

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年6月20日現在

機関番号：82636

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21500184

研究課題名（和文）声道と音源の相互作用が音声の個人性に与える影響に関する研究

研究課題名（英文）Study on the effects of acoustic interaction between the vocal tract and vocal fold on individuality of speech sounds

研究代表者

竹本 浩典（TAKEMOTO HIRONORI）

独立行政法人 情報通信研究機構 ユニバーサルコミュニケーション研究所

多感覚評価研究室・専攻研究員

研究者番号：40374102

研究成果の概要（和文）：下咽頭腔は喉頭腔と左右の梨状窩からなり、音声の個人性（声のその人らしさ）の生成要因である。本研究では、まず下咽頭腔を3次元で音響解析し、声道伝達関数に2つの深い零点を生成するメカニズムを解明した。次に声道と音源の相互作用を考慮した声帯振動モデルを構築し、下咽頭腔が音源波形に与える影響を検討した。その結果、下咽頭腔は音源より声道伝達関数により多くの個人性の要因を与えることが明らかになった。

研究成果の概要（英文）：The hypopharyngeal cavities, which consist of the laryngeal cavity and bilateral piriform fossae, might act as a source of individuality of speech sounds. In this study, three-dimensional acoustic analyses of the hypopharyngeal cavities were first performed to clarify the mechanism by which they generate two deep dips on transfer functions of the vocal tract. Then a model of vocal fold vibration that considers acoustic interaction between the vocal tract and vocal fold was developed to examine how the hypopharyngeal cavities affect the glottal source. As a result, compared with their effects on the glottal source, the hypopharyngeal cavities provide transfer functions of the vocal tract with much more factors relating to individuality of speech sounds.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,500,000	450,000	1,950,000
2010年度	1,300,000	390,000	1,690,000
2011年度	700,000	210,000	910,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：知覚情報処理・知能ロボティクス

キーワード：音声生成、音源、下咽頭腔、音声の個人性

1. 研究開始当初の背景

音声の生成過程を説明する代表的な理論は、線形のソース・フィルタ理論である。ソースとは声帯振動によって生じる音源であり、フィルタとは声道形状に固有の音響特性である。線形のソース・フィルタ理論では、

声帯振動と声道形状との間に相互作用がないと仮定し、ソースがフィルタを通過することで音声が生じるとする。この理論に沿って多くの研究がなされてきた。

下咽頭腔は声道の基部であり、喉頭腔と左右の梨状窩からなる（図1参照）。この部分

の形状の個人差は大きい、発話中の形状はほとんど変化しない[1]。また、この部分の形状は、声質に関する声道伝達関数の高い周波数領域（約 2.5 kHz～6 kHz）を特徴づける[2, 3]。これらのことから、下咽頭腔は、フィルタに関する音声の個人性（声のその人らしさ）の生成要因であると考えられている[4]。

近年、声帯振動と声道形状の相互作用に着目した研究が盛んになってきた[5-7]。そして、下咽頭腔の形状が声帯振動パターンに大きな影響を与えることが明らかにされた[5]。すなわち、下咽頭腔の形状はフィルタだけでなく、ソースにも一定した影響を与え、これが音声の個人性の生成要因になっている可能性がある。しかし、下咽頭腔の音響特性についてはまだ不明な点が多く、より詳しい音響解析が必要とされている。

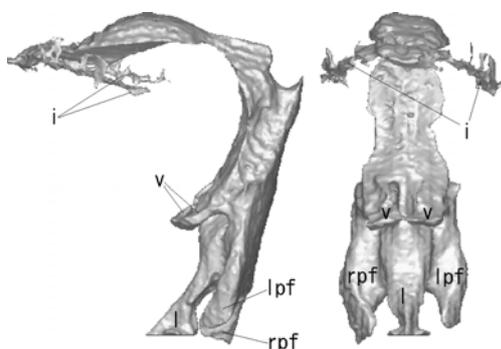


図 1. 声道形状。(左) 左側面観、(右) 正面観。
i: 歯列間隙, v: 喉頭蓋谷, l: 喉頭腔, rpf: 右梨状窩, lpf: 左梨状窩。

2. 研究の目的

背景で述べた従来の研究結果を踏まえて、下咽頭腔が音源にも音声の個人性を与えるという仮説を立てた。本研究ではこの仮説を検証する。検証に先立って、以下の目標に取り組む。

