますます発展が期待される研究分野である。しかしながら、対話型自然言語の持つ優れた実時間性に注目し、異なるモダリティ言語である音声と手話の比較を通して共通の特性とそれぞれ固有の特性を見出そうとする横断的な研究アプローチは国内外にその例がない。そこで、音声、手話というモダリティや言語体系が異なる対話型自然言語を横断的に分析することにより、各言語単独の研究では得られにくかった韻律の性質・機能が明らかになることが期待できる。

2. 研究の目的

音声など実時間で情報伝達・コミュニケーションを行う対話型自然言語では韻律（プロソディ）と呼ばれる付加的な情報が同時に表出されることにより、言語理解やコミュニケーションの向上に貢献していると考えられている。また、対話型自然言語には音声だけでなく、ろう者（聴覚障害者）が用いる手話（視覚言語）など様々な言語が存在するが、そのすべてにおいて音声同様に韻律情報が存在している。本研究では異なるモダリティ（聴覚、視覚）による対話型自然言語において、それぞれどのような韻律情報が存在し、どのような機能を有しているかを横断的に詳細に分析することを目的とする。

これまで音声（日本語）／手話（日本手話）の韻律に関して個別に研究を行ってきたが、本研究ではそれらを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ない、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) 各言語における韻律の分類
- (2) 韻律が表出される物理的特徴量の解明
- (3) 韻律の機能の解明
- (4) 音声と手話間の韻律機能の比較分析
- (5) 認知実験による検証

3. 研究の方法

(1) 各言語の対話データの収集・タグ付け

研究を進めるにあたり、まず準備段階として、韻律情報を含むデータが必要となる。すでに保有しているデータに加え、日本手話のデータを追加収録する。また、データを収録しただけでは研究に用いることはできないため、単語書き起こしや各種タグを付与する作業を行なう。

(2) 各言語における韻律の分類

各言語における韻律情報を分類し、詳細に検討を行なう。

- ・時間的情報：単語時間長変化、話速、間など。
- ・言語情報に付加される情報
- ・マルチモーダル情報：顔の動き、視線、表情、ジェスチャー等

また、言語情報に付加される情報、マルチモーダル情報はそれぞれ時間長等の情報も持っているため、時間的観点の分析についても検討を行なう。

(3) 韻律が表出される物理的特徴量の解明

上記の韻律情報がどのような物理的特徴量で表現されているかについて、モダリティの違いに考慮しながら分析を行ない、各種特徴パラメータの抽出を行なう。

(4) 韻律の機能の解明

韻律における以下の機能を解明する

- ・実時間言語理解を助ける（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）

(5) 言語間の韻律機能の比較分析

上記で得られた韻律特徴量と機能について、各言語横断的に比較分析を行ない、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説明を試みる。

(6) 韻律の認知実験

上記から得られた韻律に関する各種知見について、その韻律を変化させた言語資料を合成し、主観評価実験により、各種知見の認知的妥当性を検証する。

4. 研究成果

(1) 日本語音声の F0 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、それを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。本研究では音声に対し、非専門家でも容易に与えられる書き起こし情報とアクセント情報を与えることにより、半自動で推定する手法を提案し、その有効性を

評価した。

(2) 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出することが困難である。そこで、本研究では Microsoft 社の距離センサー Kinect を使用し、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。

(3) 日本手話における首動作の表現方法の検討

日本手話におけるプロソディとして、非常に重要な要素に首の動きが挙げられる。しかしながら、そのプロソディの機能等を分析するにあたり、実写の映像では様々な影響により、評価が困難である。そこで本研究では 3D 手話アニメーションを利用した再合成映像に首の動きを表現する方法を検討した。日本手話における首動作のうち、とくに重要である「うなずき」と「あご引き（下固定）」について、実際に測定したろう者の運動をもとに、3D アニメーションで再現する手法を提案し、実装・評価を行った。

(4) 日本手話における首動作と手話理解に関する検討

日本手話におけるプロソディである首動作を変化させた評価実験用の動画を 3D 手話アニメーションを用いて作成し、ろう者による主観評価実験を行った。その結果、このプロソディは手話の理解において必須であることが明らかとなり、逆に不適切なプロソディが付与された場合、理解の妨げとなった。これらの結果から、対話型自然言語である音声や手話では韻律情報が言語理解や円滑な会話コミュニケーションにおいて必須であり、将来の対話型自然言語処理では韻律も考慮した分析／合成が必要とされることが明らかとなった。

(5) 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り、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。

(6) 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

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。Sacksらの話者交替規則によれば、話者交替は移行適格場 (TRP: transition relevance place) 付近で生じ、TRP 以外では原則的に話者交替は生じないとされている。対話型自然言語では、この TRP を予測可能とする特性 (投射可能性: projectability) があるとされている。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することが示された。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 TRP 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 TRP 後に言換えや繰返しの表現が表出することが多いが、その場合、手や顔の動きの大きさや速さ等の韻律が消失したり、弱化したりし、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

これらの結果を総合すると、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、その存在により、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が実現可能とな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、プロソディの表出方法や制御方法には若干の違いが見られたが、本質的には言語情報に追加されたプロソディの存在により、対話型自然言語として必要な機能が実現されていることがわかった。

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)

