

音声生成における音源-フィルタ相互作用の影響に関する研究

上江洲, 安史

<https://doi.org/10.15017/1807042>

出版情報：九州大学, 2016, 博士（芸術工学）, 課程博士
バージョン：
権利関係：全文ファイル公表済

音声生成における音源-フィルタ相互作用の
影響に関する研究

A study on the effect of source-filter interaction
on the production of speech

上江洲 安史

UEZU Yasufumi

2017年3月

目次

第 1 章 序論	1
1.1 本論文の背景と目的	1
1.2 本論文の構成	5
第 2 章 音声生成の基礎知識	7
2.1 音声生成の基礎	7
2.1.1 呼吸器官	7
2.1.2 発声器官	9
2.1.3 調音器官	11
2.1.4 母音の生成過程	13
2.2 音声生成の基礎理論	14
2.2.1 音源-フィルタ理論	14
2.2.2 音源-フィルタ相互作用	16
第 3 章 音声生成の測定・分析手法	20
3.1 音源部の測定と分析	20
3.1.1 電気喉頭計	20
3.1.2 DECOM 法	23
3.2 声道部の測定と分析	24
3.2.1 外部音響励振法	24
3.2.2 ケプストラム分析	27
第 4 章 声区転換における音源-フィルタ相互作用の影響の検討	28
4.1 背景と目的	28
4.2 実験	31
4.2.1 実験参加者と実験内容	31
4.2.2 測定手法	32

4.2.3	基本周波数の分析	35
4.2.4	声道音響特性の分析	35
4.3	結果	38
4.3.1	母音/a/の分析結果	38
4.3.2	母音/i/の分析結果	42
4.3.3	全試行の分析結果	46
4.4	考察	54
4.5	まとめ	56
第 5 章	発声条件がフォルマントに与える影響の検討	57
5.1	背景と目的	57
5.2	実験	57
5.3	結果と考察	59
5.4	まとめ	63
第 6 章	声門の境界条件がフォルマントに与える影響の検討	64
6.1	背景と目的	64
6.2	音声生成の物理モデル	65
6.2.1	声門面積の多項式モデル	65
6.2.2	声道・声門・声門下部の連結音響管モデル	67
6.2.3	音声合成シミュレーション	70
6.3	合成音声の作成と分析	72
6.4	結果と考察	78
6.4.1	第 1 フォルマント周波数	78
6.4.2	第 2 フォルマント周波数の結果	84
6.4.3	声門下部インピーダンスと声道伝達関数	90
6.5	まとめ	97
第 7 章	総括	98
7.1	まとめ	98
7.1.1	声道が音源に対して影響を及ぼす影響について	99
7.1.2	音源が声道に対して影響を及ぼす影響について	99
7.2	今後の展望と課題	100

7.2.1	歌唱音声の音声生成メカニズムの解明と応用	100
7.2.2	非線形現象としての音源-フィルタ相互作用	101
	謝辞	102
	参考文献	104

第1章 序論

1.1 本論文の背景と目的

ヒトは、他者とのコミュニケーションを図る手段として、音声を発する。また、ヒトは歌唱を行う際にも、音声を発する。一般的に、歌唱における音声のことを歌唱音声と呼ぶ。これに対し、会話における音声のことを会話音声と呼ぶ。いずれの場合も、肺から送られた呼吸のエネルギーによって声帯が振動することで音源波が生成され、声帯から口唇までの管状の空間である声道を音源波が伝播することで声道で共鳴が生じ、最終的に口唇から放射されることによって、音声が生産される。歌唱音声は、会話音声と比較して、用いる声の高さや声の大きさの帯域が非常に広いことが特徴である。また歌唱音声は、会話音声では意図的に発しないような特徴的な音色を伴う場合が多々存在することも特徴である。さらに、一概に歌唱音声といえども、世界中には歌唱音声のための特殊な発声方法が多数存在するのも、会話音声とは大きく異なる点である。例えば、西洋オペラの理想的な歌唱法といわれるベルカント唱法や、ロシア連邦のトゥバ共和国に伝わる倍音唱法であるホーメイなどの伝統的な発声法から、現代の商業音楽において頻繁に用いられるスクリーム唱法やミックスボイスに至るまで、歌唱における発声法の細分化は今もなお進んでいる。一般に、会話音声の発声は自然に獲得されていくものであるが、歌唱音声はその発声の特殊性と多様性のために、専門的なトレーニングを必要とする場合が多い。このような点も、会話音声と歌唱音声の大きな違いであるといえる。

音声に関する研究は、会話音声を主な対象として、音声を生成するための器官がどのような構造になっているのか、またそのような器官がどのようなはたらきで音声を生成するのか、さらに生成された音声がどのような音響的特徴を持つのか、という点について古くから盛んに行われてきた。特に、音声のもととなる音源と、音声の音韻的な特徴を与えるフィルタが独立であるという仮定のもとで、音源がフィルタを通過することで音声が生産されると考える「音源-フィルタ理論」は、1970年代に Fant が体系化して以降、音声生成の線形モデルとして長

きにわたり音声生成の基礎理論となっている [1]。この音源-フィルタ理論を背景に、今日までに得られた音声生成に関する多大な知見と、近年における計算機の処理能力の飛躍的な発達に伴い、最近では片手に収まるほどの小さな携帯端末デバイスにおいて、音声合成や音声認識などの音声情報処理技術が実用化されるまでに至っている。その一方で、歌唱音声に関する研究は、その特殊性と多様性による複雑さのためか、近年ようやく行われ始めたのが現状である。そのため、多種多様な歌唱音声の生成メカニズム、歌唱音声の音響的特徴については、まだ十分に明らかにされていない点や、議論の余地が多い。またここ数年、ボーカロイドをはじめとする歌唱音声合成ソフトの開発や、機械学習による歌唱音声合成の研究も行われているが、歌唱音声における特殊性と多様性を網羅しているとは言い難い。このように、歌唱音声の生成メカニズムや音響的特徴を明らかにすることは、歌唱音声を理解し、技術化するにあたって重要な課題となっている。

歌唱音声を生成する上で非常に密接な関係にあるのが、声区 (register) である。声区とは、すべての音が同様の音声生成メカニズムによって生成され、同様の声質を持つような周波数の領域のことである [2]。声区の違いは声質の違いにあらわれやすいため、特に訓練をせずとも聴感上において声区の違いを聞き取ることができる。代表的な声区の種類として、地声声区 (modal voice register) と裏声声区 (falsetto voice register) の 2 声区が挙げられる。一般的に、地声声区は基本周波数が比較的低い声を発声する際の声区、裏声声区は基本周波数が比較的高い声を発声する際の声区と認識されている。会話音声は地声声区のみで発声されることが多いため、会話音声で声区を意識することはあまりない。声区を特に意識するのは歌唱を行う際である。これは、歌唱音声は音楽のメロディを担うことから、会話音声と比べて発声する音高の周波数帯域が非常に広く、地声声区だけでは到底発声できないような音高も発声する必要があるためである。

歌唱と声区にまつわる発声現象の例として、声区転換が挙げられる。これは地声声区から裏声声区もしくはその反対方向へ、声の高さを徐々に変化させながら発声すると、ある音高に達した際に声区が急に切り替わる現象のことである。一般には、声がひっくり返る、声が裏返る、換声点、などとも言われる。この声区転換の際には、声の高さである基本周波数が急に跳躍するように変化し、これを周波数ジャンプという。声区転換を自在に制御できるようになることは、古くから今日に至るまで、歌唱技術を高める上で非常に重要な課題となっている。上

述のとおり、歌唱においては地声声区と裏声声区の両方を用いることが一般的であるが、できるだけ同一の声質に聴こえるように歌唱することが、歌唱の上手さの指標の一つとされる。しかし、地声声区と裏声声区では声質が明らかに異なる上に、地声声区と裏声声区との境界付近の音高では発声が不安定になり、周波数ジャンプを伴う声区転換を生じてしまう。そのため、プロ歌手の多くは、声区転換を知覚されないように声区を滑らかに繋いで発声したり、できるだけ同一の声質に聴こえるように発声を調節したりすることで、これらの問題を回避して歌唱を行う。一方で伝統的な歌唱法に着目すると、スイス・アルプス地方のヨーデルや、日本の長唄におけるアタリなどでは、声区転換を意図的かつ高速に行うことで音楽的表現を豊かにしている歌唱法も存在する。

歌唱において重要な発声現象である声区転換がどのようにして生じるのか、その要因について様々な検討がこれまでになされてきた。音源とフィルタが相互に依存せず独立だと仮定する音源-フィルタ理論に基づいてまず検討されたのが、音源の生成機構である声帯の生理的なメカニズムである。すなわち、声帯の張力を次第に上昇させていくと、声区転換を生じるというものである。これについては、摘出喉頭を用いた声帯張力変化の実験や、発声の実測実験などが行われており、声帯の張力変化が、声区転換を引き起こす要因であることが確認されている。

一方、実際の音声の生成メカニズムにおいては、音源とフィルタは必ずしも独立ではなく、互いに影響を及ぼし合っていると考えられる。これを「音源-フィルタ相互作用」という [3]。音源-フィルタ理論が線形モデルであるのに対して、音源-フィルタ相互作用は非線形モデルであるため、線形モデルでは現れないような不安定性を生じることがある。近年、音源-フィルタ相互作用が音声の生成過程に及ぼす影響について、主として音声生成の物理モデルによるシミュレーション実験が行われており、声区転換もまた、音源-フィルタ相互作用の影響によって引き起こされることが示唆されている。特に、声道の共鳴周波数であるフォルマント周波数のうち、最も低次の第1フォルマント周波数が低い母音では、音声の基本周波数が第1フォルマント周波数に近接したときに、声道からの音響的なフィードバックが音源の生成機構に生じることで声帯振動や音源波が不安定となり、結果的に周波数ジャンプ・周波数の分岐・カオス・サブハーモニクスを伴うような非線形性の強い声区転換を生じることが、シミュレーション実験により明らかになった。このようなシミュレーション実験に対し、発声測定実験による

音源-フィルタ相互作用の影響についての検証がほとんどなされていない。従って、実際のヒトの発声について、音源-フィルタ相互作用の影響を検証する必要がある。

上述したような音源-フィルタ相互作用は、声道が音源に対して影響を及ぼしたと見ることができる。一方で、音源が声道に対して影響を与えるような音源-フィルタ相互作用も存在すると考えられる。たとえば、音源-フィルタ理論においては、声門側における声道の境界条件は完全閉鎖を仮定している。しかし、実際の発声中には声帯は準周期的に振動しているため、声帯の隙間である声門は声帯振動に伴う開放と閉鎖を交互に生じる。すなわち、音源-フィルタ理論における完全閉鎖という声門境界条件の仮定は成り立たず、境界条件が時間とともに変化することにより、声道の共鳴特性ないしフォルマントもそれによる影響を受けると考えられる。さらに、声区の違いは音声生成メカニズムの違いであることから、声区の違いは声帯振動および声道の境界条件の違いにも表れるはずである。つまり声区の違いは、声道のフォルマントにも影響を与えると考えられる。このような、音源が声道に対して影響を及ぼすような音源-フィルタ相互作用が考えられるため、このことについても検証する必要がある。

本研究では、音声生成における音源-フィルタ相互作用の影響を明らかにすることが目的である。すなわち、従来の音声生成モデルである音源-フィルタ理論において、互いに独立であり影響を及ぼし合わないとは仮定された音源と声道が、実際の音声生成メカニズムのように独立ではない場合において、互いにどのように影響を及ぼし合うのかについて検証を行う。

声道が音源に対して影響を及ぼすような音源-フィルタ相互作用の影響については、声区転換の発声測定実験による検証を行う。本研究では、声区転換時における第1フォルマント周波数と基本周波数の関係性を調べることで、声道が音源にもたらす音源-フィルタ相互作用の影響を検討する。一方で、音源が声道に対して影響を及ぼすような音源-フィルタ相互作用の影響については、発声測定実験とシミュレーション実験による検証を行う。本研究では、発声測定実験を通して、発声条件の違いが声道のフォルマントに与える影響について検討する。また、音源-フィルタ相互作用の影響を考慮した音声生成シミュレーションを行い、声門の境界条件が声道のフォルマントに与える影響について検討する。これらの実験を通して、音源-フィルタ相互作用が音声の生成のメカニズムと音声の音響的特徴に対して、どのような影響を及ぼすのかを明らかにする。

1.2 本論文の構成

以下に、本論文の構成を示す。第2章では、音声の生成機構や生成メカニズム、音声の生成理論とモデル、音声の音響的特徴について述べる。第3章では、本論文中で行われた実測実験において用いられた音声の測定・観測装置や、音声の分析手法について述べる。第4章では、音源-フィルタ相互作用が声区転換に及ぼす影響について、音響的な発声実測実験を通して検証する。声区転換を生じた際における基本周波数と第1フォルマントを同時かつ精確に計測し、それらの周波数の近接関係を調べることで、声区転換を引き起こす要因として音源-フィルタ相互作用の影響があるかどうかを確かめる。第5章では、発声条件の違いがフォルマントに与える影響について、音響的な発声実測実験を通して検証する。声門の状態が異なると考えられる複数の発声条件のもとで、フォルマント周波数を精確に計測し、発声条件とフォルマント周波数の関係性を調べる。第6章では、声門の境界条件がフォルマントに与える影響について、音声生成モデルによるシミュレーション実験を通して検証する。声門の開きの大きさや時間の長さの条件を様々に変化させて与えたときに、合成した音声のフォルマントがどのように変化するかを調べる。第7章では、本論文の総括を行う。

