

# リアルタイムMRI動画データベースプロジェクトの概要と調音音声学への応用

著者	前川 喜久雄
雑誌名	言語資源活用ワークショップ発表論文集
巻	6
ページ	96-107
発行年	2021
URL	<a href="http://doi.org/10.15084/00003483">http://doi.org/10.15084/00003483</a>

## リアルタイム MRI 動画データベースプロジェクトの概要と調音音声学への応用

前川喜久雄（国立国語研究所）<sup>†</sup>

## Overview of the Real-Time MRI Articulatory Movement Database Project and Its Application to Articulatory Phonetics

Kikuo Maekawa (NINJAL)

## 要旨

発表者らは 2017 年以来、リアルタイム MRI 動画撮像技術を用いた日本語調音運動のデータベースを構築してきており、2021 年 4 月にはその一部（標準語話者 10 名分）を試験公開した。本発表ではスペシャルセッションへの導入を兼ねて、プロジェクトの紹介をおこなった後に、リアルタイム MRI 動画が調音音声学研究にもたらす革新の可能性を論じ、これまでに得られた標準語に関する研究成果を提示する。具体的には、①日本語の発話末撥音は従来多くの教科書類に口蓋垂音と記載されてきたが、実際には直前の母音によって口蓋全体にわたって大幅に移動していること、②日本語ワ行子音には二重調音が行われている痕跡は確認できず、調音位置は両唇にあると考えられること、③日本語の「ヒ」と「シ」の子音は、正中断面に関する限りほぼ同一の調音位置を保ちながら、冠状面の特徴を変化させ、同時に流量を制御することによって、調音しわけることができること、等である。

## 1. プロジェクトの経緯と目標

われわれは 2017 年以来、主に日本学術振興会科学研究費の補助によって、また副次的には国立国語研究所コーパス開発センターと人間文化研究機構機構長裁量経費の支援によって、リアルタイム MRI 撮像技術を用いて日本語音声の調音運動を組織的に記録したデータベースである『リアルタイム MRI 調音運動データベース』（rtMRIDB）の構築を進めてきた。データベースの設計と実装ならびに計画の基盤をなす rtMRI 動画撮像技術については、先に前川他（2020）に報告した。本項では新たなデータ公開の予定と調音音声学領域における成果を紹介する。

本プロジェクトは採択された科研費の区切りに従って第 1 期と第 2 期にわけることができる。第 1 期は 2017–19 年度（科研採択前の準備期間として 2016 年の後半を含む）、第 2 期は 2020–2023 年度であり、現在は第 2 期の 2 年間に該当するが、実質的には 2017 年から現在にいたるまで、一貫性をもったデータベース構築プロジェクトとして運営されてきている。

本プロジェクトの開始時にふたつの目標をたてた。ひとつは調音音声学の研究における客観性の欠如を解消して、研究を精緻化することである。前川他（2020）、Maekawa(2021)などで繰り返し述べてきたことであるが、IPA に代表される調音音声学の

---

<sup>†</sup> kikuo@ninjal.ac.jp

研究には、20世紀初頭の成立期以来、観察の主観性という研究方法上の限界がついてまわってきており、これが研究の進展の妨げとなっている。

もちろん、この問題を解消するために、客観的・定量的な観測手法の開発が20世紀を通じて続けられてきた。しかし、音声生成時の声道の形状とその変化、とりわけ、軟口蓋・舌根・咽頭壁・喉頭・喉頭蓋などから構成される声道後部の情報の視覚化には困難が大きく、この部位における調音の正確な観察には21世紀初頭におけるrtMRI技術の開発を待つ必要があった。そのため、声道後部に関しては、現在の音声学理論および個別言語の音声学的記述は研究者の主観に大きく依存したものとなっている。

本稿の第3節では日本語の音声学的記述の問題点を列挙し、その一部について現時点までの研究成果を紹介するが、これらはともに声道後部における子音調音の問題を扱った研究であることに注意されたい。

本プロジェクトの第2の目標は、rtMRIデータを一般公開することによって、日本語の調音音声学的研究のインフラを整備することである。第1期においては、第1期終了後にデータを一般公開することを念頭においたデータ収集が進められ、第2期には公開自体を目的とした科研費の申請が行われている。

公的資金によって科学研究のために収集されたデータを研究終了後に、あるいは研究成果の発表と同時にオープンデータとして一般公開することは、近年の科学研究におけるひとつの潮流となりつつある。データ公開の目的は、当該研究の再現性と信憑性を高めることを目的としていることが普通であるが、本プロジェクトの場合は、さらに広く、日本語調音音声学の研究インフラとしてデータを提供することを目的としている。以下本稿では第2節でインフラ整備の目標の達成状況に触れ、第3節で調音音声学の精緻化の目標の達成状況に触れる。

## 2. 日本語調音音声学のインフラ整備

### 2.1 『リアルタイムMRI調音運動データベース』

インフラ整備の主要な手段として構築を進めているのが『リアルタイムMRI調音運動データベース』（rtMRIDB）である。2021年4月には試験公開版を公開した(<https://rtmridb.ninjal.ac.jp>)。標準語話者10名（男性6名女性4名）から収集した3発話クラスの発話のrtMRI動画が収録されている（表1参照）。

