

科学研究費補助金研究成果報告書

平成 23 年 5 月 11 日現在

研究種目：基盤研究 (B)

研究期間：2007～2010

課題番号：19300063

研究課題名 (和文) 実体モデルに基づく声質生成機構の構築

研究課題名 (英文) Construction of voice quality generation mechanism by a mechanical model

研究代表者

誉田 雅彰 (HONDA MASAOKI)

早稲田大学・スポーツ科学学術院・教授

研究者番号：90367095

研究成果の概要 (和文)：人間の発声・発話機構を模した機械的実体モデル (発話ロボット) を基に声質の生成・制御機構を構築し、声質の生成メカニズムを構成的に明らかにすることを目的として研究を進めた。その結果、人間に類似した実体モデルの喉頭調節を行うことによって氣息音やフライ音などの多様な声質や笑い声やしゃべり笑いなどの声の感性情報を再現できることを示すとともに、これらの声の生成時における声帯振動や空気流体音響現象を明らかにした。

研究成果の概要 (英文)：The research project was aiming to clarify quantitatively speech production process and its control strategy by using a mechanical speech production model (Talking ROBOT) which is mimicking the human mechanism. We realized to reproduce speech sounds with various voice qualities such as breathy voice and creaky voice as well as laughter and laughing voice by using the laryngeal control of the model which is similar to the human control. We also examined the vocal cord vibration behavior and aero-acoustic phenomenon in generating these voice sounds by the direct measurement.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	6,600,000	1,980,000	8,580,000
2008年度	4,600,000	1,380,000	5,980,000
2009年度	1,698,664	810,000	2,508,664
2010年度	1,001,336	0	1,001,336
年度			
総計	13,900,000	4,170,000	18,070,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報処理・知能ロボティクス

キーワード：音声情報処理

1. 研究開始当初の背景

(1) 声質研究の背景

音声の感性情報を担う声質に関する研究の

多くは音声の音響的や分析や合成に関するものが多かった。また、声質の生成的側面に関する研究としては、ファイバースコープを

用いた声帯振動の観測など、直接観測によって喉頭調整に関する知見を得るものが多く、音声生成モデルを用いて声質の生成メカニズムを構成的に明らかにする研究は少なかった。

(2) 実体モデルの研究背景

人の音声生成過程を模擬する音声生成モデルに関しては、これまで計算機シミュレーションモデルが多く研究されてきた。一方、多様な声質の音声生成される際の声帯振動や声帯音源が生成時における流体音響現象は複雑であり、計算機モデルによってそれらを完全に再現することは困難であった。一方、これらの課題を解決する方法として、人の音声生成機構を模した機械的実体モデルを用いる研究が進められてきた。これまでに、実体モデルを用いてモーダルボイス（通常発声）の音声は生成されたが、氣息音やフライ音のような多様な声質の音声は再現されていなかった。

2. 研究の目的

発声・発話系の機械的実体モデルを基に人間を模擬した喉頭制御を行うことによって多様な声質の音声と笑い声を再現するとともに、喉頭制御と声質との定量的な関係を構成的に明らかにする。

3. 研究の方法

(1) 実体モデルの構築

多様な声質の音声生成可能な発声・発話実体モデルを構築する。実体モデルは、できる限り人間の喉頭機構を模擬し、人間の喉頭調整機能に即した喉頭制御様式によって、多様な声質をもった音声の生成を目指す。

(2) 実体モデルを用いた声質音声の生成

氣息音、フライ音を対象とし、実体モデルによる声質生成を行う。また、音声の主要な感性情報の表出方法として笑い声を対象とし、実体モデルによる生成を行う。

(3) 観測実験

声質音声生成時における声帯振動を高速カメラで撮影するとともに、空気流体、及び音響信号を同時測定し、人間に対する測定データと比較することにより、モデル実験の有効性を検証する。測定では、母音/i/の声道形状をもつアクリル製の音響管を声帯部に接続し、音響管の片側の開口部から高速カメラを用いて声帯振動を測定した。また、音響管の開口部に流量測定用のマスクを設置して空気流量信号を測定し、逆フィルタを用いて声道音響管の音響特性を除去し、声門体積速度を測定した。