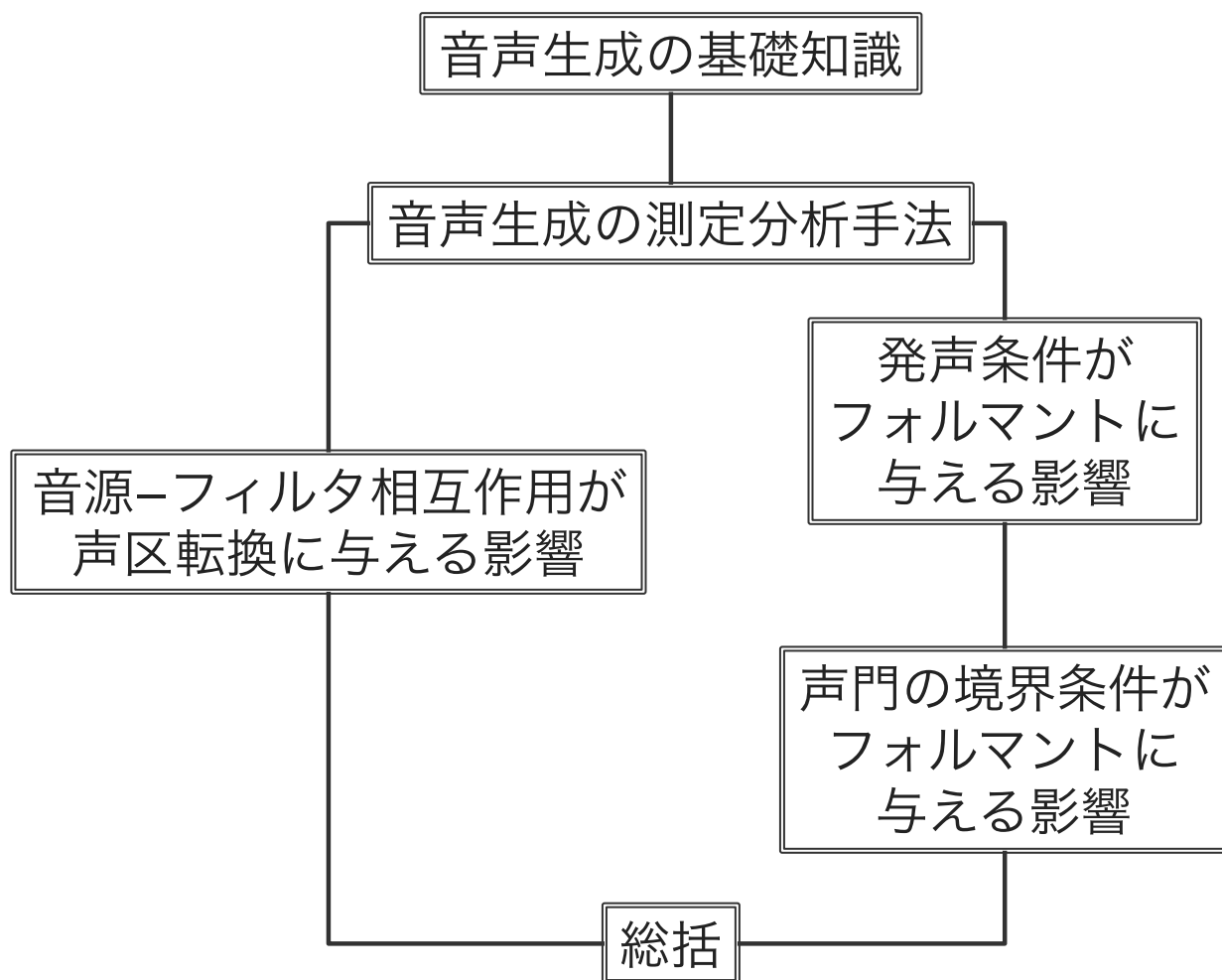


図 1.1: 本論文の構成。

第2章 音声生成の基礎知識

2.1 音声生成の基礎

音声の生成の役割を担う器官を音声器官 (speech organs) と呼ぶ。図 2.1 に、各音声器官の概略図を示す。音声器官は大きく分けて、呼吸器官 (respiratory organs)・発声器官 (phonatory organs)・調音器官 (articulatory organs) の3つの器官から構成される。音声の生成時における各器官のはたらきの概略は、次の通りである。呼吸器官は、呼気流を送り出すことで、音声を作り出すために必要なエネルギーを供給する。発声器官は、呼吸器官から送られてきた流体エネルギーを、音エネルギーに変換し、音源波を生成する。調音器官は、発声器官で生成された音源波に対して、管の共鳴による周波数の強弱を与えることで、音源波を音声に変換する。このように音声は、人体の様々な器官が有機的にはたらくことによって生成される。

2.1.1 呼吸器官

呼吸器官は、肺 (lung)・気管 (trachea)、気管支 (bronchus) から構成される。肺は胸部の左右に位置する一対の大きな“袋”であり、胸部の大部分を占める。肺は生命の維持に必要な呼吸活動、すなわち酸素の供給と二酸化炭素の排出が主な役割であるが、その際に生じる呼気の流れである呼気流は、音声の生成のためのエネルギーとして用いられる。安静時における呼気量は成人で約 0.5 l、肺圧は約 5 mmH₂O であり、発声時における肺圧は約 10 mmH₂O である [6]。歌唱時においては、肺圧はこれより更に高くなる。

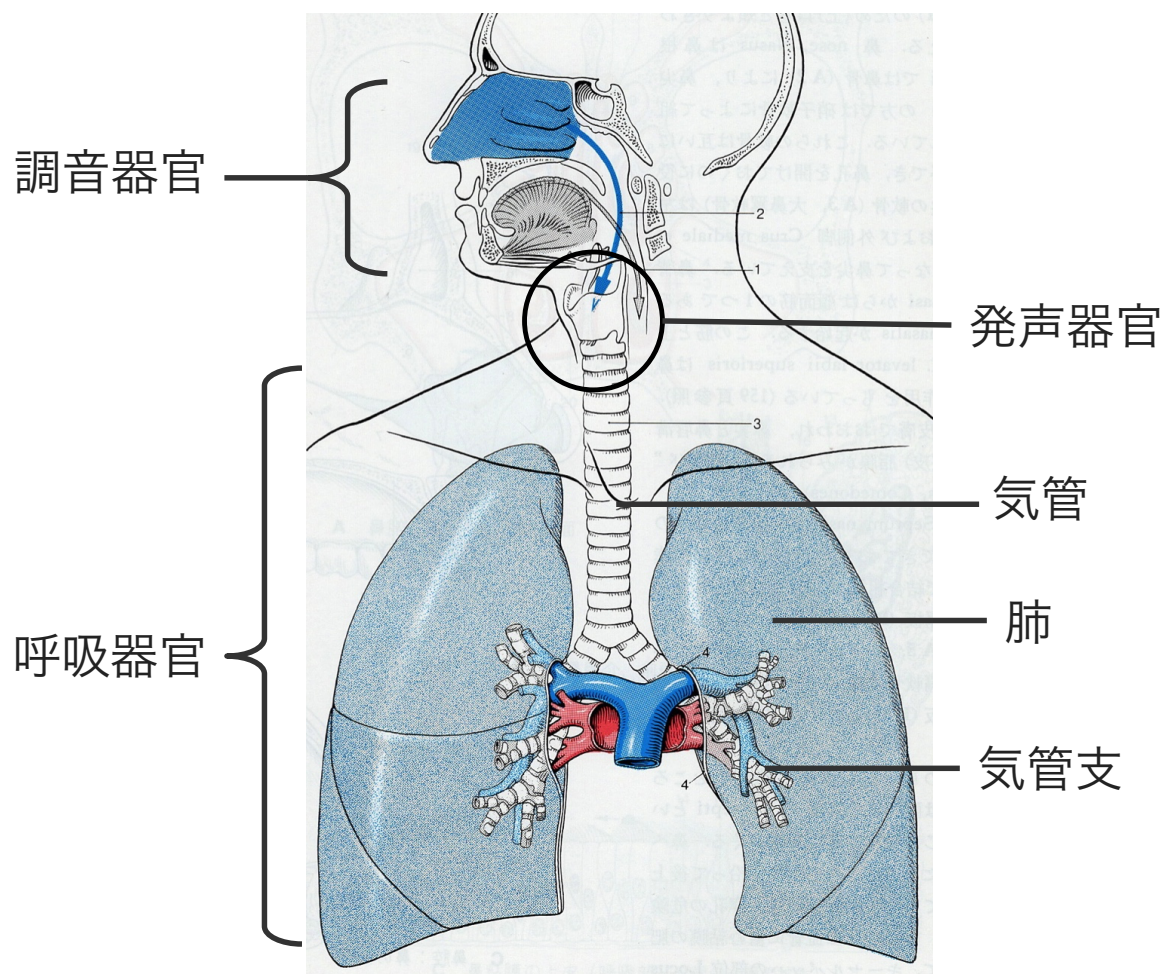


図 2.1: 音声器官の概略図 [37]。音声器官は呼吸器官・発声器官・調音器官の3つから構成される。呼吸器官は、肺と気管、気管支から構成される。肺から生じる呼気流は、音声の生成のためのエネルギーとして用いられる。

2.1.2 発声器官

発声器官は、気管の上端に位置する喉頭 (larynx) を中心に構成される。喉頭は、主に4つの軟骨組織から構成される。喉頭の外形は、成人男性の喉仏として知られる隆起が存在する甲状軟骨を中心に、下側の輪状軟骨、上側の喉頭蓋軟骨から構成されている。また甲状軟骨の内部には、一对の披裂軟骨が位置している。これらの軟骨組織は、様々な筋組織が接続することで支えられ、喉頭を形成している。特に、甲状軟骨と輪状軟骨をつなぐ輪状甲状筋は、声の高さを制御する上で重要な役割を果たしている。ヒトは歌唱を行う際に複数の声区を用いることで、会話音声よりもはるかに広い声の高さや声質で歌唱音声を発声するが、これを可能にしているのは喉頭周辺の筋肉のはたらきによるものである。

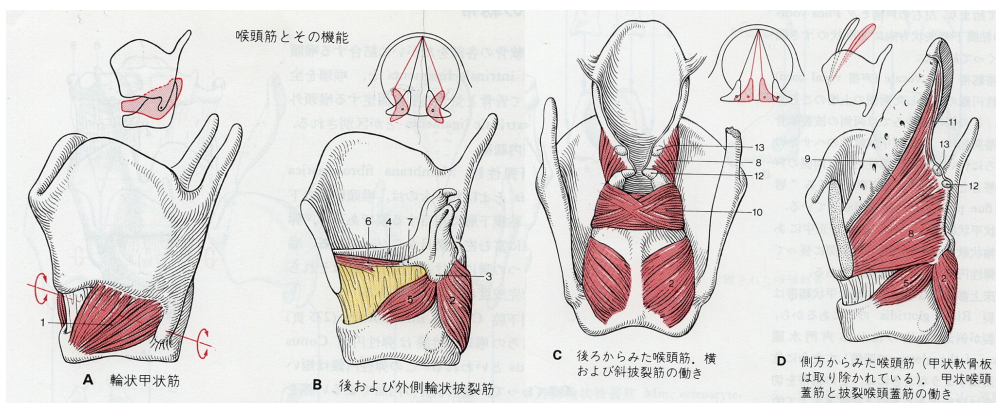
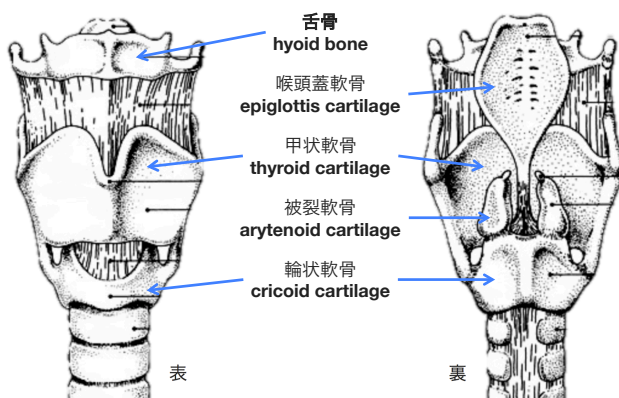


図 2.2: 喉頭の図 [37]。4つの軟骨から構成されており、それぞれが筋肉で繋がっている。これらの筋肉のはたらきにより、様々な音高や声質での発声が可能となっている。

甲状軟骨の内部には、披裂軟骨から甲状軟骨にかけて、一对の薄い膜組織である声帯 (vocal fold) が存在する (図 2.2)。甲状披裂筋は声帯筋とも呼ばれ、声帯は外側から粘膜・靭帯・声帯筋の3層構造で構成されている (図 2.3 (左))。成人男性における声帯の長さは約 15 mm、幅は約 3 mm である。成人女性の声帯はこれより一回り短くなっており、これが男女における声の高さの違いに寄与している。一对の声帯に挟まれた間隙を声門 (glottis) と呼ぶ。声門は呼気の通り道であり、安静時には開放したままであるが、発声時や咳き込む際などには、披裂軟骨の内転が生じることで声門が閉鎖する。発声の際、声門は肺から送られた呼気流によって生じる声帯の振動に伴って開閉を生じる。この声門の開閉パターンに伴って、声門を通過する呼気流は振幅変調を生じ、音源波となる。発声時において、声門が閉じている期間を閉鎖期 (closed phase)、声門が開いている期間を開放期 (open phase) と呼ぶ (図 2.3 (右))。

