

境界要素法を用いた日本語母音/a/発声時の 音響シミュレーション法の開発

氏名 白石 麻美

所属 山口大学大学院医学系研究科
医学専攻 歯科口腔外科学分野

令和1年12月

目次

1. 概要	1
2. 緒言	2
3. 研究方法	3
3.1 対象	3
3.2 シミュレーション方法	3
3.2.1 CT 撮影方法	3
3.2.2 声道モデルの作成	3
3.2.3 解析メッシュの作成	4
3.2.4 音響解析法	4
3.2.5 音響解析モデルのパラメータの設定	6
3.3 精度検証	7
3.3.1 実音声との比較	7
3.3.2 実体模型を用いた検証	7
3.3.3 評価方法	8
4. 結果	8
5. 考察	9
6. 結論	13
7. 謝辞	13
8. 参考文献	14

1. 概要

目的 声道の音響シミュレーション法を確立し、そのシミュレーション法の妥当性を検証することを目的とする。

方法 健常成人 15 名（男性 8 人、女性 7 人）を対象とした。/a/発声時の CT 撮影を行い、声道モデルを作成した。このモデルに対して境界要素法を用いた波動音響解析を行った。解析条件として、声道モデルの壁面と底面はそれぞれ剛壁と剛壁ではない壁とし、音響媒質は 37°C、音源は声帯相当部に点音源を設定した。設定した観測点において、周波数応答曲線を求め、第 1、第 2 フォルマント（F1、F2）周波数を算出した。1 名の被験者は前頭洞上端から気管分岐部の範囲を撮影し、前頭洞上端から気管分岐部までの範囲で作成した声道モデル（モデル 1）と、前頭洞上端から声門までの範囲に加えて気管部を円柱状に 12cm 延長した声道モデル（モデル 2）を作成した。本シミュレーション法の妥当性を検証するために、算出された F1 と F2 を、実音声、および CT から作製した実体模型に笛式人工喉頭にて音源を与えて産出された音声とをそれぞれ比較した。その判断根拠として母音フォルマント周波数弁別閾値を参照した。

結果 モデル 1 と 2 から算出される F1 と F2 は近似していたため、残り 14 名は前頭洞上端から声門までを撮影範囲とし、気管部を円柱状に 12cm 延長したモデル 2 をシミュレーションに使用した。シミュレーションにより算出されたフォルマント周波数と実音声を比較すると、相対弁別閾値の平均は F1 が 5.9%、F2 が 4.6%、また、実体模型と笛式人工喉頭から産出される音声と比較すると、相対弁別閾値の平均は F1 が 4.1%、F2 が 3.7%であった。

結論 CT データから/a/発声時の声道モデルを作成し、境界要素法を用いた波動音響解析を適応することにより、実音声に近似したフォルマント周波数を算出することが可能となり、本シミュレーション法は十分な妥当性を有することが示された。

キーワード：構音障害、音響シミュレーション、境界要素法、声道モデル、日本語母音/a/

2. 緒言

口蓋裂に対して行われる口蓋形成術には様々な術式の改良が加えられているが、依然術後に構音障害が残ることは少なくない。構音障害に対して、発音補正装置を用いた言語訓練や二次手術が適応されるが、訓練方法や手術術式の選択に明確な基準がなく、各施設で様々なプロトコールにより治療が実施されているのが現状である。これは、口蓋裂術後に生じる構音障害の発生機序が明らかにされていないことが最大の原因である。この構音障害の原因として、鼻咽腔の動きが大きく関わっていることは言うまでもないが、それ以外に口蓋や舌などの構音器官の形態と運動能力などが大きく影響していると考えられている[1]。過去には、口蓋の深さ、狭窄、短縮などの形態異常と口蓋化構音の出現に関与の可能性が示唆されており[2]、また反対咬合による異常な舌の位置異常が子音の口蓋化を引き起こすという報告もある[3]。しかしながら、その関連性を詳細に解析する手法がないために不明な点が多くある。

近年、有限要素法や境界要素法などの波動音響解析の発達により、三次元的な音響伝達空間の解析が容易となった。声道の三次元形状が得られれば、それらの解析方法を用いることで、形態と音響の関係を可視化し、詳細に分析することが可能となってきた。例えば、有限要素法と境界要素法による声道音響解析を比較した結果に大きな差はなかったという報告[4]や、MRI データから声道の三次元形状モデルを作成し、有限要素法を応用して鼻腔を含めた音響解析の結果、鼻腔に伴う特徴である極零対を確認したという報告がある[5]。しかしながら、これらの報告では、解析から得られる第1フォルマント (F1) と第2フォルマント (F2) は、実音声の F1、F2 と比較すると、乖離した結果となっている[4, 5]。

そこで、境界要素法による音響解析を構音器官に応用することで、声道とそこで産生される音、すなわち声道の形態と音声との関連性をシミュレーションし、両者の関連性を明らかにし、構音障害の発生機序を解明することをめざした。本研究では、声道の音響シミュレーション法の確立を目的として、その妥当性を検証した。

3. 研究方法

3.1 対象

被験者は健常成人 15 名（男性 8 人、女性 7 人）、年齢は 26 歳から 45 歳までとした。構音器官の異常は認めず、咬合関係は正常咬合であった。

山口大学医学部附属病院臨床研究審査委員会の承認を受け（H26-22-4）、被験者から同意を得て行った。

3.2 シミュレーション方法

3.2.1 CT 撮影方法

撮影には CT 装置 SOMATOM Force (Siemens Healthineers, München, Germany) を用いた。撮影条件を以下に示す。管電圧は 100(+Tin filter)kV、管電流は 96mAs、スライス厚は 0.6mm、CTDI vol は 0.23mGy とした。SOMATOM Force は被爆量が非常に少なく、通常の胸部 CT で使用する被爆線量 5~30mSv 程度に対して、胸部レントゲンと同程度の 0.1mSv とされている。仰臥位にて、母音 /a/ を持続発声状態で撮影を行った。15 名のうち 1 名の撮影範囲は、前頭洞上端から気管分岐部までとし、撮影時間は約 10 秒であった。後述する解析の結果、前頭洞上端から声門下部までのモデルを気管分岐部相当部まで仮想的に延長することで良好なシミュレーション結果が得られたため、被爆量低減のためにその他 14 名は前頭洞上端から鎖骨上とし、撮影時間は約 5 秒であった。

3.2.2 声道モデルの作成

CT データを 3D 可視化ソフト Amira® (version 5.6.0 Maxnet, Tokyo, Japan) を用いて、スライスデータの可視化と 3 次元モデルを作成した。前頭断、水平断、矢状断の CT スライスデータよりモデル作成に必要な気道を抽出し、声道モデルとした（図 1）。この声道

モデルは STL フォーマットとした。声道モデルの範囲については、これまでの報告で口腔・咽頭・喉頭・声門のモデルを作成したもの[6]、鼻腔・口腔・咽頭・喉頭・声門のモデルと鼻腔を除いたモデルを作成し、比較したものがある[5]。これらのフォルマント周波数解析は共に実音声より低い結果であった[5, 6]。そのため、本研究では前頭洞・篩骨洞・蝶形骨洞・上顎洞・鼻腔・口腔・咽頭・喉頭・声門・気管までの抽出を行い、声道モデルとした（モデル 1）。前頭洞上端から鎖骨上まで CT 撮影を行い、前頭洞・篩骨洞・蝶形骨洞・上顎洞・鼻腔・口腔・咽頭・喉頭・声門まで抽出した声道モデルをモデル 2（図 2）として、それぞれ作成し、以下の解析を行った。

