

令和 2 年 6 月 15 日現在

機関番号：62618

研究種目：基盤研究(B) (一般)

研究期間：2017～2019

課題番号：17H02339

研究課題名(和文)リアルタイムMRIおよびWAVEデータによる調音音声学の精緻化

研究課題名(英文)Elaboration of articulatory phonetics by means of realtime-MRI and WAVE data

研究代表者

前川 喜久雄 (Maekawa, Kikuo)

大学共同利用機関法人人間文化研究機構国立国語研究所・音声言語研究領域・教授

研究者番号：20173693

交付決定額(研究期間全体)：(直接経費) 13,700,000円

研究成果の概要(和文)：東京方言16名分、近畿方言5名分、モンゴル語3名の調音運動を記録したリアルタイムMRI動画を1名あたり約1時間収集した。データのブラウジング環境を構築し、音声器官(舌、唇、口蓋、咽頭壁など)の輪郭を自動抽出する技術を開発した。このデータを利用して、モンゴル語母音調和に関する舌根位置の関与を示した論文、日本語発話末に生じる撥音の調音位置が直前母音によって決まっていることを示した論文、日本語ワ行子音の調音が定説となっている二重調音ではなく、主に両唇の接近によって行われていることを示した論文を発表した。さらに日本語のラ行子音に関する分析も発表した。

研究成果の学術的意義や社会的意義

リアルタイムMRI動画を用いることで、従来、観察することが困難であった声道後半部(軟口蓋、口蓋垂、咽頭など)における調音運動を定量的かつ精密に観測することができる。このデータを用いて、モンゴル語の母音調和に舌根の前後位置が関係するとの説を支持する直接的な証拠を世界で最初に提出することができたほか、日本語の撥音、ワ行子音、ラ行子音の調音についても、従来の定説を覆す証拠を得た。これらによって日本語に関する調音音声学的な記述は今後大きく書き換えられる可能性が高い。

研究成果の概要(英文)：Realtime MRI movie of articulatory movements were recorded for 16 Tokyo Japanese, five Kinki Japanese, and three Inner Mongolian speakers; each movie is about 60 minutes long. Data browsing environment on the web and techniques for automatic contour extraction of articulatory organs (lips, tongue, jaw, palate, pharynx etc.) were developed. As for linguistic application, three papers were published with respect to issues like 1) Relevance of the tongue root in the Mongolian vowel harmony, 2) Influence of the preceding vowel for the determination of the place of articulation of the utterance-final moraic nasal in Japanese, and, 3) Exclusive relevance of labial (rather than labio-velar) articulation in the Japanese /w/. Analysis of the Japanese /r/ was also presented.

研究分野：音声学・言語学

キーワード：調音音声学 音声生成 リアルタイムMRI 日本語 モンゴル語

様式 C-19、F-19-1、Z-19（共通）

1. 研究開始当初の背景

調音音声学は音声研究全般の基礎となる分類学であるが、その知識体系のなかには、主観的な観察（内省）や推測に依拠して構築されている部分が少なからず残されている。また分類の基礎が調音諸器官の静止形状に置かれており、音声器官の運動に関する知見の反映が不十分である。このような調音音声学の根本問題を解決するための努力は、20世紀以来様々なものが試みられてきた。X線映画、超音波断層撮像法のように既存の技術を応用する以外に、X線マイクロビーム法、ダイナミックパラトグラフィーのように音声科学の領域で独自に開発された観測技術もある。しかし、これらの技術を用いても、データの蓄積は充分ではなく、特に声道後半を構成する音声器官（軟口蓋・口蓋垂・咽頭壁・奥舌面・舌根など）が関係する調音運動については正確な観測をおこなうことが極めて困難であった。

近年開発されたリアルタイムMRI法（rtMRI）は、医療用MRI装置を特殊なパラメータ設定で稼働させることによって、毎秒数十フレームで、声道の正中矢状断面形状の時間変化を記録する技術である。rtMRI法では、唇から喉頭にいたる声道の全体像が、256×256ピクセル程度の解像度をもった鮮明な画像として記録されるので、上に述べた調音音声学の根本問題の解消に大きく貢献できる。またMRI装置には被爆の危険性がないので、X線を利用する手法に比べると同一の被験者から多くのデータを収集できることも、rtMRIデータの大きな利点である。