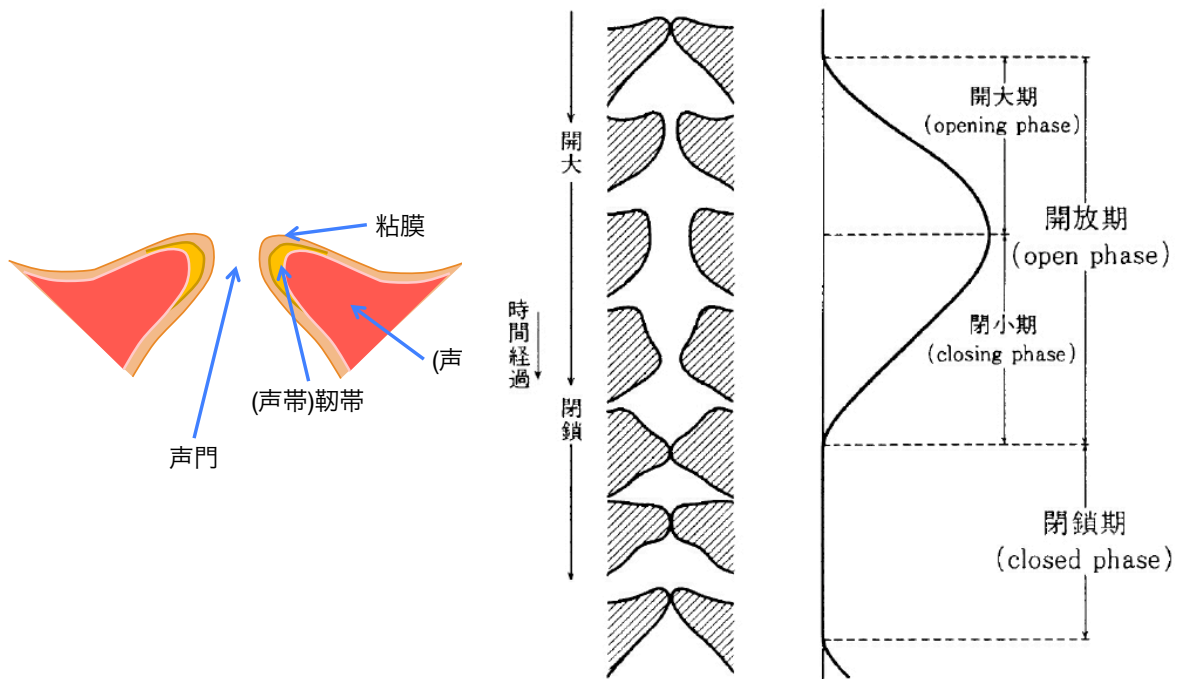


図 2.3: (左) 声帯と声門の断面図の概略図。声帯は粘膜、靭帯、筋組織の3層構造となっている。(右) 発声時における声帯振動の1周期と、それに伴う声門の開閉パタンの模式図。

2.1.3 調音器官

調音器官の概略を図 2.4 に示す。調音器官は、舌 (tongue)・下顎 (lower jaw)・口唇 (lip)・硬口蓋 (hard palate)・軟口蓋 (soft palate) などの組織から構成される。また、調音器官を構成する組織で閉ざされた空間である、咽頭腔 (pharyngeal cavity)・口腔 (oral cavity)・鼻腔 (nasal cavity) から構成される気道を、声道 (vocal tract) と呼ぶ。

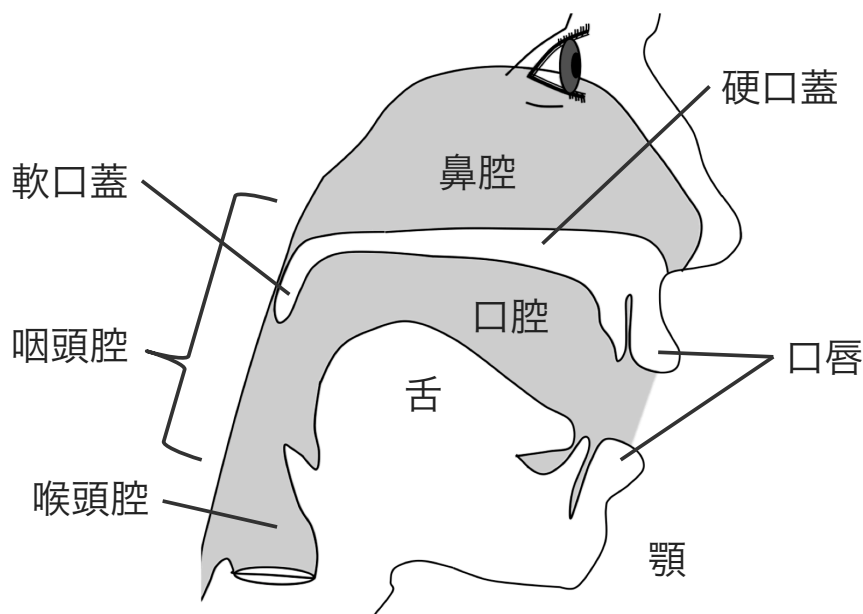


図 2.4: 調音器官の構造と声道の概略図。調音器官は舌・上顎・下顎・口唇・硬口蓋・軟口蓋などの組織からなる。調音器官を構成する組織で閉ざされた空間である、咽頭腔・口腔・鼻腔から構成される気道を、声道 (vocal tract) と呼ぶ。

音声生成における調音器官の役割は、音声の音韻性を決定することと、無声音の音源である乱流雑音の制御を行うことにある。前者に関しては、調音器官のはたらきによって口腔や咽頭腔の形状が大きく変化することにより、声道の形状が決定されることで、音声の音韻性が決定される。特に舌は、咽頭腔から口腔にかけての範囲にわたって柔軟に変形する組織であることから、調音器官の中でも特に重要な役割を担っているといえる。例えば、母音は/i/や/e/などの前舌母音、/a/や/o/などの後舌母音のように、舌の前後の位置の違いで分類されるほど、舌の位置は調音に大きく影響する (図 2.5)。

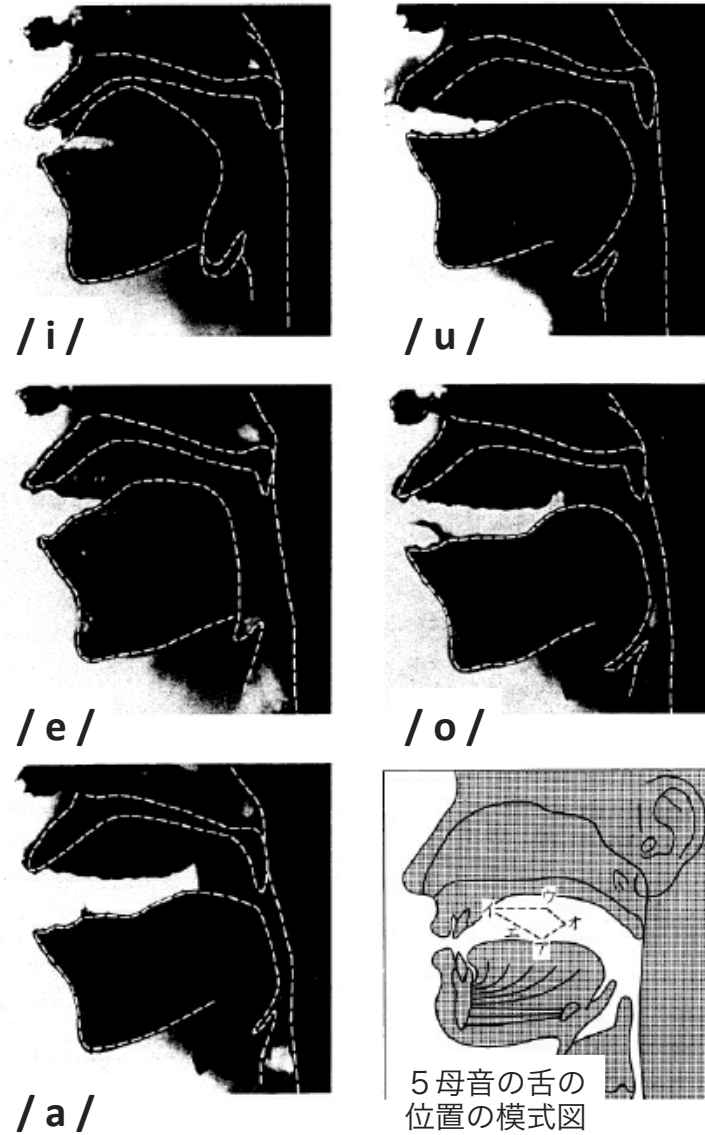


図 2.5: 日本語 5 母音をそれぞれ発声した際の声道形状の様子。母音 /i/ や /e/ などの前舌母音では、舌が比較的前方に寄っているため、口腔が狭くなり、咽頭腔が広がっていることがわかる。これに対し、母音 /a/ や /o/ などの後舌母音では、舌が比較的后方に寄っているため、口腔が広くなり、咽頭腔が狭くなっていることがわかる。

声道は成人男性で全長約 17 cm、断面積が 0.5 ~ 1.5 cm² 程度の管状の構造であるため、声道内では、音は共鳴現象を生じる。調音器官のはたらきにより、声道形状を変化させることは、共鳴管の構造を変化させることに相当する。管の構造が変化すると、共鳴のモードも変化するため、共鳴周波数と共鳴の程度も変化する。このようにして、調音器官は声道の共鳴のモードを連続的に変化させることで、音声の音韻性を変化させている。声道の共鳴の周波数特性のことを、声道共鳴特性 (vocal tract characteristics) と呼ぶ。また、声道共鳴特性における共鳴のピークをフォルマント (formant) とよび、その周波数をフォルマント周波数 (formant frequency) と呼ぶ。

2.1.4 母音の生成過程

母音の生成過程の概略図を、図 2.6 に示す。母音の生成過程は、以下に示すとおりである。呼吸器官は、音声のもととなる呼気流を発声器官へと送りこむ。呼吸器官から発声器官へと送られた呼気流のエネルギーは、喉頭の内側に位置する声帯を自励振動させる。この声帯振動に伴い、声帯のすき間である声門が周期的に開閉する。呼気流は、周期的に開閉する声門を通過することで振幅変調され、周期的な圧力変動を伴う音源波となる。音源波は声帯側から口唇側へと声道内を伝搬し、その際に声道の形状によって決定された周波数特性が付与される。最終的に、口唇から外部へと放射されることで、母音音声となる。

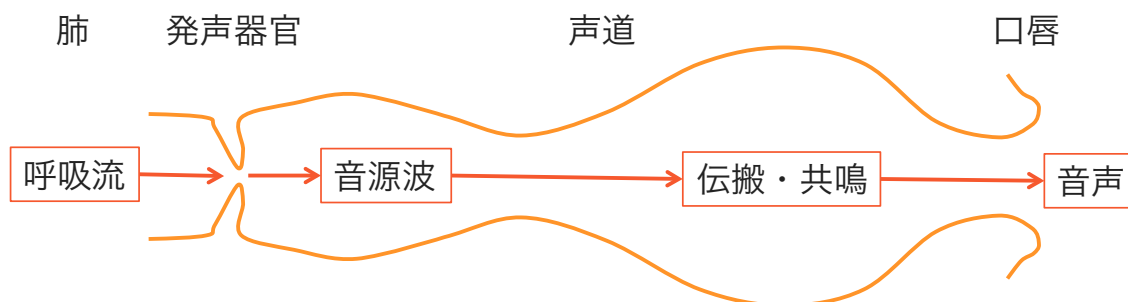


図 2.6: 母音の生成過程の概略図。呼吸器官は、呼気流を発声器官へと送りこむ。呼気流は声帯を自励振動させ、これに伴って声門が周期的に開閉する。周期的に開閉する声門を呼気流が通過することで、周期的な圧力変動を伴う音源波となる。音源波は声帯側から口唇側へと声道内を伝播・共鳴し、最終的に母音として口唇から放射される。

2.2 音声生成の基礎理論

2.2.1 音源-フィルタ理論

2.1.4項で述べた母音の生成過程を再考すると、声帯で生成された音源波が、声道の形状によって決まったある周波数特性を持つ音響フィルタを通過し、口唇から放射されて音声となる、という形で説明することができる。Fant [1]は、この音声生成過程を線形モデルとして表現する音源-フィルタ理論 (source-filter theory) を体系化した (図 2.7)。これは、発声器官で生成される音源波 (voice source) と、声道の音響フィルタ (vocal-tract filter) が独立であるという仮定のもとにおいて、音源波が声道音響フィルタを通過し、口唇の放射特性を付与されることで、音声が生産されるというものである。

音源-フィルタ理論による音声生成の線形モデルを周波数領域の形で表現すると、音源波 $G(\omega)$ 、声道音響フィルタ $V(\omega)$ 、口唇放射特性 $R(\omega)$ となり、音源波を入力として最終的に得られる出力すなわち音声 $S(\omega)$ は、 $S(\omega) = G(\omega)V(\omega)R(\omega)$ のように表すことができる。なお、それぞれの周波数領域における特徴として、音源波 $G(\omega)$ は、声の高さを決定する基本周波数 (fundamental frequency)、声の音色に寄与する倍音成分 (harmonic components)、-12dB/oct の傾斜の周波数特性を持つ。また、声道音響フィルタ $V(\omega)$ は声道共鳴特性 (vocal-tract resonance characteristics) とフォルマントを持ち、口唇放射特性 $R(\omega)$ は+6dB/oct. のハイパスフィルタの特性を持つ。

この音源-フィルタ理論の大きな特徴は、音源部とフィルタが音響的にも生理的にも独立であるという仮定である。すなわち、音源波の音響的特徴や発声器官のふるまいが、声道の音響フィルタや調音器官のふるまいに影響を及ぼすことはなく、その逆もまた同様であるという仮定である。このような仮定が非常に有用なために、Fant が音源-フィルタ理論を体系化して以来、今日に至るまで音声情報処理の分野において幅広く受け入れられており、音声生成に関する数多くの研究における理論的基盤にもなっている。