3.2.3 解析メッシュモデルの作成

本研究は境界要素法を用いた音響解析を行うため、声道モデルから解析メッシュモデルの作成を行った。STL 形式で出力した三次元声道モデルをダイレクトモデラー— Space Claim Direct Modeler (SCDM) (version 19.0 ANSYS, Canonsburg, PA, USA) を使用してメッシュサイズの調整および均一化や不要なメッシュを削除した。鼻腔と口腔を開放状態にするために開口部のメッシュを削除した。メッシュ作成の簡便さから三角形要素とし、音響解析ソフト WAON® (version 4.42 Cybernet, Tokyo, Japan) を使用し、ハードウェア環境下 (Inter® Xeon® E-1650 3.5GHz CPU) で解析可能である最小単位の 2mm メッシュでモデルを作成した (図 3)。メッシュのサイズについては、解析周波数の音波の波長の 1/6 より小さいサイズが推奨され[7]、体温 37°C (音速 352.85m/s) で 2,000Hz の音響解析を行うこととしたため、 $\text{波長}/6=352.85/2000/6=0.029 \text{ (m)} =29 \text{ (mm)}$ となり、十分に条件を満たすと考えた。

3.2.4 音響解析法

音響解析には境界要素法を利用した。音響解析ソフト WAON® (version 4.42 Cybernet,

Tokyo, Japan) を使用し、1~2,000Hz まで 1Hz 刻みでの解析を行った。

音の伝搬を表す波動方程式として、Kirchhoff-Helmholtz 積分方程式を使用し、離散化手法として境界要素法を用いた。以下に解析方法について示す。

まず、境界条件を設定する。全境界 Γ は、 Γ_0 : 剛境界、 Γ_1 : 振動境界、 Γ_2 : 吸音境界の三種類の境界から構成されるものとする。三種類の境界においては局所作用の仮定の下、以下の境界条件式が成り立つものとする。

$$\frac{\partial p(\mathbf{r}_q)}{\partial n_q} = \begin{cases} 0 & q \in \Gamma_0 \\ j\omega\rho v(\mathbf{r}_q) & q \in \Gamma_1 \\ -jkp(\mathbf{r}_q)/z(\mathbf{r}_q) & q \in \Gamma_2 \end{cases} \quad (1)$$

ここで j は虚数単位、 ρ は媒質密度、 $\omega = 2\pi f$ は角周波数、 f は周波数、 \mathbf{r}_q は境界上の点 q の位置ベクトル、 $\partial/\partial n_q$ は点 \mathbf{r}_q における境界面（領域内向き）法線方向微分、 $k = \omega/c$ は波数、 c は音速、 z は吸音境界の比音響インピーダンス比である。

次に、境界積分方程式を解くために境界を要素分解し、方程式を離散化する。Kirchhoff-Helmholtz 積分方程式に三種類の境界条件を適用して、境界要素で離散化すると以下の式を得る [8]。

$$(\mathbf{E} + \mathbf{B} + \mathbf{C})\mathbf{p} = \mathbf{A}\mathbf{v} - \mathbf{p}_d \quad (2)$$

ここで、 \mathbf{p} と \mathbf{v} はそれぞれ境界要素節点上の音圧と振動速度を成分に持つ列ベクトルである。行列 \mathbf{A} 、 \mathbf{B} 、 \mathbf{C} 、 \mathbf{E} の各要素は、それぞれ以下の通りである。

$$E_{ij} = -\frac{1}{2}\delta_{ij} \quad (3)$$

$$A_{ij} = a_j(\mathbf{r}_i) = j\omega\rho \int_{\Gamma_1} N_j(\mathbf{r}_q)G(\mathbf{r}_i, \mathbf{r}_q) dS_q \quad (4)$$

$$B_{ij} = b_j(\mathbf{r}_i) = \int_{\Gamma} N_j(\mathbf{r}_q) \frac{\partial G(\mathbf{r}_i, \mathbf{r}_q)}{\partial n_q} dS_q \quad (5)$$

$$C_{ij} = c_j(\mathbf{r}_i) = jk \int_{\Gamma_2} N_j(\mathbf{r}_q) \frac{G(\mathbf{r}_i, \mathbf{r}_q)}{z(\mathbf{r}_q)} dS_q \quad (6)$$

ここで δ_{ij} はクロネッカーのデルタ、 \mathbf{r}_i は境界要素節点 i の位置ベクトル、 N_j は境界要素節点 j の内挿関数である。 \mathbf{p}_d は点音源から放射される音圧の境界要素節点における値で、以

下の通りである。

$$\mathbf{p}_d = p_d(\mathbf{r}_i) = SG(\mathbf{r}_i, \mathbf{r}_s) \quad (7)$$

ここで S は音源の強度、 \mathbf{r}_s は音源の座標である。また基本解 G は次式で表される。

$$G(\mathbf{r}_p, \mathbf{r}_q) = \frac{\exp(jk r_{pq})}{4\pi r_{pq}} \quad (8)$$

ここで、 $r_{pq} = |\mathbf{r}_p - \mathbf{r}_q|$ 、 $r_{pq} = |\mathbf{r}_{pq}|$ である。境界要素法では、まず主処理として(2) 式を解くことにより境界要素節点上の音圧 \mathbf{p} を求める。

最後に後処理として、音場内に設定する観測点 \mathbf{r}_r における音圧 $p(\mathbf{r}_r)$ を求める。 $p(\mathbf{r}_r)$ は、境界要素節点上の音圧 \mathbf{p} と振動速度 \mathbf{v} を用いて、Kirchhoff-Helmholtz 積分方程式を離散化した次式により算出される[8]。

$$p(\mathbf{r}_r) = [\mathbf{b}^T(\mathbf{r}_r) + \mathbf{c}^T(\mathbf{r}_r)]\mathbf{p} - \mathbf{a}^T(\mathbf{r}_r)\mathbf{v} \quad (9)$$

ここで行ベクトル \mathbf{a}^T 、 \mathbf{b}^T 、 \mathbf{c}^T の各要素 $a_j(\mathbf{r}_r)$ 、 $b_j(\mathbf{r}_r)$ 、 $c_j(\mathbf{r}_r)$ は、それぞれ(4)、(5)、(6) 式で与えられ、音圧が求められる。

3.2.5 音響解析モデルのパラメータの設定

声道モデルの壁面は剛壁とし、吸音率 0%、比音響インピーダンス $=\infty$ とした。モデル 1 の開放状態としている気管分岐部は比音響インピーダンス $=0$ の剛壁ではない壁とした[9]。モデル 2 は声門から気管分岐部相当部まで仮想的に延長し、最下端部を剛壁ではない壁とした[10]。声門から気管分岐部までの長さを調査した報告の平均値は 12.6cm であった[11]。モデル 1 の声門から気管分岐部までの長さは 11.9cm であり、過去の報告と近似していた。これらを参考に、モデル 2 は声門下部から気管分岐部相当部までを仮想的に円柱状に 12cm 延長し、解析を行った。音響媒質は体温を想定した 37°C とした。37°C の音速は 352.85m/s、密度は 1.1468kg/m³ である。音源は声帯相当部に点音源を設定した。観測点は実音声録音時のマイクロフォンの位置と同様に口唇より約 10cm の位置とした。