- (1) 古谷佳大, 今村大輔, 堀内靖雄, 川本一彦, 篠崎隆宏, 黒岩眞吾:
Kinect とパーティクルフィルタによる手話認識手法の検討,
電子情報通信学会技術研究報告, 査読なし, 112(475) 巻, 2013, pp. 251-256
- (2) 白井洋輔, 柳村舞衣, 篠崎隆宏, 堀内靖雄, 黒岩眞吾, 遠藤俊樹, 宇都宮栄二:
ユーザ属性を考慮した字幕付与方法の検討

- 電子情報通信学会技術研究報告, 査読なし, 112(475)巻, 2013, pp.245-250
- (3) 西村洋介, 今村大輔, 堀内靖雄, 川本一彦, 篠崎隆宏, 黒岩眞吾:
Kinect とパーティクルフィルタを用いた HMM 手話認識手法の検討,
電子情報通信学会技術研究報告, 査読なし, 111(431)巻, 2012, pp.161-166
- (4) 村瀬優美, 堀内靖雄, 篠崎隆宏, 黒岩眞吾:
日本手話対話と日本語音声対話における話者交替現象の比較分析,
電子情報通信学会技術研究報告, 査読なし, 111(424)巻, 2012, pp.7-12
- (5) 黒田大崇, 堀内靖雄, 黒岩眞吾:
うなずきを付加した手話CGアニメーションの検討,
ヒューマンインタフェース学会研究報告集 SIG-NOI-04, 査読なし, 13(2)巻, 2011, pp.1-6
- (6) 鈴木いおり, 西村洋介, 堀内靖雄, 黒岩眞吾:
パーティクルフィルターとHMMによる動画像からの手話認識に関する検討,
電子情報通信学会技術報告, 査読なし, 110(384)巻, 2011, pp.25-30
- (7) Masafumi Nishida, Yasuo Horiuchi, Shingo Kuroiwa, Akira Ichikawa:
Automatic Speech Recognition Based on Multiple Level Units in Spoken Dialogue System for In-vehicle Appliances,
Lecture Notes in Computer Science, 査読あり, Vol.6231, 2010, pp.539--546
- (8) 齊藤涼子, 堀内靖雄, 黒岩眞吾:
話者交替規則に基づいた日本手話対話の話者交替の分析,
人工知能学会 言語・音声理解と対話処理研究会, 査読なし, 59巻, 2010, pp.13--18
- [学会発表] (計9件)
- (1) 今里圭太, 堀内靖雄, 篠崎隆宏, 黒岩眞吾:
言語情報を利用した F0 モデルパラメータの推定手法,
日本音響学会 2013 年春季研究発表会, 2013 年 3 月 15 日, 東京工科大学
- (2) 滝沢恵子, 菊池英明, 大橋浩輝, 市川熹, 堀内靖雄, 黒岩眞吾:
日本人学習者の英語対話における流暢性に関する研究—話者交替時における重複が示す学習者の流暢性の段階—,
日本音響学会 2013 年春季研究発表会, 2013 年 3 月 13 日, 東京工科大学
- (3) 古谷佳大, 今村大輔, 堀内靖雄, 川本一彦, 篠崎隆宏, 黒岩眞吾:
Kinect とパーティクルフィルタによる手話認識手法の検討,
電子情報通信学会福祉情報工学研究会, 2013 年 3 月 12 日, 福岡工業大学
- (4) 白井洋輔, 柳村舞衣, 篠崎隆宏, 堀内靖雄, 黒岩眞吾, 遠藤俊樹, 宇都宮栄二:
ユーザ属性を考慮した字幕付与方法の検討,
電子情報通信学会福祉情報工学研究会, 2013 年 3 月 12 日, 福岡工業大学
- (5) 内田正洋, 篠崎隆宏, 堀内靖雄, 黒岩眞吾:
認識対象区間を考慮した音声からの印象認識,
HCG シンポジウム 2012, 2012 年 12 月 11 日, くまもと森都心プラザ
- (6) 西村洋介, 今村大輔, 堀内靖雄, 川本一彦, 篠崎隆宏, 黒岩眞吾:
Kinect とパーティクルフィルタを用いた HMM 手話認識手法の検討,
電子情報通信学会音声研究会, 2012 年 2 月 10 日, 東北大学
- (7) 村瀬優美, 堀内靖雄, 篠崎隆宏, 黒岩眞吾:
日本手話対話と日本語音声対話における話者交替現象の比較分析,
電子情報通信学会福祉情報工学研究会, 2012 年 1 月 27 日, 名古屋工業大学
- (8) 鈴木いおり, 西村洋介, 堀内靖雄, 黒岩眞吾:
パーティクルフィルターとHMMによる動画像からの手話認識に関する検討,
電子情報通信学会 祉情報工学研究会, 2011 年 1 月 22 日, 立命館大学 びわこ・くさつキャンパス
- (9) 齊藤涼子, 堀内靖雄, 黒岩眞吾:
話者交替規則に基づいた日本手話対話の話者交替の分析,
第 59 回 人工知能学会 言語・音声理解と対話処理研究会, 2010 年 7 月 22 日, 長崎・やすらぎ伊王島

〔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) 研究代表者

堀内 靖雄 (HORIUCHI YASUO)

千葉大学・大学院融合科学研究科・准教授

研究者番号：30272347