社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

破裂子音および摩擦子音発声時の声道容積変化の計測

北村　達也†　波多野博顕†

† 甲南大学知能情報学部　〒 658–8501 兵庫県神戸市東灘区岡本 8–9–1
E-mail: †t-kitamu@konan-u.ac.jp

あらまし　有声，無声の破裂子音および摩擦子音生成時の声道容積の変化を計測した．成人男性 1 名を対象にしてこれらの子音を含む 3 モーラの無意味語生成時の 3 次元 MRI 動画を観測し，声道内に閉鎖もしくはせばめが形成されている時間区間の声道容積を計測した．その結果，有声破裂子音/d/生成時の声道閉鎖時の声道容積はほぼ単調に増加し，その値は常に無声破裂子音/t/よりも大きかった．また，有声破裂子音/d/と有声摩擦子音/z/ の生成時の声道容積を比較すると，前者の方が最大値と増加幅が大きいという結果が得られた．

キーワード　有声子音，無声子音，破裂子音，摩擦子音，声帯振動，声道容積

Measurement of temporal cange of vocal tract volume during production of plosive and fricative consonants

Tatsuya KITAMURA† and Hiroaki HATANO†

† Faculty of Intelligence and Informatics, Konan University　8–9–1, Okamoto, Higashinada, Kobe, Hyogo, 658–8501 Japan
E-mail: †t-kitamu@konan-u.ac.jp

Abstract　The volume of the vocal tract of a male speaker during production of voiced and voiceless plosives and fricatives was measured directly from magnetic resonance imaging (MRI) data. Three-dimensional cine-MRI data of three-morae non-sense words were obtained by a synchronized sampling method, and the temporal change of the vocal tract volume was measured while there was a closure or a constriction at the alveolar. The results showed that the volume of the vocal tract for the voiced plosive /d/ increased almost monotonically, and the volume was larger than that for the voiceless plosive /t/ through the closure section. The maximum value and rise range of the vocal tract volume for the voiced plosive /d/ is greater than that for the voiced fricative /z/.

Key words　voiced consonants, voiceless consonants, plosive consonants, fricative consonants, vibration of the vocal fold, volume of the vocal tract

1. はじめに

有声子音と無声子音の生成における差異は音源のみにとどまらず，調音動態にも及ぶ．その一例として，有声子音生成時に声帯振動を維持するために咽頭腔の容積が拡大するという報告がある [1] [2] [3] [4]．しかし，従来の研究は X 線透過画像に基づき咽頭後壁と舌根の距離を計測したものが多く，声道容積を直接計測した研究はほとんど見られない．そこで，本研究では磁気共鳴画像法 (magnetic resonance imaging, MRI) を用いて有声，無声の破裂子音および摩擦子音の調音動態を観測し，これらの子音が生成される際の声道容積変化を比較する．

声帯振動が持続的に生じるためには，声門の上下に圧力差が必要である [5]．通常，声帯振動が生じる際には声門下の圧力が声門上の圧力に対して高くなる．ところが，破裂子音や摩擦子音などの発話時には，口腔内や口唇に閉鎖もしくはせばめが生じ，呼気の流出が阻害される．従って，有声の破裂子音や摩擦子音などの生成時には，この状態で声門上下の圧力差が維持される必要がある．

先行研究では，これらの子音の生成中に声道容積の拡大することによって声門上下の圧力差が維持されていることを示唆する結果が報告されている．いずれも X 線動画を用いた研究で，透過画像に基づき分析を行っている．Perkell [1]，および Kent & Moll [2] は有声破裂子音の声道閉鎖区間で咽頭後壁と舌根の距離が時間とともに増加することを報告した．また，Westbury [3] は，X 線透過画像から声道容積変化を推定し，推定された声道容積の増大により声帯振動が持続可能なことを示

した．

Bell-Berti [6] は，破裂音の発声に伴う咽頭腔拡大に関連する筋活動を計測した．その結果，咽頭腔拡大には能動的および受動的な筋活動が寄与していることが確認されたが，その活動パターンは 3 名の実験協力者間でまちまちであった．

