

Title	口腔疾患を有する複雑な声道形状と音声スペクトルの関係に関する基礎的研究
Author(s)	西本, 博則
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1454
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

口腔疾患を有する複雑な声道形状と 音声スペクトルの関係に関する基礎的研究

西本 博則

北陸先端科学技術大学院大学 情報科学研究科

2001年2月15日

キーワード: 口腔疾患, 核磁気共鳴撮像法 (MRI), 有限要素法 (FEM), 声道伝達特性, フォルマント周波数.

1 緒論

本研究は口腔疾患患者に見られるような健常者とは異なる特徴を有する複雑な声道形状の構造的特徴と, 音声スペクトルなどの音声の音響的特徴の関係を導くための基礎的調査を行うことを目的とする.

口腔疾患を有する患者の音声中には, 健常者とは異なる独特の歪みが存在する. これは, 声道の中心に対して非対称な形状, 複雑な分岐等を有するといった, 患者の異常な声道形状に起因すると考えられる. その治療は PAP (Palatal Augmentation Prosthesis) と呼ばれる人工補綴物等によりなされるが, その設計には客観的指標が無く, 医療従事者の聴覚印象に頼ってきたのが現状である. 本研究では治療への応用を考慮して, 患者の声道形状と音声の歪の関係を調査し, 明らかにすることを目的としている. そのために, 声道形状の複雑さに起因する音響的特徴を含めて解析できる方法として, MR (Magnetic Resonance) 画像より 3 次元声道形状モデルを構築し, それを有限要素法 (Finite Element Method, FEM) で計算し, 声道伝達特性を推定するという手法による分析を行う. その基礎的検討として, シミュレーション結果と実音声の分析結果を比較し, シミュレーションの妥当性の検討を行う. さらに多くの症例の 3 次元声道モデルのシミュレーションや 3 次元声道形状モデルに任意の変形を与えたモデルのシミュレーションを行うことで, 声道形状の構造的特徴と音声の音響的特徴の関係について導く. 声道形状から音声スペクトルを推定する順問題を多数解くことにより, 逆問題を解く指針を得ることで, 音声スペクトルと声道形状の関係を導くことができる.

2 音響分析

本研究で対象とする被験者の詳細を表 1 に示す．Patient-A は術後の機能障害に対して PAP が適用されている．

表 1: Subjects

Subject	Age	Sex	Case
Normal-A	24	male	Normal
Normal-B	26	male	Normal
Patient-A	37	male	Tongue and mouth floor resection
Patient-B	72	male	Mouth floor resection

不偏推定法によるケプストラム平滑化対数スペクトル [2] によるパワースペクトルから各被験者の日本語母音 /i/ 発話時のフォルマント周波数の平均と標準偏差を求めた．これをシミュレーション結果の妥当性の検証に用いる．

3 健常例でのシミュレーション

ここでは，健常例でのシミュレーション結果と音響分析結果とを比較し，本シミュレーション法の妥当性を検討する．

3.1 Normal-A のシミュレーション結果

口腔部の点間隔 $0.5[mm]$ の Normal-A の 3 次元声道モデルによるシミュレーションを行った．推定された声道伝達特性の形状と，実音声のフォルマント周波数を比較したとき， F_2 以外の対応付けは良くとれている．声道伝達特性のピークと実音声のフォルマントの数が異なっているのは，フォルマントをスペクトル包絡から推定しているため，ピークが近接している場合は，なだらかな曲線となり，1 つのピークとして現れるためだと考えられる．一方で，本シミュレーション法では声道壁を剛壁とし，空気に粘性がないと仮定しているためピークが鋭く立つ．

放射球面を取り付けたことによる声道伝達特性の影響が F_3 以上に現れており，より音響分析結果に近づいている．実音声のフォルマントが $\pm 10[\%]$ の幅で変動しているということと，MR 画像より得られる測定値と実測値との間の誤差は $5[\%]$ であるという報告 [1] を考慮すれば， F_2 以外では誤差の範囲内であるとみなせる．