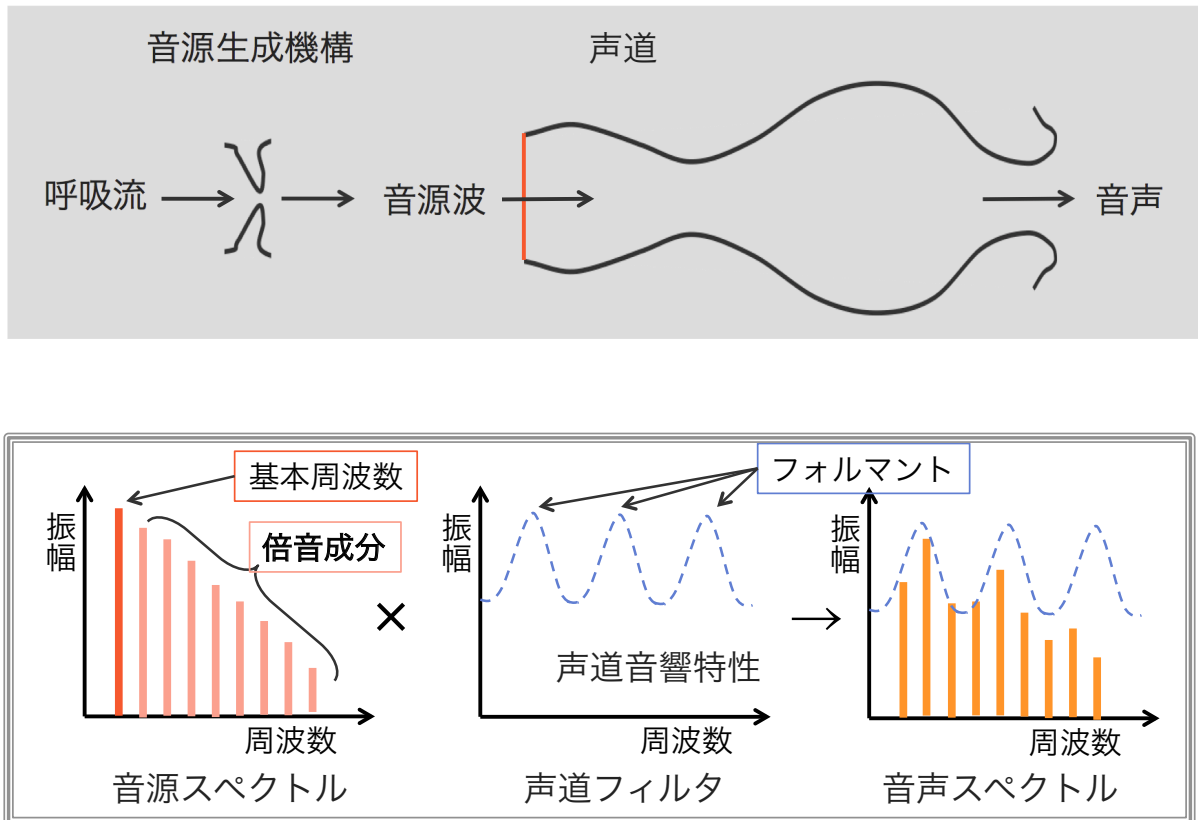


図 2.7: 音源-フィルタ理論の概略図。発声器官で生成される音源波と、声道の音響フィルタが独立であるという仮定のもとで、音源波 $G(\omega)$ が声道音響フィルタ $V(\omega)$ を通過し、口唇の放射特性 $R(\omega)$ を付与されることで、音声 $S(\omega)$ が生成されるという音声生成の線形モデル。周波数領域の形で表すと、音源波を入力として最終的に得られる出力すなわち音声 $S(\omega)$ は、 $S(\omega) = G(\omega)V(\omega)R(\omega)$ のように表される。周波数特性として、音源波は声の高さを決定する基本周波数、声の音色に寄与する倍音成分、 -12dB/oct の傾斜の周波数特性を持つ。声道フィルタは声道共鳴特性とフォルマントを持つ。口唇放射特性は $+6\text{dB/oct}$ のハイパスフィルタの特性を持つ。

2.2.2 音源-フィルタ相互作用

前述の音源-フィルタ理論では、音源と声道フィルタが音響的に独立であるという仮定が存在したが、実際にはこれらは独立ではなく、音響的に互いに影響を及ぼし合っている。これを音源-フィルタ相互作用 (source-filter interaction) と呼ぶ (図 2.8)。発声機構と声道が独立ではない場合、声道の音響負荷や、声道を伝播した音源波のフィードバックの影響を発声機構が強く受ける、すなわち音源-フィルタ相互作用の影響が強まることで、結果的に発声が不安定になったり、サブハーモニクスや分岐現象などを伴う非線形性の強い発声を生じたりする可能性があることが知られている。このように音源-フィルタ相互作用は、音源-フィルタ理論を発展させたものとして位置づけられる。

音源-フィルタ相互作用の影響として、声道の音響負荷がもたらす喉頭周辺の音場の影響が、生成される声門体積流に対して、波形が非対称となるような傾斜 (skew) を与えることが、従来より知られている [4,5]。近年、この声道から音源に対して影響を及ぼすような音源-フィルタ相互作用が、発声の不安定さや非線形性をもたらす要因として、主に音声生成の物理モデルによる音声合成シミュレーションによって検討されている。Titze [3] は音声生成シミュレーション実験によって、発声のピッチが第1フォルマント周波数に近接したとき、音源-フィルタ相互作用の影響が強まることを示した。この音源-フィルタ相互作用の影響は、epilarynx tube の断面積が狭まるとより強くなることがわかった。Tokuda [10,11] らは、4質量モデルなどの声帯物理モデルと音源-フィルタ相互作用を組み合わせた音声生成モデルによる声区転換のシミュレーション実験を行った。Kaburagi [12] は、声門の境界層解析と音源-フィルタ相互作用を組み合わせた音声生成モデルによる声区転換のシミュレーション実験を行った。いずれのシミュレーション実験においても、ピッチが第1フォルマントに近接した際に声区転換を生じることを示した。また、音源-フィルタ相互作用の影響について、測定実験もいくつか行われている。Titze ら [15] は複数のピッチパターンと母音の種類組み合わせによる声区転換の発声測定実験を行った。また、Zañartu ら [16] は実験参加者1名による上昇系列と下降系列の2種類のグリッサンド発声測定実験を行い、発声の基本周波数、声道と声門下部のフォルマント、フローマスクによる流量、声門の開閉による声門断面積の時間変化を撮像したキモグラフなどを計測した。これらの測定実験による研究は、先行していたシミュレーション実験の結果と同様に、第1フォルマント周波数の低い母音で、音響由来のメカニズムによる声区転換を生じる傾

向があることがわかった。

上述のような音源-フィルタ相互作用の影響は、声道が音源部に対して影響を及ぼす場合として考えることができる。その一方で、音源部が声道に対して影響を及ぼすような音源-フィルタ相互作用の影響も存在すると考えられる。実際の発声においては、声帯の準周期的な振動に伴って声門が時間とともに開閉することから、声道の声門境界条件も時々刻々と変化することで、フォルマントが声門や声門下部組織の影響を受けると考えられる。Barneyら [30] は、周期的に開閉する声門模型とアクリル製の一樣矩形管を組み合わせた模型実験を行った。声門の開口面積を固定した場合には、声門の開口面積が大きくなるに従って、第1フォルマントと第2フォルマントが上昇することがわかった。また、最大声門面積と声門開放率が大きくなると、第1フォルマントが上昇することがわかった。竹本ら [28] は、声道音響管モデルを用いて、声門の開放が声道の伝達関数に及ぼす影響を解析した。声門が開放すると声道が開管共鳴に近づくことから、声道の第1フォルマントが上昇することがわかった。

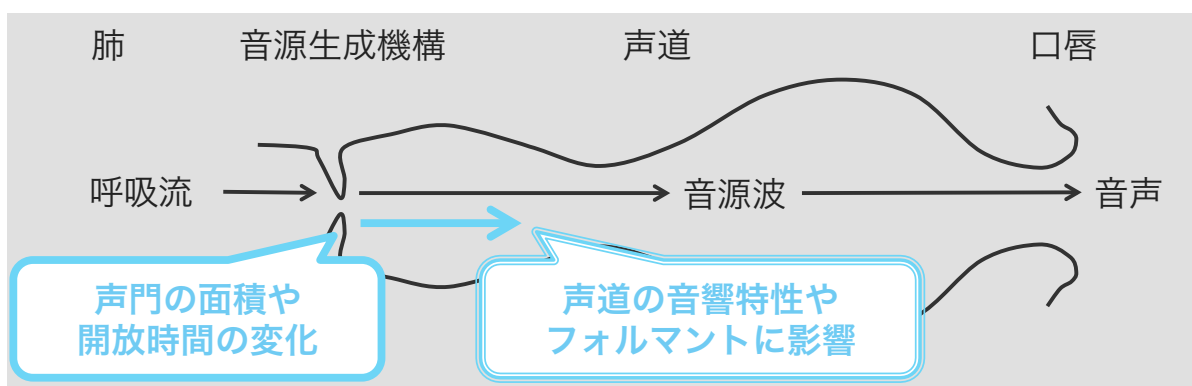
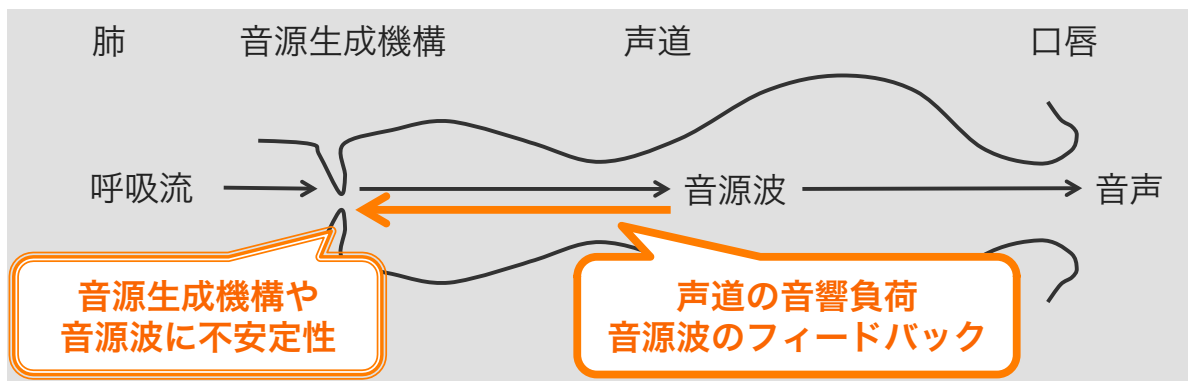


図 2.8: 音源-フィルタ相互作用の概略図。発声器官および生成される音源波と、声道の音響フィルタは、実際の発声においては独立ではない。すなわち、音源波の音響的特徴や発声器官のふるまいが、声道の音響フィルタや調音器官のふるまいに影響を及ぼし、その逆もまた同様である。

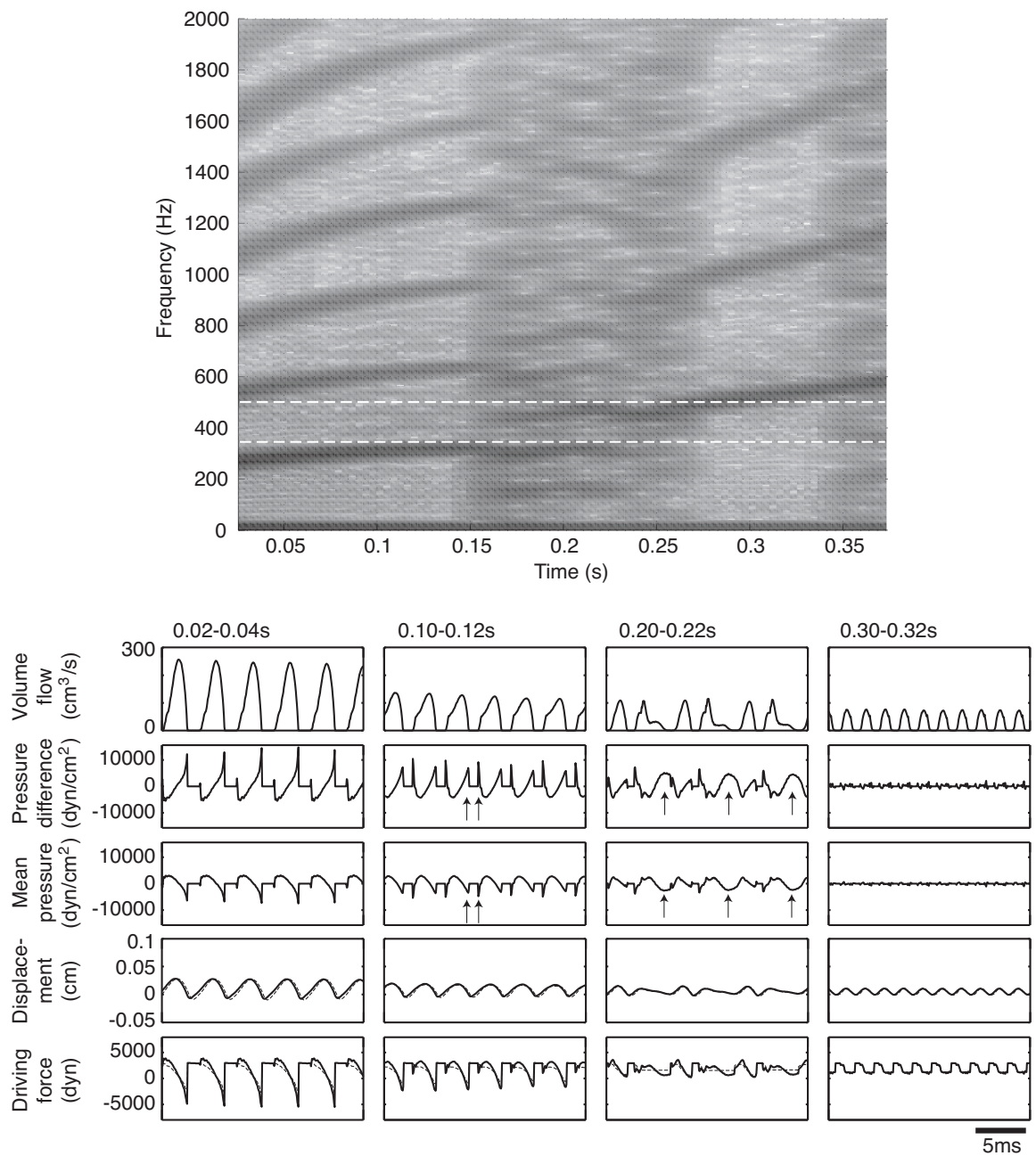


図 2.9: 音源-フィルタ相互作用の影響を考慮した音声生成モデルを用いた、声区転換のシミュレーション実験の結果の例 [12]。合成音声のスペクトログラムにおいて、0.15 sec で基本周波数が第 1 フォルマント (白点線) に近接すると、サブハーモニクスを伴うような声区転換を生じていることがわかる。