単独モーラは拡張された五十音表の音読、バイモーラ項目は26個のモーラの組み合わせを「これが\_\_型」というキャリアセンテンスに入れた音読、特殊モーラ項目は、撥音・促音・長母音などの特殊モーラを含む単語の音読である。試験公開版の詳細は前川他（2020）と試験公開版付属の解説文書に譲る。

表1: 『リアルタイムMRI調音運動データベース』試験公開版の内容

発話クラス	内容	14-fps	27-fps
MB	2モーラの組み合わせ+キャリア文	8407	0
MP	特殊モーラ（キャリア文無し）	2293	0
MU	単独モーラ（キャリア文無し）	2246	0

本2021年度には『リアルタイムMRI調音運動データベース』第1版を公開する予定である。第1版には、標準語話者17名（新規7名分を拡張）と近畿方言話者5ないし6名のデータを格納する予定であり、発話クラスも表2に示すように拡張される予定である。

表2: 『リアルタイム MRI 調音運動データベース』第1版の内容\*

発話クラス	内容	14 fps	27 fps
MB	2 モーラの組み合わせ+キャリア文	16651	284
MP	特殊モーラ (キャリア文無し)	4692	443
MU	単独モーラ (キャリア文無し)	3115	668
NS	文章朗読	33	0
PL	パラ言語情報	245	0
SR	発話速度	351	0
TT	早口言葉	64	0

The screenshot shows the rtMRIDB desktop application. At the top, it says "rtMRIDB (The real-time MRI articulatory movement database)". Below this are search filters: Phoneme (mikaN, kirinN) and Gender (M). There are buttons for "すべて選択する" (Select all), "選択を解除する" (Deselect), and "マージン" (Margin) set to 0 seconds. A search bar contains "14件の検索結果が見つかりました。" (14 search results found). Below the search bar is a table with columns: Selected, File, Start, End, Subject, Text, Phoneme, Tag, Slide2, Slide, and dialect. The table lists several entries with details like file names (s1\_9\_mp2, s2\_9\_mp2, s4\_11\_mp2), time ranges, subjects (蜜柑 (みかん), 麒麟 (きりん)), phonemes (mikaN, kirinN), and dialects (Standard).

図1: rtMRIDB 試験公開版 (デスクトップアプリ版)。

文章朗読項目と撮像速 27 fps のデータは、2018 年秋以降に収録を開始した項目で、一部の話者についてのみ提供されるデータである(朗読は 14 名、27 fps は 9 名)。文章朗読項目では、イソップ寓話の「北風と太陽」と、宇宙論をあつかった新書の導入部分を朗読した連続音声の調音運動を撮像した。前者のテキストは国際音声学協会の IPA ハンドブック (International Phonetic Association 1999) で日本語の例示に用いられたものであり、後者は『日本語話し言葉コーパス』 (Maekawa 2003) の朗読データとして利用したテキストのひとつである。撮像速度 28fps のデータは時間分解能が従来の 2 倍のデータである。撮像速度と 1 回の撮像時間の間にはトレードオフの関係があるので、14fps の撮像では約 37.3 秒までの発話が可能であるのに対し 28fps での発話は約 18.7 秒に制限される。単独モーラ項目では、前川(2019)の分析においてときに困難が感じられたラ行モーラを多く含むスライド 2 枚(MU4, MU5)を対象とした。

パラ言語情報は、種々の発話意図(発話態度)や対比の強調(focus)が調音に及ぼす影響を探ることを目的としている。SR は 14 fps の発話において疑似的に時間分解能を向上させる目的で意図的に発話速度を低下させた発話、TT は 3 種類の早口言葉の朗読である。

図1は rtMRIDB 試験公開版の Windows デスクトップアプリ版の検索画面である。画面左上に検索条件が指定されており、画面下部に検索結果が表示されている。検索結果のすべてもしくは一部を指定して新たな mp4 ファイルを作成することができ、図では右端に編集された動画がポップアップしている。この動画は mp4 ファイルとしてローカルに保存することが可能である。

## 2. 2 ツール開発

rtMRIDB は mp4 動画化された素データとその検索系を提供するが、それだけでは調音音声学研究のインフラとしては必ずしも十分ではない。これはかつて『日本語話し言葉コーパス』を公開した際にも生じた問題であるが、価値の高いデータが大量に公開されていても、それを研究用に検索するツールが提供されていない状況では、いわゆる文系研究者のなかには、コンピュータスキルの低さからデータを十分に活用することができない人がいる。

『日本語話し言葉コーパス』の場合は、コア部分の豊富なアノテーションを RDB (リレーショナルデータベース) で利用可能にした CSJ-RDB (Koiso et al. 2014) や形態論情報検索ツールである『中納言』への搭載によって、この問題を克服した。

rtMRI データの場合、具体的には以下のような問題が想定される。まず、MRI 画像の標準フォーマットである DCM は画像専用であるから、実験時に別途収録した音声信号をダビングして動画化する必要があるが、このプロセス自体が多くの研究者にとって高いハードルになる。また、完成した動画ファイルを視聴するだけでなく、分析に必要な箇所を検索し、検索された画像に対して分析用情報のアノテーションを施したりするには、特殊化されたソフトウェアが必要になる。