以上のパラメータ、条件設定のもと、音響解析を行い、観測点 \mathbf{r}_r における音圧 $p(\mathbf{r}_r)$ を求

め、音圧レベルでの周波数応答曲線を作成し、第1フォルマント (F1) と第2フォルマント (F2) を算出した。

3.3 精度検証

3.3.1 実音声との比較

防音室内で被験者の口唇から約 10cm 前方に設置したマイクロフォン SM58 (SHURE, Chicago, IL) を介し、PCM レコーダー PMD661 (Marantz, New York, NY) を用いて録音した。CT 撮影時と同様の仰臥位で行った。母音/a/を対象音とし、約 5 秒の持続音を記録した。PCM レコーダーのオーディオサンプルレートは 44.1kHz、オーディオサンプルサイズは 16 ビット、オーディオ形式は PCM とした。音声は、レコーダーに内蔵された SD メモリカード内に、WAVE audio 形式で保存した。音響分析ソフト マルチスピーチ 3700 (PENTAX Medical, Montvale, NJ, USA) を用いて、フォルマント解析を行った。オーディオサンプルレートを 44.1kHz から 11.025kHz へダウンサンプリングし、F1 と F2 を算出した。

3.3.2 実体模型を用いた検証

15 例の声道モデルのうち 4 例に対して実体模型を作製した。素材はナイロンと同じポリアミド系の PA12 (DMM, Tokyo, Japan) で、マルチジェットフュージョン方法で作製した。声道モデル作成時と同様に、前頭洞から気管分岐部までの実体模型を 1 体と、前頭洞から声門までのモデルに声門から気管分岐部相当部まで仮想的に延長した実体模型を 3 体作製した (図 4)。実体模型の気管分岐部側から音が漏れないように、気管分岐部側の開口部を超軟質の人肌[®]のゲルシート (EXSEAL, Gifu, Japan) で塞いだ。音源には笛式人工喉頭を使用した。笛式人工喉頭の使用方法はゴムチューブを模型の口腔内に挿入し、漏斗部に息を吹きかけて使用した (図 5)。この音を、前述と同じマイクロフォンとレコーダーを用いて記録し、同じマルチスピーチを用いてフォルマント解析を行った。

3.3.3 評価方法

過去に報告されている母音フォルマント周波数弁別閾をもとに妥当性を判断した。米国人は 3~5%[12]、日本人は 4.9~9.6%[13]という報告より、9%以内を良好と判断とした。シミュレーションのフォルマント周波数を F とし、実音声および実体模型それぞれのフォルマント周波数と F との差を弁別閾値 ΔF とした。 F で ΔF を割ることにより、相対弁別閾値(%)を求めた[13]。

4. 結果

同一被験者 (No. 15) から作成した 2 種類の声道モデル、すなわちモデル 1 とモデル 2 に対して上記音響解析を行い、周波数応答曲線を求め、フォルマント周波数を算出した結果、モデル 1 とモデル 2 から算出された $F1$ と $F2$ はそれぞれ、モデル 1 は 815Hz、1,641Hz、モデル 2 は 806Hz、1,629Hz であった (表 1)。 $F1$ と $F2$ はモデル 1 とモデル 2 で近似しており、日本語母音/a/の一般的に言われている $F1$ 、 $F2$ の周波数内[14]でもあった。

実音声との相対弁別閾値の $F1$ と $F2$ はそれぞれ、モデル 1 は 4.0%、2.7%、モデル 2 は 3.0%、2.0%であった。このように、シミュレーション結果は共に良好であり、解析結果に差はなかった。この結果より、CT 撮影範囲は被爆量低減目的にモデル 2 の声門下部までとし、気管分岐部相当部まで仮想的に 12cm 延長して解析を行った。

モデル 2 の声門下部から気管分岐部相当部までを仮想的に 12cm 延長した、合計 15 例の解析結果を表 2 に示す。すべての症例の $F1$ 、 $F2$ は日本語母音/a/の一般的に報告されている $F1$ 、 $F2$ の周波数内[14]であった。実音声との相対弁別閾値の平均値は $F1$ と $F2$ 、それぞれ 5.9%、4.6%であった (表 2)。15 例中 12 例が相対弁別閾値 9%以内であり、残りの 3 例は延長部の長さを 2~3cm 短くすることにより、相対弁別閾値 9%以内となった。

15 例にフォルマント周波数解析を行った結果、 $F1$ 、 $F2$ 以外に複数のピークが観察された。

全症例に一致していたのは、500Hz 付近でのピークである。このピークは、鼻腔を外したモデルでは観察されなかったため、500Hz 付近のピークは鼻咽腔結合による極零対であると考えられた（図 6, 7）。鼻咽腔結合による極零対の次のピークは、女性の多くに F1 と F2 の間に観察され（図 8, 9）、男性の多くに F2 より高周波数側に観察された（図 6, 7）。

実体模型を用いて解析を行った結果、F1 と F2 はシミュレーションした F1 と F2 と近似していた（表 3）。モデル 1 の実体模型との相対弁別閾値は F1 と F2 がそれぞれ、4.0%、5.2%で、モデル 2 の実体模型 3 体の平均は F1 と F2 がそれぞれ、4.1%、3.2%と相対弁別閾値 9%以内であった。

5. 考察

MRI のデータから作成した声道モデルを利用して、先行研究が多く行われてきた[5, 6, 15, 16]。CT と比較して MRI の最大の利点は被爆がないことである[17]。一方で、MRI にはいくつかの欠点がある。第一に、時間分解能が低いため撮像に比較的長時間を要する[6]。従って、持続発声時の画像作成は容易ではない。第二に、空間分解能が高くはないため、喉頭のような微細な構造を可視化することが容易ではないであろう[17]。第三に、MRI にて歯のアウトラインを描出することは容易ではない[18]。加えて、歯科用金属[19]や鼻咽腔閉鎖不全患者に適応される発音補正装置によるアーチファクトの問題を有する。このような点から、本研究では、持続発声時の声道モデルを高精度に作成することを優先し、被爆量を軽減するように CT データを得た。今後、上記の MRI の欠点が克服されれば、CT から MRI に変更されることとなるであろう。

今回の研究で作成した声道モデルは前頭洞、篩骨洞、蝶形骨洞、上顎洞、鼻腔、口腔、咽頭、気道から気管分岐部までの範囲を声道モデルとした。過去の報告では口腔、咽頭から声門までのモデル[6]や鼻腔、口腔、咽頭から声門までのモデル[5]であった。鼻腔の有無は後述する極零対の発生により、母音/a/に大きく影響を与えることが知られている[20]。

副鼻腔および気道まで、声道モデルの範囲を広げたのは下記の報告を参考とした。母音の発生は肺から送られた空気が声帯を振動させ、口唇や舌などの構音器官によって形作られる空洞を共鳴させて生成される[20]。声帯で生じた音声波は鼻咽腔から鼻腔に入り、外鼻孔から放射される経路には前頭洞、篩骨洞、蝶形骨洞、上顎洞が開口しており、その共鳴が周波数スペクトルに影響する[15]。これらより本研究では声道モデルの範囲を副鼻腔から気道まで広げることとした。声道モデルの範囲を広げるにあたって、副鼻腔や鼻腔、口腔は個人差があるのに対して、気道は比較的単純な円柱状である。CT撮影による被爆量低減目的に声門下部から気管分岐部相当部までを仮想的に延長したモデルと解析結果の比較することとした。延長するにあたって、声門から気管分岐部までの長さを調査した過去の報告を参考とした。最短で男性が 10.6cm、女性が 8.8cm で、平均は 12.6cm、最長で 16cm であった[11]。今回の症例 No. 15（モデル 1）の声門から気管分岐部までの長さも 11.9cm と報告の平均値と近似していた。よって声門から気管分岐部までの長さを 12cm 延長したモデルを作成して解析を行った。解析の結果、モデル 1 とモデル 2 の F1 と F2 は非常に近似した結果であった。よって CT 撮影範囲は前頭洞上端から声門下部までとし、声門下部から気管分岐部までは仮想的に延長し、解析を行うこととした。