2. 研究の目的

本研究の主要な目的は以下の三点である。①日本語を主要な対象として、様々な子音・母音の組み合わせを組織的に網羅し、統計的な分析に耐える量をそなえたrtMRI動画のデータベースを構築すること、②そのデータを活用するための環境を整備し、データから重要な情報を自動抽出する技術を開発すること、そして、③実際にデータを分析し、調音音声学上の未解決問題を検討すること。

3. 研究の方法

㈱ATR-Promotion（京都府相楽郡）が保有するMRI装置を利用して、主に日本語東京方言話者の調音運動データを収録した。これと対照する目的で、近畿方言話者のデータも収録した。収集にあたっては、可能な限り話者の年齢と性別に配慮した。また、声道後半部に位置する音声器官の運動が言語学上重要な対立に関与している可能性が報告されているモンゴル語（内モンゴル方言）の母音調和に関するデータも収録した。

日本語（東京方言と近畿方言）の場合、収録したデータの内訳は、①単独モーラ（音節）142項目、特殊モーラ関係151項目、2モーラ連鎖の包括的記録676項目を中心とし、必用に応じていくつかの追加項目を収録した。また一部の話者については、5母音生成時の声道形状データと歯列形状のデータも収録した。モンゴル語については母音調和現象を中心に約600単語のデータを収録した。

データの大部分は毎秒14フレームの速度で撮像したが、一部については試験的に毎秒25フレームでの撮像も実施した。rtMRI画像データはDCMファイル形式で提供される。これと別途収録された音声を結合して、一般的な動画ファイル（AVI, MP3）を構築した。そして、構築された動画ファイルをウェブ上でブラウジングし、音声波形やスペクトログラムなどの音声情報と比較しながら、声道形状を測定を実施するためのウェブアプリを開発した。さらに動画に記録された音声器官（唇、舌、歯茎、口蓋、咽頭壁、喉頭、顎など）の輪郭を自動抽出するための技術を開発した。

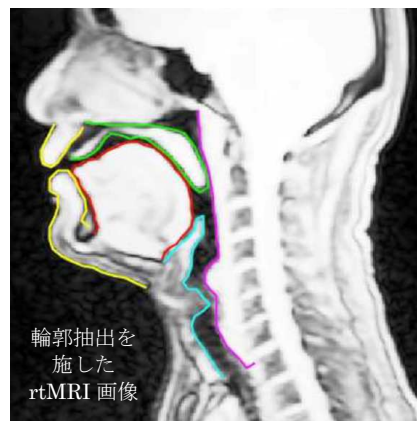
rtMRIデータのブラウジング環境は、Java script技術を用い、公開されている音声処理用ライブラリを活用しながら、マルチプラットフォームのウェブアプリケーションとして構築した。連続する3フレームを比較しながら、分析に必要とするフレームを選択したり、選択したフレームに対して種々の測定をおこなって、結果をローカルに保存する機能が実装されている。このアプリケーションの構築ではGitHubを利用しており、将来的には、rtMRIデータとともに広く一般公開する予定である。

調音器官の輪郭抽出技術の開発では、顔抽出用に開発されたdlibライブラリを利用した。あらかじめ人間が作成した輪郭データを教師データとして学習することにより、実用的なレベルの自動抽出技術を開発した（次節参照）。

4. 研究成果

最初にrtMRI動画データベースについてまとめる。東京方言話者16名（男性10名女性6名）、近畿方言話者5名（男性3名、女性2名）のデータを収録した。データ量は1名あたり約1時間である。またrtMRIと同時に収録した音声データ（16bit, 48kHzサンプリング）にはMRI装置が発する動作ノイズを軽減するためのノイズリダクション処理をほどこしたうえで、両者を結合した動画ファイルをAVIとMP3の両形式で作成した。このデータに、各発話項目の開始・終了時刻などの情報を付与し、さらに話者と発話内容に関するメタデータを付与した。現在の検索システムは、PC上の表計算ソフトで稼働しており、必要な発話を単語文字列、音素列、話者性別などによって指定すると、検索にヒットした発話を自動的に切り出した後、ひとつのファイルに再結合して、検索結果動画ファイルを生成する。このファイルを前節で説明したrtMRI動画のブラウジング用ウェブアプリで閲覧しながら分析を進めるのがひとつの定番的分析手法である。