党ら [7] は，咽頭腔容積の増大と軟口蓋のスティフネスの変化を考慮することによって有声破裂子音の生成をシミュレートできることを示した．このシミュレーションにより得られた声道各部の音圧波形は彼らの実測データとよく一致することが確かめられている．

X 線による透過画像を用いていた先行研究に対して，Proctor ら [4] は MRI を用いて声道の 3 次元形状を計測した．彼らは有声および無声の摩擦子音を持続的に生成したときの咽頭腔および喉頭腔の形状を比較し，有声摩擦子音生成時には咽頭腔が拡大することを示した．

以上の先行研究により，有声子音生成時に声道容積が増加している可能性は高いと考えられる．もしこの声道容積変化を実測することができれば，それが声帯振動の維持に十分か否かを声帯モデルを用いて検証し，有声子音生成メカニズムの理解を深めることができる．そこで，本研究では有声，無声の破裂子音および摩擦子音生成時の 3 次元 MRI 動画を撮像し，声道容積変化を計測する．

2. 方　　法

2.1 実験協力者

東京方言の 40 代の成人男性 1 名である．

2.2 発 話 資 料

発話資料は表 1 に示す 4 種類で，これらは有声，無声の歯茎破裂子音もしくは歯茎摩擦子音を含む 3 モーラの無意味語である．筆者らが以前撮像した MRI 動画 [8] において有声軟口蓋破裂子音/g/は閉鎖調音が弱化しがちであり，破裂子音というよりむしろ摩擦子音と呼ぶべき調音をしていたため，本研究の分析対象から軟口蓋破裂子音 (/g, k/) を除外した．なお，/g/における閉鎖調音の弱化に関しては前川 [9] も同様の指摘をしている．また，MRI では歯列が造影されないため，閉鎖の手前 (声門側) に歯列が含まれる両唇破裂子音 (/b, p/) も分析対象から外した．

表 1　発 話 資 料
Table 1　Speech materials

	Voiced	Voiceless
Plosive	/adada/	/atata/
Fricative	/azaza/	/asasa/

2.3 MRI 撮像条件

本研究では，実験協力者が上記の発話資料を生成している最中に，彼の頭頸部の右半側を MRI を用いて撮像した．

撮像に用いた MRI 装置は，ATR-Promotions 脳活動イメージングセンタに設置されている Siemens 社製 MAGNETOM Verio である．

表 2　MRI の撮像条件
Table 2　Conditions of MRI measurement.

Echo time (TE)	1.62 ms
Repetition time (TR)	11 ms
Flip angle (FA)	15°
Field of view (FOV)	256×256 mm
Pixcel size	1×1 mm
Slice	Sagittal
Slice thickness	5.0 mm
Number of average	1
Frame rate	90 frames/s

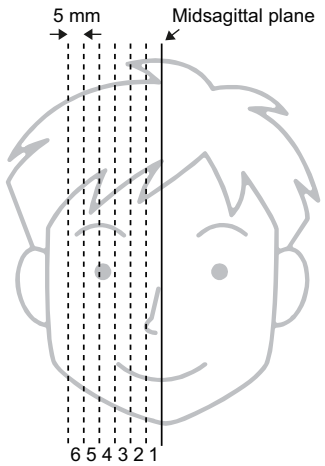


図 1　撮像対象の矢状断面
Fig. 1　Sagittal planes for measurement.

撮像は発話同期撮像法 [10] にて行った．この撮像法は，実験協力者が同じタイミングかつ同じ動作で発話を繰り返す必要がある．そこで，実験協力者には MRI 装置内において MRI 用イヤホンにより発話ガイド音を提示した．このガイド音は，0.1 s の純音 1 回の後に 0.1 s のバーストノイズ 2 回が各 0.3 s 間隔で接続されたものである．このガイド音は，撮像の間，実験協力者に繰り返し提示される．実験協力者には，純音の後の 1 つ目のバーストノイズのタイミングで第 1 モーラを発話し，2 つ目のバーストノイズのタイミングで第 3 モーラを発話するよう指示した．すべての発話資料は平板アクセントで発話された．