- (1) 音源と声道の相互作用を考慮した声帯振動モデルを構築する。
- (2) 3次元の音響解析によって声道の精密な音響特性を求める。
- (3) 下咽頭腔の音響特性を詳細に検討する。

3. 研究の方法

まず、音源と声道の相互作用を考慮した声帯振動モデルを構築する。これは、声門上部の系である声道と、声門下部の系である気管や肺からの反射関数（系の入り口で発生した音が系の内部で反射することによって得られる応答）[8]に基づいて、声門を通過する体積流や圧力損失を計算し、声帯振動をシミュレートするモデルである。反射関数はそれぞれの系の入力インピーダンスを計算することで求めることができる。

次に、3次元の音響解析手法の一つである、時間領域差分法によって、声道の精密な音響特性を求め、上記の声帯振動モデルを駆動するために必要な反射関数を計算する。まず、MRIを用いて声道の精密な形状を計測し、時間領域差分法によって音響特性を計算する。次に、計測した声道形状を光造型法によって実体化して声道模型を作成し、音響特性を実測する。そして、計算した音響特性を実測したものと比較してその精度を検証する。

次に、下咽頭腔の音響特性を詳細に検討する。下咽頭腔の音響特性は、これまで縦続音響管モデルを用いて検討されてきた。しかし、縦続音響管モデルは平面波伝播のみを仮定した近似的なモデルであり、平面波伝播が仮定できなくなる高い周波数領域では有効でない。また、実際の音声スペクトルでは、左右の梨状窩に由来すると考えられる2つの零点の深さは母音によって著しく異なっているが、これは縦続音響管モデルによるシミュレーションでは再現されない。そこで、3次元の音響解析手法を用いて下咽頭腔の音響特性を詳細に検討する。

そして最後に、声帯振動モデルと3次元の音響解析によって得られた反射関数を用いて、下咽頭腔が声帯振動に与える影響を検討する。ある声道形状に対して、下咽頭腔の形状に微小変動を与えて音源波形とスペクトルの変化を観測し、下咽頭腔が音源に与える影響を特定する。そして、声帯音源に個人性が含まれるかどうか、出力される音声の個人性にどのような影響があるか検討する。

4. 研究成果

まず、声門上下の系の音響的影響を考慮した声帯振動モデルを構築した（雑誌論文⑥）。このモデルでは、従来の声帯の2質量モデルや2次元モデルを改良し、声道からの反射関数だけでなく、気管や肺からの反射関数も考慮できるようにした。また、従来のモデルでは、声道を縦続円筒管で近似して反射関数を求めていた[8]。しかし、このモデルでは、3次元の音響解析手法である時間領域差分法によって求めた反射関数を組み込むことができる。これによって、下咽頭腔の共鳴と反共鳴はもとより 10 kHz までの高い周波数領域で反射関数を精密に計算することができる。また、気管や肺からの反射関数は、Harperらの実測値に基づくモデル[9]を用いて計算した。これによって、従来は声門下部の圧力は一定であると仮定されていたが、反射関数の影響により、500 Pa以上の圧力変動が生じ、その結果、例えば母音/a/では声門体積流がRMSで17%も減少することが明らかになった（図2）。なお、これらの結果は、Titzeの結果[5]とほぼ一致した。