次に、音声器官の輪郭自動抽出技術に触れる。これは基本的に画像処理の研究であり、従来、難易度の高い研究として認知されてきた。しかし、顔抽出用の機械学習ライブラリである dlib を活用したことがブレークスルーとなり、実用的な処理技術を開発することができた。1 名分の rtMRI データは数万枚の画像データから構成されているが、このうち適切に選択した数十枚に対して人手で輪郭抽出作業を実施すれば、それを教師データとして、人手と変わらない精度での輪郭自動抽出が可能になった(図参照)。この成果は音声情報処理研究の領域で最大の国際会議である INTERSPEECH の予稿集に査読を経て採録された[1]。現在は複数の話者を一括して処理できる技術の開発を継続中である。



最後に、調音音声学領域での成果に触れる。日本語の撥音「ん」は直後の子音ないし母音と同じ調音器官の構えで調音されることが知られている。そして、直後に子音も母音も存在しない環境、つまり語末(発話末)では、口蓋垂(口蓋のもっとも奥よりで咽頭壁と接する部分)と奥舌面が接触する子音(口蓋垂音)として発音されると教科書類には記載されている。しかし、東京方言話者 8 名分の rtMRI データの解析により、この定説が成立しないこと、発話末の撥音の調音位置は、その直前に位置する母音の構えによってほぼ全面的に決定されていることが明らかになった[2]。この成果は音声科学領域の代表的国際会議である ICPhS の予稿集に査読を経て採録された後、現在、音声学領域で定評のある国際誌で査読を受けている最中である。

日本語東京方言については、ほかにワ行とラ行の子音についても従来の定説を覆す研究成果を発表することができた。ワ行子音は従来、両唇と軟口蓋の二箇所ですべて声道に狭めが発生する二重調音子音として記述されてきたが、rtMRI データでみると、東京方言に関するかぎり、軟口蓋に狭めを発生させている話者は一人もおらず、両唇音として記述するのが妥当であることが判明した。この成果は大学の紀要論文として英文で発表した[4]。

ラ行子音については、これを後部歯茎の flap(はたき音)とみる記述が 1990 年代から国際音声学協会(IPA)のハンドブックに掲載されている。しかし rtMRI データを検討すると、この子音の調音位置は歯茎から後部歯茎におよぶ広い範囲に分布しており、また舌尖の調音運動も、flap ではなく tap(弾き音)であることが明らかになった。この成果は 2019 年度の日本音声学学会全国大会で発表し[5]、現在、国際誌への論文投稿を準備中である。

最後に現代モンゴル語の母音調和については、これを舌背の上下方向位置の調和とみる説と舌根の前後方向位置の調和とみる説とが永年対立してきた。我々のデータの解析結果は、舌根位置に母音調和と相関して明瞭な差が発生していることを示しており、モンゴル語の母音調和に関する舌根説を支持する世界最初の直接的証拠となった。この成果は ICPhS の予稿集に査読を経て掲載され[3]、現在、国際誌への投稿を準備中である。

<引用文献>

- [1] Takemoto, Tsubasa Goto, Yuya Hagihara, Sayaka Hamanaka, Tatsuya Kitamura, Yukiko Nota and Kikuo Maekawa. Speech organ contour extraction using real-time MRI and machine learning method. *Proc. INTERSPEECH 2019*, pp.904-908, 2019.
- [2] Kikuo Maekawa. A real-time MRI study of Japanese moraic nasal in utterance-final position. *Proc. International Congress of Phonetic Sciences (ICPhS) 2019*, pp. 1987-1991, 2019.
- [3] Yoshio Saito, Yurong, Kikuo Maekawa. An investigation into modern Mongolian vowel harmony using real-time magnetic resonance imaging. *Proc. International Congress of Phonetic Sciences (ICPhS) 2019*, pp. 1431-1434, 2019.
- [4] Kikuo Maekawa. Remarks on Japanese /w/. *ICU Working Paper in Linguistics*. 10, pp. 45-52, 2020.
- [5] 前川喜久雄「日本語ラ行子音の調音：リアルタイム MRI による観察」日本音声学学会第 33 回全国大会予稿集, 2019.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 1件/うちオープンアクセス 4件）