撮像パラメータを表 2 に示す．この条件では，1 度の撮像で 3 断面の動画を得ることができる．そこで，図 1 のように正中矢状断面から 5 mm 間隔で 6 枚の断面を配置し，2 回の撮像により実験協力者の頭頸部の右半側 30 mm の動画を得た．これらの断面を正中矢状断面に近い側から断面 1 から断面 6 と呼ぶ．1 回目の撮像で偶数番の断面を撮像し，2 回目の撮像で奇数番の断面を撮像した．

1 回の撮像における発話の繰り返しは 128 回で，撮像には約 4 分を要した．

2.4 分 析 方 法

本研究では 2 つの分析を行った．

第 1 に，調音運動のタイミングの分析を行った．この分析では，断面 1 の画像から目視により舌尖と歯茎部の距離を計測し

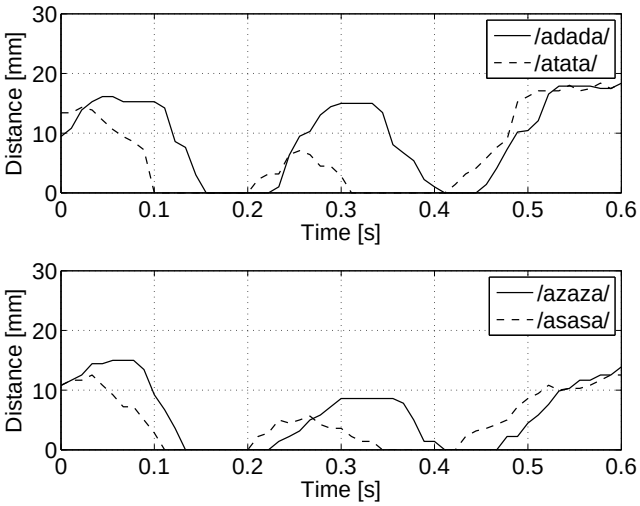


図 2　舌尖と歯茎部の距離．上段：/adada/ (実線)，/atata/ (破線)，下段：/azaza/ (実線)，/asasa/ (破線)
Fig. 2　Distance from the tongue tip to the teethridge in mm.

た．時刻の原点は発話ガイド音の 1 つ目のバーストノイズが開始する時点とした．

第 2 に，声道内に閉鎖もしくはせばめが形成されている時間区間における声道容積変化を計測した．上述の 6 断面のうち声道が含まれる断面を対象にして，声門から閉鎖もしくは狭めまでの声道領域を目視にてトレースした．本研究で得られた MRI 動画には声道の境界があいまいな部分があったため，時間方向と断面方向のトレースの一貫性を保つよう配慮した．MRI データの画素サイズは 1×1 mm，スライス厚は 5 mm であるので，トレースした声道領域の画素数に 5 mm^3 を乗じて声道容積を得た．

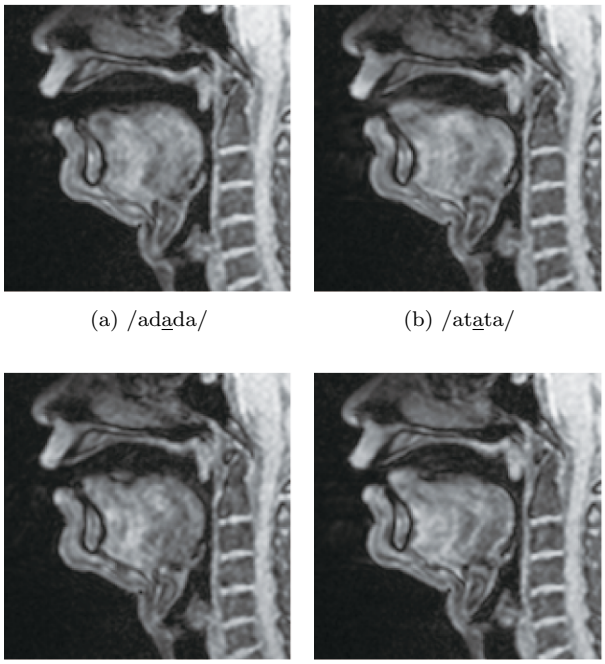
3. 結　　果

3.1　調音運動の比較

4 つの発話資料を発話した際の舌尖と歯茎部の距離を図 2 に示す．上述のように，時刻の原点は発話ガイド音の 1 つ目のバーストノイズが開始する時点である．なお，摩擦子音 (図 2 下段) のせばめ区間で舌尖と歯茎部の距離が 0 になっているのは，MR 画像の空間分解能の問題である．/s/生成時の高空間分解能 3 次元 MRI データによると，声道の最狭窄部の高さは 1 mm を切っている．そのため，本研究の撮像条件では舌尖と歯茎部の距離が正確に計測できなかった．