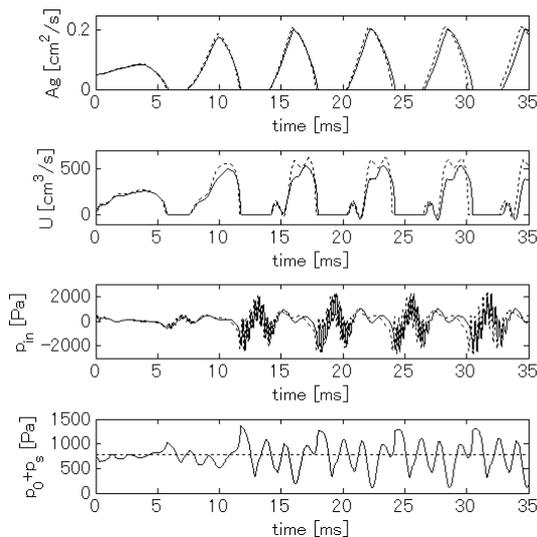


図 2. 上段から声門面積, 声門体積流, 声道入り口の圧力, 声門下圧. 実線: 声門下の系を考慮した場合, 点線: 声門下の系を考慮しない場合.

次に、3次元の時間領域差分法によって、MRIで計測した声道の音響特性を求め、同じ声道形状を光造型法で実体化した声道模型の実測値と比較した。その結果、計算した音響特性は、実測したものと10 kHzまでよく一致した(雑誌論文④)。そして、声道の3種類の分岐管(歯列間隙、喉頭蓋谷、梨状窩)が伝達関数に与える影響の大きさを定量的に求め、下咽頭腔の一部である梨状窩の影響がもっとも大きいことを示した(表1)。また、その過程で、左右の梨状窩に音響的な相互作用があることを発見した。左右の梨状窩は伝達関数上に2つの零点を生じさせるが、どちらか一方の梨状窩の形状を変動させると、2つの零点の周波数と深さが共に変動した。これは、左右の梨状窩が独立してそれぞれ1つの零点を生成しているのではなく、相互に影響を及ぼしながら2つの零点を生成することを示す結果となった。

表 1. 梨状窩(P)、喉頭蓋谷(V)、歯列間隙(I)をそれぞれ閉塞したときと開放したときの平均スペクトル間距離(dB)。/a/と/o/では歯列間隙は存在しない。

	/a/	/i/	/u/	/e/	/o/
P	6.0	5.9	7.1	6.5	6.0
V	3.1	1.8	2.9	2.3	3.4
I	-	0.4	3.3	1.2	-

さらに、3次元の時間領域差分法による音響解析によって、左右の梨状窩の相互作用を詳しく追及した(雑誌論文②、⑤)。その結果、(1)左右の梨状窩が生成する2つの零点の周波数は、左右の梨状窩がそれぞれ単独で生成する2つの零点のいずれより一方は高く、

他方は低い、(2)左右の梨状窩は、周波数が低い方の零点では逆相で振動し、高い方の零点では同相で振動する、(3)周波数が高い方の零点のほうが低い方の零点より深い、(4)左右の梨状窩がそれぞれ単独で生成する零点の周波数が同じ場合は、左右の梨状窩が同相で振動する零点が1つだけ生成される、ということが明らかになった。これは、梨状窩に由来する2つの零点は、左右の梨状窩を2つの振動子とする連成振動系の2つの基準振動(同相と逆相の振動)によって生成されることを示している(図3参照)。なお、従来の縦続音響管モデルでは、左右の梨状窩はそれぞれ音響的に独立した2つの分岐管として近似的に実装されるため、このような相互作用を表現することができない。そのため、下咽頭腔に由来するスペクトル構造、特に梨状窩による2つの零点の周波数や深さが実際と大きく異なっていることが明らかになった。すなわち、下咽頭腔がソースやフィルタに及ぼす影響を検証するためには、縦続円筒管モデルでは不十分であり、3次元の音響解析手法を用いる必要がある。

