

Title	有限要素法による声道伝達特性推定の有効性に関する検討
Author(s)	西本, 博則; 赤木, 正人; 北村, 達也; 鈴木, 規子
Citation	日本音響学会誌, 62(4): 306-315
Issue Date	2006-04-01
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/4627
Rights	日本音響学会, 西本 博則, 赤木 正人, 北村 達也, 鈴木 規子, 日本音響学会誌, 62(4), 2006, 306-315.
Description	

論 文

有限要素法による声道伝達特性推定の有効性に関する検討*

西本博則^{*1} 赤木正人^{*1} 北村達也^{*2} 鈴木規子^{*3}

【要旨】 有限要素法と声道等価回路モデルの声道伝達特性の推定精度の調査を行うため、MR 計測から得られた声道モデルより推定された伝達特性のピークと実音声のホルマントの比較を行った。その結果、健常な被験者では両推定法とも精度良く推定された。一方、複雑で左右非対称な声道形状の被験者では、有限要素法を用いれば音声ホルマントの個数が同じであり、ホルマント周波数と伝達特性のピーク周波数に大きな差がなかったが、等価回路モデルでは伝達特性のピークの数と音声ホルマントの数が一致しなかった。複雑で左右非対称な声道の伝達特性推定には、等価回路モデルは不適切であり、有限要素法が有効であることが示された。

キーワード 声道伝達特性, ホルマント周波数, 有限要素法, MRI

Vocal tract transfer function, Formant frequency, Finite element method, MRI

1. はじめに

口腔疾患等により構音障害を持つ患者の音声には歪が生じており、発語明瞭度が低くなる傾向がある。この問題に対して、各症例に対する声道形状計測 [1, 2] や音声の音響分析 [3-5] がなされている。口腔疾患患者の声道形状は健常者と比較して非常に複雑で、声道断面が左右非対称であるという特徴を有しており、それが音声に歪を与えていると考えられている。我々は声道形状の構造的特徴に起因する異常音声生成メカニズムの原因解明を志向して、声道の構造的特徴と音声の関係の調査をしており、これが発話訓練や音声の歪の改善のために口腔内に装着する舌接触補助床等の人工補綴物の設計など、構音障害の治療に役立つことが期待される。この調査を行う際に、声道形状の計測法及び声道伝達特性の推定法の精度が重要となる。

声道形状の計測を行う方法として X 線画像から正中断面図を得る方法 [6], 3 次元 CT による声道計測法 [7], 核磁気共鳴画像法 (MRI) での計測により 3 次元形状データを得る方法 [1, 2, 8] などが挙げられる。X 線写真による計測方法では短時間で計測が行え、過渡的な運動を取り扱えるという利点があるが、直接的には 2

次元の正中断面画像しか得られず、また被曝という問題点がある。3D CT を用いると 3 次元形状データが得られるが、X 線被曝量が更に多くなるという問題がある。一方、MRI による計測では、精密な 3 次元形状データが得られ、人体に対して被曝の問題がないという利点を有する。撮像に比較的長時間を要し、歯列形状を計測できないという問題点があるが、歯列形状の計測の問題については歯冠プレートを装着することで MR 画像に歯列を写しこむ方法 [9] や、口腔内にブルーベリージュースを含ませて MR 撮像し、MR 画像の歯列形状を計測し、それを発話時の画像にはめ込む歯列補填法 [10] によりこの問題の回避が可能であり、母音や摩擦子音のような構音が変化しない音声であれば高精度な 3 次元形状データを得ることが可能となる。

声道形状データから声道伝達特性の推定を行う方法として、人体の声道そのものの応答を音響計測により推定する方法 [11], 声道模型の伝達特性を音響分析により計測する方法 [12-15], 声道の断面積を計測し、等価音響管の電気回路モデル (等価回路モデル) に置き換えて伝達特性を推定する方法 [16], 3 次元声道モデルを計算機上に構築し、有限要素法 (FEM) により伝達特性を推定する方法 [17-23] などがある。声道模型等を用いる音響分析法は、声道形状が精密に再現されていれば正確な声道伝達特性が計測できる方法であるが、声道形状の影響を調査するための変形に煩雑さを有する方法である。声道の断面積から声道断面積関数を決定し、それを等価回路モデルに置き換えて声道伝達特性を計測する方法は音声分析、合成に広く用いられており、声道内部の音響的特徴の分析や、声道形状の変形が音響特徴に与える影響の調査も行うことができる。健常者の場合であれば声道断面の計測法が多数提

* Investigation of effectiveness to estimate vocal tract transfer functions by FEM, by Hironori Nishimoto, Masato Akagi, Tatsuya Kitamura and Noriko Suzuki.

^{*1} 北陸先端科学技術大学院大学情報科学研究科

^{*2} ATR 人間情報科学研究所

^{*3} 昭和大学歯学部

(問合先: 西本博則 〒 923-1292 能美市旭台 1-1 北陸先端科学技術大学院大学情報科学研究科 e-mail: nhiro@jaist.ac.jp)

(2005 年 5 月 20 日受付, 2005 年 11 月 21 日採録決定)

案されており、分析法の有効性の検証もなされているが、この手法を構音障害による音声の歪の分析に用いる場合には、複雑な形状を有する声道の音響的特徴の分析に対して有効であるかどうかの検討が必要となる。

FEMによる推定法は声道の3次元構造をそのままモデル化したものであり、より厳密な推定が行えると期待できる。FEMによる声道伝達特性の推定精度の検証として、音響管模型の音響分析結果との比較 [24, 25] がされており、推定値と比較して良好な推定精度であることが示されている。今後は声道のような複雑な形状に対する検証が期待されている。一方、FEMにより推定される声道伝達特性と音声スペクトルの比較 [22] がされている。この報告ではMR撮像時の発話音声と防音室のようなMR撮像室と異なる環境での発話音声の違いの検討がなされていないため、声道の伝達特性と発話音声の比較による検討を行うことが困難であった。また、口腔疾患患者のような複雑な形状を有する声道が伝達特性に影響を与えるのかどうかについて検討はなされていなかった。

本稿では2名の被験者に対してMR撮像と同時に、その音声の録音と音響分析を行い、その音声のホルマントとMR計測により得られる3次元声道形状モデルのFEM及び声道の等価回路モデルによる伝達特性のピーク周波数の比較を行い、推定精度の検討を行う。更に、健常者の声道形状と比較して声道形状が複雑で声道断面が非対称である1名の口腔疾患患者を被験者とし、特に音声の歪の著しい音声/i/発話時の声道のMR計測を行い、FEMと声道の等価回路モデルによる声道伝達特性の推定を行う。その伝達特性の推定結果と音声の音響分析結果を比較し、各手法による口腔疾患患者のような複雑な形状を有する声道の伝達特性の推定精度の検討を行う。更に、複雑で左右非対称な声道断面形状が声道伝達特性に与える影響について調査するために、断面形状の異なる声道モデルを作成し、FEMでそれらモデルの伝達特性を推定することでFEMによる複雑な声道形状の伝達特性推定の有効性を検証する。