(4) 声質生成メカニズムの解明

多様な実体モデルの喉頭制御と生成される声帯振動および流体音響信号のデータから、喉頭制御と声質との関係を定量的に明ら

かにする。

4. 研究成果

(1) 発声・発話系の実体モデルの構築

人間の発声・発話機構を模擬した機械的実体モデル（発話ロボット）を構築した。発話ロボットは、図1に示すように、人間の音声生成機構を模擬し、肺、声帯、舌、顎、唇、口蓋帆から構成される3次元モデルであり、肺から送りだされる空気流を駆動源として声帯が自励振動して声帯音源が生成され、声道内での共鳴現象を介して音声は口唇あるいは鼻孔から放射される。声質制御においては、特に肺機構と声帯機構が中心的な役割を果たす。声帯モデルは、図2に示すような左右のヒダ状の形態を有し、人間の声帯と同等のヤング率をもつセプトン素材で形成されている。喉頭の調整機構としては、図3に示すように、声帯の張力を可変として音声のピッチ周波数を調整する機構、声門の開きを可

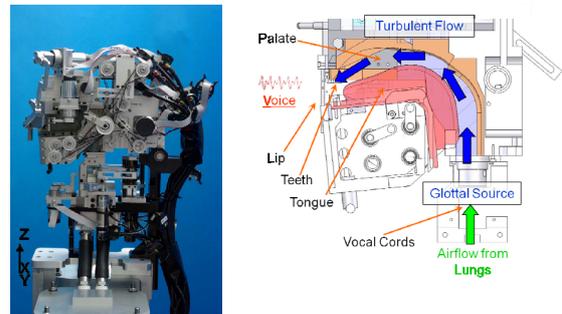


図1 発声・発話実体モデルの構成

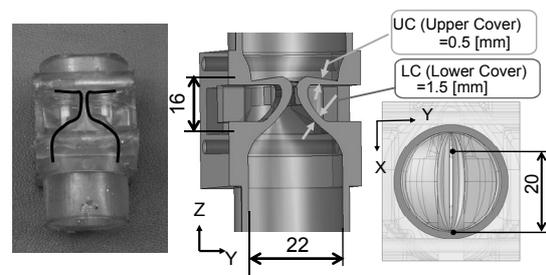


図2 声帯モデル

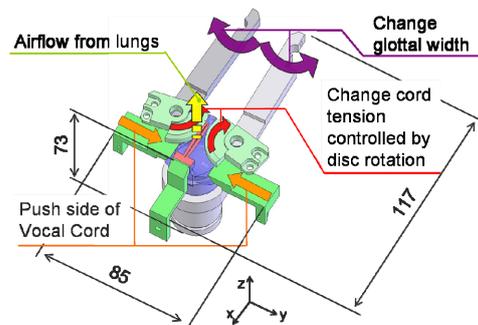


図3 喉頭調整機構

変として有声・無声を調整する機構、及び声帯の振動質量を可変とする横押し機構を構築した。また、声道部は舌、唇、顎を3次元構造を模擬した3次元モデルであり、全ての母音と子音の生成することが可能な構造となっている。

(2) 声質の生成

声の声質として、モーダル音、氣息音、フライ音を対象として、実体モデルによりこれらの声質を生成した。モーダル発声は通常の発声様式であり、声帯を閉じた状態にし、適度な声帯張力を与えることにより生成される。一方、氣息音は声帯の開きをわずかに開き気味にし、声帯張力と肺圧はモーダル発声と同程度に調整した。また、フライ発声は、会話音声における語尾によく現れる声質であり、肺圧はモーダル発声より低く、声帯張力を極めて低く、また声帯の振動質量が小さくなるように調整した。これらの喉頭調整様式は人間の発声時にも見られるものであり、人間の喉頭調整様式がほぼそのまま実体モデルにおいても用いられている。