DCM 形式ファイルに対する計測・アノテーションツールとしては、ImageJ というフリーウェアが普及しており (Schneider, Rasband & Eliceiri 2012)、高機能で知られているが、動画をあつかうことはできず、当然音声も聴取することもできない。談話研究などで利用されるフリーソフトの ELAN (<https://archive.mpi.nl/tla/elan>) は音声を含む動画のアノテーションツールであるが、スペクトログラムの表示機能がなく、また、基本的にかなり広い時間区分 (数百ミリ秒から数秒程度) を対象としたテキストアノテーションを行うことが目的となっていて、動画を構成する個々の静止画像に対する画像アノテーションや位置計測の機能は実装されていない。そこで本プロジェクトでは、mp4 形式で動画化された rtMRI データを対象としたブラウジング・アノテーションツールを独自開発することとし、MRI Viewer と呼んでいる (浅井・菊池・前川 2018)。ユーザーは rtMRIDB の検索結果を mp4 ファイルに保存した後に MRI Viewer でアノテーションを施すことになる。

現在公開されている MRI Viewer Ver.2 は、mp4 動画を読み込むと、波形とサウンドスペクトログラムを表示した状態で、連続する 3 フレームの MRI 画像を表示する。この状態で音声に対する多層のアノテーション (時系列アノテーション) が可能であり、Praat (Boersma & Weenink 2021) の interval tier と point tier の両方を用いることができる。

また、モードを切り替えると特定の MRI 画像に対するアノテーション (フレームアノテーション) を実施できる (図2参照)。図2で背景化している時系列アノテーション画面の右上隅にはデータに付与されたアノテーションテキスト (図2では1層のみ) とその開始終了時刻が表示されている。このウィンドウではテキストの検索や選択的視聴が可能である。一方、図2で前景となっているフレームアノテーションウィンドウでの作業結果は、時系列アノテーション画面の MRI 画像にも反映され、作業中は動画をスクロールさせても保持される。

MRI Viewer は Chrome や Firefox などのウェブブラウザ上で稼働する Java script アプリケーションであり、ソースコードは公開されている (<https://github.com/kikuchiken-waseda/mri-vuewer.ver2>)。本スペシャルセッションでは開発者である浅井拓也氏らによる MRI Viewer のより詳細な解説が予定されている。

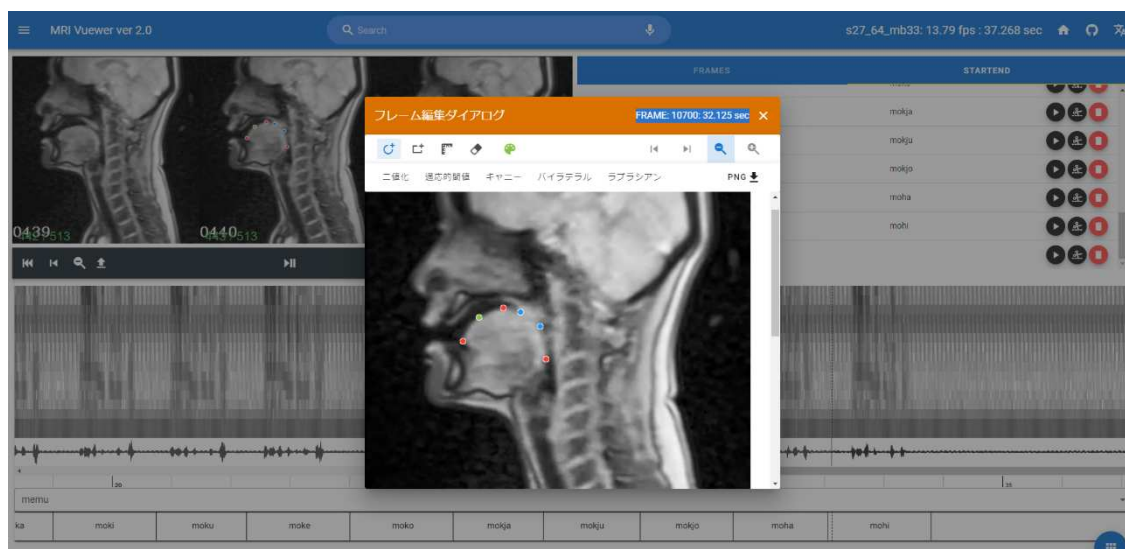


図2: MRI Viewer Ver.2 の作業画面。背景が時系列アノテーション画面、前景となっているのがフレームアノテーション用ウィンドウ

### 2. 3 アノテーション情報の自動付与

MRI Viewer はユーザーが rtMRIDB のデータにアノテーションを施し、計測するためのツールである。研究者が自分でデータを作りそれを自分で解析するという、古典的な実験研究の手順に則ったツールである。一方、近年では、ある程度の汎用性をそなえた基礎的なアノテーション情報が素データとともに公開されていて、ユーザーはそのアノテーション情報を活用して自分の研究目的を遂げるといった形の研究が広く行われるようになってきている。勿論、すべての研究がこのような形で実施できるわけではないが、ユーザーが独自にアノテーションを実施する場合であっても、様々な出発点となるアノテーションデータが公開されていれば、研究の効率が飛躍的に向上する。『日本語話し言葉コーパス』や『現代日本語書き言葉均衡コーパス』における形態論情報や分節音情報はそのような機能を担ったアノテーション情報である。