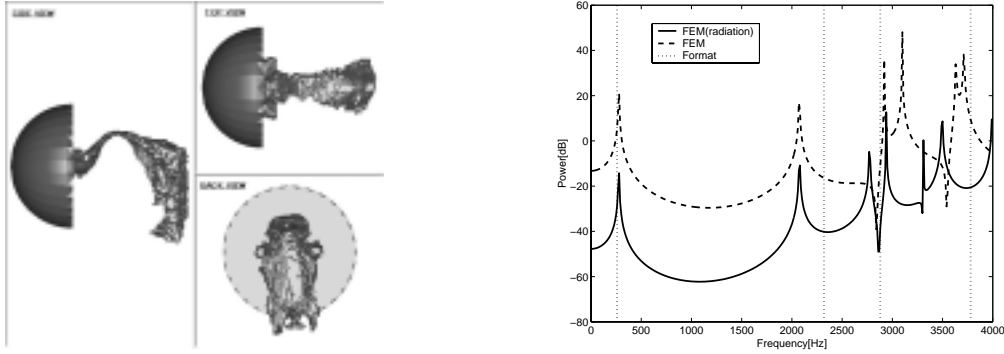


図 1: A 3-D vocal tract model of Normal-A with hemispherical surface of radiation and transfer functions of a Normal-A vocal tract model with hemispherical surface of radiation and without hemispherical surface of radiation.

3.2 Normal-B のシミュレーション結果

Normal-B の 3 次元声道モデルによるシミュレーションを行った。Normal-B でのシミュレーションの場合、 F_1 、 F_2 以外の比較では MR の計測誤差や音声の揺れを考慮すると誤差の範囲内であると考えられる。 F_1 の誤差は Normal-B の音声の /i/ 発話時の基本周波数が $160[Hz]$ であるため、その倍ピッチ成分が音響分析の結果出ていることが考えられ、そのため大きな誤差が現れると考えられる。

以上のことから、本シミュレーション法では F_2 以外は音響分析結果と近似がとれており、本シミュレーション法の妥当性が言えた。

3.3 梨状窩がフォルマントに与える影響

ここでは、これまでのシミュレーション結果と音響分析結果で誤差の範囲を超える F_2 の検討を行うため、声道モデルの梨状窩を取り除いたモデルによる声道伝達特性の変化を調べる。

梨状窩を除去した 3 次元声道モデルを用いたことで F_2 が音響分析結果に近づいた。梨状窩は発声中に膨張、もしくは収縮を起こす器官である。また、MR 撮像の姿勢は仰臥位であり、MR 画像には写らないような唾液等の体液により、梨状窩が実際には声道の分岐としての役割を果たしていないことが考えられる。そのため、本モデルのように梨状窩を切除した状態のほうが F_2 に関して実音声との対応がよくなるという可能性がある。

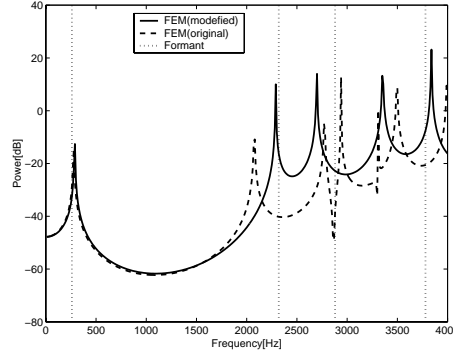


図 2: A transfer function of a Normal-A vocal tract removed pyriform fossas.

4 舌・口底切除症例のモデルによるシミュレーション

ここでは、舌・口底切除症例の3次元声道モデルによるシミュレーションを行った。各症例とも、放射球面を取り付けている方が推定された声道伝達特性の形状が音響分析結果に近いことが分かる。しかし、 F_3 以上を比較すると、音響分析結果とうまくあっていないことが分かる。Patient の場合、構音が一定しないため MR 画像と音声スペクトルの一致が困難になるためである。

4.1 1次元近似モデルとの比較

ここでは、Patient-A (PAP (-)) のデータを用いて1次元近似モデル [3] による分析を行い、その結果と FEM による分析結果と比較した。音響分析結果では $0 - 4000 [Hz]$ の間にフォルマントは4つ存在するが、FEM により得られた声道伝達特性は同じ周波数帯域にピークが同様に4つ存在していることが分かる。一方、1次元近似モデルでは実音声に見られる F_2, F_3 近辺にピークが1つしかない。これは、口腔内の異常な形状、特に舌による狭めの位置のずれや前室の水平方向の広がり起因する音声の歪みを表現できていないためと考えられ、FEM による声道の3次元的特徴を含めたシミュレーションを行うことの必要性が示された。

