

平成21年 6月 23日現在

研究種目：若手（B）

研究期間：2006～2008

課題番号：18700193

研究課題名（和文）発話運動の連続性と音声の離散性に関する研究

研究課題名（英文）Study on the relationship between continuous speech articulation and the quantal nature of speech sounds

研究代表者

竹本浩典（TAKEMOTO HIRONORI）

株式会社国際電気通信基礎技術研究所 認知情報科学研究所・研究員

研究者番号：40374102

研究成果の概要：本研究では、緩やかな発話運動から急峻に変化する音声が生産される非線形過程について、声道内の音響感度分布と時間領域で音声を生成する発話シミュレータを用いて検討を行った。その結果、母音の連続発話では、特定のホルマントに影響が大きい部分がそうでない部分に先行して運動することで急峻なホルマント遷移を実現していること、歯列間隙による零点がホルマント遷移に不連続を生じさせていることが明らかになった。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,500,000	0	1,500,000
2007年度	1,200,000	0	1,200,000
2008年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,600,000	270,000	3,870,000

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：音声生成，ホルマント，声道形状，MRI，量子性

1. 研究開始当初の背景

緩やかな発話運動から、急峻に変化する音声が生産される非線形的な過程は、音声の量子性と呼ばれ[1]、音声の生成における重要な性質であるとされる。しかし、音声生成は体内で行われる運動であり、可視化や計測が困難であるため、音声の量子性に関する研究は、簡単なモデルを用いたシミュレーションによるものが主体であり[2]、十分に検討されているとはいえない。また、音声の量子性の程度や範囲については、長い間議論されてきたが、未だに決着を見ていない。

声道の音響解析の手法は高度化され、さま

ざまな現象を扱うことが可能になった。例えば、声帯から口唇までの断面積の変化の関数である声道断面積関数から、微小な断面積変化がスペクトルに与える変動量の大きさを予測する、音響感度特性（Sensitivity Function）を計算する手法がある。これを用いることにより、特定のホルマントに対して、声道のどの部分が音響的に重要であるかを定量的に評価できる。また、声道断面積関数の時間パターンを入力値として、連続音声を合成する手法も開発され、発話運動と出力される音声の関係をシミュレートできるようになった。

しかしその一方で、発話運動の観測技術の進歩は遅れていた。そのため、声道の精密な形状データや、喉頭動態まで含めた発話運動の観測データは不足しており、音声生成に関する研究のボトルネックとなっていた。

近年、磁気共鳴画像法(MRI)を用いて、母音などの持続発声可能な音韻の声道立体形状を精密に観測する手法(発話同期撮像法)と、単語発話中の発話運動を断面の動画として観測する手法(MRI 動画)が開発された[3]。前者によって、声道の音響特性のより精密な解析が可能になり、例えば喉頭腔が一つのホルマント(主としてF4)を誘導するなどの新しい知見が得られた[4]。後者では、喉頭も含めた声道全体の運動を動画として記録することが出来るため、発話運動に関する多様な時間パラメータを抽出することが可能になった。また、MRI 動画撮像を多断面で行うことにより、声道立体形状の時間パターンを記録する3次元MRI 動画撮像法が開発され、声道共鳴の時間パターンも実測できるようになった。このように、音声の量子性を検討するための観測データが得られるようになった。

2. 研究の目的

本研究は、緩やかな発話運動から急峻に変化する音声が生産される非線形過程、すなわち音声の量子性を、MRI による発話運動の観測と音響シミュレーションによって検討する。現在、量子論の効果が見られるといわれているのは、母音の/a/から/i/, および子音の/s/から/ʃ/への移行である。これらの音韻それぞれは、発話器官の位置に対して比較的安定している。しかし、例えば/a/から/i/に移行するときは、発話運動に比して急激な音響的变化が生じ、その移行は量子的であるとされている。本研究では、このような量子的な移行が存在するのか、そのメカニズムは何なのか、を検討する。また、それ以外の音韻で量子的な移行が存在するかどうかとも検討する。