解析時の各種設定パラメータについて考察する。声帯の振動を音源、声帯から口唇までを声道、音源が声道を通り、空気中に発せられた音を音声と呼ぶ[21]。このように音源と声道によって音声が生成されるという考え方を音源-フィルタ理論という[21]。母音生成の音源-フィルタ理論は線形理論に基づいている[21]。音源-フィルタ理論では母音は音源に関係なく、口腔、鼻腔、咽頭などの声道の空洞の共鳴にのみ特徴付けられ、声帯の特性には依存しないという仮定のもと、成り立っている[21]。しかしながら、実際の音声生成過程は非線形であり、主に声帯の振動の非線形効果が示唆される[22]。音源を点音源とすれば、非線形学の考えを考慮しても、音源に非線形性は生じない。よって、音源は、声帯相当部に点音源を設定した。過去の報告では声道モデルは鼻腔、口腔、声門までで、声門部は

開放しておらず、声道モデルの底面が音源となっている[4, 5, 23]。本研究では声道モデルの底面を声門から気道まで延長したため、声帯相当部に点音源を設定することとした。声道モデルの壁面の設定は剛壁とした。声道壁が剛壁か剛壁でない壁かが解析結果のフォルマント周波数に影響し[5, 6]、剛壁より剛壁ではない壁が実音声に近似した結果となるとされている[5]。しかし、壁面を剛壁より剛壁ではない壁とした場合、過去の報告では、帯域幅が増加し、1~2,000Hzの範囲では、極零対とF1との結合が認められるという報告がある[5]。また、本研究では精度検証のために、実体模型を作製した。実体模型を作製する際に、軟性材料では困難であるため、剛壁とした。開口している気管分岐部の設定を示す。気管分岐部は肺へとつながっており、モデルの底面を剛壁ではない壁とした。音響媒質は体温を想定し、37°Cの音速 352.85m/s、密度 1.1468kg/m³とした。

本研究で使用した境界要素法は波動音響解析の一つである。波動音響解析とは音を波として捉える解析法であり、他に有限差分法と有限要素法がある。有限差分法は空間領域を節点格子に置換して支配方程式を離散化し[16]、閉空間を対象とする[24]。有限要素法は空間領域を要素に分割して領域積分方程式を離散化し、有限差分法と同様に閉空間を対象とする[24]。有限要素法による声道の音響解析報告は散見されるが、メッシュ作成や解析に時間を要する問題があり、時間退縮を目的に声道形状の簡略化が考えられてきた[25]。しかし、本来存在すべき声道の音響特性が失われてはならない。今回使用した境界要素法は、空間領域の境界を要素に分割して境界積分方程式を離散化することから、開・閉空間を問わず対象となる。有限要素法と比較して、モデルの境界部分のみをモデル化する手法であるため、時間短縮が可能となった[26]。過去の報告では有限要素法と境界要素法それぞれで解析を行い、その結果には大きな差はないといわれている[4]。

シミュレーションの結果の妥当性を判断する根拠として、日本人における母音フォルマント周波数弁別閾は 4.9~9.6%[13]、米国人における母音フォルマント周波数弁別閾は 3~5%[12]という報告を用いた。日本語と英語の弁別閾値の差については、定かではないが、

日本語の母音数が英語に比べて少ないことが関与している可能性も挙げられている[13]。

実音声を用いた検証では、15例中12例は弁別閾値が9%以内となり、波動音響解析法を用いて、良好なシミュレーションが実現できた。弁別閾値が10%以上となった3例についても声門からの長さを9~13cmの範囲で調整を行うことで9%以内の良好な結果となった。長さを調整するにあたって、性別や身長などの基準は見いだせなかったため、今後も症例数を増やして研究を継続していきたい。

実体模型を用いた検証は、15例中4例で行った。実体模型の素材と声道モデルの壁面の設定がなるべく近似するようにと、硬質なナイロン樹脂とした。音源は実音声に近いと言われている[24]笛式人工喉頭を使用した。過去の報告で、音源に使用されているものは、人工喉頭以外にホーンドライブユニットである[6, 24]。しかし、ホーンドライブユニットを使用する場合は模型との接続が非常に複雑である[6, 24]。また、周波数解析の結果は実音声から解離しているため、本研究では人工喉頭を使用した。気管分岐部側の開口部は音が漏れないためと声道モデルでの軟壁設定を再現するために、シリコンゲルで塞いだ。その結果、4例中4例で弁別閾値は9%以内であった。よって、本研究は十分な有効性があることを示唆している。

境界要素法によるシミュレーションの結果、周波数応答曲線にフォルマント周波数以外のピークを複数認められた(図6~8)。このピークは極零対と呼ばれ、500Hz付近に生じる極零対は母音/a/のような広母音の場合、鼻腔結合により出現する[20]。周波数応答曲線において、極零対の後方に位置するF1は極零対の影響を受け、高値となることが知られている[20]。その他の極零対は、副鼻腔の共鳴により複数生じ、周波数応答曲線が複雑となることが報告されている[15]。鼻腔および副鼻腔は顔面頭蓋骨の形状を反映して個人差が大きいため、鼻音は高い個人性を有しており、鼻音により話者識別率が向上することが報告されている[27]。今回の研究では女性の多くにF1とF2の間に極零対が観察され、男性の多くにF2より高周波数側に極零対が観察された。性差の因子として、声道の長さや喉頭など

の発声器官の差や発声時の鼻腔へ抜ける空気の量などの報告がある[28]。また、過去の報告では少ない被験者ではあるが、極零対の発生には性差があり、男女間の母音の特徴の一つではないかと考えられており[28]、今回の研究結果もこの特徴を示唆していると考えられる。これを明らかにするためには、さらなるデータの蓄積が必要である。

6. 結論

今回の研究において、CT データから日本語母音/a/発声時の声道モデルを作成し、音源を与えて、境界要素法を用いた波動音響解析を行うことにより、実音声に近似したフォルマント周波数を算出することが可能となった。

7. 謝辞

本研究は科学研究費補助金（基盤研究(B)18H03001）の援助を受けた。サイバネットシステム株式会社の竹川真弘さま、秋山誠さまから助言を頂いたことに感謝する。

8. 参考文献

1. J. E. Bernthal, N. W. Bankson, *Articulation and Phonological Disorders*, fourth ed., Pearson, Boston, 1998.
2. K. Okazaki, M. Kato, T. Onizuka, Palate morphology in children with cleft palate with palatalized articulation, *Ann. Plast. Surg.* 26 (1991) 156-163.
3. C. W. Lawrence, B. J. Philips, A telefluoroscopic study of lingual contacts made by persons with palatal defects, *Cleft. Palate. J.* 12 (1975) 85-94.
4. Y. Kagawa, R. Shimoyama, T. Yamabuchi, T. Murai, K. Takarada, Boundary element models of the vocal tract and radiation field and their response characteristics, *J. Sound. Vibr.* 157 (1992) 385-403.
[https://doi.org/10.1016/0022-460X\(92\)90523-Z](https://doi.org/10.1016/0022-460X(92)90523-Z)
5. H. Matsuzaki, K. Motoki, Study of acoustic characteristics of vocal tract with nasal cavity during phonation of Japanese /a/, *Acoust. Sci. & Tech.* 28 (2007) 124-127. <https://doi.org/10.1250/ast.28.124>
6. T. Kitamura, H. Takemoto, S. Adachi, K. Honda, Transfer functions of solid vocal-tract models constructed from ATR MRI database of Japanese vowel production, *Acoust. Sci. & Tech.* 30 (2009) 288-296.
<https://doi.org/10.1250/ast.30.288>
7. S. Marbure, Six boundary elements per wavelength: Is that enough?, *J. Comput. Acoust.* 10 (2002) 25-51. <https://doi.org/10.1142/S0218396X02001401>
8. T. Sakuma, S. Sakamoto, T. Otsuru, *Computational Simulation in Architectural and Environmental Acoustics - Methods and Applications of Wave-Based Computation*, Springer, Japan, 2014.
9. B. H. Story, A-M. Laukkanen, I. R. Titze, Acoustic impedance of an artificially

