

# $F_0$ 変化に伴う声道形状変化の観測

北村 達也<sup>†</sup> パーハムモクタリ<sup>†</sup>

<sup>†</sup> ATR 人間情報科学研究所 〒 619-0288 「けいはんな学研都市」光台 2-2-2

E-mail: †{kitamura,parham}@atr.jp

あらまし  $F_0$  の変化にともなう母音/a/および/i/発声時の声道形状の変化を 3 次元 MRI を用いて観測した。成人男性 1 名に 110, 123, 130, 146, 164 Hz の純音を呈示し, それに  $F_0$  を合わせて持続発声させて MRI 撮像を行った。声道断面積関数を分析した結果, 母音/a/において  $F_0$  変化に伴う口腔の断面積関数の変化が顕著であり,  $F_0$  とこの領域の体積との間に強い負の相関があることを示した。また, 母音/i/においては  $F_0$  の変化に伴って咽頭腔の断面積関数の変化が顕著であり,  $F_0$  と咽頭腔の体積との間に負の相関があることを示した。

キーワード 声道, 声道断面積関数, ピッチ周波数 ( $F_0$ ), 主成分分析

## Measurement of changes of vocal tract shape by $F_0$ shift

Tatsuya KITAMURA<sup>†</sup> and Parham MOKHTARI<sup>†</sup>

<sup>†</sup> ATR Human Information Science Laboratories 2-2-2 Hikaridai, “Keihanna Science City” Kyoto 619-0288 Japan

E-mail: †{kitamura,parham}@atr.jp

**Abstract** Effects of  $F_0$  shift in vocal tract shape were analyzed by volumetric magnetic resonance imaging (MRI). One male subject performed sustained productions of Japanese vowel /a/ and /i/ with being asked to adjust these  $F_0$  to 110, 123, 130, 146, and 164 Hz pure tone. The results of comparison of vocal tract area functions extracted from the MR images revealed that  $F_0$  and area function of the oral cavity show a strong negative correlation for vowel /a/ and  $F_0$  and area function of the pharyngeal cavity show a negative correlation for vowel /i/.

**Key words** vocal tract, vocal tract area function, pitch frequency ( $F_0$ ), principle component analysis

### 1. はじめに

音声中には話者ごとに異なる情報, すなわち個性が含まれており, 我々は音声を聞くことにより話者を識別することができる。音声の個性は声帯音源と声道の音響特性に多く含まれるといわれているが, 発声される音声のピッチ周波数 ( $F_0$ ) や音量などが変化すればこれらの音響特性は変化してしまう。このような話者内変化は音声認識技術や話者認識技術の高精度化を困難にする要因の一つであると考えられる。

従って, 話者内変化が生じるメカニズムを解明し, その影響の範囲を予測することは重要な意味を持つと考えられる。しかし, 音声生成系における話者内変化に関してこれまで十分な検討が行われてきたとは言い難い。そこで, 本研究では種々の話者内変化のうち  $F_0$  の変化が声道形状に与える影響に焦点を絞り, 3 次元核磁気共鳴画像法 (MRI) を用いて観測を行う。

$F_0$  の変化と声道形状の関係については, 平井ら [1] が MRI

による観測結果を示している。彼らは,  $F_0$  を約 1.5 オクターブの範囲で下降させて母音/a/を発声したときの声道断面形状を示している。また, Tom ら [2] は, electron beam computed tomography (EBCT) を用いて声区 (register) を 2 種類,  $F_0$  を 5 段階, ラウドネスを 2 段階に変化させた場合の声道断面積関数を比較した。しかし, いずれの研究でも母音が 1 種類に限定されている。母音によって  $F_0$  変化に伴うフォルマント周波数への影響が異なることが指摘されているので [3],  $F_0$  による影響が現れる声道上の部位が異なる可能性がある。

そこで, 本研究では 3 次元 MRI を用いて  $F_0$  変化に伴う母音/a/および/i/発声時の声道形状変化を観測する。被験者に 110, 123, 130, 146, 164 Hz の純音を呈示し, それに  $F_0$  を合わせて持続発声させて MRI 撮像を行う。得られた 3 次元 MR 画像から声道断面積関数を抽出し,  $F_0$  変化との関連について分析する。