2. 声道モデルの構築

声道形状の計測のために、MR撮像を行った。被験者の内訳は日本人成人男性2名(以下Subject-1, Subject-2とする)及び、口腔疾患により舌、口底切除手術を受けた日本人成人男性1名(以下Subject-Pとする)である。被験音はSubject-1, Subject-2については日本語5母音/a/, /i/, /u/, /e/, /o/とし、Subject-Pについては特に著しい音声の歪が聴感上認められる日本語母音/i/とした。

Subject-1, Subject-2のMR画像はATR脳活

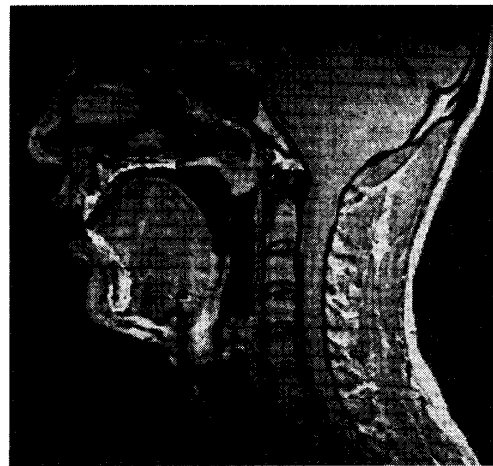


図-1 歯列補填された Subject-1 の/i/発話時の矢状断面のMR画像

動イメージングセンターに設置された Shimadzu-Marconi ECLIPSE 1.5T により撮像された。撮像シーケンスは Fast Spin Echo を用い、矢状方向、スライス厚 2.0 mm, 撮像範囲 256 mm×256 mm, 分解能 512 pixel×512 pixel, 51 スライス, TE 11 ms, TR 3,000 ms とした。Subject-1, Subject-2 の MR 画像には歯列補填法 [10] により歯列補填を施した。その画像を図-1 に示す。また、被験者の口元から約 200 mm 離れた位置にマイクロホン (SONY ECM-77) を設置し、MR 撮像室外の録音装置 (TEAC RW-02USB) で CD-R に音声データの収録を行った。MR 撮像に要する時間はおよそ 3 分であった。被験者には撮像開始の約 4 秒前から発話を開始させ、更に撮像終了後もしばらく発話させ、その間に音声の収録を行った。

Subject-P の MR 画像は昭和大学に設置された MAGNETOM VISION Ver31B (シーメンス社製) により撮像された。撮像シーケンスは Fisp3D を用い、矢状方向、スライス厚 2.5 mm, 撮像範囲 250 mm×250 mm, 分解能 256 pixel×256 pixel, 32 スライス, TE 2 ms, TR 5 ms とした。被験者への負担を考慮し、MR 撮像に要する時間をおよそ 21 秒とした。なお、この画像には歯列補填を施しておらず、被験者に歯冠プレートを着用させる方法 [9] で歯列形状の抽出を行った。Subject-P の音声は撮像と同時に収録が行えなかったため、撮像直後に別室の防音室で仰臥位で録音を行った。マイクロホンは SONY C-536P を使用し、SONY TCD-D10PRO II により DAT に音声データを収録した。

以上の方法で得られた MR 画像から声道形状を計測し、それを基に 3 次元声道形状モデルの作成を行う。この MR 画像から計測された声道形状を基に、3 次元声道モデル (サーフェスモデル) を作成する。それに

は Geomagic 社のサーフェスモデル作成ソフト Wrap を使用する。更に、このサーフェスモデルから Altair 社の汎用プリポストシステム HyperMesh により、ソリッドモデルである 3 次元声道モデルを作成する。モデルの表面の節点間隔は 1.0 mm である。また、口唇からの放射を表現するために口唇部に放射球面を取り付ける [26]。その半径は 40 mm とし、中心位置は左右の口角を結ぶ線分の中心とする。以上の手法により作成された Subject-1 の 3 次元声道モデルを図-2 に示す。また、Subject-P の 3 次元声道モデルを図-3 に示し、以下、このモデルを Model-O と呼ぶ。図-3 のように本モデルは口腔疾患の影響により /i/ 発声に必要な舌の持ち上げが見られず、口腔部は左右に広がりを持っており、複雑で左右非対称な形状となっている。また、咽頭部も舌切除の影響により非対称な断面形状となっている [2]。これらのモデルの要素数は最大は Subject-2 の /a/ 発話時の 229,242、最小は Subject-1 の /i/ 発話時の 128,974 であった。節点数は最大は同じく Subject-2 の /a/ 発話時の 75,012、最小は同じく Subject-1 の /i/ 発話時の 29,206 であった。

声道断面積関数を用いた等価回路モデルによる伝達

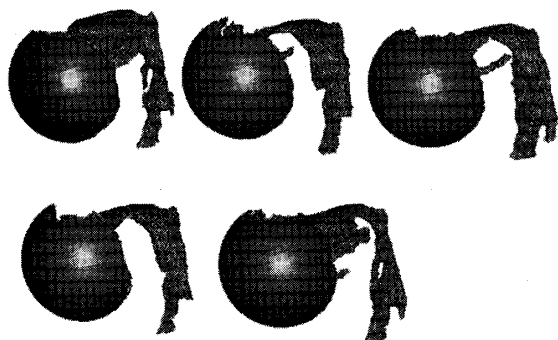


図-2 Subject-1 の 3 次元声道モデル
上段左から /a/, /i/, /u/ 発話時のモデル。下段左から /e/, /o/ 発話時のモデル。

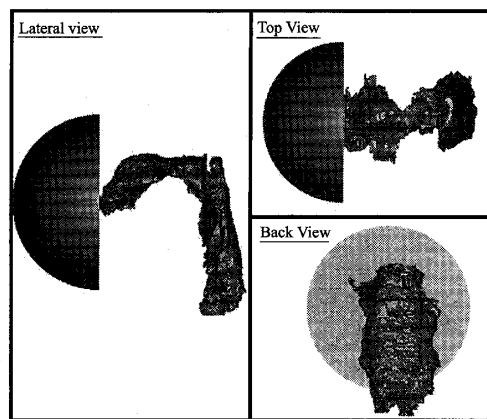


図-3 3 次元声道モデル
Model-O とする (Subject-P の /i/ 発話時のもの)。

特性の推定を行うため、MR 計測結果から声道の断面積を正中矢状断面の基準点を基に断面を決定する Tiede and Honda の手法 [27] を用いて計測し、声道断面積関数を求めた。各断面の間隔は 5 mm とした。