rtMRIDB の場合、何がそのような基礎アノテーションに該当するかを考えると、最も基礎的であるのは発話時刻（発話の開始・終了時刻）の情報であり、もうひとつは発話を構成する音素列の情報であろう。これらは疑似的な形であれば既に rtMRIDB に付随して公開されている。今後 rtMRIDB に付与すべきアノテーション情報としては、音素境界情報と音声器官の輪郭情報のふたつが挙げられる。

このうち音素境界の情報を得るためには、音声自動認識の研究で培われた音素強制アライメントの技術（所与の音声信号を時間軸上で所与の音素列に分割する技術）を活用することが考えられる。rtMRIDB の音声信号には、MRI 装置の稼働ノイズが重畳しているので、その影響が心配されたが、Julius( Lee & Kawahara 2009) の segmentation toolkit

(<https://julius.osdn.jp/index.php?q=ouyoukit.html>) を用いた予備実験の結果は、ほぼ実用レベルに近い精度の自動アライメントが可能であることが示唆されている。

図3に、Juliusのsegmentation toolkitでIPA音響モデル(モノフォン・16状態・男声)を用い、男性話者が発話した特殊モーラ項目の発話に音素強制アライメントを施した結果を示す。細かな問題は指摘できるが、全体としてほぼ満足のゆく結果が得られていることがわかる。

次に、音声器官の輪郭情報について。本プロジェクトでは、前川他(2020)でも言及したように、これを自動抽出する技術を第1期から継続して開発してきており、現在では、8名程度の教師データがあれば、多数(20名程度)の話者のrtMRIデータを安定して処理できるレベルに達している(Takemoto et al. 2019, 後藤他 2019, 2020, 天野他 2021)。声道は舌、口唇・下顎、口蓋、咽頭後壁・披裂部、および喉頭蓋・声帯の5部位に分割されて、それぞれの輪郭情報が40個の点情報の連鎖として推定されるので、MRI画像1フレームに対して200個の点情報が提供される。本スペシャルセッションではこの技術の開発を指導した竹本浩典氏らによる解説が予定されている。

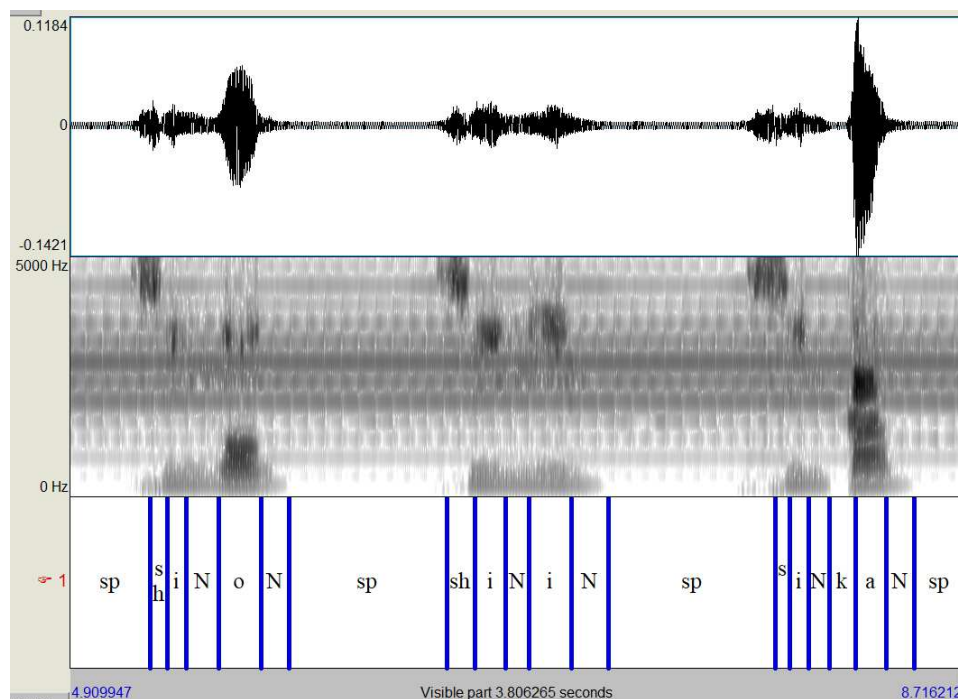


図3: Julius segmentation toolkit を用いた音素自動アライメントの結果。

発話は左から/siNoN/「心音」、/siNin/「心因」、/siNkaN/「新刊」。

記号 sp は無音区間に該当している

音素境界情報と音声器官輪郭情報が提供されると、これらをリレーショナルデータベース上で自由に組み合わせてデータを自動抽出することが可能になる。例えば或る発話に含まれる音素区分ごとに舌の最高点や最奥点の平均値を計算するようなことはSQL言語で十数行のスクリプトを書くだけで実施できるようになるだろう。このようなデータ分析