実験協力者は発話ガイド音にタイミングを合わせて発話を繰り返したにもかかわらず，舌尖と歯茎部の距離の変化は有声子音と無声子音とで大きく異なっている．

第 1 に，無声子音生成時は閉鎖やせばめが形成されはじめるタイミングが有声子音生成時より早い．/atata/と/adada/の 1 つ目の子音において閉鎖が形成されはじめる時刻には約 56 ms (5 フレーム分) の差があり，2 つ目の子音においては約 100 ms (9 フレーム分) の差がある．同じく，/asasa/ と/azaza/の 1 つ目の子音においてせばめが形成されはじめる時刻には約 33 ms (3 フレーム分) の差があり，2 つ目の子音においては約 67 ms (6 フレーム分) の差がある．加えて，無声子音発話時は舌尖が歯茎部から離れはじめるタイミングも有声子音発話時より早い．結果的に，無声子音を含む発話資料では母音の継続時間長が短くなっている．



(a) /adada/　(b) /atata/　(c) /azaza/　(d) /asasa/

図 3　各発話資料の 2 つ目の母音/a/を生成している時の MR 画像 (断面 1)
Fig. 3　MR images of the sagittal plane 1 during production of the second /a/ in /aCaCa/.

第 2 に，/a/発話区間において音素環境により舌尖と歯茎部の距離が異なる．そこで，各発話資料における子音にはさまれた/a/，すなわち 2 番目の/a/発話時の断面 1 の MR 画像を図 3 に示す．/adada/以外では，舌尖の下降が弱化した状態で後舌母音/a/の発話を行っていたことがわかる．

3.2　声道容積の変化

本研究の実験協力者に関しては，断面 1 から 3 に声道が含まれていた．そこで，本研究ではこの 3 断面から声道領域を抽出した．その結果得られた声道の右半側の容積変化を図 4 に示す．時刻の原点は声道内に閉鎖もしくはせばめが形成された時点である．図中の円や四角のマーカーが MRI 動画のフレームに対応している．図 4 におけるフレーム数と図 2 における声道閉鎖またはせばめが形成されたフレーム数が必ずしも一致しないのは，声道内の閉鎖もしくはせばめが正中側からではなく右側から開放される場合があったためである．

以下，本研究の実験協力者の声道が左右対称であると仮定する．/d/生成における声道閉鎖時の声道容積は，約 24 cm^3 から最大約 32 cm^3 までほぼ単調に増加し，その差は約 8 cm^3 であった．一方，/t/生成における声道閉鎖時の声道容積は，約 20 cm^3 から最大約 25 cm^3 までほぼ単調に増加し，その後，平坦になっている．声道閉鎖開始時点と最大値の差は約 5 cm^3 であった．/d/の声道容積は常に/t/よりも大きく，声道閉鎖開始