3. 研究の方法

まず、母音生成時の声道形状をMRI で計測し、声道軸上に沿う断面積の微小変化とスペクトルの変化量との対応関係を示す音響感度特性を計算する。また、声道形状を3次元的に抽出し、これに壁をつけて光造形法によってエポキシ樹脂で成型した「声道模型」を作成し、この模型の伝達関数を実測することにより、厳密な声道の音響特性を取得する。

次に、時間領域で音声生成をシミュレートする発話シミュレータを構築する。これは、発話運動の量と出力される音響信号の変化を対比するためである。

そして、短い単語(例えば/aka/, /ai/, /ama/, /aiueo/)の発話運動をMRI 動画撮像法により観測し、音響感度特性が高い部分がどのよ

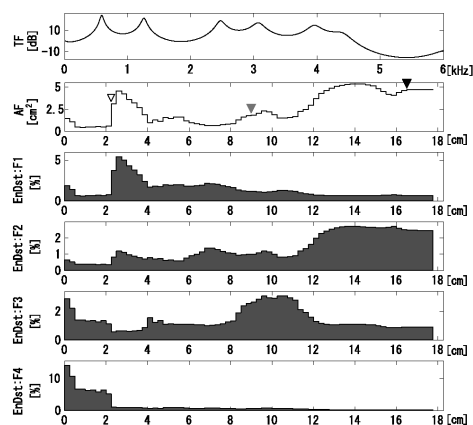


図1. 母音/a/の音響感度特性。

うに運動しているかを観測する。そして、構築した発話シミュレータを用いて、緩やかな発話運動から急峻な音響変動が起こる要因を追求する。そして、音声の量子性を検討する。

4. 研究成果

母音発話時のMRI データに歯列を補填した後、声道断面積関数を抽出して、音響感度特性を計算した。図1は母音/a/の例である。上段から声道伝達関数、断面積関数、第1~4ホルマント(F1~F4)周波数における音響感度特性である。この図はF1では咽頭下部、F2では口腔、F3では軟口蓋、F4では喉頭腔に、それぞれ音響感度が高い部位が局在していることを示す。このような音響感度特性を全ての母音のF1~F4で計算した。その結果、母音/e/では、どのホルマントでも音響感度が局在せず、一様管的であった。それ以外の母音についてまとめると以下ようになる。F1では一般に咽頭腔で感度が高く、前舌母音では狭めの位置、円唇母音では口唇も感度が高かった。F2では、前舌母音では咽頭腔、中・後舌母音では口腔、円唇母音では口唇で感度が高かった。F3では、前舌母音では喉頭腔、中舌母音では咽頭腔上部、後舌母音では軟口蓋近辺で感度が高かった。F4では喉頭腔のみ感度が局在していた。

これらの結果を踏まえて、音響感度特性の

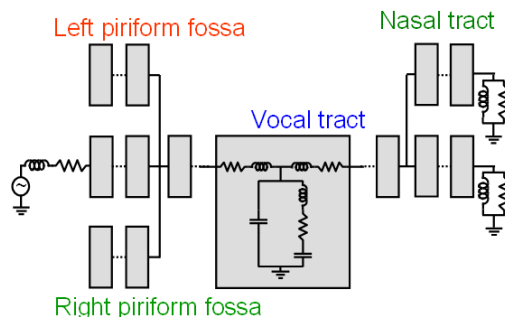


図2.タイムドメイン合成器

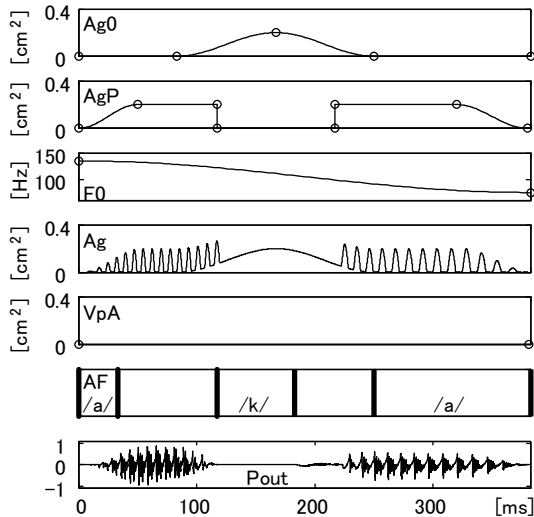


図 3. 発話運動/aka/のパラメータの例.