3. 音響分析

MR 撮像と同時に収録された Subject-1, Subject-2 の音声の分析を行った。MR 撮像中は撮像装置の騒音が非常に大きいので音声の録音は困難であるが、非撮像時は音声の録音が可能である。そこで、MR 撮像の直前と直後に収録された音声のスペクトル包絡を求めた。音響分析には対数スペクトルの不偏推定法 [28] を用いた。CD-R に録音された音声 (サンプリング周波数 44.1 kHz) を 11.025 kHz にダウンサンプリングし、分析区間として撮像直前、直後の音声区間から、有声区間の中心から前後 0.5 s の音声区間をそれぞれ切り出し、フレーム長 60 ms, フレームシフト 2 ms で分析を行った。以上の方法により得られた MR 非撮像時の MR 撮像室内の騒音と MR 撮像直後の /i/ 発話時の音声のスペクトル包絡を図-4 に示す。MR 撮像室内は冷却装置や空調などが常に作動しているために雑音が常に存在しているが、音声よりも雑音の振幅が小さい区間ではホルムント周波数の推定は可能である。分析の結果、第 1 ホルムントから順に F_1 , F_2 と表すと、Subject-2 の /a/, /u/ 及び /e/ の F_4 の推定ができなかったが、その他の音声では F_1 から F_4 までの推定が可能であった。

MR 撮像室内で録音された撮像直前と直後の音声のスペクトログラム及び、MR 撮像後に防音室内で収録された音声のスペクトログラムを図-5 に示す。防音室内での音声収録は MR 撮像時と同じく発話者は仰臥位である。図-5 より、撮像直前と直後の間にホルムント

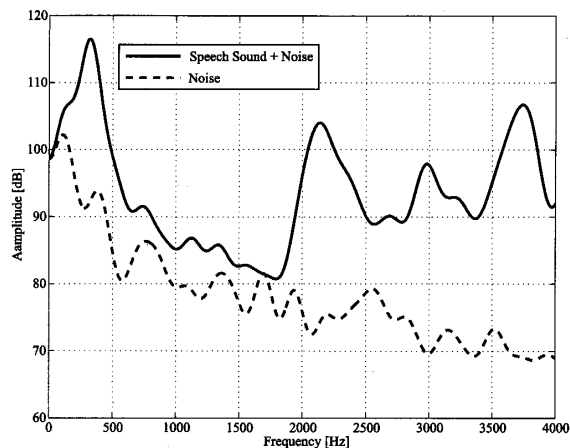


図-4 MR 撮像室内で録音された Subject-1 の /i/ 発話時のスペクトル包絡
実線：音声、破線：非発話時の騒音。

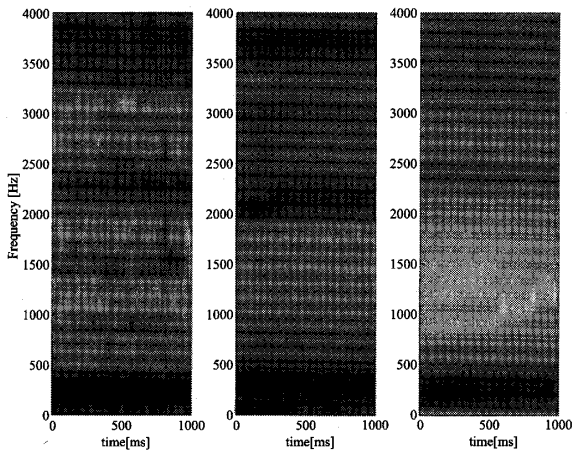


図-5 Subject-1 による/i/発声時のMR撮像直前(左)、直後(中)及び防音室内(右)で録音された音声スペクトル

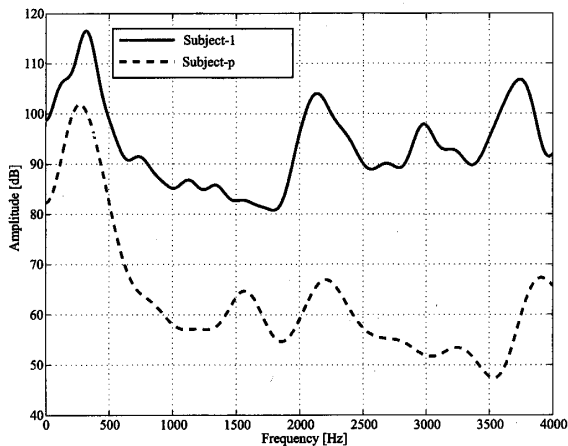


図-6 MR撮像室内で録音されたSubject-1の/i/発話時のスペクトル包絡と、防音室内で録音されたSubject-Pの/i/発話時のスペクトル包絡

周波数に変化していることが分かる。他の音声についても変化しており、それには一定の傾向や再現性が見られない。このようにホルマント周波数が大きく変動する場合には、撮像中に構音を一定に保つことができず、調音器官が動いていることが考えられるので、声道伝達特性との比較にはこのことを考慮する必要があると考えられる。

MR撮像室内の収録音声と防音室内での収録音声を比較すると、ホルマントの数は一致していた。しかし、それぞれのホルマント周波数には違いが見られ、特に F_3 以上では差が顕著になっている。これは、MR撮像室内と防音室内の雑音等の環境の違いによる影響[29]や、閉塞感や圧迫感等が原因となり構音が変わるためと考えられる。

Subject-Pの音声についても同様の分析を行った。Subject-1, Subject-Pの/i/発話時の音声スペクトル包絡を図-6に示す。Subject-Pの F_2 は1,550 Hz付近に存在しているが、これは健常者には見られないホ

ルマントである。これがSubject-Pの音声の歪の原因であると考えられる。

以上の結果から、音声のホルマントの数はどの録音環境でも変わらないので、伝達特性のピークの数とホルマントの数の比較を声道伝達特性のピーク周波数の推定精度の検討に用いることができると言える。また、MR撮像と同時に録音された音声であれば、そのホルマント周波数と声道伝達特性のピーク周波数の比較により、推定精度の検証を行うことができると言える。よって、Subject-1, Subject-2の声道伝達特性の推定精度の検討は、ホルマント周波数とその数との比較によって行い、Subject-Pの音声はMR撮像と同時に行われたものではないため、ホルマントの数と伝達特性のピークの数との比較により推定精度の検討を行う。