5 口腔内の声道形状を変化させたモデルによるシミュレーション

ここでは口腔疾患の患者の口腔内の異常な声道形状が音声スペクトルに与える影響を調べるためのシミュレーションを行った。Patient-A(PAP(-)) の口腔内の声道管経路は声道の中心軸に対してやや右手側を通っている。そこで、Patient-A (PAP(-)) の3次元

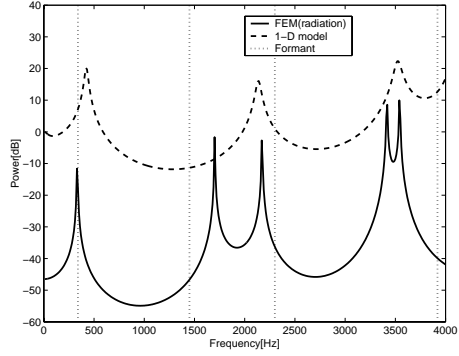


図 3: A transfer function of a Patient-A(PAP(-)) vocal tract model simulated by FEM and 1-D model.

声道モデルの口腔内に相当する部分に変形を与え，声道の中心軸を基準としてスライドさせたモデルを構築した．

4000[Hz] 以上では，オリジナルの声道モデルによるシミュレーション結果との間で違いが見られる．この周波数領域では声道長よりも声道表面の滑らかさ，連続性の影響を受けやすいため，と考えられる．一方，4000[Hz] 以下の範囲ではほとんど変化が見られなかった．これは，Patient-A (PAP(-)) の口腔内の声道の中心軸に対するずれが大きくななく，モデルの口腔内の声道をずらしても，実質的な声道管経路に変化がないため，声道長にほとんど変化を与えなかったためであると考えられる．

よって，以上の結果からは Patient-A (PAP(-)) の PAP の設計では，声道管経路以外の部分 (舌による狭めの幅，狭めを作る位置) に着目した設計が重要であると言える．

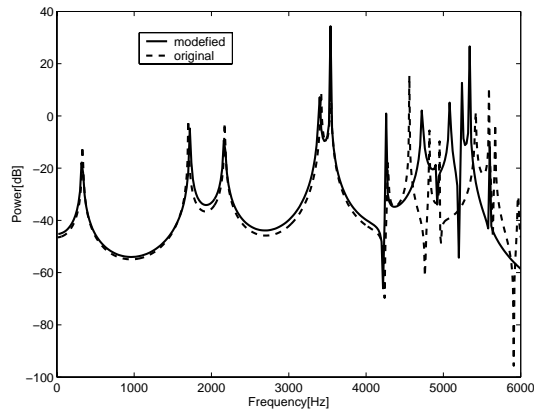


図 4: A transfer function of a modified vocal tract model of Patient-A (PAP(-)).

6 結論

構音が一定しており、フォルマント周波数の時間的な揺れが少ない健常者を対象とすることで、実音声の分析結果と比較した結果本研究で用いた手法の妥当性を示すことができた。また、舌・口底切除症例の場合、1次元近似モデルと比較した結果、本シミュレーション法の優位性が示された。これは、本手法が声道の複雑な形状、特に舌・口底切除患者の声道形状が音声スペクトルに及ぼす影響を表現できているためである。また、声道形状に変形を与えることで音声スペクトルの変化が現れることがわかった。舌・口底切除例の3次元声道モデルに対して変形を加えることで、変形を加えた部位が音声に及ぼす影響を調査することができた。これにより、言語障害の治療や PAP の設計への応用が可能であると言える。

参考文献

- [1] 斎藤 浩人, 鈴木 規子, 藤田 幸弘, 道 健一, 高橋 俊行, "MR 撮像法を用いた 3 次元声道形状の計測 -研究方法および健常人の声道断面積の検討-", 日本口腔科学会雑誌, 49, pp.92-101, 2000
- [2] 古市 千枝子, 今井 聖, "対数スペクトルの不偏推定," 信学論, Vol.J70-A, No.3, pp.471-480, 1987
- [3] Man Mohan Sondhi, JUERGEN SCHROETHER, "A hybrid time-frequency domain articulatory speech synthesizer," IEEE Transactions on acoustics, speech, and signal processing, Vol. ASSP-35, No. 7, July, pp955-967, 1987