Ag0: 声門の中立面積, AgP: 声帯振動の振幅, F0: 声の高さ, Ag: 声門開口面積, VpA: 鼻腔との結合面積, AF: 声道断面積関数

高い部位をどのように利用して緩やかな発話運動から急峻な音声を生成するのかを検討するために、時間領域で音声生成をシミュレートするタイムドメイン合成器を作成した(図2)。この合成器の原理は、Maeda (1982) [5]に基づいており、時々刻々変化する声道断面積関数から、声道内の音圧・粒子速度の分布を逐次計算して、出力される音声信号を得る。Maeda (1982)では声道の分岐管として鼻腔のみが考慮されていたが、タイムドメイン合成器では音声の個人性に重要な役割を果たしている梨状窩も考慮できるように改良した(発表論文5)。これにより、個人性を備えた連続音声を合成することが可能になった。従来の音声合成器は、事前に録音した音声の素片を結合して目的の音声を合成する。そのため、結合部分に不自然さが残る。しかし、この合成器は、時間領域で発話を模擬して音声を生成するので、継ぎ目のない、自然な音声を合成できる。そのため、次世代の音声合成器として発展すると思われる。しかし、パラメータ設定が複雑であること、計算時間を要することなど、実用化に当たっては解決すべき問題が残されている。タイムドメイン合成器に、MRI 動画データから抽出した発話運動のパラメータ(図3)を入力して、連続音声を生成できるかどうか検討した。その結果、目的とする音声を高い品質で合成することが出来た(発表論文6)。このタイムドメイン合成器を用いて、連続発話(短い単語)のシミュレーションを行い、音声の量子性が見られるか検討した。

まず、子音に関しては否定的であった。子音は、例えば破裂子音では、声道閉鎖による圧力上昇と、急峻な開放によって生成される。

音韻の変化は急峻でも、この運動そのものが急峻であるため、量子論の対象ではない。さまざまな音韻について検討したが、量子性を見出すことは出来なかった。また、/s/から/j/への移行に関しては、明確な判断を下せなかった。これは、/j/の生成では、舌の下に小さな腔(sublingual cavity)が形成される場合が多かったが、解析に用いたタイムドメイン合成器ではこの腔を精密に扱うことが出来なかったためである。

連続母音の生成に関しては、肯定的な結果が得られた。母音の連続発話では、子音に比べて発話運動そのものは緩やかである。しかし、音響感度の高い部分がそうでない部分よりも速く動くことにより、ホルマント遷移が急峻になる可能性が示唆された。

まず、/iu/, /eo/の連続発話について述べる。/u/や/o/のF2では、唇の音響感度が高く、また/iu/, /eo/の連続発話では、ともにF2の下降が見られる。このとき、唇の運動が舌の運動よりも早く開始され、早く終了する。これは、唇の運動が舌の運動より相対的に小さいこと、調音結合が現れやすいことが原因かもしれないが、結果的には音響感度の高い部分が速く動くことにより、F2の遷移が急峻になる。これが意図的な操作であるかどうかは、被験者数を増やすなどして更なる追求をする必要がある。

次に、量子性が見られるといわれる/ai/について述べる。/ai/の連続発話ではF2が上昇する。/a/のF2では口腔、/i/のF2では咽頭腔で音響感度が高い。/a/から/i/への発話運動は、口が閉じる運動と同時に舌が後下方から前上方へ移動する。この運動は、口腔の体積を減少させると共に咽頭腔の体積を増大させる。つまり、結果として/a/と/i/の両方のF2の音響感度の高い部分が早く変化することにより、F2のホルマント遷移が急峻になる。

さらに、/a/から/i/の発話運動では、ホルマントの移動を急峻にする新たな要因を発見した。これは、歯列間隙の影響により、ホルマントの移動が不連続になる現象である(発表論文1)。

歯列間隙とは、上下は歯列、外側は頬、内側は舌によって囲まれた細長い空間である。歯列間隙は、/a/, /o/などの低母音では口腔に融合するために存在しないが、/i/, /u/, /e/では長さが約4.5 cm、断面が0.2 cm²の腔である。