4. 声道伝達特性の推定法

FEMによる声道モデルの伝達特性の推定法について述べる。音響問題の速度ポテンシャル ϕ の定常状態における波動方程式は以下に示される。

$$\nabla^2 \phi = -k^2 \phi \quad (1)$$

ここで、 $k (= \omega/c, \omega: \text{角周波数 } c: \text{音速})$ は波数である。速度ポテンシャル ϕ から計算される音圧 p 及び粒子速度 i は、それぞれ

$$p = j\omega\rho\phi \quad (2)$$

$$i = -\nabla\phi \quad (3)$$

となる。ここで ρ は大気密度である。以上の式を用いて3次元声道モデルの伝達特性を計算する。FEMの計算にはNIT (Numerical Intergration Technologies)社の音響解析プログラムSYSNOISEを使用する。入力端では正弦波振幅が1.0の体積速度で励振させ、出力端となる放射球面上のインピーダンス Z は次式を用いた。

$$Z = \rho c \frac{jkr}{1 + jkr} \quad (4)$$

r は放射球面の半径を表す。それ以外の境界面は無損失(無損失)とするため、インピーダンスは ∞ を与えた。声道内の空気温度を考慮して、 36°C のときの音速 $c=353.46 \text{ m/s}$ 、気体密度 $\rho=1.1421 \text{ kg/m}^3$ で一定とした。声道モデルの伝達特性 H は次の式で求めた。

$$H = \frac{n_{in} A_{out} \sum I_{out}}{n_{out} A_{in} \sum I_{in}} \quad (5)$$

ここで、 $\sum I_{in}$ は入力面(声門部)の各節点における粒子速度の和であり、 $\sum I_{out}$ は出力面の各節点における粒子速度の和である。 A_{in} 及び A_{out} は入力面及

び出力面の表面積, n_{in} 及び n_{out} は入力面及び出力面の節点の数を表す。周波数の範囲は 20~4,000 Hz で 20 Hz 刻みとした。CPU が Pentium4 2 GHz, メモリが 1,024 MB の計算機により計算を行った際の計算時間は, 1 モデル当たり平均で 91 時間を要した。

次に, MR 画像より計測された断面積関数から声道伝達特性を推定する方法について述べる。本稿では Sondhi and Schroeter の手法 [16] を用いた。これによると入出力端の音圧と体積速度は次のようになる。

$$\begin{pmatrix} P_{out} \\ U_{out} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} P_{in} \\ U_{in} \end{pmatrix} \quad (6)$$

ただし, P_{in} , U_{in} はそれぞれ声門での音圧と体積速度であり, P_{out} , U_{out} はそれぞれ口唇での音圧と体積速度である。マトリックスの各構成要素 A , B , C , D は断面積関数, 断面間隔, 各気体定数により決定される。伝達特性 $H(\omega)$ は次の式で求められる。

$$H(\omega) = \frac{Z_L}{A - CZ_L} \quad (7)$$

Z_L は口唇での放射インピーダンスであり, 口唇部の半径を R mm とすると次の式で求められる。

$$Z_L = \frac{(\omega R)^2}{2c^2} + \frac{j\omega(8R)}{3\pi c} \quad (8)$$

分析範囲は 20~4,000 Hz であり, 各被験者の R の最大値は 17 mm であることから, 本手法による分析は適切であると言える。

5. 健常者の声道伝達特性とホルマント周波数の比較

FEM による 3 次元声道モデルの伝達特性の推定, 及び声道の等価回路モデルによる伝達特性の推定を行った。Subject-1 の /i/ 発話時の結果を図-7 に, Subject-2 の /i/ 発話時の結果を図-8 に示す。また, 各被験者の各音素の撮像直前, 直後のホルマント周波数, FEM 及び等価回路モデルにより推定された声道伝達特性のピークの値をまとめたものを図-9 に示す。

最初に, MR 撮像直前と直後の音声の音響分析より音声の各フレームで求めたホルマント周波数の平均値の比較を行う。撮像直前と直後ではホルマント周波数の平均値が変化する場合があり, 最大の差は Subject-2 の /u/ の F_2 で 300 Hz である。そこで, 声道モデルの伝達特性との比較には, 撮像直前と直後のホルマント周波数に対して F -検定による分析を行い, 危険率が 5% 以下のものだけを採用する。その結果, 棄却されるのは Subject-1 では /a/ の F_4 , /i/ の F_2 , そして /u/, /e/, /o/ の F_3 となり, Subject-2 では /a/ の F_3 , /u/ の F_1 及び F_3 となった。

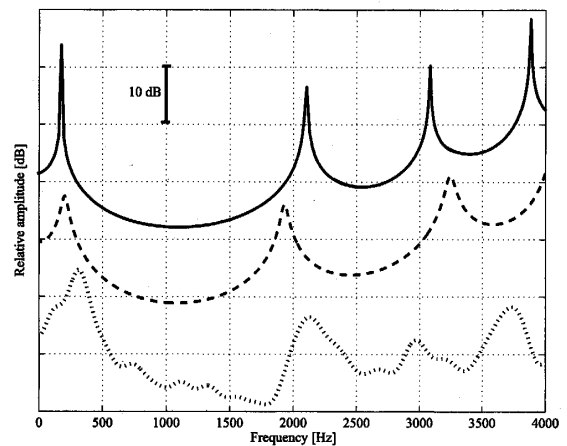


図-7 Subject-1 の /i/ 発声時の声道伝達特性と音声のスペクトル包絡
実線: FEM による推定結果。破線: 等価回路モデルによる推定結果。点線: 音声のスペクトル包絡。

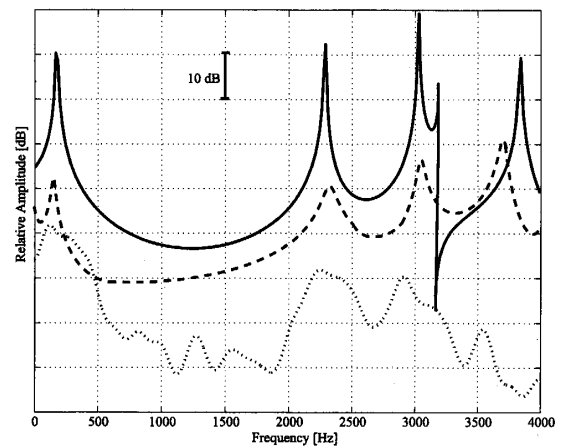


図-8 Subject-2 の /i/ 発声時の声道伝達特性と音声のスペクトル包絡
実線: FEM による推定結果。破線: 等価回路モデルによる推定結果。点線: 音声のスペクトル包絡。

FEM 及び等価回路モデルによる声道伝達特性の推定精度を検証するために, 推定された声道伝達特性のピーク周波数と, 撮像直前と直後の音声の各ホルマント周波数の平均値の比較を行った。