- lengthened and constricted vocal tract, *J. Voice*. 14 (2000) 455–469.
[https://doi.org/10.1016/S0892-1997\(00\)80003-X](https://doi.org/10.1016/S0892-1997(00)80003-X)
10. C. Zwikker, C. W. Kosten, *Sound Absorbent Material*, Elsevier, New York, 1949.
 11. D. Y. C. Chong, K. B. Greenland, S. T. Tan, M. G. Irwin, C. T. Hung, The clinical implication of the vocal cords–carina distance in anaesthetized Chinese adults during orotracheal intubation. *Br. J. Anaesth.* 97 (2006) 489–495. <https://doi.org/10.1093/bja/ae1186>
 12. J. Lyzenga, J. W. Horst, Frequency discrimination of bandlimited harmonic complexes related to vowel formants, *J. Acoust. Soc. Am.* 98 (1995) 1943–1955.
<https://doi.org/10.1121/1.413314>
 13. S. Eguchi, Difference limens for the formant frequencies: Normal adult values and their development in children, *J. Am. Audiol. Soc.* 1 (1976) 145–149.
 14. T. Hirahara, R. Akahane–Yamada, Acoustic characteristics of Japanese vowels, *Proc INTERSPEECH2004–ICSLP*. (2004) 3287–3290.
 15. J. Dang, K. Honda, H. Suzuki, Morphological and acoustical analysis of the nasal and the paranasal cavities, *J. Acoust. Soc. Am.* 96 (1994) 2088–2100.
<https://doi.org/10.1121/1.410150>
 16. M. J. Crocker, *Handbook of acoustics*, John Wiley & Sons, New York, 1997.
 17. B. H. Story, I. R. Titze, E. A. Hoffman, Vocal tract area functions for an adult female speaker based on volumetric imaging, *J. Acoust. Soc. Am.* 104 (1998) 471–487. <https://doi.org/10.1121/1.423298>
 18. H. Takemoto, T. Kitamura, H. Nishimoto, K. Honda, A method of tooth superimposition on MRI data for accurate measurement of vocal tract shape and dimensions, *Acoust. Sci. & Tech.* 25 (2004) 468–474.

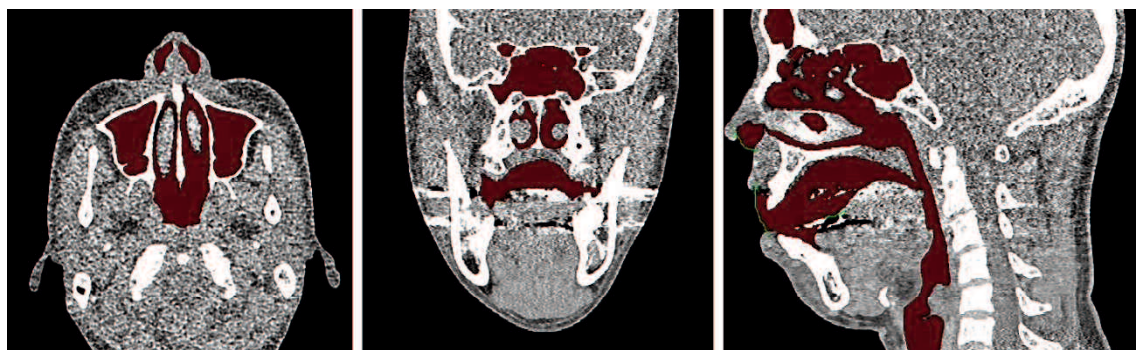
<https://doi.org/10.1250/ast.25.468>

19. D. B. Hinshaw, B. A. Holshouser, H. I. M. Engstrom, A. H. L. Tjan, E. L. Christiansen, W. F. Catelli, Dental material artifacts on MR imaging. *Radiology*. 166 (1988) 777-779. <https://doi.org/10.1148/radiology.166.3.3340777>
20. K. N. Stevens, *Acoustic Phonetics*. MIT Press, Cambridge, 2000.
21. G. Fant, *Acoustic Theory of Speech Production with Calculations based on X-Ray Studies of Russian Articulations*. Mouton, The Hague, 1970.
22. M. Rothenberg, S. Zahorian, Nonlinear inverse filtering technique for estimating the glottal-area waveform, *J. Acoust. Soc. Am.* 61 (1977) 1063-1071. <https://doi.org/10.1121/1.381392>
23. K. Motoki, H. Matsuzaki, Computation of the acoustic characteristics of vocal-tract models with geometrical perturbation, *Proc INTERSPEECH2004-ICSLP*. (2004) 521-524.
24. T. Arai, Visualizing vowel-production mechanism using simple educational tools, *J. Acoust. Soc. Am.* 118 (2005) 1862-1862. <https://doi.org/10.1121/1.4778955>
25. S. Sakamoto, H. Tachibana, Calculation of impulse responses in 3-D sound field with absorptive boundary by the finite difference method, *Proc Inter-Noise*. 97 (1997) 1597-1600.
26. C. A. Brebbia, J. C. F. Telles, L. C. Wrobel, *Boundary element techniques: theory and applications in engineering*. Springer, Berlin, 1984.
27. K. Amino, H. Makinae, T. Kitamura, Nasality in speech and its contribution to speaker individuality, *Proc Interspeech2014*. (2014) 1688-1692.
28. D. H. Klatt, L. C. Klatt, Analysis, synthesis, and perception of voice quality variations among female and male talkers, *J. Acoust. Soc. Am.* 87 (1990) 820-