## 2. MRI 撮像条件

撮像に用いた MR 装置は、ATR 脳活動イメージングセンタに設置された島津 Marconi 社製 MAGNEX ECLIPSE 1.5T Power Drive 250 である。

撮像にはネックコイルを用いた。シーケンスは Fourier Acquired Steady State (FAST) 法, TE = 3 ms, TR = 10 ms, FA = 10 度。撮像領域は 256 × 256 mm, 分解能は 512 × 512 pixels である。スライスは矢状方向で, 厚み 3 mm, 間隔 3 mm, 18 スライスで撮像した。加算回数は 1 回で, 撮像時間は約 20 秒である。

被験者は成人男性 1 名 (26 歳) である。タスクは母音/a/ および/i/の持続発声である。被験者にヘッドフォンにて 110, 123, 130, 146, 164 Hz の純音を呈示し, その周波数に  $F_0$  を合わせて発声させた。

この他, MR 画像に歯列画像を補填するため, Takemoto ら [4] による方法で口腔造影剤 (ブルーベリージュース) を用いた撮像も行った。その際のシーケンスは Fast Spin Echo 法, TE = 9 ms, TR = 4900 ms, FA = 90 度。撮像領域は 256 × 256 mm, 分解能は 512 × 512 pixels である。スライスは矢状方向で, 厚み 1.5 mm, 間隔 1.5 mm, 51 スライスで撮像した。加算回数は 1 回で, 撮像時間は約 3 分である。

## 3. 正中矢状断面の分析

$F_0$  が 110 Hz と 164 Hz のときの正中矢状断面の MR 画像を対象にして声道形状の差異を調査した。これらの  $F_0$  で母音/a/ および/i/を発声したときの MR 画像をそれぞれ図 1, 2 に示す。

声道の変化のうち最も顕著なものは声門位置の変化であり,  $F_0$  が 164 Hz のときは 110 Hz のときと比較して母音/a/で 5 mm, 母音/i/で 6 mm 上昇した。この上昇は声道長を短くさせる効果がある。

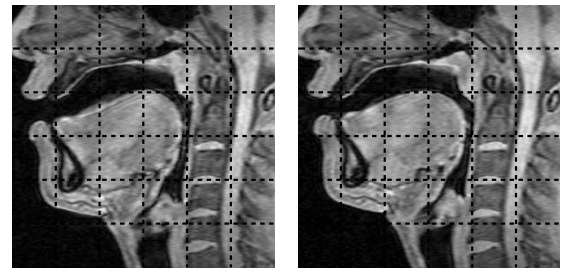
また, 母音/a/では  $F_0$  の上昇によって次のことが生じている。第 1 に下顎が上昇する。前鼻棘 (Anterior Nasal Spine: ANS) と下顎骨の下端との距離を比較したところ,  $F_0$  の上昇に伴って 3 mm 上昇していた。第 2 に, 舌背が上方に持ち上げられている。これらの影響によって,  $F_0$  が 164 Hz の場合に前室 (front cavity) の体積が減少している。この点は平井ら [1] が示した結果と一致する。

母音/i/では  $F_0$  の上昇により咽頭腔の幅 (MR 画像の  $y$  軸方向) が変化する。第 II・III 椎間円板を通る水平断面における咽頭腔の幅は, ピッチ周波数が 110 Hz の場合で 13 mm, 164 Hz の場合で 11 mm であった。