図4に、実体モデルによるモーダル発声、氣息音発声、フライ発声の声帯振動パターンを示す。それぞれは、1周期分の声帯振動のスナップショットを示したものである。これらの結果は、人間に対して測定された声帯振動パターンの特徴を良く再現している。モーダル発声の振動パターンでは、声帯の上部と下部において振動の位相が異なっている。ま

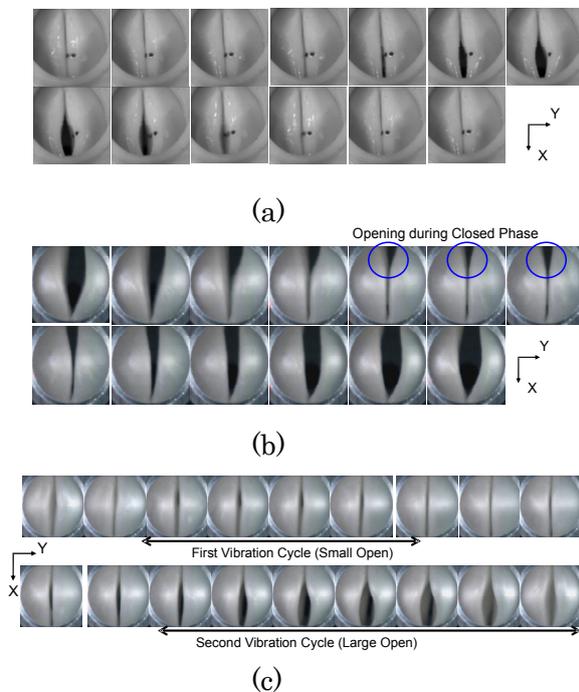


図4 実体モデルにおける声帯振動

- (a) モーダル発声 (b) 氣息発声
(b) フライ発声

た、氣息音発声では、声帯振動の閉鎖区間においても声帯の上隅の部分に隙間が存在して不完全閉鎖が生じ、その隙間を通して空気が流れることにより氣息音特有の摩擦性の成分が声帯音源が重畳される。このような氣息音の音源特徴は、声門堆積速度や音源スペクトル上においても人における観測結果と一致する特徴として確認された。フライ発声は、極めて低いピッチとして知覚される声質であり、このような低ピッチの音声は声帯振動における倍ピッチ振動によって生じることが知られている。実体モデルによるフライ発声においても、このような声帯の倍周期振動が再現されており、前半の周期区間では声門の開きが小さく、後半の周期区間で声門が大きく開く振動が確認された。実体モデルにおけるフライ発声では、肺圧と声帯張力に調整に加えて、振動する声帯の質量を調整する横押し機構が用いられた。このような機構は人間の声帯には存在しない機構である。人間に声帯は、声帯内部に筋構造が存在し、声帯張力を調整すると同時に声帯の振動する部分(振動質量)が変化することが知られている。モデルにおいても声帯張力の調整によって振動する声帯部分が若干変化するが、モデルでは声帯内部に筋構造を有していないため質量変化が十分でなく、これを振動質量を調整するための新たな機構が必要となった。今後は、声帯内部の構造をも模擬し、より一層人間の声帯構造にモデルを近づけるには、必要があると考えられる。

図5は、声帯張力と声帯振動周波数の関係を示したものである。モーダル発声では、声帯張力を下げるとともに緩やかに振動周波数が低下していくが、肺圧が低い状況では、声帯張力をある程度以上上げると振動周波数が急激に約半分に低下する現象が明らかになった。この振動周波数が急激な低下した状態がフライ発声となる。このように、実体モデルでは、声帯振動の振動モードの急激な変化を再現しており、フライ発声極めて限定された条件でのみ生じることが定量的に示された。この結果は、人間にとってフライ発声は他の発声に比べて難しいことを裏付