857. <https://doi.org/10.1121/1.398894>

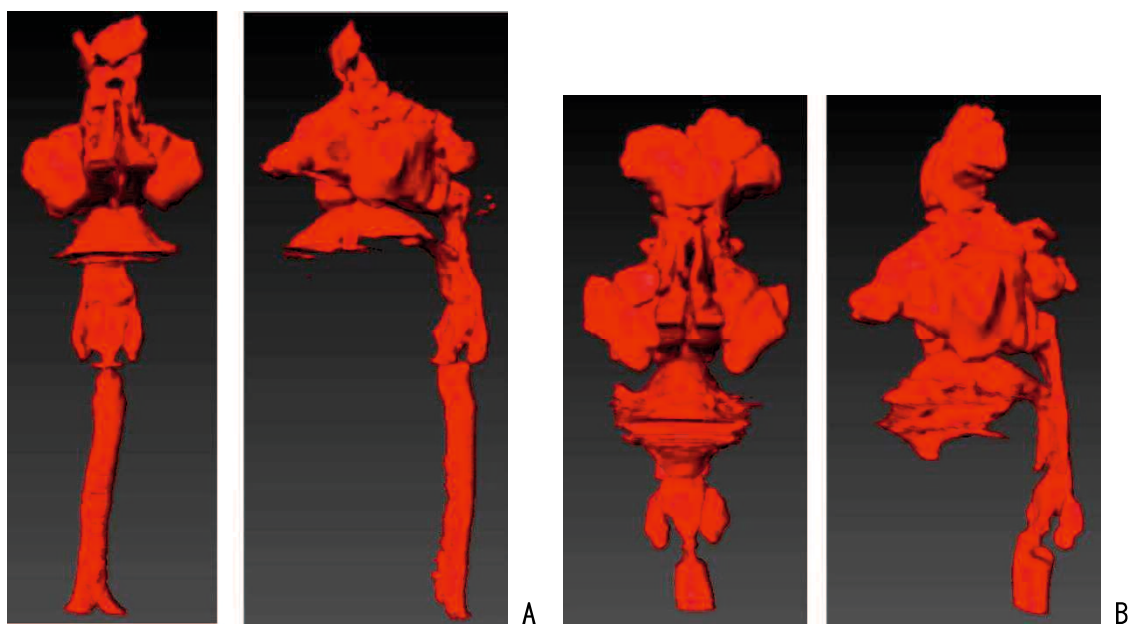
29. M. Shiraishi, K. Mishima, H. Umeda, Development of an Acoustic Simulation Method during Phonation of the Japanese Vowel /a/ by the Boundary Element Method, *J. Voice*. in press. <https://doi.org/10.1016/j.jvoice.2019.11.022>