はじめに, 音声のホルマントの数と声道伝達特性のピークの数との比較を行ったが, それらは同一であった。

次に, 撮像前後の音声のホルマント周波数の平均値と FEM により推定された伝達特性のピーク周波数の比較を行う。 F_2 - F_4 の範囲では撮像前後の音声ホルマント周波数と伝達特性のピーク周波数の差は 6% 以内となっている。各ホルマント周波数の分散との比較を行った結果, その範囲内であり, また伝達特性の計算ステップが 20 Hz ごとであることも考慮すると伝達特性のピーク周波数は良好な精度で分析できていると言える。

F_1 では伝達特性のピーク周波数が音声ホルマント周

波数の平均値から約 55~210 Hz の差 (最大で約 30% の差) があつた。原因として、声道壁を剛壁として扱っているため、伝達特性の低域部のピーク周波数が音声のホルマントに対して低くなるという影響が現れていると考えられる。例えば、Subject-1 の /i/ の F_1 では音声のホルマントの平均値が 302 Hz であるのに対して FEM により推定された伝達特性のピーク周波数は 240 Hz であり、低域に移動していることが分かる。また、放射インピーダンスの設定法が影響を与えていることが考えられる。更に、MR 撮像室内では、撮像直前、直後にも図-4 にあるような雑音があり、特に低域に大きなエネルギーを持っているため、それがホルマント周波数の推定に影響を与えており、ホルマント周波数が正しく推定できていないとも考えられる。

等価回路モデルによる各ピーク周波数の推定結果は、Subject-1 の /i/ の第 4 ピーク周波数が 4,080 Hz となり、20~4,000 Hz という分析範囲から外れているが、それを除けば F_2 - F_4 の範囲での音声ホルマントとの

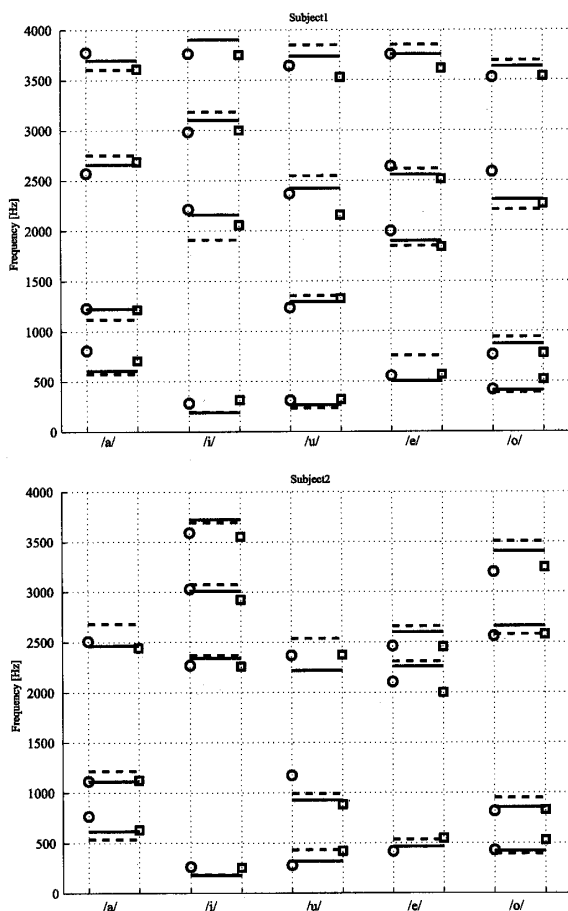


図-9 ホルマント周波数と伝達特性のピークの比較
上図が Subject-1、下図が Subject-2 の結果。各記号は実線が FEM により推定された声道伝達特性のピーク周波数、破線が等価回路モデルによる伝達特性のピーク周波数、“○” が撮像直前の音声の各フレームごとに求めたホルマント周波数の平均値、そして“□” が撮像直後の音声の各フレームごとに求めたホルマント周波数の平均値を表す。

差は 8% 以内であつた。 F_1 については最大で 30% の差になった。これは FEM 分析結果と同様理由が考えられる。

以上のことから、健常者の声道モデルの伝達特性の推定を行う場合、FEM による手法、及び等価回路モデルによる手法の双方で、20~4,000 Hz の範囲内では音声のホルマントの数と推定された伝達特性のピークの数と一致することが分かった。また、 F_1 以外についてはホルマント周波数に対応した声道伝達特性のピーク周波数が得られることが分かった。

6. 口腔疾患患者の声道伝達特性の推定

本節では、Subject-P の声道伝達特性の推定結果を示す。Model-O の FEM により推定された伝達特性と、声道の等価回路モデルにより推定された伝達特性、及び Subject-P の音声スペクトル包絡を図-10 に示す。音声のホルマントは 4,000 Hz までの帯域に四つ存在する。また、FEM により推定された伝達特性のピークも同様に四つ存在し、それらの周波数は比較的近い値であり、対応が取れていると考えられる。一方、声道の等価回路モデルによる伝達特性のピークは三つしかなく、音声のホルマントの数との対応が取れていないことが分かる。

そこで、FEM で推定された声道伝達特性のピーク数が音声のホルマントの数と一致し、等価回路モデルで推定された声道伝達特性のピーク数が音声ホルマントの数と一致しない原因を調査する。

FEM と等価回路モデルの違いの一つとして声道断面の形状的特徴を反映した推定手法であるかどうかという点が挙げられる。Model-O の声道形状、特に口腔部では左右非対称で複雑な形状をしており、これが推定結果や音響特徴に関係しているかを検証する。その

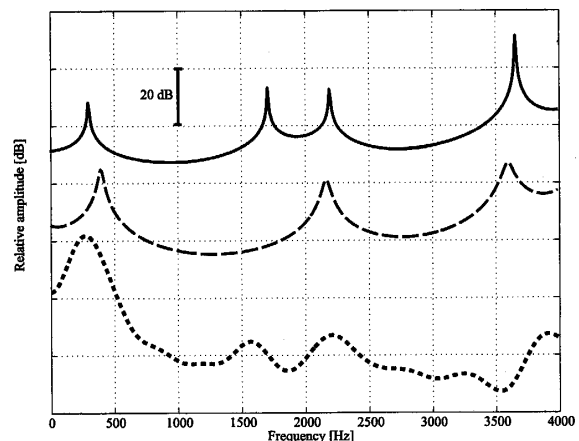


図-10 Subject-P の /i/ 発声時の声道伝達特性と音声のスペクトル包絡
実線：FEM による推定結果。破線：等価回路モデルによる推定結果。点線：音声のスペクトル包絡。