第3章 音声生成の測定・分析手法

音声生成において特に重要な役割を果たす発声器官と調音器官が、発声時においてどのような振る舞いを行うのかを観測することは、音声の生成メカニズムを解明する上で重要である。また、そのような振る舞いによって決定される音源や声道フィルタの音響的特徴を分析することは、生成される音声の音響的な特徴を解明する上で重要である。これらの測定分析を組み合わせることで、研究対象である音声がどのように生成され、どのような音響的特徴を持つのかを関連付けて検討することが可能となる。

本章では、本研究で実際に用いた音声生成に関する測定手法と分析手法について説明する。ここでは、音源部（発声器官）とフィルタ部（調音器官）に大別して、それぞれの測定/分析手法を述べる。

3.1 音源部の測定と分析

基本周波数や声区を頻繁に変化させる歌唱音声において、発声時の声帯振動の様子を観測することは、歌唱音声の発声メカニズムを解明する上で重要である。声帯振動を直接的に観測する方法として、ファイバースコープやストロボコピー、ハイスピードカメラを用いた撮像法が挙げられる。しかし、これらの装置は医療機関などの限られた場所でしか使用することができず、また、被験者の口腔ないし鼻腔に測定機器を挿入するため、侵襲的な側面も持つ。そこで本研究では、声帯振動の様子を間接的ではあるが非侵襲的に測定できる装置として、電気喉頭計を用いる。

3.1.1 電気喉頭計

電気喉頭計 (electroglottography; EGG) は、声帯振動によって生じる声門の開閉パターンを電氣的アドミタンスの形で測定する装置である。電気喉頭計には一対の電極が付属しており、測

定時はこの電極を被験者の喉頭付近に装着する。測定中に電極がずれないように、マジックテープのバンドなどを用いて固定する。電気喉頭計の電極は、人体に影響のない程度の微弱な電流が流れる構造となっている。これを装着した被験者が発声すると、声門の開閉に伴って電極間を流れる電流が変化する。これを計測することによって、発声時の声門の開閉パターンを測定することができる仕組みとなっている。



図 3.1: (上) 電気喉頭計装置 (Glottalenterprises 社製 EG-2) (下) EGG 電極 (Glottalenterprises 社製)

電気喉頭計によって測定された発声時の声門の開閉パタンの電気信号を、EGG 信号と呼ぶ。また、EGG 信号を微分したものを DEGG 信号と呼ぶ。図 3.2 に、成人男性の典型的な地声声区における EGG 信号と DEGG 信号の波形の例を示す。EGG 信号の振幅は声門の開閉の程度を示しており、正方向は閉鎖、負方向は開放をそれぞれ表している。また、DEGG 信号におけ

る正負のピークは、EGG 信号における声門の開閉が開始する時点を示しており、正のピークを声門閉鎖時点 (glottal closure instance)、負のピークを声門開放時点 (glottal opening instance) と呼ぶ。これより、声門閉鎖時点から次の声門閉鎖時点までの区間を基本周期 (fundamental period)、声門閉鎖時点から声門開放時点までの区間を閉鎖期 (closed phase)、声門開放時点から声門閉鎖時点までの区間を開放期 (open phase) とそれぞれ呼ぶ。このように、DEGG 信号から得られる情報を用いることで、元の EGG 信号の状態を知ることができる。

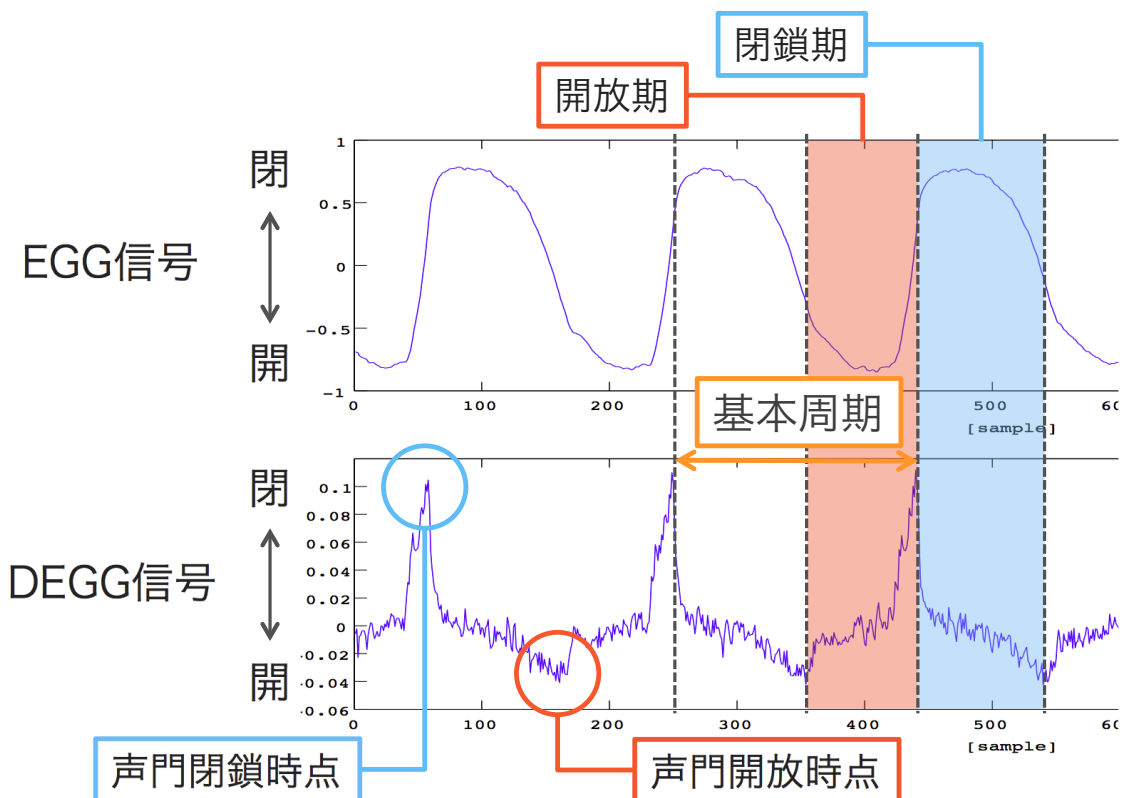


図 3.2: 典型的な地声声区における EGG 信号とその DEGG 信号の波形例。DEGG 信号から基本周期、開放期、閉鎖期を検出することができる。

EGG 信号の波形は、声区による違いを現す。図 3.3 に、男性の地声声区と裏声声区における典型的な EGG 信号の波形の例を示す。2つの声区における EGG 信号を比較すると、異なる波形パターンを示していることがわかる。地声声区における EGG 信号の波形パターンは、急な上昇と緩やかな下降からなる非対称な形状が特徴である。このことは、地声声区における声門が急な閉鎖と緩やかな開放を生じていることを意味している。一方、裏声声区における EGG 信

号の波形パターンは、三角波のような対称な形状と、地声声区と比べて波形の振幅が明らかに小さい点が特徴である。裏声声区の発声時においては、声帯の表層部である粘膜組織のみが小さく振動しており、また声帯同士が接触しないために、声門が閉鎖しない。

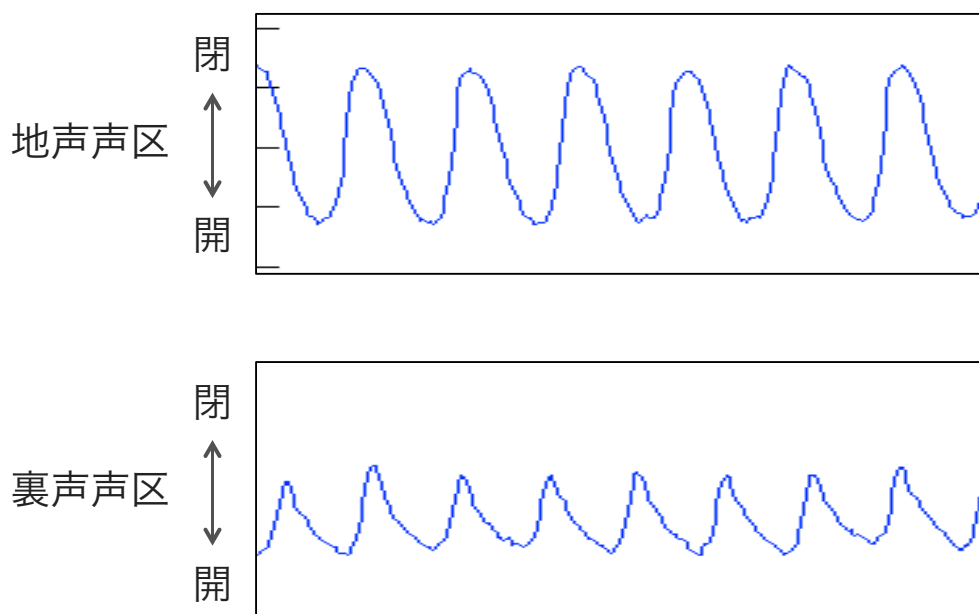


図 3.3: 成人男性の典型的な 2 声区における EGG 信号の時間波形の例。(上) 地声声区での発声時における EGG 信号波形。(下) 裏声声区での発声時における EGG 信号波形。

3.1.2 DECOM 法

電気喉頭計で声門の開閉パターンを測定することで、声門の開閉の周期や、声門の開放期と閉鎖期を分析することが可能となる。これらを分析することで、発声時における音源の基本周波数や声門開放率を明らかにすることができる。本論文では、電気喉頭計によって測定した EGG 信号から、基本周波数や声門開放率を求める手法として、DECOM 法 (DEgg Correlation-based Open quotient Measurement) [22] を用いている。以下に、DECOM 法の原理について述べる。

はじめに、EGG 信号を微分フィルタでフィルタリングすることで、DEGG 信号を得る。ここでは、微分フィルタとしてハイパスフィルタを用いたが、DEGG 信号に生じる微細構造を除去するために、700 Hz 以上の周波数成分を弱めるように設定した。得られた DEGG 信号に

において、正のピークは声門閉鎖時点を、負のピークは声門開放時点をそれぞれ示す。このことから、正のピークから次の正のピークまでの時間は EGG 信号の基本周期を、負のピークから正のピークまでの時間は声門の開放期をそれぞれ表している。DEGG 信号を正の波形と負の波形に分割し、正の波形における自己相関を求めることにより、基本周期を得る。また、負の波形と正の波形の相互相関を求めることにより、声門開放期を得る。最終的に、得られた基本周期の逆数を求めることで基本周波数を、基本周期に対する声門開放期の比を求めることで声門開放率を得られる。

相関を求める際に用いる分析フレームの長さは前のフレームで推定した基本周期の 4 倍長に、分析フレームのシフト幅は基本周期の 2 倍長にそれぞれ設定した。また、フレームで切り出した波形にはハミング窓をかけた。もし、前のフレームで基本周期を推定できなかった場合は、デフォルト値としてフレーム長を 40 msec、シフト幅を 5 msec に設定し、分析を行った。

3.2 声道部の測定と分析

本研究では、声道の共鳴特性やフォルマントを測定・分析する手法として、外部音響励振法とケプストラム法を組み合わせた手法を用いる。

3.2.1 外部音響励振法

外部音響励振法とは、発声中の被験者の声道を、外部から入力した励振信号で駆動し、放射される応答信号を分析することで、声道音響特性を得る手法である。

外部音響励振装置 (図 3.4) は、スピーカユニット (FF165WK; Fostex)、全長 1095 mm のエクスポネンシャルホーン、ホーン先端に取り付ける全長 300 mm 内径 7mm のビニルチューブで構成される。エクスポネンシャルホーンは 3D CAD で設計し、3D プリンターで作成したものである。この装置の低域カットオフ周波数は 150 Hz となっており、すべての母音の第 1 フォルマント周波数を十分に満たすように設計されている。