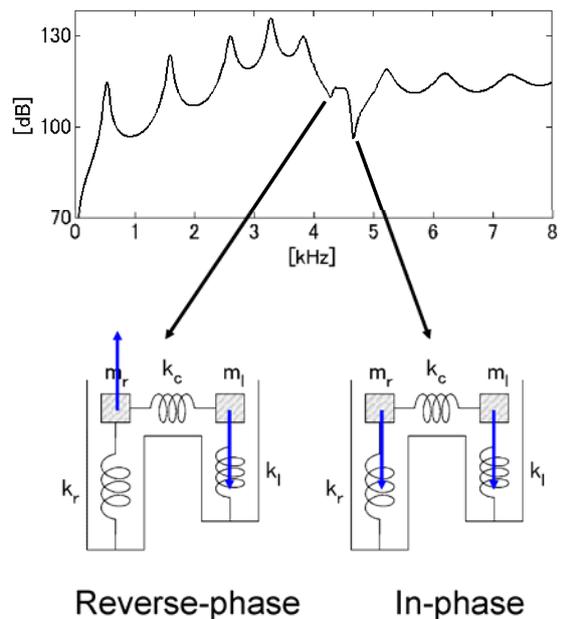


図 3. 左右の梨状窩が生成する2つの零点(上)と、各零点における振動モードを示す模式図(下)。 m_r 、 m_l は質量、 k_r 、 k_l 、 k_c はバネ定数。低い周波数の零点では、左右の梨状窩が逆相で振動し、高い周波数の零点では同相で振動する。すなわち、2つの零点は基準振動によって生成される。

以上の研究成果を踏まえて、下咽頭腔が音源に音声の個人性を与えるという仮説の検証を試みた。下咽頭腔の形状を共通にして5母音の声道形状を作成し、下咽頭腔の形状に変動を与えたときの反射関数を3次元の時間領域差分法によって計算した。そしてこの反

射関数を用いて、声門上下の系の音響的影響を考慮した声帯振動モデルを駆動し、音源波形とスペクトルの変化を検討した。その結果、確かに下咽頭腔は音源波形に一定した影響を与えるが、その影響は下咽頭腔以外の声道が与える影響に比べてきわめて小さく、聴感上ほとんど知覚できなかつた。これは、少なくとも声道と声帯との音響的相互作用が比較的小さい、すなわち基本周波数が低い音声の生成過程では、仮説が棄却されることを示す結果となった。

参考文献

- [1] Kitamura, T., Honda, K., and Takemoto, H., "Individual variation of the hypopharyngeal cavities and its acoustic effects," *Acoust. Sci. & Tech.* 26, pp.16-26, 2005.
- [2] Dang, J. W., and Honda, K., "Acoustic characteristics of the piriform fossa in models and humans," *J. Acoust. Soc. Am.* 101, pp.456-465, 1997.
- [3] Takemoto, T., Adachi, S., Kitamura, T., Mokhtari, P., and Honda, K., "Acoustic roles of the laryngeal cavity in the vocal tract resonance," *J. Acoust. Soc. Am.* 120, pp.2228-2238, 2006.
- [4] Honda, K., Takemoto, H., Kitamura, T., Fujita, S., and Takano, S., "Exploring human speech production mechanisms by MRI," *IEICE Trans. Inf. Syst.* E87-D, pp.1050-1058, 2004.
- [5] Titze, I. R. "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* 123, pp. 2733-2749, 2008.
- [6] Titze, I. R., Riede, T., and Popolo, P. "Nonlinear source-filter coupling in phonation: Vocal exercise," *J. Acoust. Soc. Am.*, 123, pp.1902-1915, 2008.
- [7] Titze, I. R., and Worley, A. S. "Modeling source-filter interaction in belting and high-pitched operatic singing," *J. Acoust. Soc. Am.*, 126, pp.1530-1540, 2009.
- [8] Adachi, S., and Yu, J. "Two-dimensional model of the vocal fold vibration for sound synthesis of voice and soprano singing," *J. Acoust. Soc. Am.*, 117, pp.3213-3224, 2005.
- [9] Harper, P., Kraman, S. S., Pasterkamp, H., and Wodica, G. R. "An acoustic model of the respiratory tract," *IEEE Transactions on Biomedical Engineering*, 48, pp.543-550, 2001.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計7件)