## 4. 声道断面積関数の分析

### 4.1 声道断面積関数の抽出

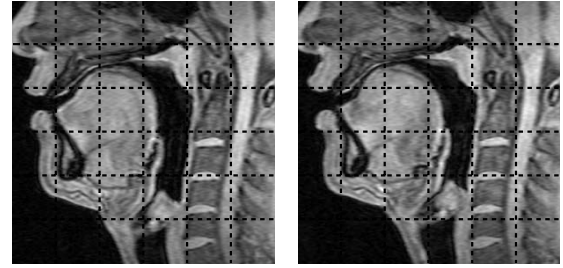
3 次元 MR 画像から声道断面積関数を抽出した。まず, MR 画像を 0.5 × 0.5 × 0.5 mm/voxel のポリウムデータに変換した。このポリウムデータに歯列のポリウムデータを Takemoto ら [4] による方法で補填した。次に, 正中矢状断面において声道中心線をひき, その線に直交する断面を 2.5 mm



(a)  $F_0=110$  Hz

(b)  $F_0=164$  Hz

図 1 母音/a/を  $F_0=110$  Hz および 164 Hz にて発声したときの正中矢状断面の MR 画像。



(a)  $F_0=110$  Hz

(b)  $F_0=164$  Hz

図 2 母音/i/を  $F_0=110$  Hz および 164 Hz にて発声したときの正中矢状断面の MR 画像。

間隔で抽出した。そして, 閾値処理に目視による修正を加えて声道領域を抽出し, 断面積を求めた。

### 4.2 声道断面積関数の比較

母音/a/および/i/の  $F_0$  変化に伴う声道断面積関数の変化をそれぞれ図 3, 4 に示す。母音/a/の声道断面積関数では  $F_0$  変化に伴う口腔の断面積関数の変化が顕著であり, 母音/i/では咽頭腔の断面積関数の変化が顕著であった。

そこで, これらの領域の体積を定量的指標として  $F_0$  変化との関係を調査した。得られた結果を図 5 に示す。母音/a/に関しては  $F_0$  と口腔領域の体積の間に強い負の相関がみられる ( $R^2=0.89$ )。母音/i/についても負の相関がみられる ( $R^2=0.65$ )。

また,  $F_0$  変化に伴う声道長の変化を表 1 に示す。従来の報告と同様に  $F_0$  の上昇とともに声道長が短くなる傾向が見られた。

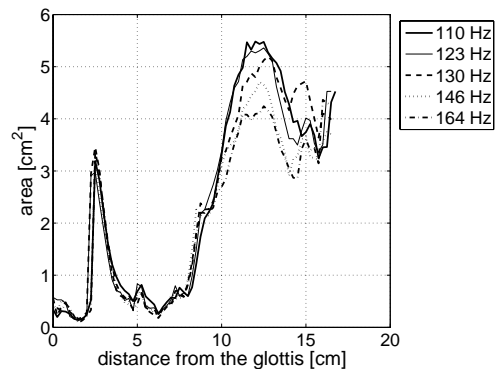


図 3  $F_0$  と声道断面積関数の関係 (母音/a/)。

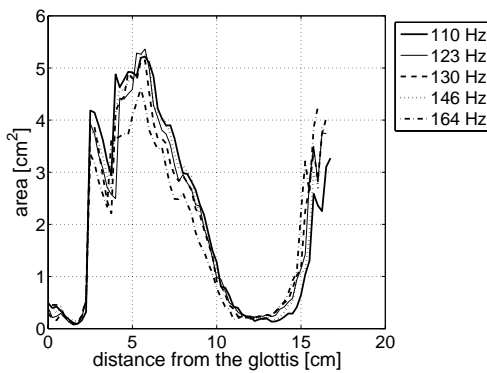


図 4  $F_0$  と声道断面積関数の関係 (母音/i/).

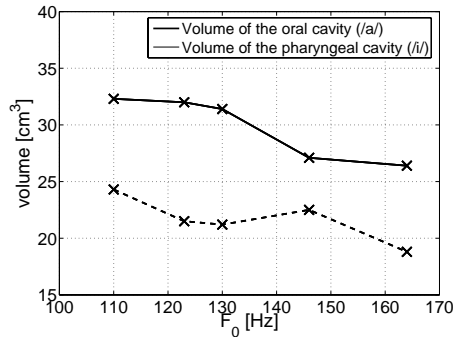


図 5 母音/a/発声時の  $F_0$  と口腔領域の体積との関係 (実線) および母音/i/発声時の  $F_0$  と喉頭腔領域の体積との関係 (波線)

表 1  $F_0$  と声道長の関係.