この腔の音響的な影響を調べるために、声道模型の実測とシミュレーションを行った。まず、母音発話時のMRIデータに歯列を補填し、声道立体形状を抽出した。次に厚さ3 mmの壁をつけ、光造形法によりエポキシ樹脂で声道模型(図4(左))を作成し、声道伝達関数を実測した(発表論文2)。声道模型の実

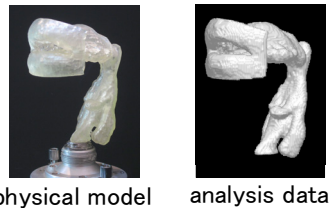


図4. 声道模型 (左) と音響解析用データ (右)

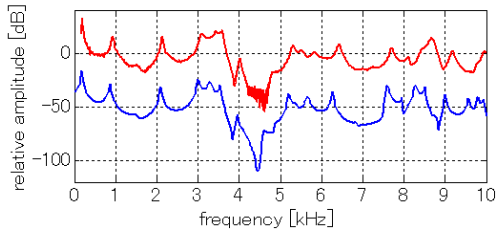


図5. 母音/u/の伝達関数. 赤: 声道模型の実測値, 青: 計算値

測は、声道の音響特性を精密に計測する一つの有力な方法である。実音声では、音源となる声帯振動が周波数的には離散的であるため、十分な解像度を持つ伝達関数が得られない。しかし、声道模型の伝達関数を TSP 信号で計測すれば、高い解像度を持つ伝達関数を計測できる。これによって得られた伝達関数が図5の赤線である。これは母音/u/の声道模型の伝達関数である。

そして、声道模型と同じデータから作成した、3次元の音響解析用形状データ (図4. 右) から、数値計算により伝達関数を求めた。計算は、時間領域差分法 (Finite-difference Time-Domain method: FDTD 法) で行った (発表論文3, 7)。この手法は、音響の2つの支配方程式 (運動方程式と連続の式) を時間的・空間的に差分式で表現し、時々刻々の音圧と粒子速度の分布を計算してゆく手法である。これにより得られた伝達関数が図5の青線である。伝達関数は実測値と計算値でよく一致する。これは、声道内部の音響現象を、高い精度でシミュレートできていることを示す。

この音響解析用の形状データから歯列間隙を除去すると、1.5 kHz 付近にある極零対が消失した。すなわち、歯列間隙は、口唇付近に開口する声道の分岐管として作用し、1.5 kHz 付近に極零対 (零点) を生成する。なお、この零点は実測した伝達関数にも存在していることから、実音声においても歯列間隙は低い周波数帯域に零点を作ると推定される。

/ai/の連続発話においては、歯列間隙による零点が高い周波数帯域から、低い周波数帯域へと移動する。低母音である/a/では歯列間隙が存在しないが、/a/から/i/へ発話運動が進むに従い、次第に歯列間隙が形成され、

伸張する。すなわち、口を閉じながら舌を後下方から前上方へと移動させるので、歯列間隙は長くなりながら次第に断面積が狭くなる。これにより、歯列間隙は声道の分岐管として機能するようになって零点が生じ、その周波数は歯列間隙が長くなると下降する。

一方、/ai/の連続発話では、F2, F3 がともに上昇する。これらのホルマントの上昇の過程において、下降してくる歯列間隙の零点と干渉することが予想される。図6は、/ai/の連続発話のシミュレーション結果である。歯列間隙の零点 (赤, 緑の○) が生じて下降し、声道伝達関数の上昇する F2, F3 と干渉の様子を 297 ms の間、表示している。予想通り、上昇する F2, F3 と歯列間隙の零点は相互に打ち消しあう。これは、/ai/の発話において、F2, F3 の移動に不連続が生じることを示唆している。

そこで、/ai/の音声スペクトルを解析したところ、F2, F3 の移動に不連続が生じている被験者がいた (図7)。図の赤丸で囲んだ部分で、F2, F3 のホルマント遷移に高い周波数領域から下降してきた零点が干渉して不連続点が生じている。シミュレーションと同じパターンを示すことから、この不連続は、歯列間隙の零点によって生じたと考えられる。すなわち、この被験者の/a/の連続発話では、歯列間隙が緩やかな発話運動から急峻な音響変化を生じる一つの要因となっている。