環境は先述した CSJ-RDB に該当するものであり、これを実現できれば、日本語調音音声学のインフラ整備は飛躍的に進捗したことになる。<sup>1</sup>

### 3. 日本語調音音声学の精緻化

ここからは rtMRIDB が日本語調音音声学の精緻化にどのように貢献したかの実例を紹介する。日本語の調音音声学には 100 年に近い研究の蓄積があるが、本稿 1 節に指摘した理由によって、ごく基本的なことから関しても現在でも多くの未解決問題が残されている。以下では、そのうち、これまでに研究成果を出版することのできた研究をふたつ紹介する。発話末撥音の調音位置の問題とワ行子音の調音位置の問題である。

#### 3. 1 発話末撥音の調音位置

発話末（語末）撥音の調音位置の問題は、rtMRI データの長所が発揮される課題と考えられたので、プロジェクトの初年度から取り組んだ（前川他 2018, Maekawa 2019, Maekawa 2021）。撥音は語中であって閉鎖音や閉鎖鼻音が後続する場合は、後続子音と調音位置を共有する鼻音として実現される。一方、後続子音がない発話末の撥音は口蓋垂鼻音として実現されると記述している文献が多い（服部 1951, 川上 1977, Vance 2008 など）が、これらはすべて主観的な観察による記述である。吐師他(2014)は客観的な観測データに基づいて、口蓋垂以外に調音点がある発話が多いことを示唆しているが、明瞭な結論は下していない。これは吐師らが利用した X 線マイクロビーム装置では、測定用マーカーを口蓋垂に接着することができないために口蓋垂で調音が行われているかどうかを直接確認できないからである。国立国語研究所(1990)は、X 線映画の分析に基づいて発話末撥音の調音位置に大きな変異が観察されることを報告しているが、発話末撥音の調音位置がカ行子音よりも奥よりであることを根拠に、口蓋垂鼻音であるとの記述にも合理性が認められるという結論を述べている。

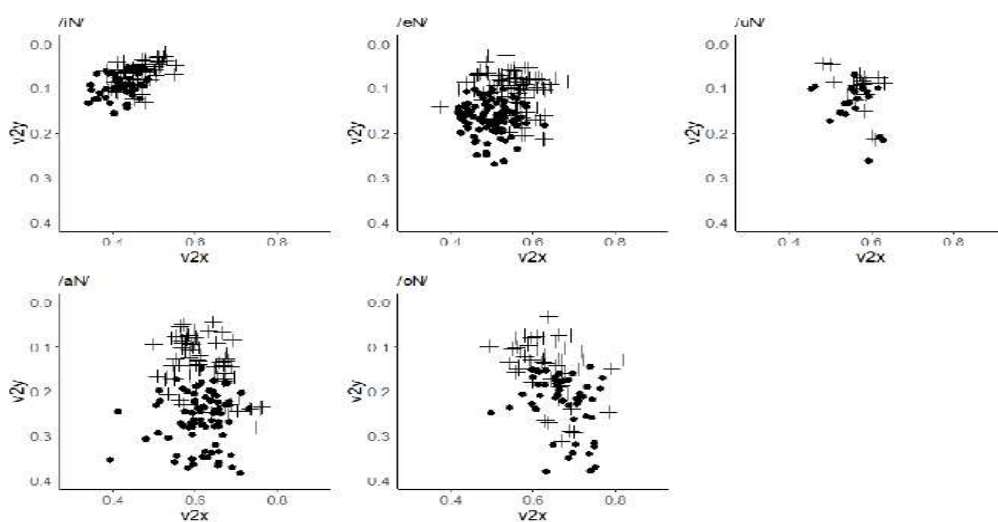


図4: 発話末撥音の声道閉鎖位置（閉鎖の後端、+記号）と先行母音の最高点（●記号）。

<sup>1</sup>ただし第2期科研費研究の申請書で設定した目標は rtMRIDB と MRI Viewer の公開までであり、ここに述べたような RDB 化の計画までは含まれていない。



Maekawa(2021)では、標準話者 11 名（男性 7 名女性 4 名）による各 211 語の発話末撥音の調音位置を、声道サイズの差を正規化したうえで比較した。その結果、正規化された調音空間において先行母音の最高点が分布する領域の直上に、撥音による声道閉鎖の後端が分布していることが確認できた（図 4 参照）。先行母音の種別を説明変数（/a/,/i/,/u/,/e/,/o/のいずれかをとり離散変数）とし、話者と単語をランダム変数とした GLMM（一般化線形混合効果モデル）で発話末撥音の声道閉鎖位置を予測すると、予測値と実測値の間にはほぼ直線的な関係が観察され、両者の相関係数は声道閉鎖の前端位置に関して 0.887、同じく後端に関して 0.868 であることが確認できた（相関係数は leave-one-out 交差検証による値）。一方、語中の撥音に対して後続子音を説明変数として同様の予測を行うと、相関係数は閉鎖区間の前端位置で 0.94、後端位置で 0.95 に達することも判明した。