$F_0$	/a/	/i/
110 Hz	166 mm	168 mm
123 Hz	164 mm	167 mm
130 Hz	164 mm	167 mm
146 Hz	162 mm	166 mm
164 Hz	160 mm	162 mm

### 4.3 声道断面積関数の主成分分析

$F_0$  変化に伴う声道断面積関数の変化の主要因を明らかにするため、Story and Titze [5] の方法によって声道断面積関数の主成分分析を行った。彼らはこの手法で声道断面積関数の音韻間の変化について分析したが、本研究では 5 段階の  $F_0$  による変化について分析する。

彼らの手法では、声道断面積関数を以下のように平均と変化成分で表すことができると仮定する。

$$A(x, f) = A_0(x) + \alpha(x, f) \quad (1)$$

ここで、 $A(x, f)$  はある  $F_0 = f$  における声道断面積関数、 $A_0(x)$  は全ての  $F_0$  に関する平均声道断面積関数、 $\alpha(x, f)$  は変化成分である。ただし、声道断面積関数の要素数  $N$  は一定にしておく。本研究では  $N = 44$  となるように声道断面積関数をキュービックスプライン法で内挿した。そして、断面積の平方根の領域で主成分分析を行う。

以上の方法で母音ごとに分析を行った結果、母音/a/では第 1 主成分の寄与率が 55.2 %、第 2 主成分の寄与率が 27.7 % となった。また、母音/i/では第 1 主成分の寄与率が 58.6 %、第

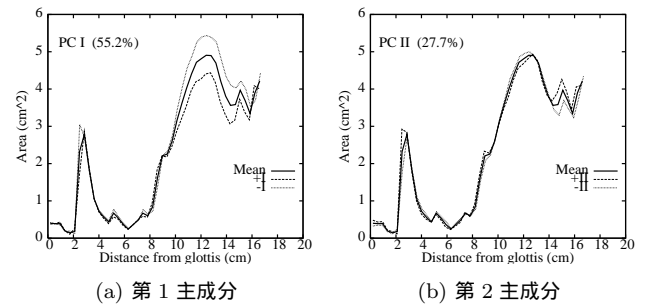


図 6 母音/a/の第 1, 第 2 主成分の変化に伴う声道断面積関数の変化.

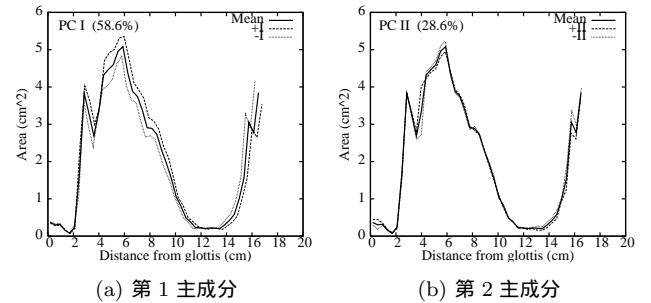


図 7 母音/i/の第 1, 第 2 主成分の変化に伴う声道断面積関数の変化.

2 主成分の寄与率が 28.6 % となり、いずれの母音でも累積寄与率が第 2 主成分までで 80 % を超えた。

母音/a/の第 1, 第 2 主成分の変化に伴う声道断面積関数の変化を図 6 に示す。これらの図から、第 1 主成分は口腔領域の変化、第 2 主成分は口唇領域の変化に対応すると考えられる。

同様に、母音/i/の第 1, 第 2 主成分の変化に伴う声道断面積関数の変化を図 7 に示す。これらの図から、第 1, 第 2 主成分はともに咽頭腔と口唇の領域の変化に対応すると考えられる。また、第 1 主成分の変化はせばめの位置にも影響を与えている。