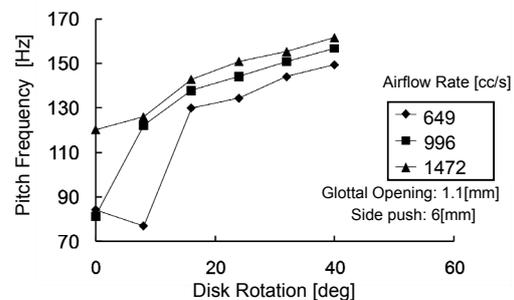


図5 声帯張力と声帯振動周波数の関係

けている。

以上述べたように、質量調整の機構を除けば、人の喉頭調整機構と同様な機構を用いて声質固有の声帯振動を再現できることが明らかとなった。

(3) 笑い声の生成

音声における感性情報の表出手段としてよく用いられる笑い声の生成メカニズムをの実体モデルを用いて検討した。笑い声の生成は、喉頭調整が主要な働きをもち、声道の調整は比較的定常的であると考えられる。また、笑い声発声時における声帯観測より、声帯の開閉が周期的に生じること、また、人の笑い声発声時における口唇部における流体計測により、通常発声時と比べて発声開始時において空気流量が大きく、徐々に空気流量が減衰する特徴が観測された。また、笑い声の音響分析において、笑い声のピッチ周波数は通常発声よりも高く、1モーラ内において山形のピッチ変化が生じることが観測された。

これらの人の笑い声の分析結果を基に、実体モデルの肺圧、声帯張力、および声門の開きを調整することにより笑い声の生成を試みた。図6は、「ははは」の笑い声について、人の発声と実体モデルによる発声を比較したものである。図の上段は、口唇部における空気流量と音声波形、下段はスペクトルグラムを示す。これより、流体音響レベルにおいて笑い声の特徴が実体モデルによって比較的良く再現できているのが分かる。ただし、発声開始時において、モデルによる発声の空気流量は人の笑い声に比べて小さくなっている。これは、肺モデルの最大供給流量が不十分なためであり、今後は肺モデルの駆動アクチュエータのパワーを増大することにより解決できるものと考えられる。

また、笑い声の生成に元に、しゃべり笑い

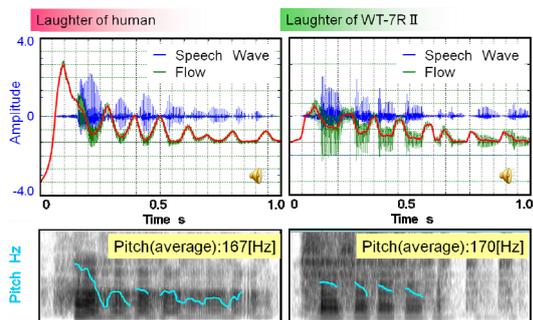


図5 実体モデルによる笑い声の生成

左：人の笑い声 右：モデルの笑い声

の生成を試みた。しゃべり笑いでは、笑い声を生成するための喉頭の調整と発声のための喉頭と声道の調整が必要となる。この場合、

喉頭の調整は、笑い声と発話の両方が競合し、その競合の度を調整することにより、しゃべり笑いにおける笑い成分の度を調整することとなる。ただし、現状におけるしゃべり笑いの音声品質は、人の発声に比べて未だ不十分であり、さらに喉頭と声道の制御方法についてさらに検討を行う必要がある。

(4) 人の発話動作を模した実体モデルの発話制御

より巧みな発話動作を実体モデルを用いて実現するための第一歩として、人の発話動作を計測し、その発話動作データから実体モデルの制御パラメータを決定するいわゆる発話の見真似制御について検討を行った。

人の発話動作は、磁気センサーシステムを用い、連続音声発話時における顎、舌、唇の発話動作データを測定した。顎に関しては、下顎の測定位置から実体モデルの顎の回転角度を求め、唇に関しては、上下唇の測定位置から実体モデルの唇の開きと突き出しパラメータを算出した。また、舌に関しては、図6に示すように、舌上の4点の測定値から舌の輪郭線を求め、実体モデルの舌の制御点

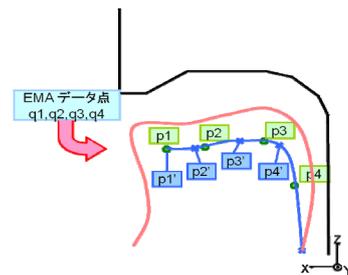


図5 見真似発話制御における舌パラメータの決定方法

の位置が輪郭線上に乗るように逆キネマティックスを用いて決定した。

図6は、従来の実体モデルの制御法による舌の発話動作の運動軌道と見真似発話における運動軌道を比較して示す。見真似発話制御を導入することにより、人の発話動作を模したより巧みな動きを生成できることが確認された。