語中撥音の調音位置に関するデータと比較すると、発話末撥音の調音位置は、硬口蓋から軟口蓋を経て口蓋垂までの広い位置に分布していることが判明した。また子音/k/の調音位置と発話末撥音の調音位置を母音ごとに比較すると（/k/においては後続母音、発話末撥音においては先行母音）、両者の分布には大幅な重複が生じており、国立国語研究所（1990）が論拠とした事実は確認できなかった。結論として、日本語の発話末撥音の調音位置を口蓋垂鼻音とする従来の記述は不正確であることが判明した。なお Maekawa(2021)では、分散分析の結果と GLMM の予測精度の比較に基づいて、発話末撥音の調音位置の変異を語中撥音と同様の条件異音として記述することは合理的でないことも主張した。

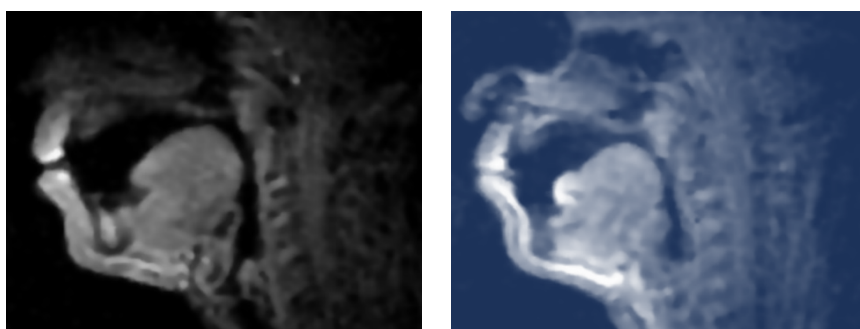


図 5: SPAN-IPA に収録されている両唇軟口蓋接近音[w]のスナップショット。  
話者は左が John Esling（男性）, 右が Pat Keating（女性）。

### 3. 2 ワ行子音の調音位置

IPA（国際音声記号）による日本語の転記には他にも多くの問題がある。前川(2019)ではラ行子音の問題をとりあげた。ここではワ行子音の問題に触れる。日本語ワ行子音の IPA 表記としては、有声両唇軟口蓋接近音の記号[w]が用いられることが多いが、有声軟口蓋接近音の記号[u̟]が用いられることもある。これらはいずれも批判的に検討する必要がある。前者にはワ行子音が本当に二重調音子音であるかという疑問が生じ、後者には軟口蓋と奥舌面だけによるワ行子音の調音が実際に観察されるのかという疑問が生じるからである。

Maekawa (2020)ではこれらの問題を rtMRIDB を用いて予備的に検討した。この研究では、南カリフォルニア大学の研究グループが公開した SPAN-IPA データベースを比較対象として利用した。このデータベースは著名な音声学者 4 名による IPA 分節音調音を記録したデータベースである ([https://sail.usc.edu/span/rtmri\\_ipa/](https://sail.usc.edu/span/rtmri_ipa/))。図 5 に SPAN-IPA に収録されている IPA の[w]の rtMRI 動画のスナップショットを 2 例示す。音声信号から[w]の区間と判断される時点を選択した。両唇と軟口蓋における二重調音が明瞭に観察できる。