しかし、このような歯列間隙によると思われるホルマント遷移の不連続は、全ての被験者で明確に見られるわけではなかった。これは声道形状の個人差に由来すると考えられ、この点に関しては更なる検討を要する。

本研究では、音声の量子性について、音響感度特性とタイムドメインシミュレータを用いて検討した。その結果、連続母音では、特定のホルマントに対して音響感度特性が高い部位が先行して運動することにより、急

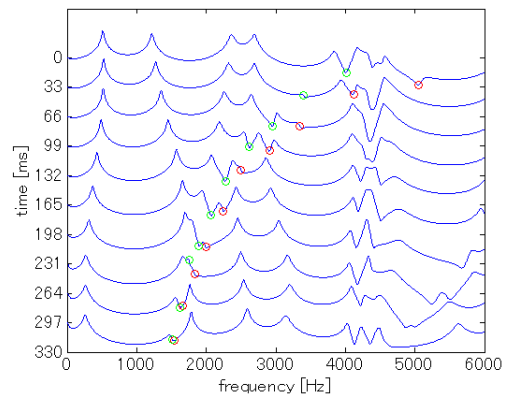


図6. /ai/発話において、下降する歯列間隙の零点と上昇する F2, F3 との相互作用のシミュレーション。

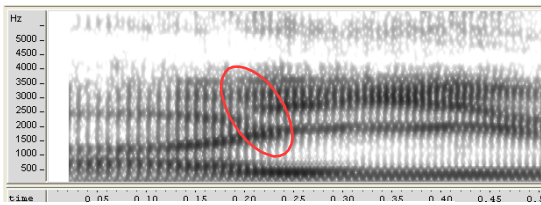


図7. スペクトログラムに現れるホルマント遷移の不連続 (赤丸)

峻なスペクトル変動が生じていることが示唆され、歯列間隙の零点がホルマント遷移に不連続を生じさせる場合があることが明らかになった。

近年の量子性に関する研究では、声門下の共鳴 (subglottal resonances) がホルマント遷移に不連続性を生じさせているという報告がある [6]。これは声門下の共鳴が F2 のホルマントのピークを弱めることにより、ホルマント遷移に不連続性を生じさせ、聴覚的にも影響するというものである。この声門下の共鳴も、本研究で明らかになった歯列間隙の影響も、ホルマント遷移に不連続性を生じさせるという点で同様の効果があり、ともに音声の量子性に関与する重要な発見であると思われる。

- [1] Stevens, K. N. "On the quantal nature of speech," *Journal of Phonetics*, 17, pp. 3-45, (1989).
- [2] Ladefoged, P. and Lindau, M. "Modeling articulatory-acoustic relations: a moment on Steven's Quantal nature of speech," *Journal of Phonetics*, 17, pp. 99-106, (1989).
- [3] Masaki, S., Tiede, M. K., Honda, K., Shimada, Y., Fujimoto, I., Nakamura, Y., and Ninomiya, N. "MRI-based speech production study using a synchronized sampling method," *J. Acoust. Soc. Jpn. E* 20, 375-379, (1999).
- [4] H. Takemoto, S. Adachi, T. Kitamura, P. Mokhtari, and K. Honda, "Acoustic roles of the laryngeal cavity in vocal tract resonance," *J. Acoust. Soc. Am.* 120 (4). pp. 2228-2238 (2006).
- [5] Maeda, S. "A digital simulation method of the vocal-tract system," *Speech Comm.* 1 (3-4), pp. 199-229, (1982).
- [6] Lulich, S.M. Bachrach, A. and Malyska, N. "A role for the second subglottal resonance in lexical access," *J. Acoust. Soc. Am.* 122 (4), pp. 2320-2327, (2007).