図 1. CT データからの気道抽出



前頭断、水平断、矢状断の CT スライスデータよりモデル作成に必要な気道を抽出する。

図2 声道モデル

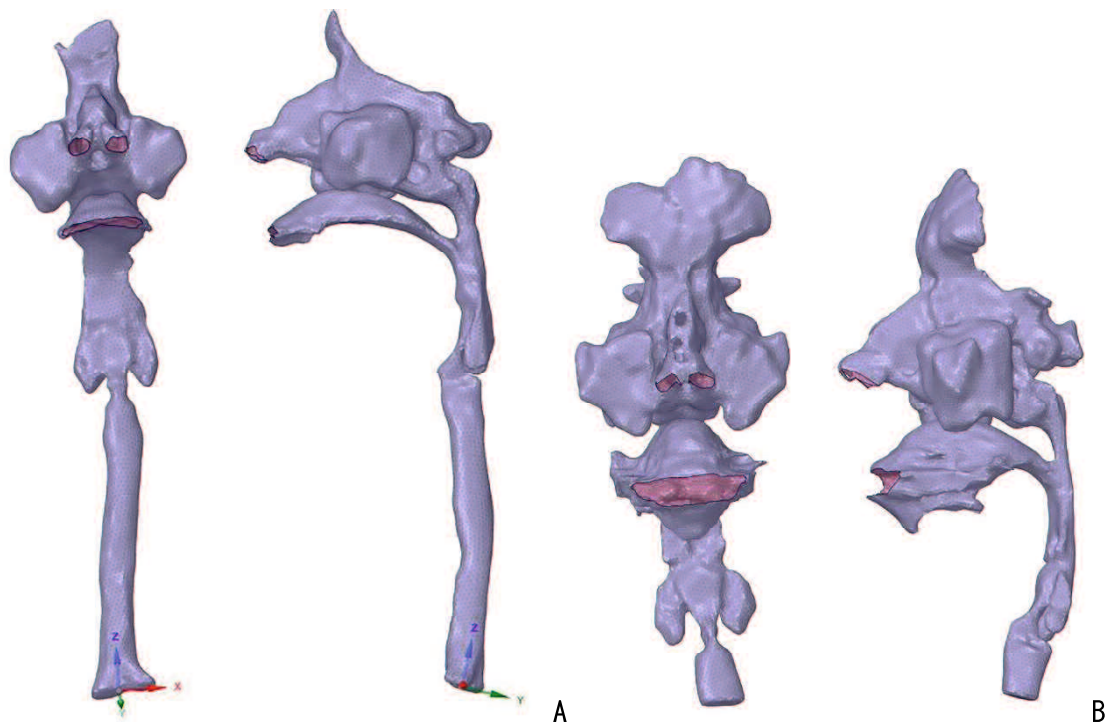


二種類の声道モデルを示す。

A: モデル1 前頭洞・篩骨洞・蝶形骨洞・上顎洞・鼻腔・口腔・咽頭・喉頭・声門・気管・気管分岐部を含む。

B: モデル2 前頭洞・篩骨洞・蝶形骨洞・上顎洞・鼻腔・口腔・咽頭・喉頭・声門下部を含む。

図3 解析メッシュモデル

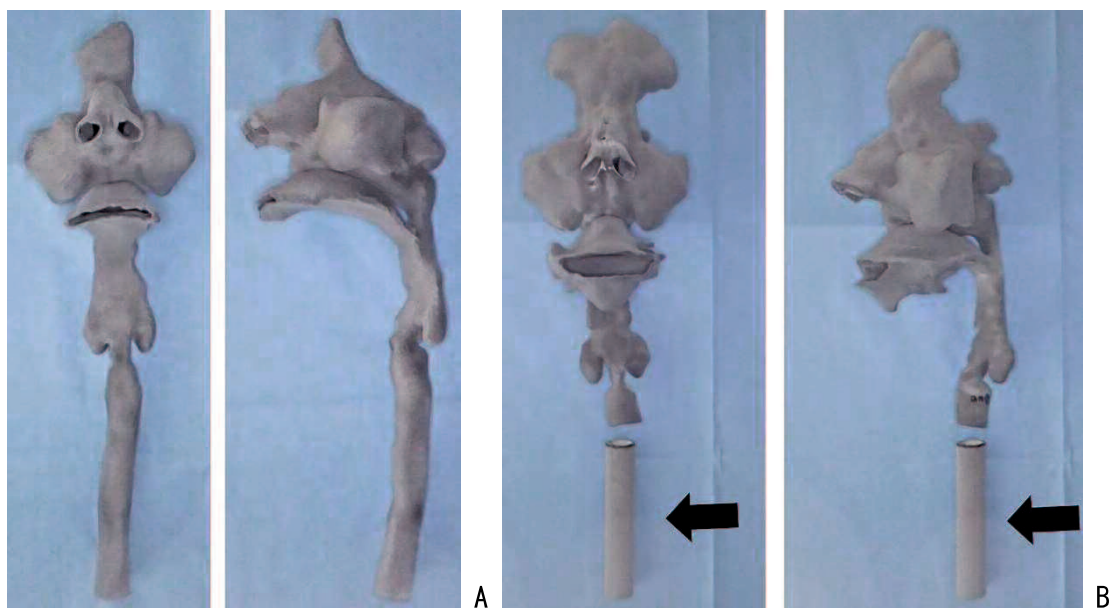


三次元声道モデルを三角形要素の解析メッシュモデルへ変換する。メッシュサイズを均一化し、鼻孔と口には開口部を設ける。

A : モデル 1

B : モデル 2

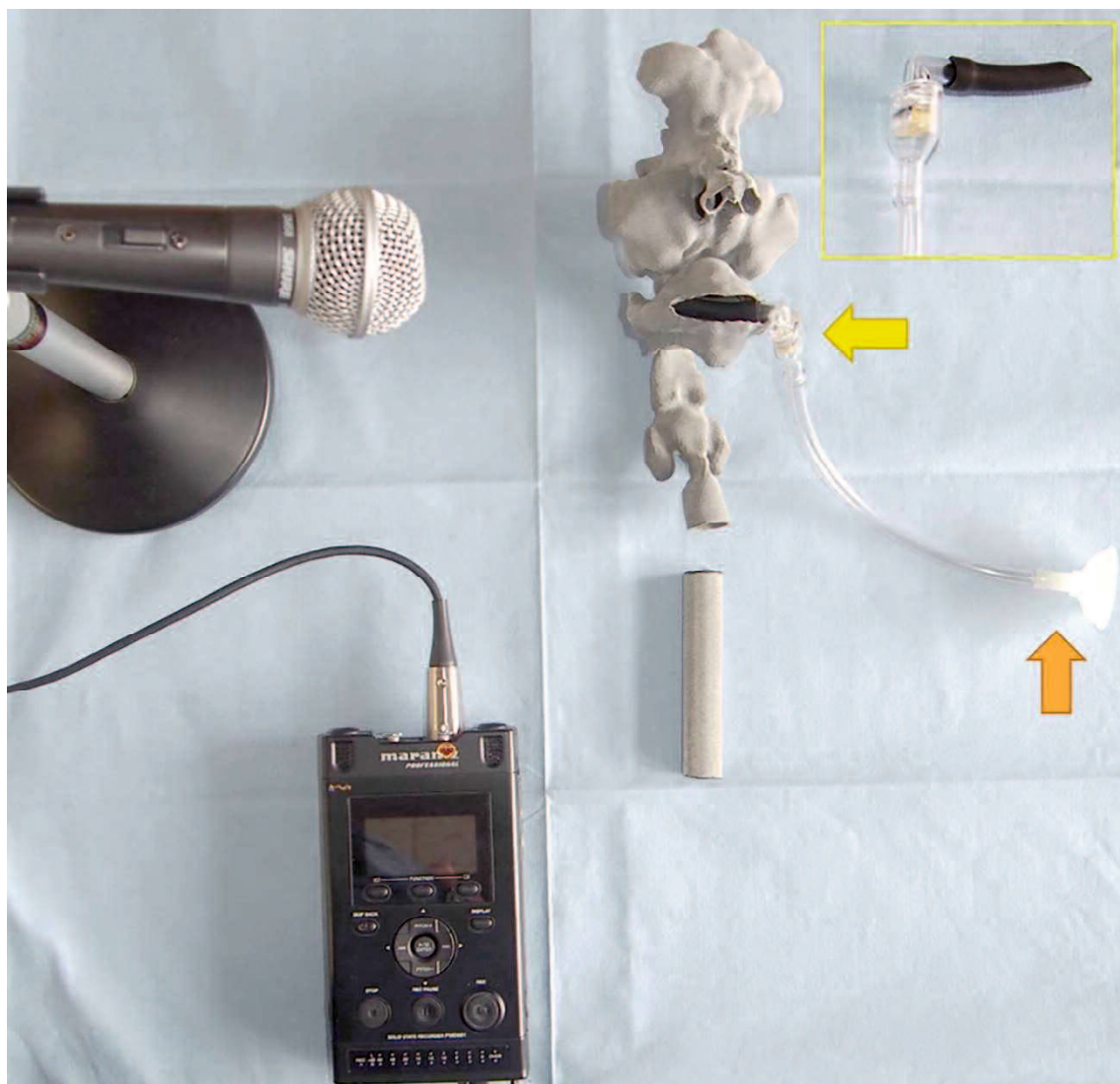
図4 実体模型



A : モデル1の実体模型

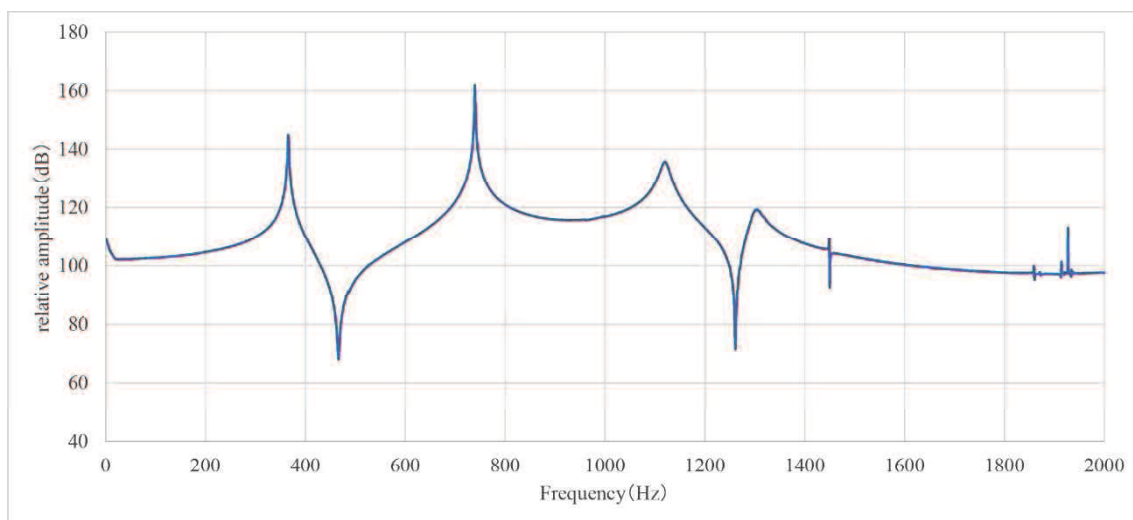
B : モデル2の実体模型 黒い矢印部分は、筒状の円柱で仮想的に延長した部分である。

図5 実体模型と笛式人工喉頭を使用した検証実験



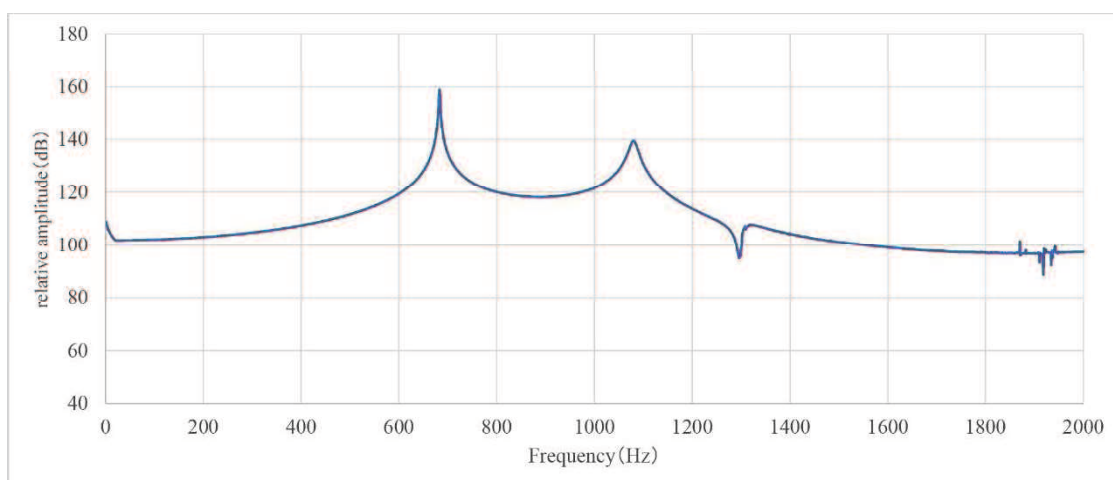
実体模型を用いた検証方法を示す。笛式人工喉頭の白色部が気管口にあてるあてざら（橙色矢印）で、ゴム製のチューブで口腔内へ挿入して使用する。黄色矢印部が音源となり、ゴム膜の振動を引き起こす。今回の研究では、実体模型の口腔内に笛式人工喉頭のゴム製のチューブを挿入し、あてざらより息を吹き込むと音が発生する。音声の録音にはマイクロフォンを使用する。