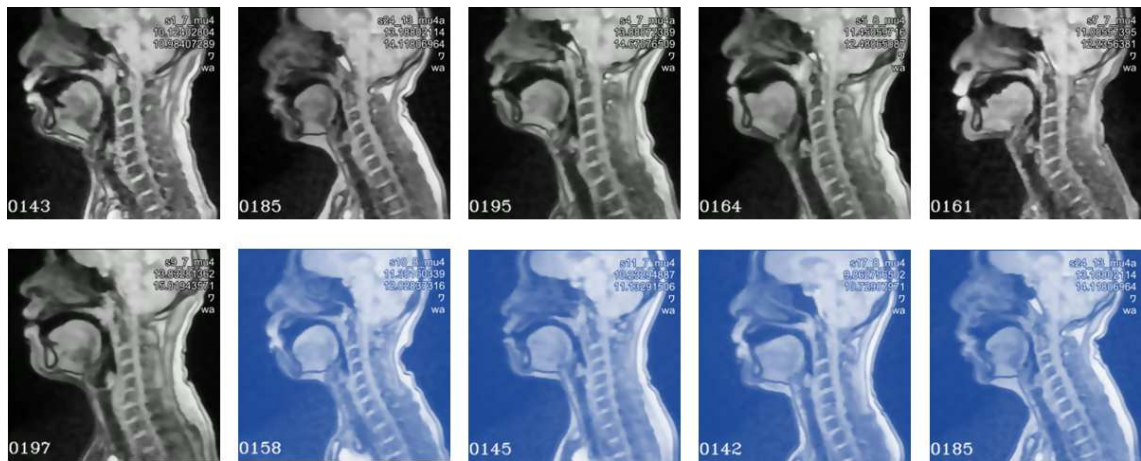


図6: rtMRIDB 試験公開版に収録された標準語話者 10 名の単独モーラ「ワ」の子音部スナップショット。右下の 4 枚 (色調が青のもの) は女性話者



図7: rtMRIDB 試験公開版に収録された標準語話者 10 名の単独モーラ「ファ」の子音部スナップショット。話者の配置は図6と同一



図8: rtMRIDB 試験公開版に収録された標準語話者 10 名の単独モーラ「フ」の母音部スナップショット。話者の配置は図6,7と同一

一方、図6は、rtMRIDB 試験公開版に収録されている標準語話者10名が発音した単独モーラ「ワ」の子音部である。両唇が接近し口蓋垂が挙上されたタイミングを選択した。図6を図5と比較すると、前者（「ワ」）では後者（IPA[w]）におけるほど明瞭な軟口蓋と奥舌面による狭めが観察されないように思えるが、個人差も大きく必ずしも明瞭でない。両者の差を検討するために、同じ話者が発音した「ファ」の子音部（図7）および「フ」の母音冒頭部（図8）との比較を行う。「ファ」の測定タイミングは「ワ」と同一であり、「フ」では音響信号で母音の始端を過ぎた直後のフレームを選択した。

図6-8の比較から以下の結論が得られる。まず「ワ」の子音部と「ファ」の子音部の声道形状はきわめて類似性が高く、視覚的に両者を識別することは困難である。「ファ」の子音が無声両唇摩擦音[ɸ]であることを前提として受け入れるならば、「ワ」の子音においても、軟口蓋と奥舌面による狭めは形成されていないとみることが合理的である。

次に口蓋と奥舌面の距離を「ワ」の子音部と「フ」の母音部とで比較すると、前者の方が距離が大きい。ワ行子音が二重調音接近音であるならば、母音よりも強い狭めが生じるはずであり、口蓋と奥舌面の距離は母音よりも小さくなるはずであるから、やはりワ行子音では軟口蓋での狭めは形成されていない（二重調音ではない）とみることが合理的である。また同じ理由によって、ワ行子音を軟口蓋接近音[wɥ]とみなすことはきわめて不合理であると判断できる。

結論として、日本語（標準語）のワ行子音は、両唇にだけ狭めをもつ接近音（すなわち有声両唇接近音）とみることが最も合理的であると考えられる。ただし、残念ながら、現在のIPA子音チャートにはそのような分節音を表す記号は準備されていないので、実務に際しては、記号[w]を「軟口蓋の狭めはほとんど形成されない」というような注釈とともに利用するのが現実的だろう。<sup>2</sup>

#### 4. その他の研究

最後に2節、3節では紹介しきれなかった本プロジェクトの成果に触れる。調音音声学の理論上の問題点のひとつは、その分類体系が声道正中矢状面の情報に強く依存して構成されており、冠状断面や横断面の情報を、気流通路のcentral / lateralの区別を例外として、ほぼ無視している点にある。

本プロジェクトでは、無声摩擦子音の生成における冠状面形状の関与を示すための物理実験および空力シミュレーションをおこない、冠状断面形状と呼気流量を調整することで、同じ調音位置をもつ摩擦音の音韻カテゴリを変更できることを示した(Yoshinaga, Maekawa & Iida 2021)。

また個人差が大きいことが知られている口蓋の立体形状を定量的に把握して、音声生成への影響を考察するために、歯科用印象剤を用いた口蓋印象を作成して3Dスキャナで数値データ化する作業も進めている。本スペシャルセッションでは、担当者である能田由紀子氏らによる発表が予定されている。

さらに rtMRI データを音声合成に利用する可能性も検討しており（丹治・大村・桂田 2021）、本スペシャルセッションでは桂田浩一氏らによる発表が予定されている。rtMRI データに含まれるのは正中矢状面の情報だけであるから、そこからどの程度までの音声合

<sup>2</sup> もちろん国際音声学協会に新しい記号の導入を要請することも考えられる。

成が可能であるかを探ることは、とりもなおさず、上に述べた調音音声学の理論上の問題点を定量的・実験に評価する試みとしても位置づけられると考えている。

## 5. おわりに

2021年3月1日に米国南カリフォルニア大学の Speech Production and Articulation Knowledge Group (SPAN) が話者 75 名のデータからなるリアルタイム MRI データのデータベースを公開した (<https://sail.usc.edu/span/75speakers/>)。3.2 節で利用した SPAN-IPA と同じ研究グループによる成果である。昨年の本ワークショップのために執筆した前川他(2020)の結論として、rtMRIDB について「組織的に収集された多人数の rtMRI データがデモンストレーション目的を超えて研究用に公開されるのは、おそらくこれが世界初の機会になると思われる」という予想を述べたが、見事に先を越されてしまったわけである。もちろん、これによって日本語音声データとしての rtMRIDB の価値が失われるわけではないから、計画どおりに rtMRIDB の開発は進めていく予定である。

思うに、SPAN グループによるデータ公開は、調音音声学を含む音声研究が新しいステージに入りつつあることを示すモニュメンタルな業績になるのではなかろうか。今後、日本語・英語以外の多くの言語の rtMRI データが収集されて公開されれば、調音音声学は確実に研究の新しいステージに入ることになる。そしてそこでは、調音運動の主観的観察だけに依拠した検証不可能な議論は、その有効性を大幅に制約されることになると思われる。