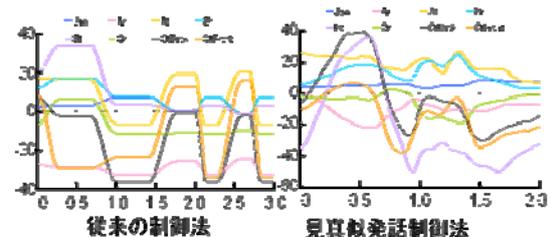


図6 見真似発話制御による発話動作の生成

5. 主な発表論文等

〔雑誌論文〕(計6件)

- ① K.Fukui, Y.Ishikawa, E.Shintaku, M.Honda, A. Takanishi, Production of Various Vocal Cord Vibrations Using a Mechanical Model for an Anthropomorphic Talking Robot, *Advanced Robotics*, 査読有、2011 (in press)
- ② 福井孝太郎、草野世大、高西淳夫、菅田雅彰、発話運動データに基づく人間形発話ロボットの制御、音声研究、査読有、Vol.14、No.2、2010、pp.57-64.
- ③ 福井孝太郎、菅田雅彰、高西淳夫、人間形発話ロボットによる人間の発話メカニズムの解明、査読有、*バイオメカニズム学会誌*、33巻、4号、2009、pp.237-242
- ④ 磯崎英治、飛澤晋介、西沢美里、中山英人、福井孝太郎、高西淳夫、人工声帯をもちいた実験の声帯外転障害-多系統萎縮症における非侵襲的陽圧換気療法の圧設定-、*臨床神経学*、査読有、49巻、7号、2009、pp.407-413
- ⑤ 福井孝太郎、新宅英滋、下村彰宏、榊原菜々、石川優馬、菅田雅彰、高西淳夫、NN順逆モデリングを用いた人間形発話ロボットの声帯振動制御、*日本ロボット学会誌*、査読有、26巻、7号、2008、pp.76-82
- ⑥ 福井孝太郎、新宅英滋、菅田雅彰、高西淳夫、生理学的構造に基づく人間形発話ロボットの声帯機械モデル、査読有、*日本機械学会論文集 (C編)*、73巻、734号、2007、pp.2750-2756.
〔学会発表〕(計16件)
- ① K. Fukui, T. Kusano, Y. Mukaeda, Y.Suzuki, A.Takanishi, M.Honda, Speech Robot Mimicking Human Articulatory Motion, *INTERSPEECH 2010*, 2010年9月28日、幕張.
- ② 福井孝太郎、鈴木悠人、草野世大、迎田美和、菅田雅彰、高西淳夫、光距離センサを用いたロボットの舌形状の計測、第28回日本ロボット学会学術講演会、2010年9月22日、神戸.
- ③ 福井孝太郎、草野世大、鈴木悠人、迎田美和、高西淳夫、菅田雅彰、発話運動データを用いた人間形発話ロボットの制御、*日本音響学会 2010年秋季研究発表会*、2010年9月16日、大阪.
- ④ 菅田雅彰、発音発語-工学から見た発音発語、招待講演、第68回日本矯正歯科学会大会、2009年11月17日、福岡.
- ⑤ K.Fukui, Y.Ishikawa, K. Ohno, N. Sakakibara, M.Honda, A.Takanishi, Three Dimensional Tongue with Liquid Sealing Mechanism for Improving Resonance on an Anthropomorphic Talking Robot, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009年9月14日、St. Louis.
- ⑥ 福井孝太郎、迎田美和、榊原菜々、草野世大、菅田雅彰、高西淳夫、人間形発話ロボットの開発-子音生成のための新型口唇機構の設計・製作-、第27回日本ロボット学会学術講演会、2009年9月16日、横浜.
- ⑦ 迎田美和、榊原菜々、福井孝太郎、草野世大、菅田雅彰、高西淳夫、人間形発話ロボットの新型口唇機構による両唇閉鎖子音の生成、*日本音響学会 2009年秋季研究発表会*、2009年9月17日、福島.
- ⑧ M.Honda, K.Fukui, E.Shintaku, Y.Ishikawa, N.Sakakibara, Y.Mukaeda, A.Takanishi, A., Production of various voice qualities using mechanical vocal code model, 1st Int. Workshop on Dynamic Modeling of the Oral, Pharyngeal and Laryngeal Complex for Biomedical Applications, Invited Talk, 2009年6月27日、Vancouver.
- ⑨ K.Fukui, E.Shintaku, Y.Ishikawa, N.Sakakibara, Y.Mukaeda, A.Takanishi, M.Honda, Three-dimensional mechanical tongue and vocal cord models for human speech production, Meeting of Acoustical Society of America, Invited Talk, 2009年5月18日、Portland.
- ⑩ K.Fukui, N.Sakakibara, E.Shintaku, A.Shimomura, Y.Ishikawa, M.Honda, A.Takanishi, Three-dimensional Tongue Mechanism of Anthropomorphic Talking Robots for Human-like Voice Production, 17th CISM-IFTtoMM Symposium on Robot Design, Dynamics and Control, 2008年7月6日、Tokyo.
- ⑪ K.Fukui, E.Shintaku, A.Shimomura, N.Sakakibara, Y.Ishikawa, M.Honda, A.Takanishi, Control Methods Based on Neural Network Forward and Inverse Models for a Biomechanical Structured Vocal Cord Model on an Anthropomorphic Talking Robot, The 2008 IEEE International Conference on Robotics and Automation, 2008年5月19日、Pasadena.
- ⑫ 榊原菜々、新宅英滋、下村彰宏、福井孝太郎、石川優馬、菅田雅彰、高西淳夫、人間形発話ロボットにおける3次元舌実