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 10 件)

1. Kiyoshi Honda, Sayoko Takano, and Hironori Takemoto, "Effects of side cavities and tongue stabilization: Possible extensions of the quantal theory," *Journal of Phonetics*, (in press), [査読有]

2. Tatsuya Kitamura, Hironori Takemoto, Seiji Adachi, and Kiyoshi Honda, "Transfer functions of solid vocal tract models constructed from ATR MRI database of Japanese vowel production," *Acoustical Science & Technology*, (in press), [査読有]

3. 竹本浩典, モクタリ・パーハム, 北村達也, "3次元時間領域差分法による母音生成時の声道の音響解析," *日本音響学会講演論文集 (春)*, pp.351-354, (2009), [査読無]

4. Hironori Takemoto, Tatsuya Kitamura, Kiyoshi Honda, and Shinobu Masaki, "Deformation of the hypopharyngeal cavities due to F0 changes and its acoustic effects," *Acoustical Science & Technology*, 29(4), pp. 300-303, (2008), [査読有]

5. Parham Mokhtari, Hironori Takemoto, and Tatsuya Kitamura, "Single-matrix formulation of a time domain acoustic model of the vocal tract with side branches," *Speech Communication*, 50, pp.179-190, (2008), [査読有]

6. 竹本浩典, モクタリ・パーハム, 北村達也, 井上謙次, 正木信夫, "MRI動画データの分析に基づく日本語音節の合成," *日本音響学会講演論文集 (春)*, pp. 317-318, (2008), [査読無]

7. 竹本浩典, パーハム・モクタリ, "時間領域差分法による声道の音響解析," *日本音響学会講演論文集 (秋)*, pp. 476-478, (2007), [査読無]

8. 竹本浩典, 北村達也, パーハム・モクタリ, 足立整治, 本多清志, "声門解放時の声道共鳴特性," *信学技報 (SP2006-165)*, pp. 31-36, (2007), [査読無]

9. Hironori Takemoto, and Tatsuya Kitamura, "Changes in the vocal tract resonance during a pitch cycle," *The Journal of the Acoustical Society of America*, 120(5), p. 3375, (2006), [査読有]

10. Hironori Takemoto, Seiji Adachi,

Tatsuya Kitamura, Perham Mokhtari, and Kiyoshi Honda, "Laryngeal cavity resonance and the fourth formant," Proceedings of the 5th International Conference on Voice Physiology and Biomechanics, pp.11-12, (2006), [査読有]

[学会発表] (計 7件)

竹本浩典, モクタリ・パーハム, 北村達也, "3次元時間領域差分法による母音生成時の声道の音響解析," 日本音響学会 2009年春季研究発表会, 2009.03.19, 東京工業大学

Hironori Takemoto, and Tatsuya Kitamura, "Mechanical models of the vocal tract constructed from 3D MRI data," Speech Production Workshop Instrumental-based Approach, 2008.07.05, ILPGA, Paris.

竹本浩典, モクタリ・パーハム, 北村達也, 井上謙次, 正木信夫, "MRI動画データの分析に基づく日本語音節の合成," 日本音響学会 2008年春季研究発表会, 2008.03.18, 千葉工業大学

竹本浩典, パーハム・モクタリ, "時間領域差分法による声道の音響解析," 日本音響学会 2007年秋季研究発表会, 2007.09.18, 山梨大学

竹本浩典, 北村達也, パーハム・モクタリ, 足立整治, 本多清志, "声門解放時の声道共鳴特性," 音声研究会, 2007.03.26, 東京大学

Hironori Takemoto, and Tatsuya Kitamura, "Changes in the vocal tract resonance during a pitch cycle," Forth Joint Meeting of ASA and ASJ, 2006.12.02, Honolulu, USA

Hironori Takemoto, Seiji Adachi, Tatsuya Kitamura, Perham Mokhtari, and Kiyoshi Honda, "Laryngeal cavity resonance and the fourth formant," The 5th International Conference on Voice Physiology and Biomechanics, 2006.07.12, 東京大学

6. 研究組織

(1) 研究代表者

竹本浩典 (TAKEMOTO HIRONORI)
株式会社国際電気通信基礎技術研究所
認知情報科学研究所・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

パーハム・モクタリ (MOKHTARI PARHAM)
株式会社国際電気通信基礎技術研究所
認知情報科学研究所・研究員

本多清志 (HONDA KIYOSHI)
株式会社国際電気通信基礎技術研究所
認知情報科学研究所・客員研究員

北村達也 (KITAMURA TATSUYA)
甲南大学知能情報学部知能情報学科・准教授