### 4.4 声道伝達特性の分析

図 3 および 4 に示した声道断面積関数から等価回路モデル [6] [7] を用いて伝達特性  $H(\omega) = U_{out}(\omega)/U_{in}(\omega)$  を計算し、ホルマント周波数を求めた。ここで、 $\omega$  は周波数、 $U_{out}(\omega)$  は放射端の体積速度、 $U_{in}(\omega)$  は入力端の体積速度である。伝達特性を 4.5 kHz までの周波数帯域で計算し、第 1 から第 4 ホルマント周波数をピークピッキングにより求めた。

放射インピーダンス  $Z_R$  は Causséら [8] により提案された以下の式で近似した。

$$Z_R/\rho c = z^2/4 + 0.0127z^4 + 0.082z^4 \ln z - 0.023z^6 + j(0.6133z - 0.036z^3 + 0.034z^3 \ln z - 0.0187z^5), \quad (2)$$

$$z = kr. \quad (3)$$

ここで、 $c$  は音速、 $\rho$  は空気密度、 $k$  は波数、 $r$  は放射端の半径である。ただし、この式は  $kr < 1.5$  で有効である。母音/a/および/i/の全ての声道断面積関数の放射端の半径および上限周波数 (4.5 kHz) はこの条件を満たす。本研究では、音速  $c$  および空気密度  $\rho$  としてそれぞれ  $\rho = 1.12 \text{ kg/m}^3$ 、 $c = 352.9 \text{ m/sec}$  を用いた。

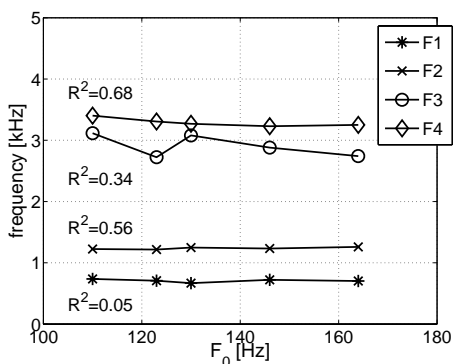


図 8  $F_0$  とホルマント周波数との関係 (母音/a/).

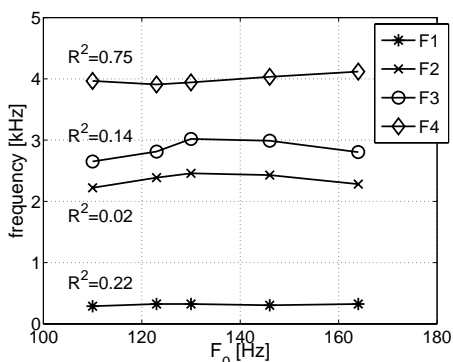


図 9  $F_0$  とホルマント周波数との関係 (母音/i/).

母音/a/および/i/発声時の  $F_0$  の変化に伴う伝達特性の変化をそれぞれ図 8, 9 に示す. 母音/a/では  $F_0$  と  $F_2$  との間に正の相関,  $F_0$  と  $F_4$  との間に負の相関がみられ, 母音/i/では  $F_0$  と  $F_4$  との間に正の相関がみられるが, それ以外のホルマント周波数では  $F_0$  との相関は弱かった.

## 5. 考察

母音/a/の声道断面積関数では  $F_0$  変化に伴う口腔の断面積関数の変化が顕著であり, 母音/i/では咽頭腔の断面積関数の変化が顕著であった. この結果は,  $F_0$  変化に伴って変化が顕著な声道上の部位は母音によって異なることを示している.

Story and Titze [5] の方法にもとづく声道断面積関数の主成分分析の結果から, 母音/a/の第 1 主成分は口腔領域の変化, 第 2 主成分は口唇領域の変化に対応することが示された. このことは, 母音/a/に関する  $F_0$  変化に伴う声道形状の変化は, 口腔領域の変化をもとらず下顎および舌の変位と, 口唇の変位で表されることを示唆している. また, 母音/i/では第 1, 第 2 主成分が咽頭腔と口唇の領域の変化に対応することが示された. 咽頭腔の変形は喉頭の変位と関連があるのではないかと推察されるが, これを明らかにするのは今後の課題である.