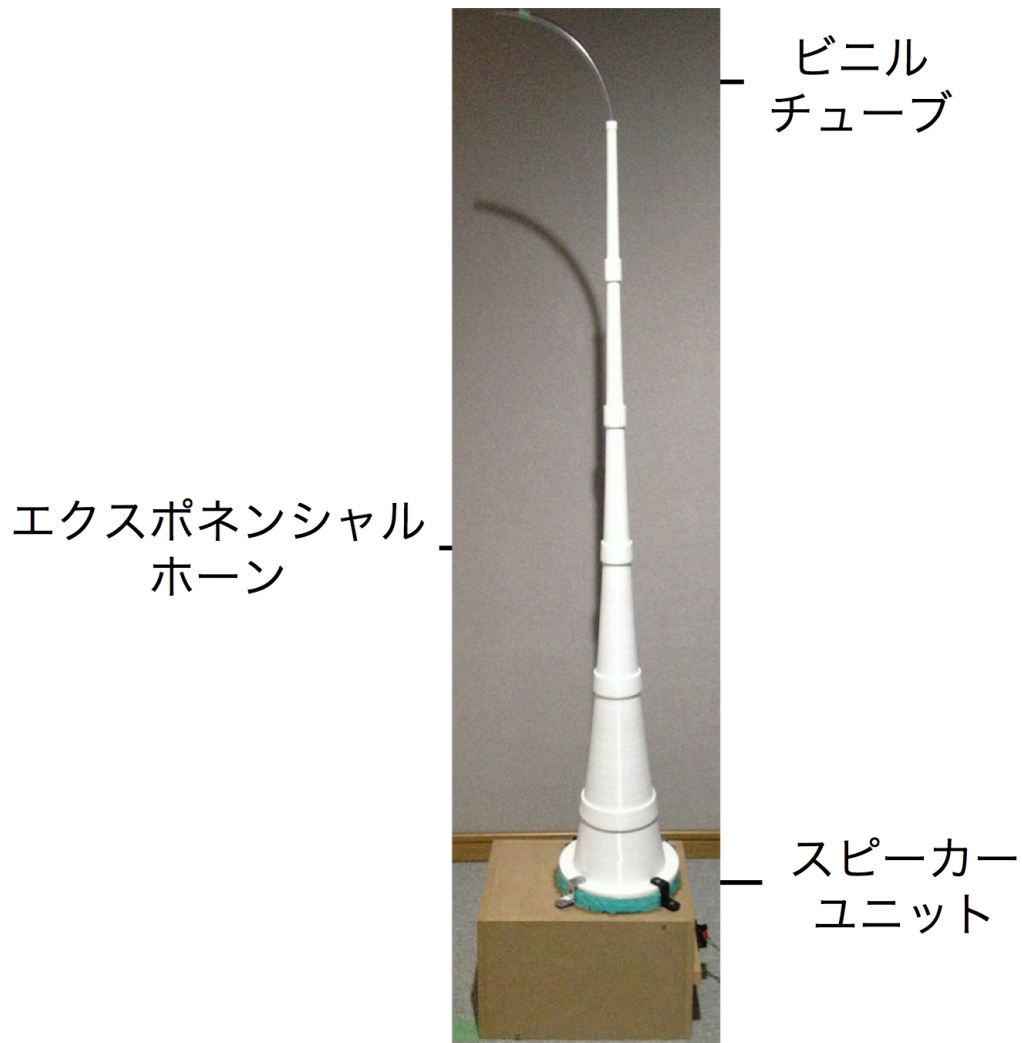


図 3.4: 外部音響励振装置 (研究室製作)。エンクロージャーによって格納されたスピーカーユニット、エクスポネンシャルホーン、ビニルチューブから構成される。エクスポネンシャルホーンは 3D CAD で設計したものを、3D プリンタによって 6 つのセクションに分割した形で樹脂から作製した後、それらを接続して製作した。エンクロージャーにはパイン材を用いた。

励振信号は計算機によって次のように生成する。まず、170 Hz から 6000 Hz までの広帯域の M 系列信号を生成する。サンプリング周波数は 16000 Hz とする。次に、外部音響励振装置の周波数特性を校正する。作成した M 系列信号を外部音響励振装置に入力し、装置に接続されたビニルチューブの先端から出力される信号をマイクロホンで収録する。この出力信号には、声道の音響特性を計測する上で不要な外部音響励振装置の周波数特性が含まれている。そこで、線形予測分析によってこの出力信号の周波数特性を推定した後、その逆特性を持つ線形フィルタではじめに作成した M 系列信号をフィルタリングすることで、励振装置の周波数特性をキャンセルする。これにより、外部音響励振装置から出力された際に、広帯域でフラットな周波数特性を持つ励振信号を生成できる。

励振信号は実験参加者が持続発声を行う間、口から声道に入力する。声道内に入力された励振信号は声道を駆動・伝搬したのち、声道の音響特性を付加した応答信号として、被験者の自発発声とともに放射される。口唇付近に設置したマイクロホンによって、この音響応答信号を収録する。この応答信号の周波数特性を分析することで、声道音響特性およびフォルマント周波数を得ることができる (図 3.5)。

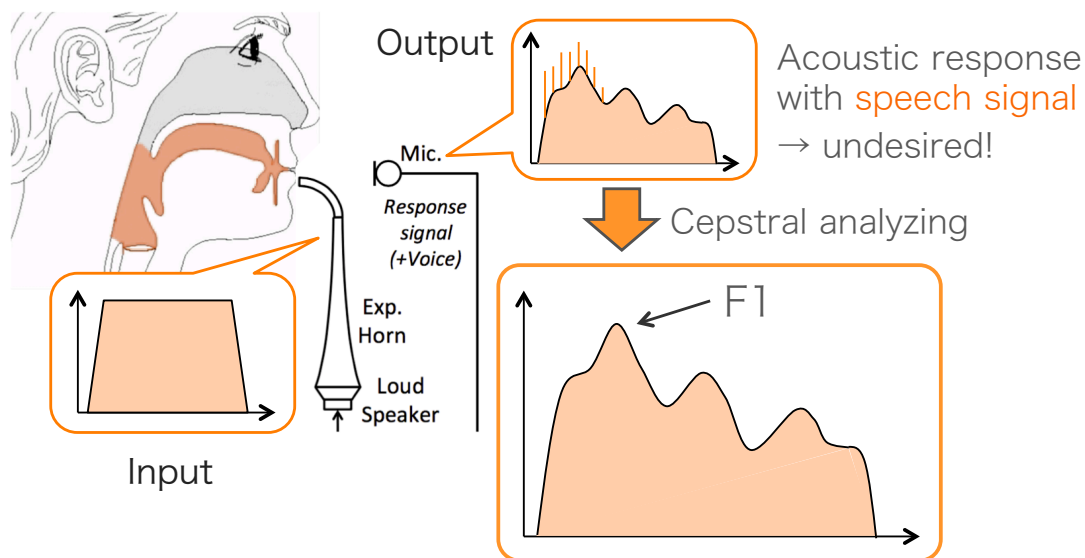


図 3.5: 外部音響励振法による声道音響特性の測定イメージ。周波数特性がフラットな励振信号を声道に入力することで、声道音響特性を付加した応答信号が、自発発声成分とともに出力される。自発発声成分（特に音源成分による線スペクトル）が不要なため、ケプストラム分析などで声道音響特性を抽出する。

3.2.2 ケプストラム分析

声道の音響特性は、測定した音響信号から得られる。しかし、この信号には被験者の自発発声が含まれている。特に自発発声成分の基本周波数と倍音成分は、主として音源に由来すること、声道音響特性と比べて鋭いピークを持つスペクトルとして現れることから、これを取り除くことが望まれる。ケプストラム分析は、音声スペクトルの大局的な情報である声道特性と、音声スペクトルの微細情報である音源特性を分離することができる。ケプストラム領域において、低次に声道情報、高次に音源情報があらわれるので、リフトリングによってこれらの情報を分離することで、所望の情報を得ることができる。

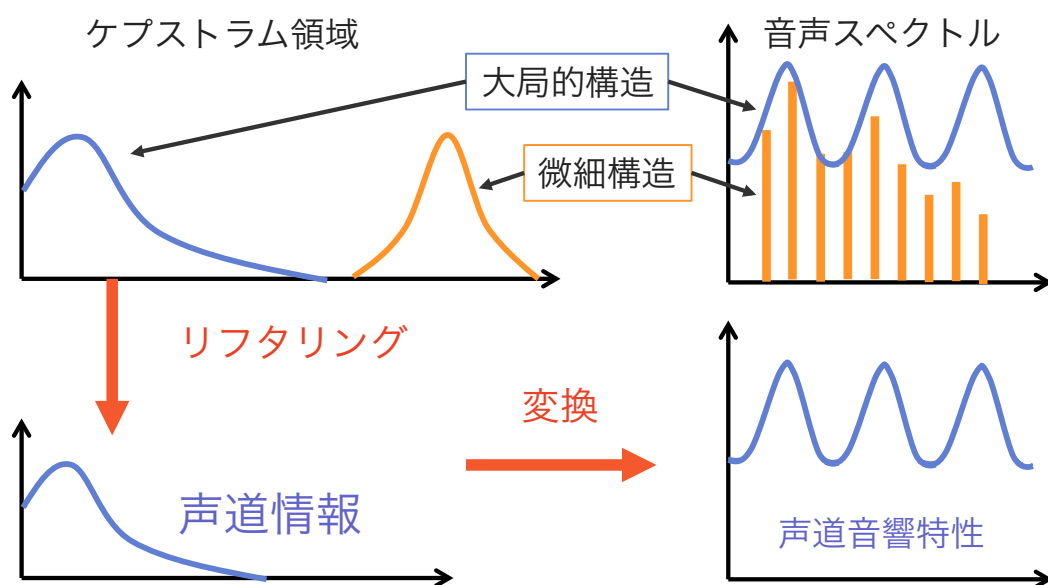


図 3.6: ケプストラム分析のイメージ。音声スペクトルにおいて、音源情報は微細構造、声道情報は大局的構造であるため、ケプストラム領域に変換することで、それぞれが低次成分と高次成分に分離できる。リフトリングによってこれらの情報を分離することで、所望の情報を得ることができる。

第4章 声区転換における音源–フィルタ相互作用の影響の検討

4.1 背景と目的

音声のピッチを、地声声区から裏声声区へと次第に上昇させたとき、もしくは裏声声区から地声声区へと次第に下降させたとき、ピッチがある音高に達すると、声区が急に切り替わる。また声区が切り替わる際、周波数ジャンプと呼ばれるピッチの急激な変化も伴うことが知られている。このような発声現象を声区転換 (Voice Register Transition) と呼ぶ。

声区転換を生じる要因として、これまで2つの異なるメカニズムについて検討されている。一つは、声帯の張力変化と、それに伴った声帯振動における有効質量の変化によるものである。Van den Vergら [7] はイヌの摘出喉頭を用いた測定実験で、声帯の張力を滑らかに変化させたとき、基本周波数が徐々に上昇したのち、より高い周波数へと急激な周波数ジャンプを生じることを観測した。Švecら [8] と Millerら [9] は、ヒトの摘出喉頭を用いた同様の実験で、ピッチのジャンプが生じることを示した。さらに、実験参加者を用いた実際の発声測定実験において、周波数ジャンプを伴う声区転換が生じることを観測した。このように、声帯のなだらかな張力変化は声区転換を引き起こすことが知られており、これを音源由来のメカニズムと呼ぶ。

一方、声区転換を生じるもう一つの要因として検討されているものが、音源–フィルタ相互作用である。これは、喉頭周辺の音源機構と声道の音響負荷が違いに影響を及ぼしあうことで声帯振動が不安定になり、結果としてピッチの周波数ジャンプやサブハーモニクスなどを生じるというものである。Titze [3] は音声生成シミュレーション実験によって、発声のピッチが第1フォルマント周波数に近接したとき、音源–フィルタ相互作用の影響が強まることを示した。また、この音源–フィルタ相互作用の影響は、epilarynx tubeの断面積が狭まるとより強くなることがわかった。Tokuda [10, 11] らは、4質量モデルなどの声帯物理モデルと音源–フィル

タ相互作用を組み合わせた音声生成モデルによる声区転換のシミュレーション実験を行った。Kaburagi [12] は、声門の境界層解析と音源-フィルタ相互作用を組み合わせた音声生成モデルによる声区転換のシミュレーション実験を行った。いずれのシミュレーション実験においても、ピッチが第1フォルマントに近接した際に声区転換を生じることを示した。このように、発声の基本周波数が第1フォルマント周波数に近接したとき、音源-フィルタ相互作用の影響が強まることで発声が不安定になり、声区転換を生じることが示唆されている。これを音響由来のメカニズムによる声区転換と呼ぶ。

音響由来のメカニズムによる声区転換について、音源-フィルタ相互作用の影響を直接的に測定することは非常に難しい。そこで、声区転換を生じる際における、基本周波数と第1フォルマント周波数の関係性を調べることによって、音源-フィルタ相互作用が声区転換に影響を及ぼすかどうかを間接的に調べることが行われている。Titzeら [15] は複数のピッチパターンと母音の種類を組み合わせたによる声区転換の発声測定実験を行った。また、Zañartuら [16] は実験参加者1名による上昇系列と下降系列の2種類のグリッサンド発声測定実験を行い、発声の基本周波数、声道と声門下部のフォルマント、フローマスクによる流量、声門の開閉による声門断面積の時間変化を撮像したキモグラフなどを計測した。これらの測定実験による研究は、先行していたシミュレーション実験の結果と同様に、音源由来のメカニズムと音響由来のメカニズムの両方が、声区転換を引き起こす要因であることを示した。また、母音/i/のような比較的低い第1フォルマント周波数の母音で、音響由来のメカニズムによる声区転換を生じる傾向があることがわかった。

しかしながら、これらの先行研究における問題点として、フォルマント周波数の精確さという点が残っている。Titzeら [15] はヴォーカルフライで発声した音声信号から第1フォルマント周波数を計測した。しかしヴォーカルフライでの声門の開閉の挙動は、地声声区や裏声声区、さらには声区転換を生じる際のものとは異なっている。声道の末端である声門の開閉の挙動の違いは、気管支や肺などの声門下部の組織が声道に与える音響的な影響にも違いを生じる。そのため、フォルマント周波数にも違いを与えられられることから、ヴォーカルフライ発声によって測定したフォルマント周波数が、必ずしも精確なフォルマントであるとはいえない。また、Zañartuら [16] は、基本周波数が上昇していく音声信号に対して線形予測分析を適用することにより、第1フォルマント周波数を推定した。しかし、音声に含まれる倍音成分