- ① 竹本浩典、本多清志、斉藤毅、羽石英里、大前拓也、島田育廣、藤本一郎、西本博則、岸本宏子、MRIで計測したソプラノ歌

手の声道の音響解析の試み、日本音響学会講演論文集、査読無、2011年9月、2011、pp.313-316

- ② Hironori Takemoto, Seiji Adachi, Parham Mokhtari, Tatsuya Kitamura, Acoustic interaction between bilateral piriform fossae, Proceedings of the 9th International Seminar on Speech Production (ISSP 2011)、査読有、2011、pp.305-312
- ③ 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、電子情報通信学会誌、査読有、94、2011、pp.585-590
- ④ Hironori Takemoto, Parham Mokhtari, Tatsuya Kitamura, Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method, Journal of the Acoustic Society of America, 査読有、128 (6), 2010, pp.3724-3738.
- ⑤ 竹本浩典、足立整治、Parham Mokhtari、北村達也、左右の梨状窩の音響的相互作用、電子情報通信学会技術報告書(音声)、査読無、110 (297)、2010、pp.7-12
- ⑥ 竹本浩典、足立整治、田部洋祐、声門上下の系の音響的影響を考慮した声帯振動モデル、電子情報通信学会技術報告(音声)、査読無、109 (451)、2010、pp.63-68
- ⑦ 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(招待論文)、電子情報通信学会技術報告(音声)、査読無、109 (99)、2009、pp.11-16

[学会発表] (計5件)

- ① 竹本浩典、本多清志、斉藤毅、羽石英里、大前拓也、島田育廣、藤本一郎、西本博則、岸本宏子、MRIで計測したソプラノ歌手の声道の音響解析の試み、日本音響学会、2011年9月21日、島根大学松江キャンパス(島根県)
- ② Hironori Takemoto, Seiji Adachi, Parham Mokhtari, Tatsuya Kitamura, Acoustic interaction between bilateral piriform fossae, the 9th International Seminar on Speech Production (ISSP 2011)、2011年6月21日、ケベック大学モントリオール校(カナダ)
- ③ 竹本浩典、足立整治、Parham Mokhtari、北村達也、左右の梨状窩の音響的相互作用、電子情報通信学会音声研究会、2010年11月18日、愛知県立大学(愛知県)
- ④ 竹本浩典、足立整治、田部洋祐、声門上下の系の音響的影響を考慮した声帯振動モデル、電子情報通信学会音声研究会、2010年3月5日、芝浦工業大学(東京都)

- ⑤ 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（招待講演）、電子情報通信学会音声研究会 2009年6月24日、北海道大学（北海道）

〔その他〕

ホームページ等

<http://www2.nict.go.jp/univ-com/mcc/people/takemoto/>

6. 研究組織

(1) 研究代表者

竹本 浩典 (TAKEMOTO HIRONORI)

独立行政法人 情報通信研究機構

ユニバーサルコミュニケーション研究所

多感覚・評価研究室・専攻研究員

研究者番号：40374102

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

北村 達也 (KITAMURA TATSUYA)

甲南大学知能情報学部知能情報学科・教授

研究者番号：60293594

足立 整治 (ADACHI SEIJI)

Fraunhofer institute for building

physics, Department of Acoustics・

Researcher

研究者番号：90305516

モクタリ パーハム (MOKHTARI PARHAM)

独立行政法人 情報通信研究機構

ユニバーサルコミュニケーション研究所

多感覚・評価研究室・専攻研究員

研究者番号：00395089

田部 洋祐 (TANABE YOSUKE)

株式会社 日立製作所 日立研究所

機械研究センター・研究者

研究者番号：なし