$F_0$  の変化が声道伝達特性におけるホルマント周波数に与える影響の分析では, 母音/a/において  $F_0$  と  $F_2$  との間に正の相関がみられた. 母音/a/の前室の体積の減少は  $F_2$  の上昇を引き起こすので, この結果は妥当なものといえる.

また, 母音/a/に関して  $F_0$  と  $F_4$  との間に負の相関, 母音/i/において  $F_0$  と  $F_4$  との間に正の相関がみられた.  $F_4$  の周波数

は喉頭腔共鳴との関連が指摘されているので [9] [10],  $F_0$  の変化に伴う喉頭腔の形状の変化がこの相関に関連している可能性がある.

## 6. おわりに

本研究では,  $F_0$  の変化にともなう声道形状の変化を 3 次元 MRI を用いて観測した. 成人男性の被験者 1 名に 110, 123, 130, 146, 164 Hz の純音を呈示し, それに  $F_0$  を合わせて母音/a/および/i/を持続発声させて MRI 撮像を行った. MR 画像から作成したボリュームデータに歯列のボリュームデータを補填した後, 声道断面積関数を抽出した. そして, 声道断面積関数および声道伝達特性の  $F_0$  の変化による影響を分析した.

その結果, 本研究の被験者に関しては, 母音/a/において  $F_0$  変化に伴う口腔の断面積関数の変化が顕著であり,  $F_0$  とこの領域の体積との間に強い負の相関があることを示した. また, 母音/i/においては  $F_0$  の変化に伴って咽頭腔の断面積関数の変化が顕著であり,  $F_0$  と咽頭腔の体積との間に負の相関があることを示した.

さらに, Story and Titze [5] の方法によって声道断面積関数の主成分分析を行ったところ, 母音/a/の第 1, 第 2 主成分はそれぞれ口腔, 口唇の変化に対応することが示された. また, 母音/i/では第 1, 第 2 主成分が咽頭腔と口唇の領域の変化に対応することが示された.

謝辞 本研究は総務省平成 16 年度戦略的情報通信研究開発推進制度により実施したものである. 歯列画像補填および声道断面積関数の抽出には当研究所 竹本浩典氏により作成されたプログラムを利用した. ここに感謝の意を表す.

## 文献

- [1] 平井, 党, 本多, “舌と喉頭との相互作用を考慮した発話器官の生理学的モデル,” 日本音響学会誌, Vol. 51, No. 12, pp. 918–928, 1995.
- [2] Tom, Titze, Hoffman, and Story, “Three-dimensional vocal tract imaging and formant structure: Varying vocal register, pitch, and loudness,” *J. Acoust. Soc. Am.*, Vol. 109, pp. 742–747, 2001.
- [3] 本多, “基本周波数変化に依存するホルマント周波数のシフト,” 信学技報, SP86-122, pp.55–62, 1987.
- [4] Takemoto, Kitamura, Nishimoto, and Honda, “A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions,” *Acoust. Sci. Tech.*, Vol. 25, No. 6, pp. 468–474, 2004.
- [5] Story and Titze, “Parameterization of vocal tract area functions by empirical orthogonal modes,” *J. Phonetics*, Vol. 26, pp. 223–260, 1998.
- [6] Flanagan, *Speech analysis synthesis and perception 2nd Edition* Springer-Verlag, Berlin Heidelberg New York, 1972.
- [7] Adachi and Yamada, “An acoustical study of sound production in biphonic singing, Xöömij,” *J. Acoust. Soc. Am.*, Vol. 105, pp. 2920–2932, 1999.
- [8] Caussé, Kergomard, and Lurton, “Input impedance of brass musical instruments – comparison between experiment and numerical models,” *J. Acoust. Soc. Am.*, Vol. 75, pp. 241–254, 1984.
- [9] 竹本, 本多, 正木, 島田, 藤本, “3 次元 MRI 動画データに基づく声道下部構造のモデル化,” 音講論 (秋), pp. 281–282, 2003.
- [10] 本多, 北村, 竹本, 藤田, モクタリ, “下咽頭腔の共鳴特性,” 音講論 (秋), pp. 235–236, 2004.