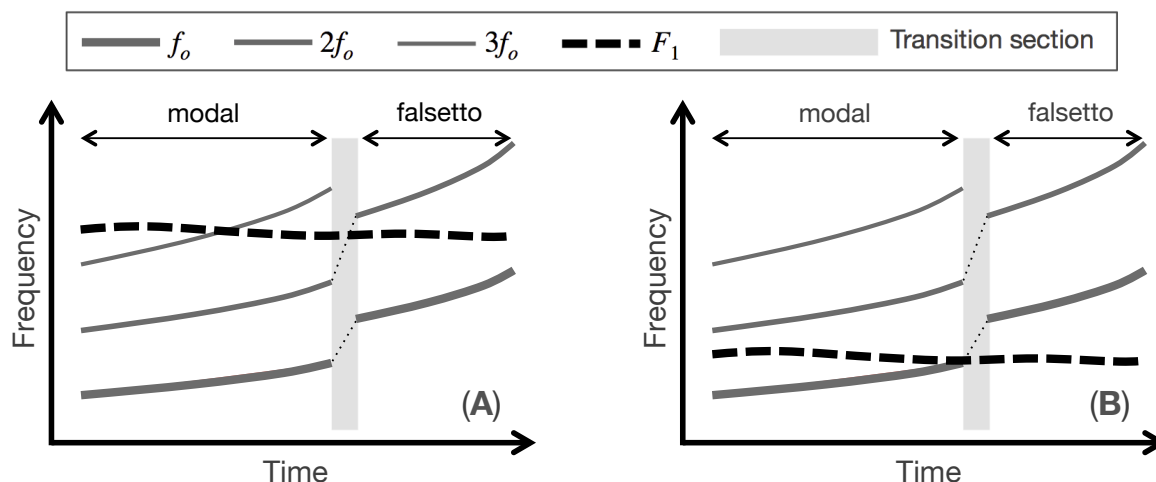


図 4.1: これまでに検討されている、声区転換時における基本周波数および低次の倍音成分と第 1 フォルマント周波数の近接関係と、声区転換を生じる要因との関係性。(A) 声区転換を生じる際において、基本周波数も低次の倍音成分も第 1 フォルマント周波数に近接していない場合。このとき、音源由来の声区転換を生じていると考えられる。(B) 声区転換を生じる際において、基本周波数が第 1 フォルマント周波数に近接している場合。このとき、音源-フィルタ相互作用の影響が強まったことで、音響由来の声区転換を生じていると考えられる。

は基本周波数が上昇するほど疎になるため、高い基本周波数の音声信号から信号処理によって精確なフォルマント周波数を推定することは難しいことが知られている。またキモグラフや流量を測定するために、エンドスコープの内挿やフローマスクの装着を行っているために、これらが発声のしづらさや、実際の声道音響特性に何らかの影響を及ぼしている可能性もある。

これらの問題を解決した上で、フォルマント周波数をより精確に測定するために、本研究では外部音響励振法を用いた。この手法では、発声中の声道を外部から音響的に励振することで、声道の音響特性を測定するものである。はじめに、雑音信号を口内に入力して、声道を励振する。この雑音信号は、低次のフォルマント周波数を十分に含むように広帯域で平坦な周波数振幅特性を持つように設計されている。口内への入力の際には、声道を十分に駆動させるために、出力の音圧パワーをできるだけ高めた状態で口内に入力する。声道を励振した信号は、声道の音響特性を付加した応答信号として口から出力されるため、これを分析することでフォルマント周波数を得ることができる。本手法は、音声や歌唱音声の発声時、または管楽器の吹奏時における声道の特性を計測するための手法として既に幅広く使用されている [18-21]。

本研究の目的は、声区転換を引き起こす要因として考えられる 2 種類のメカニズムが存在す

ることを明らかにすることである。特に、音源-フィルタ相互作用による声区転換への影響について明らかにするために、発声の基本周波数と倍音成分、および第1フォルマント周波数をより精確に測定・分析し、両者にどのような周波数の関係性が存在するのかを調べる。

4.2 実験

4.2.1 実験参加者と実験内容

実験参加者は歌唱のトレーニングを受けていない成人男性5名とした。声区転換の発声測定実験を行う前に、地声声区と裏声声区の両声区で発声することが可能な音域であるオーバーラップレンジを、各実験参加者ごとに測定した。はじめに、地声声区で発声しやすい典型的な音高で母音/a/を持続発声してもらい、音高を確認した。その音高を開始点として、半音だけ上昇させた音を実験参加者に呈示し、母音/a/で地声発声できるか確認した。これを繰り返し、地声声区で発声可能な上限の音高を確認し、オーバーラップレンジの最高音とした。続いて、裏声声区で発声しやすい典型的な音高で母音/a/を持続発声してもらい、音高を確認した。その音高を開始点として、半音だけ下降させた音を実験参加者に呈示し、母音/a/で裏声発声できるか確認した。これを繰り返し、裏声声区で発声可能な下限の音高を確認し、オーバーラップレンジの最低音とした。表4.1に、各実験参加者のオーバーラップレンジを示す。これより、いずれの実験参加者も裏声声区で発声可能であり、オーバーラップレンジを持つことから、地声声区から裏声声区への声区転換の発声測定実験に適していると判断した。

発声測定実験は防音室内で行った。本実験において実験参加者には、地声声区から裏声声区へ、基本周波数が上昇するようなグリッサンド発声を行うように指示した。これは一般的に、

表 4.1: 各被験者の年齢およびオーバーラップレンジ。

Subject	Age	Overlap range
S1	27	B3(246.9 Hz) — C5(523.3 Hz)
S2	26	C4(261.6 Hz) — F4(349.2 Hz)
S3	24	A3(220.0 Hz) — F \sharp 4(370.0 Hz)
S4	23	C4(261.6 Hz) — E5(659.3 Hz)
S5	23	D4(293.7 Hz) — E4(329.6 Hz)

基本周波数を下降させるようなグリッサンド発声と比べて、上昇グリッサンド発声のほうが、よりピッチを滑らかに変化させて発声しやすいためである。なお、実験参加者にはできるだけ自然な発声を行ってもらうこと、意図的な声区転換を伴わせないようにすることから、声区転換を伴ったグリッサンドを発声するかどうかなどの指示は特に与えなかった。実験条件として、実験参加者には日本語母音の/a/または/i/の単母音で発声すること、それぞれの母音について20試行以上の上昇グリッサンド発声を行うこと、またグリッサンド発声のタイミングをなるべく一定にするために、イヤホンから呈示されるガイド音の音高にできるだけ沿ってグリッサンドを発声することを指示した。図4.2に、実験時にイヤホンから呈示したガイド音のスペクトログラムを示す。このガイド音は、50 msec 長で500 Hzの高さの純音を断続的に4つ並べた合図音と、1500 msecの間に100 Hzから500 Hzまで対数的に上昇するようなチャープ信号を連結することで作成した。また、チャープ信号の周波数帯域は、すべての実験参加者のオーバーラップレンジを満たすように設定した。

4.2.2 測定手法

図4.3は本研究における測定系のブロックダイアグラムを示す。声道の音響特性は外部音響励振法 [17-21] を用いて計測した。本実験において、マイクロホンは実験参加者の口唇から10 cmの位置に設置した。発声測定実験は立位で行った。実験参加者の喉頭付近に電気喉頭計の電極を装着した。外部励振装置のビニルチューブの先端部およそ3cmを実験参加者の口内に挿入して、励振信号によって声道を駆動した。得られたEGG信号と音響信号を同時計測し、オーディオインターフェースを介して計算機に同時収録した。この音響信号には声道の応答信号と実験参加者の自発発声の両方が含まれている。なお、オーディオインターフェースは広帯域励振信号とガイド音の出力も行っている。

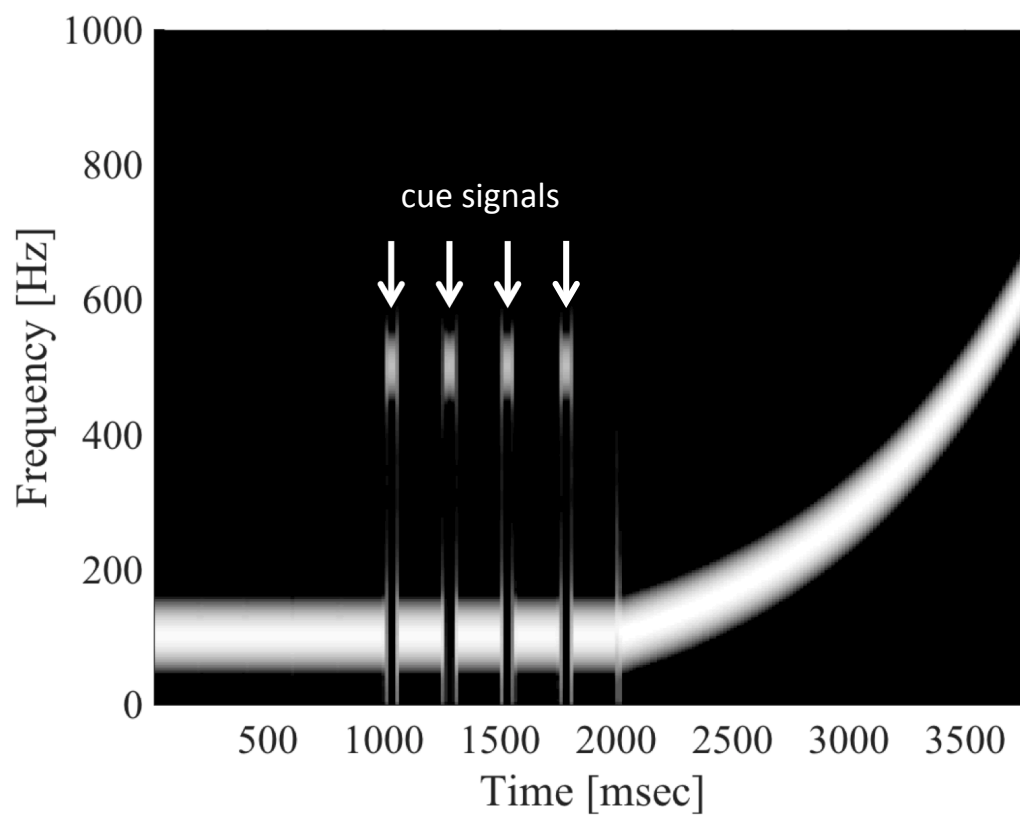


図 4.2: 本実験で用いた実験参加者に呈示するガイド音のスペクトログラム。50 msec 長 500 Hz の純音を断続的に 4 つ並べた合図音と、1500 msec の間に 100 Hz から 500 Hz まで対数的に上昇するようなチャープ信号を連結している。

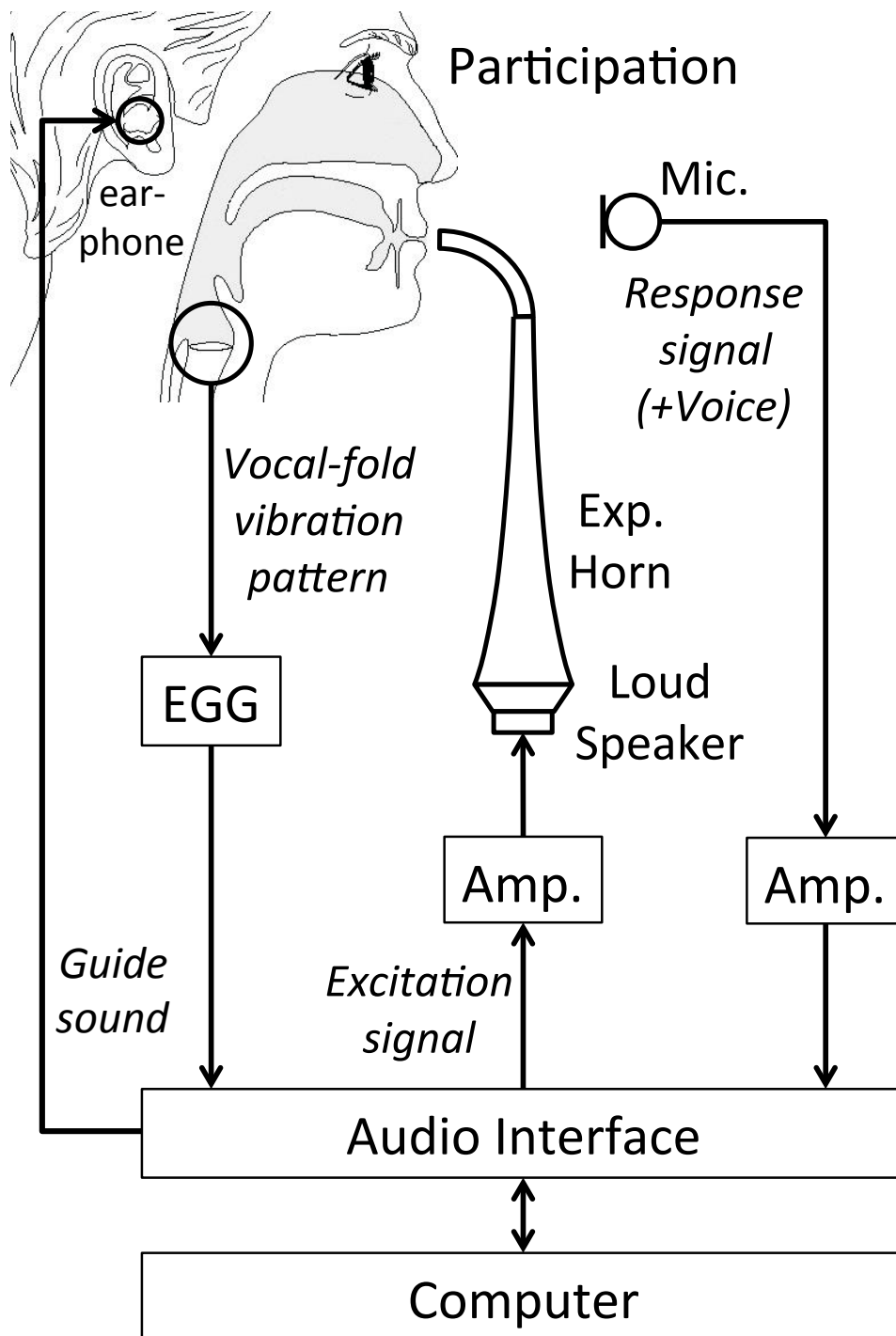


図 4.3: 本実験に用いた測定系のブロックダイアグラム。電気喉頭計と外部音響励振法を用いて、声門の開閉パターンと声道の音響応答信号を同時に計測し、オーディオインターフェースを介して計算機に収録した。また、実験参加者にはイヤホンを通じて、グリッサンド発声のためのガイド音を呈示した。