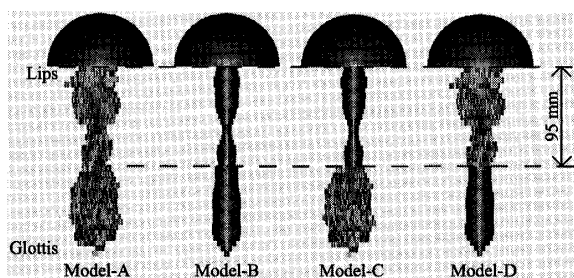


図-11 声道断面を元に作成した 3 次元声道モデル
左から断面形状を変えずに並べなおしたもの (Model-A)、
断面積一定のままで円筒管に置き換えたモデル (Model-B)、
Model-A の口唇部から 95 mm までを Model-B で
置き換えたモデル (Model-C)、Model-A の口唇部から
95 mm の断面から声門までを Model-B で置き換えたモ
デル (Model-D)。

ために、Model-O の一部、又はすべてを断面積を変えずに左右対称な断面形状に変形したモデルを作成し、それらのモデルの声道伝達特性を FEM により計算することで、声道の非対称性を有する部位である口腔部の形状が音響特徴に影響を与えるかどうかを調査する。

変形モデルは、変形操作を簡単化するために、Model-O の曲がりの影響をなくし、声道中心線を直線に変換したモデルにする。Model-O の断面が直線と直行するように 5 mm 間隔で並べた 3 次元モデル (Model-A)、及び、断面積はそのままで、円筒管の縦続接続に置き換えた 3 次元モデル (Model-B) を作成した。また、本被験者の声道形状の特徴である複雑で左右非対称な断面形状が見られる口腔部の影響を調査するため、Model-A の口唇部から 95 mm までの口腔部を Model-B と置き換えたモデル (Model-C) を作成した。更に、Model-A の口唇部から 95 mm の場所から声門までの部分を Model-B と置き換えたモデル (Model-D) を作成した。なお、これらモデルの口唇部での放射を表現するために、口唇部に半径 40 mm の放射球面を取り付けた。これらのモデルを図-11 に、伝達特性を図-12 に示す。これらのモデルの要素数は最大は Model-A の 143,282、最小は Model-B の 128,741 であった。節点数は最大は同じく Model-A の 38,611、最小は同じく Model-B の 33,451 であった。各モデルの計算時間の平均は 90 時間であった。

Model-O と Model-A の伝達特性を比較すると、ピークの数は一貫しており、声道中心線を直線に置き換えてもピーク周波数はホルマント周波数に対して大きく変化していないと言える。Model-A と Model-B の伝達特性を比較すると Model-A の伝達特性の 2 番目のピークが Model-B の伝達特性ではなくなっており、断面積関数が同じでも伝達特性が変わりえるということが分かる。更に、Model-C の伝達特性は Model-B の伝

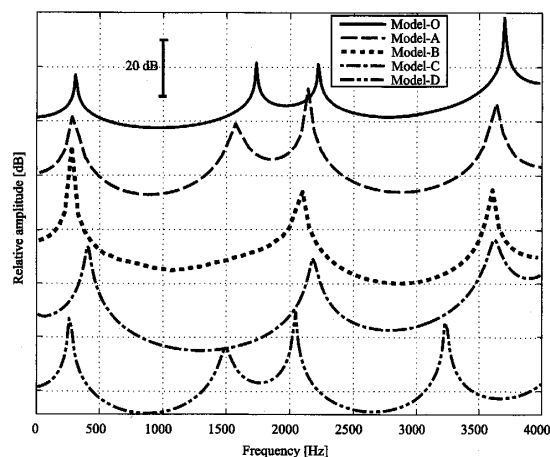


図-12 Model-O, Model-A, Model-B, Model-C 及び Model-D の FEM で推定した声道伝達特性

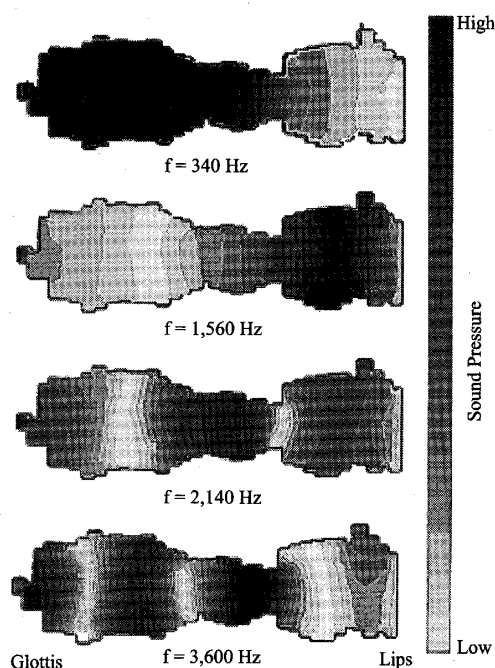


図-13 Model-A のピーク周波数での音圧分布

達特性と同様に 2 番目のピークが消えていることが分かる。また、Model-D の伝達特性は Model-A の 2 番目のピークに対応するピークが存在している。Model-B, -C, -D とともに断面積は変わらないことから、本被験者の音響的特徴である 2 番目のピークは軟口蓋から口唇部にかけての本被験者の声道の形状の特徴、特に舌・口底切除の影響による口腔内の断面形状の非対称性に起因して生成されていると考えられる。

Model-A, Model-B の各ピーク周波数での音圧分布を図-13, 図-14 に示す。また、Model-A の第 2 ピーク周波数での位相の分布を図-15 に示す。Model-B の等圧線は中心線に対して直交しており、FEM による声道伝達特性の推定法でも等価回路モデルと同じく平面波の音圧分布と同じ分布となることが分かった。一

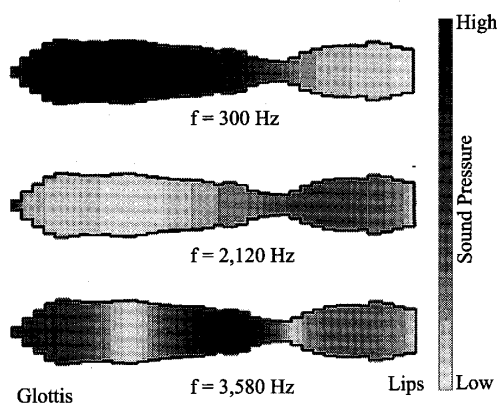


図-14 Model-B のピーク周波数での音圧分布

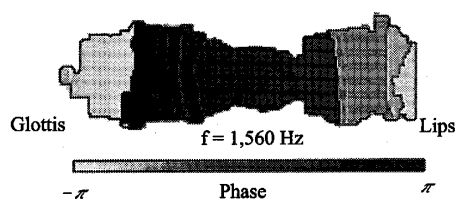


図-15 Model-A の第 2 ピーク周波数での位相の分布

方, Model-A の等圧線は口腔部では平面波とは異なる分布であった。特に, 伝達特性に違いが見られる周波数である 1,560 Hz では図中の口腔部上下に音圧が高い部分が見られ, Model-B のものとは全く異なる分布であった。Model-A では口腔内での等価回路モデルのような伝搬方向とほぼ直交するように等圧面が現れていることから, 声道の断面形状が本被験者のように複雑で左右非対称であると, その特徴が声道内での音波の伝搬に影響を与えていると考えられる。図-15 の Model-A の第 2 ピークでの位相の分布を見ると, 口腔内部の位相は同一であることから, 図中の口腔部上下間でのモードが発生していると考えられる。