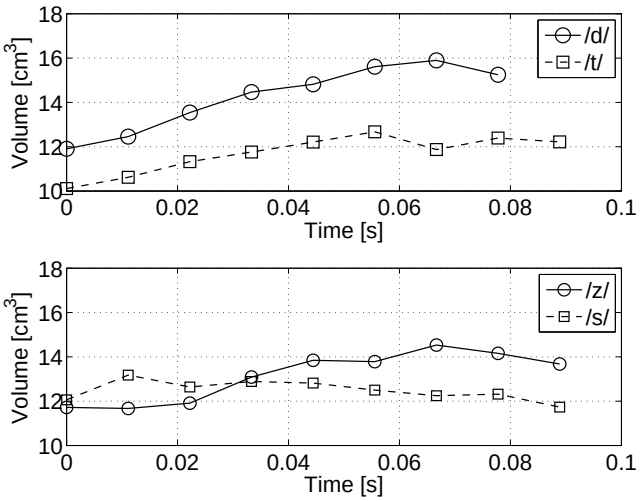


図 4　有声，無声の破裂子音および摩擦子音発話時の声道右半側容積の変化 (上段: /atata/の 1 つ目の/t/ (実線) と/adada/の 1 つ目の/d/ (破線)，下段: /azaza/ (実線) の 1 つ目の/z/と/asasa/の 1 つ目の/s/ (破線))

Fig. 4　Volume change of the right half of vocal tract during production of voiced and unvoiced plosive and fricative consonants.

時点で約 4 cm^3，最大値で約 7 cm^3 の差があった．さらに，/d/の方が声道容積の増加幅も大きかった．なお，/d/生成時の声道容積増加に寄与するのは，主に咽頭腔であった．

/z/生成におけるせばめ形成時の声道容積は，約 24 cm^3 から最大約 29 cm^3 まで増加し，その差は約 5 cm^3 であった．一方，/s/生成におけるせばめ形成時の声道容積は，約 24 cm^3 から最大約 26 cm^3 まで増加し，その後ゆるやかに減少した．

4. 考　　察

本研究の実験協力者においては，有声子音と無声子音とでは調音運動のタイミングが異なり，無声子音では閉鎖やせばめが形成されるタイミングが有声子音に対して先行していた (図 2)．このような差異は他の子音でも報告されている．例えば，Tachibana ら [8] は，軟口蓋破裂子音/k/と/g/の発話において前者の調音のタイミングが早いことを報告している．母音に無声子音が後続する際に母音の時間長が短くなる現象は多くの言語で確認されており [11]，pre-fortis clipping と呼ばれることもある [12]．

声道の左右対称性を仮定すると，本研究の実験協力者の声道容積は約 20 cm^3 から約 32 cm^3 の範囲を変化した．Stevens は，その著書 [5] の 1.1.3.4 節 (24 ページ) にて，下顎を閉じた状態の成人女性の声道容積は約 40 cm^3，成人男性の声道容積は約 60 cm^3 と記しており，上記の値とは異なる．しかし，本研究の実験協力者の頭部が成人男性としては細いことに加え，声道閉鎖形成中の声道容積であることを考慮すると，おおむね妥当な値ではないかと考えられる．

Westbury [3] は破裂子音生成時の声道閉鎖中の X 線動画から声道容積を推定している．それによると，無意味語/padap/中の/d/の生成時には 80 ms の間に約 6 cm^3 の容積増大があると推定されている．また，/batab/中の/t/の生成時には 80 ms の間に容積が増減し，最大で約 3 cm^3 の容積増大があるとされている．これらの値は本研究で得られたものに近い値となっている．

Proctor ら [4] は，/VCV/の文脈で有声，無声摩擦音を生成したときの咽頭腔の容積を比較している．歯茎摩擦子音を/aCa/の文脈で生成した場合，実験協力者 4 名の平均で，有声子音の容積は無声子音の約 1.5 倍であった．しかし，本研究の結果 (図 4 下段) ではそこまでの差異はない．Proctor ら [4] の研究では実験協力者は 36 s 間の持続的な子音生成を行っており，over-articulation が生じた可能性もある．

Rothenberg [13] は，咽頭腔の容積が 1 cm^3 増加すると声帯音源の持続時間が約 10 ms 延びると述べている．これを本研究の/d/の結果にあてはめると，声道容積の増加によって声帯音源の持続時間が約 80 ms 延びることになる．この値は，/d/生成時の声道閉鎖区間の時間長 (図 4 上段) とほぼ一致する．

また，Rothenberg [13] は鼻咽腔結合により声道内の圧力が低下する可能性を指摘した．しかし，他の報告 [1] [2] [3] と同様に，本研究の実験協力者でも閉鎖もしくはせばめの形成中には断面 1〜3 において軟口蓋と咽頭後壁が広い面積で接触しており，鼻腔への空気流出はないと考えられる．