図 6 周波数応答曲線 モデル 2 (症例 No. 1)



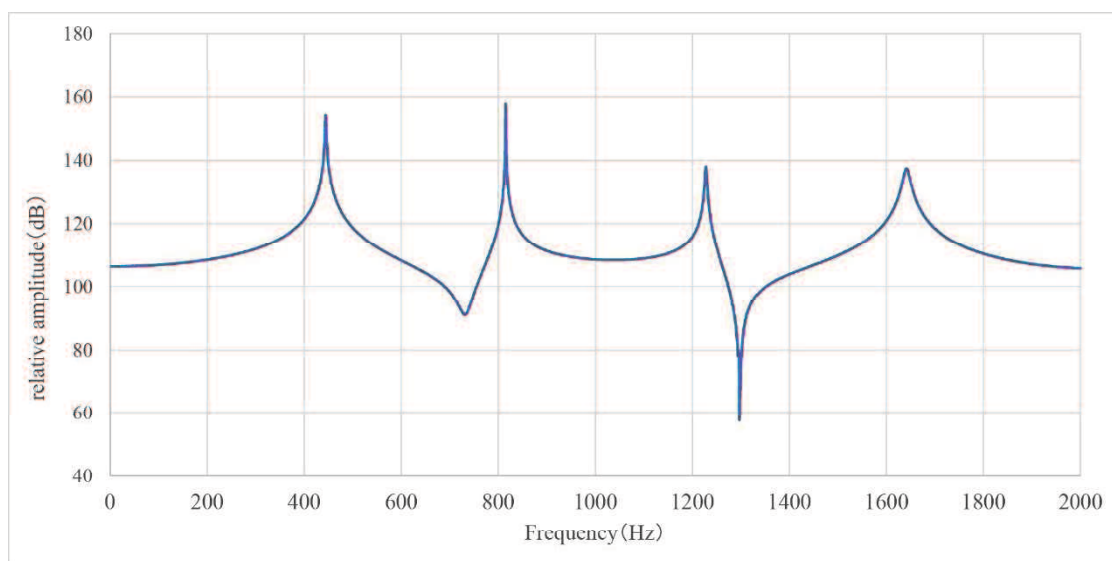
最初のピークは鼻咽腔結合によるもの、2 番目のピークは F1、3 番目のピークは F2、4 番目のピークは副鼻腔に結合することにより生じると考えられる極零対を示す。

図7 周波数応答曲線 モデル2 鼻腔を含めないモデル (症例 No. 1)



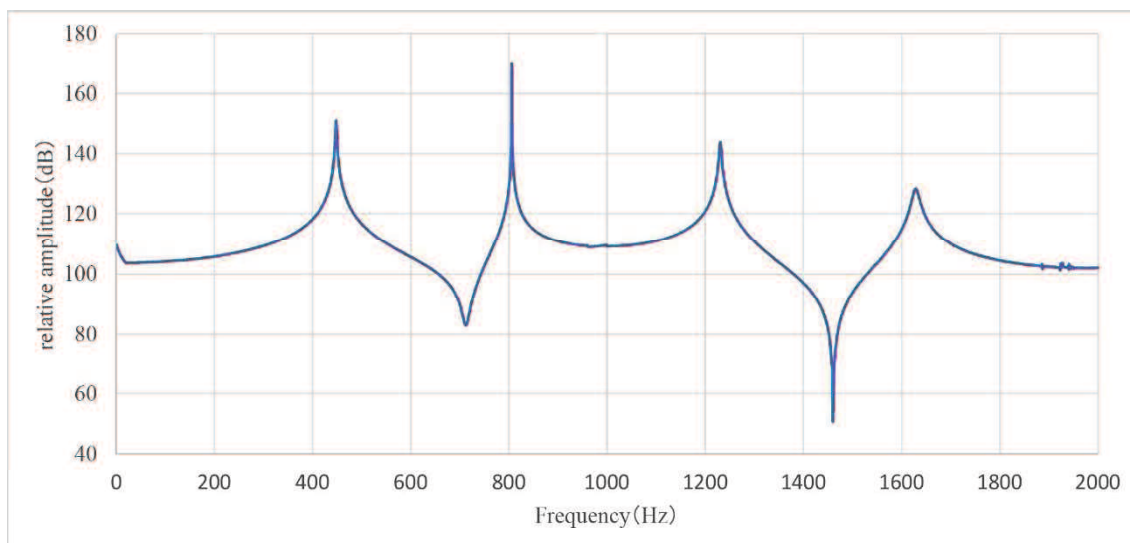
鼻腔を外して、解析した結果を示す。鼻腔を外すことにより 500Hz 付近の鼻咽腔結合による極零対や副鼻腔に結合することにより生じると考えられる極零対も発生しない。

図8 周波数応答曲線 モデル1 (症例 No. 15)



モデル1の解析結果を示す。最初のピークは鼻咽腔結合によるもの、2番目のピークはF1、3番目のピークは副鼻腔に結合することにより生じると考えられる極零対、4番目のピークはF2である。

図9 周波数応答曲線 モデル2 (症例 No. 15)



モデル2の解析結果を示す。最初のピークは鼻咽腔結合によるもの、2番目のピークはF1、3番目のピークは副鼻腔に結合することにより生じると考えられる極零対、4番目のピークはF2である。

表 1

同一被験者によるシミュレーション値と実音声におけるフォルマント周波数の比較（症例 No. 15）（Hz）

	モデル 1	モデル 2	実音声
F1	815	806	782
F2	1641	1629	1597

表 2

声道モデル 2 を使用したシミュレーション値と実音声のフォルマント周波数および弁別閾値

性別	症例 No.	F1			F2		
		シミュレーション値 (Hz)	実音声 (Hz)	相対弁別閾値 (%)	シミュレーション値 (Hz)	実音声 (Hz)	相対弁別閾値 (%)
男性	1	739	784	6.1	1120	1215	8.5
	2	635	615	3.1	1115	1123	0.7
	3	712	743	4.4	1110	1206	8.6
	4	712	740	3.9	1097	1131	3.1
	5	616	709	15.1	1100	1167	6.1
	6	723	745	3.0	1172	1122	4.3
	7	711	795	11.8	1094	1271	16.2
	8	696	687	1.3	1137	1157	1.8
女性	9	794	732	7.8	1559	1590	2.0
	10	880	933	6.0	1761	1717	2.5
	11	795	834	4.9	1740	1712	1.6
	12	770	898	16.6	1561	1437	7.9
	13	732	738	0.8	1111	1115	0.4
	14	758	760	0.3	1557	1598	2.6
	15	806	782	3.0	1629	1597	2.0

表 3

声道モデルのシミュレーション値と人工喉頭を使用した実体モデルのフォルマント周波数および弁別閾値

	性別	症例 No.	F1			F2		
			シミュレーション 値 (Hz)	実体模型 (Hz)	相対弁別閾値 (%)	シミュレーション 値 (Hz)	実体模型 (Hz)	相対弁別閾値 (%)
モデル1	女性	15	815	782	4.0	1641	1555	5.2
	男性	1	739	790	6.9	1120	1221	9.0
モデル2	女性	9	794	789	0.6	1559	1568	0.6
		10	880	923	4.9	1761	1763	0.1