以上の結果から, 声道断面形状も伝達特性に影響を与えており, これが Model-O の伝達特性のピークの数と音声ホルマントの数を比較したときに, FEM で推定された伝達特性では同じであるが等価回路モデルで推定されたものは異なる原因であると考えられる。

7. 考 察

本章では, FEM 及び等価回路モデルによる各被験者の声道伝達特性の分析結果をまとめ, それらの推定精度, 特徴について考察する。

声道の断面形状がほぼ左右対称である健常者 Subject-1, Subject-2 の声道伝達特性の分析を行った結果, FEM による方法, 及び等価回路モデルによる方法により得られた声道伝達特性のどちらも音声のホルマント周波数, 及び同一帯域内でのホルマントの数が一致していることが分かった。このことから, 健常

者の場合は, 3次元声道モデルの FEM による声道伝達特性の推定, 及び断面積関数を用いた等価回路モデルによる声道伝達特性の推定法のどちらでも良好な精度で行うことができると言える。

一方, 声道断面形状が複雑で左右非対称になっている Subject-P の声道伝達特性を健常者の場合と同様に 2通りの方法で推定した結果, FEM による推定結果は音声のホルマント周波数と伝達特性のピーク周波数がよく対応しており, 更にホルマントの数が伝達特性のピークの数と一致していた。一方, 等価回路モデルの場合では音声ホルマントの数が伝達特性のピークの数と一致していなかった。このことから, 口腔疾患患者のような複雑な形状を有する声道の伝達特性の推定には等価回路モデルによる方法を用いることは不適切であり, FEM のような平面波伝搬を仮定していない推定法を用いなくてはならないことが示された。また, Model-O の口腔内の左右非対称性に着目して, 断面積関数を変えずに断面形状を円形で置き換えたモデルを作成し, 断面形状と伝達特性の関係を調査した結果, 口腔部の形状が声道伝達特性のピークの数に影響を与えていることが分かった。Model-A と Model-B の等圧線を調査すると, Model-B の等圧線は中心線に対して直交しており, 平面波の音圧分布と同じ分布となっているが, Model-A の等圧線は口腔部では平面波とは異なる分布であった。声道の断面形状が本被験者のように複雑で左右非対称であると, その特徴が声道内での音波の伝搬が平面波とは異なる音波の伝搬が生じることがあることが分かった。

以上のことから, 等価回路モデルは口腔疾患患者のような複雑な声道形状の伝達特性の推定には適切ではないことが分かった。また, FEM による声道伝達特性の推定は, 声道の形状の特徴, 特に口腔疾患患者のような複雑な声道形状が音響特性に及ぼす影響を調査する際に有効な手法であると考えられる。FEM を用いることで, 口腔疾患患者の口腔形状と音響的特徴の関係の調査ができ, 異常音声の治療に役立てることができると考えられる。

8. おわりに

本稿では, 2名の健常者を被験者として MR 撮像と同時に音声の収録を行い, 撮像直前, 直後の音声のホルマント周波数を音響分析により推定した。更に MR 画像から構築した 3次元声道モデルの FEM による伝達特性の推定結果, 声道断面積関数を用いた等価回路モデルによる伝達特性の推定結果と音響分析の結果を比較した。その結果推定された伝達特性のピーク周波数とホルマント周波数は F_1 以外は非常に近い値であ

り、精度良く伝達特性が推定されており、FEM 又は声道の等価回路モデルの伝達特性と実音声のホルマントを対応付けられることが確認された。よって声道断面形状がほぼ左右対称な健常者の声道伝達特性の推定は FEM 及び等価回路モデルのどちらでも良好な精度で行えることが定量的に示された。

更に、声道形状が複雑で断面形状が複雑で左右非対称な 1 名の口腔疾患患者の MR 撮像も行い、その伝達特性を健常者の場合と同様の方法で推定した。音声のホルマントと伝達特性のピークの数と比較すると、FEM による推定法では一致したが、等価回路モデルによる方法では、それらの数が一致せず、口腔疾患患者の音声の特徴であるホルマントに対応するピークが見られなかった。更に、この被験者の声道をもとに、声道断面積関数は同じで、断面形状の異なる声道モデルを作成し、その伝達特性を FEM で計算したところ、声道伝達特性に違いが見られ、音圧分布が平面波とは異なっていることから、声道断面形状が声道伝達特性に影響を与えることが分かった。このことから、複雑な声道形状を有する声道形状の声道伝達特性の分析には声道の構造的特徴を捉えた分析法が必要であると言え、FEM による伝達特性推定法が有効であると言える。口腔疾患患者のような複雑な声道の音響分析を行う際に FEM による声道伝達特性の推定法を用いることで、これまで医療従事者の聴観印象に頼ってきた構音障害の治療や発話補助具の設計に対して一定の指針を与える等の応用が期待される。