有声子音同士の比較をすると，/d/生成時の方が/z/生成時に対して声道容積の最大値と増加幅が大きかった．/z/生成時にはせばめから呼気が流出するため，声道を拡大しなくても声帯振動が維持できるためであろう．

5. おわりに

本研究では，MRI 動画を用いて有声，子音の破裂子音および摩擦子音生成時の声道容積の時間変化計測を試みた．MRI は非侵襲的に 3 次元計測が可能という大きな利点があるが，欠点も存在する．まず，MRI 動画の撮像においては，実験協力者が正確なタイミングで同じ発話運動を繰り返す必要がある．この条件が満たされない場合には，画像にぶれが生じてしまう．

次に，本研究ではスライス厚を 5 mm に設定したが，この厚さでは断面の厚み方向に声道形状が変化することがある．MRI で得られる画像は断面内の構造を平均したものとなるため，画像のコントラストが低くなることがある．別の言い方をすれば，5 mm の断面のうち例えば 2 mm が声道だった場合，その部分は声道として造影されないことが起こりうる．なお，スライスを厚くすると観測対象からの MR 信号が強くなるという利点があるため，一概にスライス厚を薄くすればよいというものではない．

以上 2 つの理由により，今回得られた MRI 動画には声道の境界部のコントラストが低い部分がある．従って，得られた結果には若干の誤差が含まれると考えられる．本研究の結果において，声道容積が Stevens [5] の記述より小さいのはこのようなことが関係しているかもしれない．今後，このような誤差を見積もる方法の考案や，横断面の撮像による検証などが必要である．

謝辞

本研究は科研費基盤研究 (B)(21300071) にて実施された．発話資料に関するご助言をいただいた北海道医療大学の榊原健一先生，摩擦子音生成時の高空間分解能 3 次元 MRI データをご提供いただいた NICT の竹本浩典氏に感謝いたします．

文　　献

[1] J.S. Perkel, *Physiology of speech production: Results and inplications of a quantitative cineradiographic study*, MIT Press, Cambridge, Massachusetts, London, England (1969).

[2] R.D. Kent and K.L. Moll, Vocal-tract characteristics of the stop cognates, *J. Acoust. Soc. Am.*, 46, 1549–1555 (1969).

[3] J.R. Westbury, Enlargement of the supraglottal cavity and its relation to stop consonant voicing, *J. Acoust. Soc. Am.*, 73, 1322–1336 (1983).

[4] M.I. Proctor, C.H. Shadle and K. Iskarous, Pharyngeal articulation in the production of voiced and voiceless fricatives, *J. Acoust. Soc. Am.*, 127, 1507–1518 (2010).

[5] K.N. Stevens, *Acoustic phonetics*, MIT Press, Cambridge, Massachusetts, London, England (1998).

[6] F. Bell-Berti, Control of pharyngeal cavity size for English voiced and voiceless stops, *J. Acoust. Soc. Am.*, 57, 456–461 (1975).

[7] 党建武, 中井孝芳, 鈴木久喜, 破裂子音における口腔内圧及び放射音の測定とシミュレーション, 音響誌, 49, 313–320 (1993).

[8] R.O. Tachibana, T. Kitamura and M. Fujimoto, Differences in articulatory movement between voiced and voiceless stop consonants, *Acoust. Sci. & Tech.* (in printing).

[9] 前川喜久雄, 日本語有声破裂音における閉鎖調音の弱化, 音声研究, 14, 1–15 (2010).

[10] S. Masaki, M.K. Tiede, K. Honda, Y. Shimada, I. Fujimoto, Y. Nakamura and N. Ninomiya, MRI-based speech production study using a synchronized sampling method, *J. Acoust. Soc. Jpn.(E)*, 20, 375–379 (1999).

[11] K.J. Kohler, Phonetic explanation in phonology: The feature fortis/lenis, *Phonetica*, 41, 150–174 (1984).

[12] J.C. Wells, *Accents of English* Cambridge University Press, Cambridge, New York (1982).

[13] M. Rothenberg, The breath-stream dynamics of simple-released-plosive production, *Bibliotheca Phonetica*, 6, 1–117 (1968).