- 体モデルの開発、日本音響学会 2007 年秋季研究発表会、2007 年 9 月 20 日、山梨。
- ⑬ 福井孝太郎、新宅英滋、下村彰宏、榊原菜々、石川優馬、高西淳夫、蒼田雅彰、人間形発話ロボットにおける多様な音声の生成を目的とした声帯制御、日本音響学会 2007 年秋季研究発表会、2007 年 9 月 20 日、山梨。
- ⑭ 下村彰宏、新宅英滋、榊原菜々、福井孝太郎、石川優馬、蒼田雅彰、高西淳夫、人間形発話ロボットの開発－多様な音声の生成が可能な新声帯機構の設計製作－、日本ロボット学会第 25 回学術講演会、2007 年 9 月 14 日、千葉。
- ⑮ 福井孝太郎、新宅英滋、下村彰宏、榊原菜々、石川優馬、蒼田雅彰、高西淳夫、人間形発話ロボットの開発－NN 順逆モデリングを用いた声帯振動制御－、日本ロボット学会第 25 回学術講演会、2007 年 9 月 14 日、千葉。
- ⑯ K.Fukui、Y.Ishikawa、T.Sawa、E. Shintaku、M.Honda、A.Takanishi、New Anthropomorphic Talking Robot having a Three-dimensional Articulation Mechanism and Improved Pitch Range, Proceeding of IEEE International Conference on Robots and Automations, 2007 年 4 月 12 日、Roma.

6. 研究組織

(1) 研究代表者

蒼田 雅彰 (HONDA MASA AKI)
早稲田大学・スポーツ科学学術院・教授
研究者番号：90367095

(2) 研究分担者

高西 淳夫 (TAKANISHI ATSUO)
早稲田大学・理工学学術院・教授
研究者番号：50179462

小林 哲則 (KOBAYASHI TETSUNORI)
早稲田大学・理工学学術院・教授
研究者番号：30162001

匂坂 芳典 (SAGISAKA YOSHINORI)
早稲田大学・国際情報通信研究科・教授
研究者番号：70339734

(3) 連携研究者

福井 孝太郎 (FUKUI KOUTARO)
早稲田大学・理工学学術院・客員研究員
研究者番号：60547074