1. 著者名 Takemoto Hironori, Goto Tsubasa, Hagihara Yuya, Hamanaka Sayaka, Kitamura Tatsuya, Nota Yukiko, Maekawa Kikuo	4. 巻 NA
2. 論文標題 Speech Organ Contour Extraction Using Real-Time MRI and Machine Learning Method	5. 発行年 2019年
3. 雑誌名 Proc. INTERSPEECH 2019	6. 最初と最後の頁 904-908
掲載論文のDOI（デジタルオブジェクト識別子） 10.21437/Interspeech.2019-1593	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Kikuo Maekawa	4. 巻 NA
2. 論文標題 A real-time MRI study of Japanese mora ic nasal in utterance-final position	5. 発行年 2019年
3. 雑誌名 Proc. ICPhS 2019	6. 最初と最後の頁 1987-1991
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Yoshio Saito, Yurong, Kikuo Maekawa	4. 巻 NA
2. 論文標題 An Investigation into Modern Mongolian Vowel Harmony Using Real-time Magnetic Resonance Imaging	5. 発行年 2019年
3. 雑誌名 Proc. ICPhS 2019	6. 最初と最後の頁 1431-1434.
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Kikuo Maekawa	4. 巻 10
2. 論文標題 Remarks on Japanese /w/	5. 発行年 2020年
3. 雑誌名 ICU Working Papers in Linguistics	6. 最初と最後の頁 45-52
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 能田由紀子・籠宮隆之・前川喜久雄
2. 発表標題 磁気センサシステムをもちいた計測による座位・仰臥位・腹臥位における舌運動の差異の検討
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集, pp.823-824, 2019.03.05.
4. 発表年 2019年

1. 発表者名 後藤翼・萩原裕也・濱中彩夏・竹本浩典・北村達也・能田由紀子・前川喜久雄
2. 発表標題 機械学習によるrtMRI動画における発話器官の輪郭抽出精度の評価
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集, pp.821-822, 2019.03.05.
4. 発表年 2019年

1. 発表者名 浅井拓也・菊池英明・前川喜久雄
2. 発表標題 リアルタイムMRI動画を利用したパラ言語発話生成時調音運動の動体検出
3. 学会等名 日本音響学会2019年春季研究発表会講演論文集, pp.1135-1136, 2019.03.05.
4. 発表年 2019年

1. 発表者名 浅井拓也・菊池英明・前川喜久雄
2. 発表標題 調音運動動画アノテーションシステムの開発と応用
3. 学会等名 日本音声学会第32回全国大会予稿集, pp.201-206, 2018.09.15.
4. 発表年 2018年

1. 発表者名 浅井拓也・菊池英明・前川喜久雄
2. 発表標題 調音運動動画アノテーションシステムの開発
3. 学会等名 日本音響学会2018年秋季研究発表会講演論文集, pp.1235-1238, 2018.09.13.
4. 発表年 2018年

1. 発表者名 後藤翼・萩原裕也・濱中彩夏・竹本浩典・北村達也・前川喜久雄
2. 発表標題 機械学習によるrtMRI動画における発話器官の輪郭抽出方法の検討
3. 学会等名 日本音響学会2018年秋季研究発表会講演論文集, pp.813-814, 2018.09.13.
4. 発表年 2018年

1. 発表者名 前川喜久雄・能田由紀子・北村達也・竹本浩典・石本祐一
2. 発表標題 日本語撥音の調音音声学的記述の精緻化：rtMRIデータによる試み
3. 学会等名 日本音響学会2018年春季研究発表会講演論文集, pp.1247-1248, 2018.03.14.
4. 発表年 2018年

1. 発表者名 前川喜久雄・能田由紀子・北村達也・竹本浩典・石本祐一
2. 発表標題 日本語撥音の調音音声学的記述の精緻化：rtMRIデータによる試み
3. 学会等名 日本音響学会2018年春季研究発表会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

リアルタイムMRI動画で見る日本語の調音運動 https://www2.ninjal.ac.jp/openhouse2018/pdf/P2_r.pdf 講義「音声学」国立国語研究所言語学レクチャーシリーズ Vol.2 https://www.youtube.com/watch?v=xeV3hU1aorU
--

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	斎藤 純男 (Saito Yoshio) (10225740)	拓殖大学・外国語学部・教授 (32638)	
研究分担者	藤本 雅子 (Masako Fujimoto) (30392541)	早稲田大学・人間科学学術院・その他(招聘研究員) (32689)	
研究分担者	竹本 浩典 (Hironori Takemoto) (40374102)	千葉工業大学・先進工学部・教授 (32503)	
研究分担者	北村 達也 (Tatsuya Kitamura) (60293594)	甲南大学・知能情報学部・教授 (34506)	
研究分担者	菊池 英明 (Hideaki Kikuchi) (70308261)	早稲田大学・人間科学学術院・教授 (32689)	

6. 研究組織（つづき）

	氏名 (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	籠宮 隆之 (Takayuki Kagomiya) (10528269)	大学共同利用機関法人人間文化研究機構国立国語研究所・言語変異研究領域・特任助教 (62618)	