謝 辞

本研究は一部 NICT の支援を受けて実施したものである。MR 撮像に関して協力いただいた ATR 脳活動イメージングセンタに感謝する。

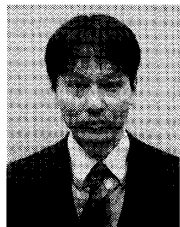
文 献

- [1] 齋藤浩人, 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—,” 口腔科学会雑誌, 49, 92-101 (2000).
- [2] 北村達也, 鈴木規子, 齋藤浩人, 道 健一, 高橋俊行, 赤木正人, 和久本雅彦, “MRI による舌・口底切除患者の 3 次元声道形状の分析,” 信学技報, SP2000-153 (2001).
- [3] 高木直子, 北村達也, 赤木正人, 鈴木規子, 藤田幸弘, 道 健一, “側音化構音の音響特性について,” 音講論集, pp. 363-364 (1996.3).
- [4] 齋藤浩人, 鈴木規子, 北村達也, 赤木正人, 道 健一, “舌・口底切除における異常構音の音響的特徴—スペクトルのピーク分析の試み—,” 信学技報, SP98-149 (1999).
- [5] M. Akagi, N. Suzuki, K. Hayashi, H. Saito and K. Michi, “Perception of lateral misarticulation and its physical correlates,” *Folia Phoniatr. Logop.*, 53, 291-307 (2001).
- [6] S. Kiritani, K. Itoh and O. Fujimura, “Tongue-pellet tracking by a computer-controlled X-ray microbeam system,” *J. Acoust. Soc. Am.*, 57, 1516-1520 (1975).
- [7] 片岡竜太, 秋月弘道, 齊藤健一, 大野康亮, 吉田 広, 道 健一, 関 健次, 岡野友宏, 岡部 馨, “3 次元 CT を用いた声道, 鼻咽腔断面積の計測—口腔癌疾患術前/i/発話時について—,” 音講論集, pp. 253-254 (1992.3).
- [8] 松村雅史, 杉浦 淳, “磁気共鳴画像法 (MRI) による声道形状の計測,” 信学論, J75-A, 1115-1123 (1992).
- [9] 和久本雅彦, 正木信夫, 党 建武, 本多清志, 藤本一郎, 中村裕二, 島田育廣, “MRI 用歯冠造影剤プレートの摩擦音構音への影響—X 線マイクロビームを用いた観察—,” 信学技報, SP97-123 (1998).
- [10] T. Takemoto, T. Kitamura, H. Nishimoto and K. Honda, “A method of tooth superimposition of MRI data for accurate measurement of vocal tract shape and dimensions,” *Acoust. Sci. & Tech.*, 25, 468-474 (2004).
- [11] O. Fujimura and J. Lindqvist, “Sweep-tone measurements of vocal tract characteristics,” *J. Acoust. Soc. Am.*, 49, 541-558 (1971).
- [12] M.M. Sondhi and B. Gopinath, “Determination of vocal tract shape from impulse response,” *J. Acoust. Soc. Am.*, 49, 1867-1873 (1971).
- [13] K. Motoki, “Measurement of sound-pressure distribution in replicas of the oral cavity,” *J. Acoust. Soc. Am.*, 92, 2577-2585 (1992).
- [14] J. Dang and K. Honda, “New method for measuring anti-resonance details of the vocal tract transmission characteristics — An experimental study of acoustic tubes —,” *J. Acoust. Soc. Jpn. (E)*, 17, 93-99 (1996).
- [15] 持田岳美, 菅田雅彰, 横山 徹, 大谷 猛, 三木信弘, 和久本雅彦, 正木信夫, 本多清志, “口腔復元模型を用いた声道断面積関数の音響的計測,” 音講論集, pp. 307-308 (1999.9).
- [16] M.M. Sondhi and J. Schroeter, “A hybrid time-frequency domain articulatory speech synthesizer,” *IEEE Trans. Acoust. Speech Signal. Process.*, 35, 955-967 (1987).
- [17] C. Lu, T. Nakai and H. Suzuki, “Finite element simulation of sound transmission in vocal tract,” *J. Acoust. Soc. Jpn. (E)*, 14, 63-72 (1993).
- [18] 松崎博季, 三木信弘, 小川吉彦, “音響管内の 3 次元の音響的特徴量の有限要素法による解析,” 信学論, J80-A 1351-1359 (1997).
- [19] T. Niikawa, T. Ando and M. Matsumura, “Frequency dependence of vocal-tract length,” *Proc. IC-SLP 2000*, pp. 1525-1528 (2000).
- [20] 中井孝芳, 森川将太, 杉山哲也, “2 次元声道の有限要素法解析に基づく断面積の推定,” 信学技報, SP2002-57 (2002).
- [21] 北村達也, 竹本浩典, 本多清志, “ボクセルメッシュを用いた有限要素法による声道の音響解析,” 信学技報, SP2004-78 (2004).
- [22] 西本博則, 赤木正人, 北村達也, 鈴木規子, 齋藤浩人, 道 健一, 高橋俊行, “舌・口底切除患者の 3 次元声道形状モデルの有限要素法による分析,” 信学技報, SP2000-154 (2001).
- [23] 西本博則, 赤木正人, 北村達也, 鈴木規子, “舌・口底切除患者の声道モデルを変形させたモデルの声道伝達特性の分析,” 信学技報, SP2002-58 (2002).
- [24] 山崎 憲, 瀬谷浩一郎, “開放型段付き管内音圧分布の有限要素計算と実験,” 音響学会誌, 46, 711-715 (1989).
- [25] 北村達也, 西本博則, 藤田 覚, 本多清志, “音響計測, 電気回路理論および有限要素法により求めた声道模型の伝達関数の比較,” 信学技報, EA2003-7, SP2003-7 (2003).
- [26] 松崎博季, 三木信弘, 小川吉彦, “3-D FEM による声道の放射空間付き楕円近似音響管モデルに関する一考察,” 信学技報, SP95-113 (1996).

[27] M.K. Tiede and K. Honda, "Improved area function estimation from midsagittal distances on 3D MRI DATA," 音講論集, pp. 239-240 (1999.9).

[28] 今井 聖, 古市千枝子, "対数スペクトルの不偏推定," 信学論, **J70-A**, 471-480 (1987).

[29] J.B. Tufts and T. Frank, "Speech production in noise with and without hearing protection," *J. Acoust. Soc. Am.*, **114**, 1069-1080 (2003).



西本 博則

平 11 年大阪工業大・工・機械工卒。平 13 年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同大同研究科博士後期課程在学中。音声の生成の研究に従事。日本音響学会会員。



赤木 正人

昭 54 名工大・工・電子卒。昭 59 東工大大学院博士課程情報工学専攻了。工博。同年電電公社（現 NTT）研究所入社。以来、ATR 視聴覚機構研究所、NTT 基礎研究所を経て、現在、北陸先端科学技術大学院大学情報科学研究科教授。音声信号処理、聴覚機構のモデル化の研究に従事。電子情報通信学会、IEEE、ASA、ISCA 各

会員。



北村 達也

平 4 年山形大・工・情報工卒，平 6 年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了，平 9 年同博士後期課程修了。その間，平 7 年より平 9 年まで日本学術振興会特別研究員。博士（情報科学）。同年静岡大学情報学部助手。平 14 年 ATR 人間情報科学研究所。音声の知覚・生成の研究に従事。日本音響学会，電子情報通信学会，日本音声学会，ASA 各会員。



鈴木 規子

東京医科歯科大学歯学部卒業。歯学博士。現在昭和大学歯学部顎口腔疾患制御外科学教室助教授。口腔外科指導医，外国人修練指導医，顎顔面補綴学会理事。言語病理学，音声科学，音声生理学の研究に従事。