**謝辞：**本研究は、日本学術振興会科学研究費（17H02339 および 20H01265、いずれも代表者は前川）、国立国語研究所「コーパスアノテーションの拡張・統合・自動化に関する基礎研究」（代表者：浅原正幸）、および令和 2・3 年度人間文化研究機構・機構長裁量経費により実施しました。

## 文献

- Boersma, Paul & Weenink, David (2021). *Praat: doing phonetics by computer* [Computer program]. Version 6.1.51, retrieved 22 July 2021 from <http://www.praat.org/>
- International Phonetic Association (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press.
- Koiso, Hanae, Yasuharu Den, Ken'ya Nishikawa, Kikuo Maekawa (2014). "Design and development of an RDB version of the Corpus of Spontaneous Japanese". *Proceedings LREC 2014*, 1471-1476.
- Lee, Akinobu & Tatsuya Kawahara (2009). "Recent Development of Open-Source Speech Recognition Engine Julius". *Proceedings : APSIPA ASC 2009 : Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*, 131-137.
- Maekawa, Kikuo (2003). "Corpus of Spontaneous Japanese: Its Design and Evaluation". *Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003)*, Tokyo, 7-12.
- Maekawa, Kikuo (2019). "A real-time MRI study of Japanese moraic nasal in utterance-final position." *Proceedings of ICPHS 2019*, Melbourne, 1987-1991.
- Maekawa, Kikuo (2020). "Remarks on Japanese /w/". *ICU Working Papers in Linguistics (ICUWPL)*, 10, 45-52. (<http://id.nii.ac.jp/1130/00004625/>).
- Maekawa, Kikuo (2021). "Production of the utterance-final moraic nasal in Japanese: A real-time MRI study". *Journal of the International Phonetic Association*. DOI:10.1017/S0025100321000050.

- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. “NIH Image to ImageJ (2012). “25 years of image analysis”. *Nature Methods*, 9(7), 671–675. DOI:10.1038/nmeth.2089.
- Takemoto, Hironori, Tsubasa Goto, Yuya Hagihara, Sayaka Hamanaka, Tatsuya Kitamura, Yukiko Nota & Kikuo Maekawa (2019). “Speech organ contour extraction using real-time MRI and machine learning method.” *Proc. INTERSPEECH 2019*, Graz, 904-908. DOI: 10.21437/Interspeech.2019-1593
- Vance, Timothy (2008). *The sounds of Japanese*. Cambridge: Cambridge University Press.
- Yoshinaga, Tsukasa, Kikuo Maekawa & Akiyoshi Iida (2021). “Aeroacoustic differences between the Japanese fricatives [e] and [ç]”. *The Journal of the Acoustical Society of America*, 149, 2426; <https://doi.org/10.1121/10.0003936>
- 浅井拓也・菊池英明・前川喜久雄(2018).「調音運動動画アノテーションシステムの開発」日本音響学会 2018 年秋季研究発表会講演論文集, 1235-1238.
- 天野沢海・並木崇宏・宮川翔多・後藤翼・竹本浩典・北村達也・能田由紀子・前川喜久雄 (2021).「日本語話者 20 名の rtMRI 動画における発話器官の輪郭抽出」日本音響学会 2021 年春季研究発表会講演論文集, 745-746.
- 川上泰(1977).『日本語音声概説』桜楓社.
- 国立国語研究所(1990).『日本語の母音, 子音, 音節: 調音運動の実験音声学的研究』国立国語研究所報告 100, 秀英出版.
- 後藤翼・荻原裕也・濱中彩夏・竹本浩典・北村達也・能田由紀子・前川喜久雄(2019).「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」日本音響学会 2019 年春季研究発表会講演論文集, 822-823.
- 後藤翼・天野沢海・竹本浩典・北村達也・能田由紀子・前川喜久雄(2020).「rtMRI 動画から抽出した発話器官の輪郭データに基づく平均声道の生成と分析」日本音響学会 2020 年秋季研究発表会講演論文集, 821-822.
- 丹治涼・大村英史・桂田浩一(2021).「real-time MRI 動画像を用いた音声合成システムの作成」日本音響学会 2021 年春季研究発表会講演論文集, 1-8-14.
- 吐師道子・児玉明菜・高倉祐樹・林良子(2014).「日本語語尾撥音の調音実態: X 線マイクロビーム日本語発話データベースを用いて」音声研究, 18 (2), 95-105.
- 服部四郎(1951)『音声学』岩波書店.
- 前川喜久雄・能田由紀子・北村達也・竹本浩典・石本祐一(2018).「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」日本音響学会 2018 年春季研究発表会講演論文集,1247-1248.
- 前川喜久雄(2019)「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 会第 33 回全国大会予稿集, 98-103.
- 前川喜久雄・西川賢哉・浅井拓也・能田由紀子・正木信夫・島田育廣・竹本浩典・北村達也・斎藤純男・籠宮隆之・石本祐一・菊池英明・藤本雅子・八木豊(2020).「リアルタイム MRI 動画日本語調音運動データベースの設計」言語資源活用ワークショップ 2020 発表論文集, 国立国語研究所コーパス開発センター, 209-230. DOI/10.15084/00003139