

MRI を用いた音声生成系の観測

北村 達也

甲南大学



1. はじめに

音声研究会では、音声に関する技術的、科学的な研究について議論している。本研究会は日本音響学会との共催ではほぼ毎月開催されており、その半数以上は他研究会とも共催になっている。研究会では毎回テーマが設定され、その研究動向に関する招待講演及び一般講演で研究会が構成される。2009年度は、「感情音声、韻律、声質、音声生成、知覚、脳機能 (2009年6月)」、「福祉と音声処理 (2009年10月)」、「クロスモーダル (2010年1月)」などのテーマが設定された [1]。本稿では、2009年6月のテーマに関連して、磁気共鳴画像法 (Magnetic Resonance Imaging, MRI) を用いた音声生成系の観測技術 [2], [3] について概説し、観測例を紹介する。

2. 音声生成系の観測における MRI の利用

MRI は水素の原子核 (プロトン) の核磁気共鳴現象を利用して生体を構成する部位や要素を画像化する技術である。X線を利用する技術と異なり、被曝の心配なく身体の任意断面を可視化できる点が MRI の長所である。この非侵襲性は、健常者を対象にして研究を行う際には必要不可欠な条件である。

MRI の音声生成系観測への利用は 1980 年代後半から始まり、以来 20 年以上が経過する。その過程で、様々な技術が開発されてきた。以降ではそれらの技術について紹介する。

2.1 動態観測

MRI は時間分解能が低いいため、一般に動画を撮像するのが難しい。しかし、心臓の拍動のよう

に同じ運動が周期的に繰り返される場合は、位相をずらしながら撮像することによって一連の運動の動画を得ることができる。発話運動の場合も、被験者に同じ発話を繰り返させ、位相をずらしながら撮像することによって、発話動態を観測できる。この手法を発話同期撮像法という。ただし、被験者は発話資料を同じタイミング、同じ動作で 100 回程度繰り返す必要があり、その負担は小さくない。

この手法は、連続発話における声道形状及び声道断面積関数 (声門から口唇までを結ぶ声道中心線に直交する断面の面積を示したデータ) の時間変化など、従来知り得なかった多くの知見をもたらしている。現在、この撮像法は発話の観測のみならず、臨床研究、リハビリテーションなどにも利用されている。

この撮像法は静止画撮像にも応用された。一般に MRI による静止画撮像には数分を要するが、被験者がその間に吸気することによって発話器官が動き、画像にぶれが生じてしまう。そこで、動態観測と同様に被験者に持続発話を繰り返させて発話中のみ撮像を行う手法が開発され、画像の S/N 向上に大きく寄与した。

2.2 高分解能撮像

一般の MRI 撮像の空間分解能は $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ 程度である。これは通常の分析には十分であるが、喉頭のような粘膜に囲まれた微細な軟骨を含む部分や、口唇のように筋繊維が複雑に絡み合う構造を有する部分の観測では、通常よりも細かい空間分解能が求められる。一般に、

撮像対象の空間分解能を上げようとする、観測対象のボクセルサイズは小さくなり、そこから出力される信号強度は弱まる。これを補うためには信号を受信するアンテナ (コイル) の感度を高くする必要がある。

そこで、ATR (国際電気通信基礎技術研究所) にて高感度小型表面コイルが開発された。このコイルを用いることにより、 $0.19\text{mm} \times 0.19\text{mm}$ の画像が取得できるようになった。このコイルは、音声の基本周波数 (ピッチ周波数) の変化に伴う甲状軟骨の動きの観測、口唇の突き出しに伴う筋肉の動きの観測、下咽頭腔形状の個人内変動の観測などを可能とし、それぞれの検討から多くの知見が得られている。

2.3 歯列形状観測

発話器官において歯は重要な要素の一つであるが、MRI では歯が空気と同じく低輝度に撮像され、歯の境界を抽出できない。この問題点を克服するために、マンガ含有量の多い飲料を口腔内に含んだ状態で MRI 撮像を行い、その飲料と歯の信号値の差を利用して歯列形状を抽出する方法が提案されている。そして、別に撮像しておいた発話時の MRI データに歯列をスーパーポーズすることによって、歯列を考慮した声道形状が得られるようになった。この手法を利用して、摩擦子音の発音原理を探る研究などが行われている。

3. 観測例：感情音声発話時の声道形状

ここでは、MRI を用いた研究の一例として、激怒、喜び、悲しみといった感情を込めた発話における声道形状を MRI で観測した結果を示す [4]。近年、音声科学では音声の中の非言語情報 (文字に書き起こせない情報) に注目が集まってきており、この研究もその一環である。声道形状を知ることによって、感情による声質の差異が生じるメカニズムを知ることができる。

この MRI 観測においては、上記の感情が生じ

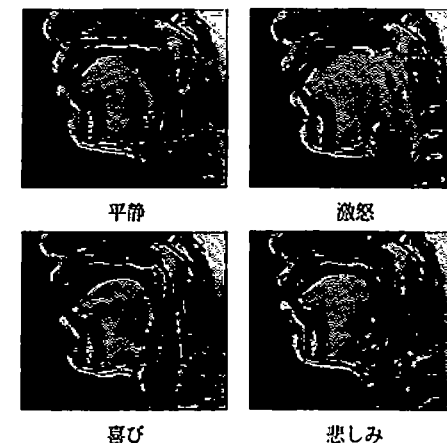


図 1. 平静、激怒、喜び、悲しみの感情下における「ええ」発話時の正中矢状断面 (身体を左右に分ける断面) における MRI データ

るようなシチュエーションの対話文を見せ、その中の「ええ」の部分が発話するよう話者に依頼している。図 1 は女性話者の観測結果であるが、同じ「ええ」でも声道形状が大きく異なることがわかる。この声道形状の違いが声質の違いを生み出す一つの要因となっている。

4. おわりに

音声生成系の研究は、医学、生物学、物理学、言語学、脳科学、工学など様々なバックグラウンドを持つ研究者がそれぞれの知識を活かしつつ協力して研究を進めている。ぜひ多くの人にこの分野に参入していただき、新しい研究を始めていただきたいと考えている。本稿がそのきっかけとなれば著者にとって望外の喜びである。

参考文献

- [1] 音声研究会 (<http://www.ieice.org/iss/sp/jpn/>)
- [2] 竹本浩典, 北村達也: "MRI に基づく音声生成の研究手法の概要," 信学技報 (SP), Vol.109, No.99, pp.11-16 (2009).
- [3] 筒本時彦ら: 音声生成の計算モデルと可視化, コロナ社 (2010).
- [4] 北村達也: 情動による声道形状変化の MRI 観測, 日本音声学会全国大会, pp.45-50 (2009).