4.2.3 基本周波数の分析

発声時の基本周波数と声門開放率は、EGG 信号の微分信号である DEGG 信号に DECOM 法 [22] を適用して分析した。はじめに、EGG 信号を微分フィルタでフィルタリングすることで、DEGG 信号を得る。ここでは、微分フィルタとしてハイパスフィルタを用いたが、DEGG 信号に生じる微細構造を除去するために、700 Hz 以上の周波数成分を弱めるように設定した。得られた DEGG 信号において、正のピークは声門閉鎖時点を、負のピークは声門開放時点をそれぞれ示す。このことから、正のピークから次の正のピークまでの時間は EGG 信号の基本周期を、負のピークから正のピークまでの時間は声門の開放期をそれぞれ表している。DEGG 信号を正の波形と負の波形に分割し、正の波形における自己相関を求めることにより、基本周期を得る。また、負の波形と正の波形の相互相関を求めることにより、声門開放期を得る。最終的に、得られた基本周期の逆数を求めることで基本周波数を、基本周期に対する声門開放期の比を求めることで声門開放率を得られる。

相関を求める際に用いる分析フレームの長さは前のフレームで推定した基本周期の 4 倍長に、分析フレームのシフト幅は基本周期の 2 倍長にそれぞれ設定した。また、フレームで切り出した波形にはハミング窓をかけた。もし、前のフレームで基本周期を推定できなかった場合は、デフォルト値としてフレーム長を 40 msec、シフト幅を 5 msec に設定し、分析を行った。

4.2.4 声道音響特性の分析

外部音響励振法によって測定した音響信号に対して、ケプストラム分析とリフタリング処理を行い、自発音声に由来した不要な信号成分を取り除くことで、声道音響特性を抽出した。はじめに、分析フレームで音響信号を切り出し、窓掛け処理を行った上で対数パワースペクトルを計算した後、ケプストラムパラメータを計算した。ここでは、分析フレーム長は 30 msec、シフト幅は 5msec とし、分析窓にはハミング窓を用いた。また、リフタリングの閾値は 2.5msec とした。最終的に、各フレームの声道音響特性からピークピッキングによって第 1 フォルマント周波数を推定し、第 1 フォルマントの時間変化パターンを得た。図 4.4 および 図 4.5 は、母音/a/または母音/i/で、声区転換を含むような上昇系列のグリッサンドを発声した際に、外部音響励振法を用いて測定した声道音響特性の結果を示す。それぞれの曲線が、各分析フレーム

においてケプストラム分析により得られた声道音響特性である。また、下から上へ進むにつれて、分析フレーム時間が経過している。赤い曲線部は、声区転換を生じている区間における声道音響特性を示す。

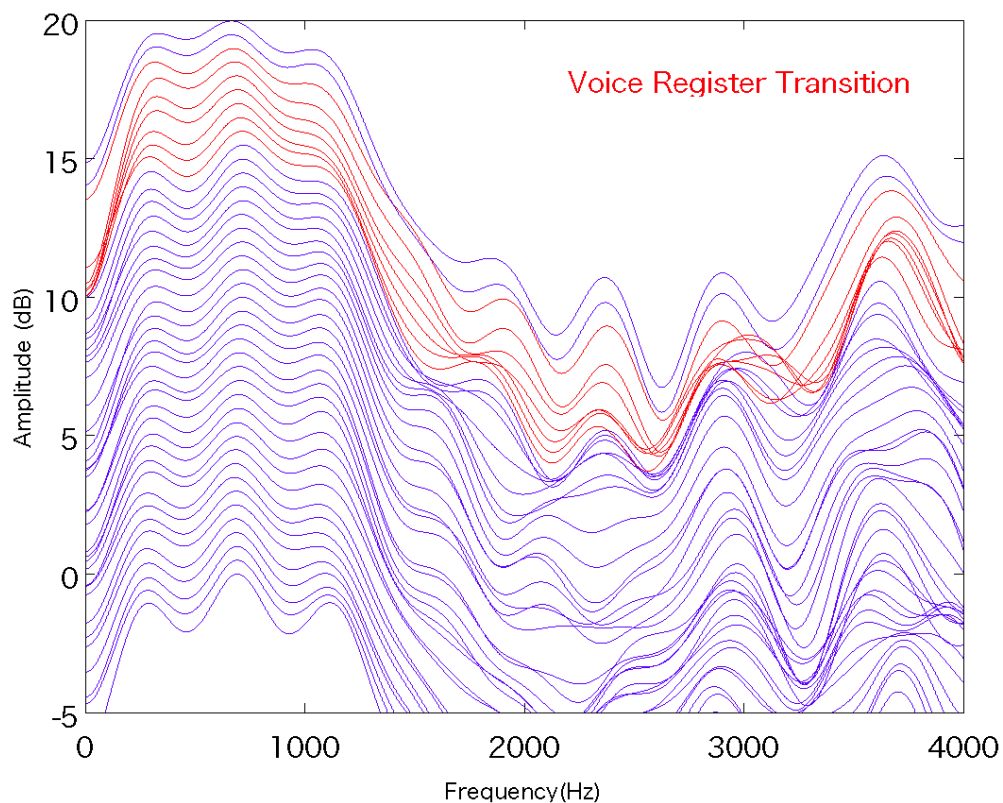


図 4.4: 母音/a/で声区転換を含む上昇系列のグリッサンドを発声した際における、声道音響特性の時間変化の様子。上方方向へ進むほど、時間が経過している。特に赤い曲線部は、声区転換を生じている区間における声道音響特性を示す。

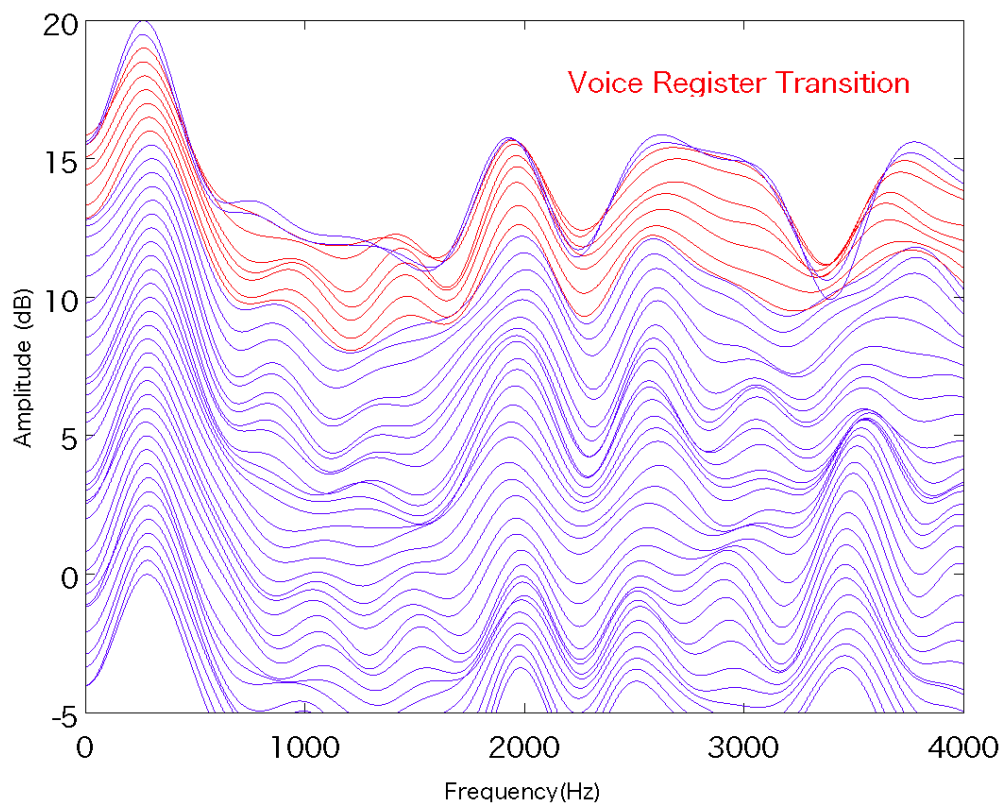


図 4.5: 母音/i/で声区転換を含む上昇系列のグリッサンドを発声した際における、声道音響特性の時間変化の様子。上方向へ進むほど、時間が経過している。特に赤い曲線部は、声区転換を生じている区間における声道音響特性を示す。

4.3 結果

4.3.1 母音/a/の分析結果

図 4.6 ~ 4.8 に、母音/a/の上昇グリッサンド発声時における分析結果の例を示す。図 4.6 は声区転換を生じた際における基本周波数と低次の倍音成分、および第 1 フォルマント周波数の時間変化を示す。これより、基本周波数の周波数ジャンプがおおよそ 1100 msec 付近で生じていることがわかる。また、このジャンプを生じる直前、基本周波数および倍音成分は第 1 フォルマントに近接していないことがわかる。

図 4.7 は声区転換を生じた際における EGG 信号の時間波形と、得られた声門開放率の時間変化を示す。EGG 信号の波形に着目すると、図 4.6 で周波数ジャンプを生じたおおよそ 1100 msec 付近において、波形の振幅が急に低下していることがわかる。また、声門開放率の時間変化に注目すると、同じタイミングで値が急に低下していることがわかる。以上のことから、この分析結果の例では、おおよそ 1100 msec 付近で周波数ジャンプを伴った声区転換を生じているといえる。

図 4.8 は EGG 信号のスペクトログラムの分析結果である。声区転換を生じた 1100 msec 付近を拡大して示している。これより、基本周波数とその倍音成分が急な周波数ジャンプを生じていることがわかる。一方で、周波数ジャンプの際に、明瞭なサブハーモニクス成分はみられないことがわかる。このような特徴を持つ声区転換は、他の実験参加者による母音/a/の声区転換の分析結果においても同様に見られた。

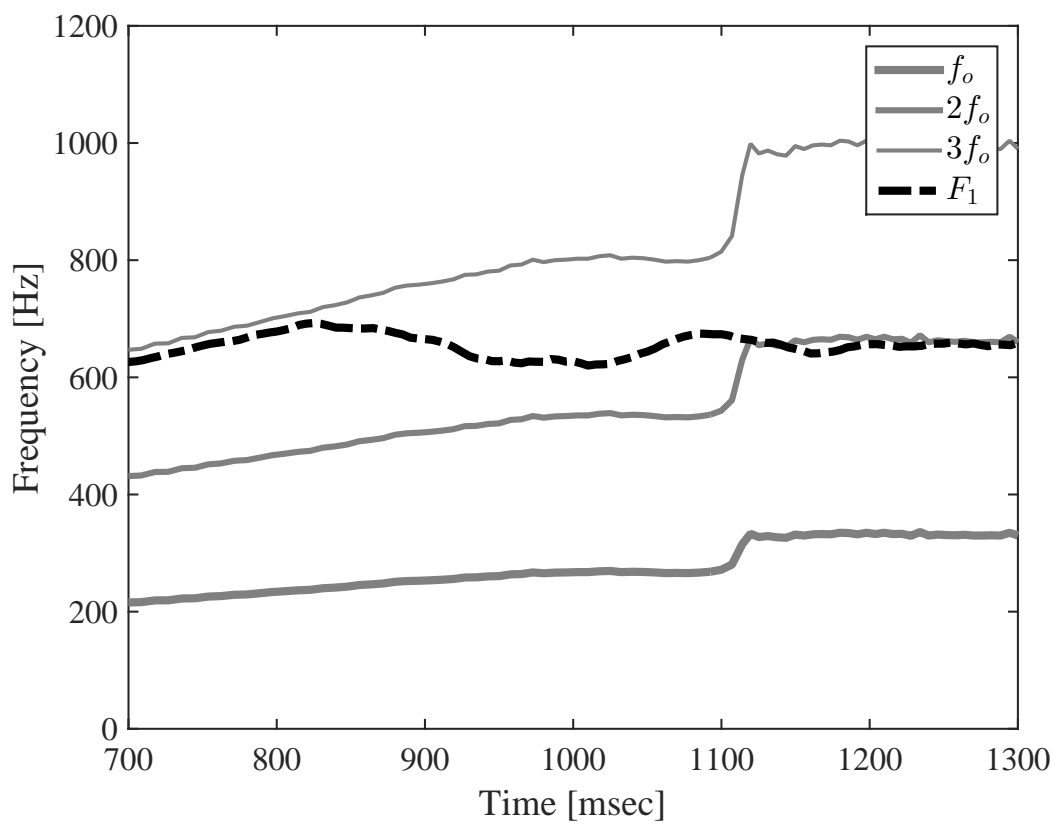


図 4.6: 母音/a/で声区転換を生じる際における、基本周波数と低次の倍音成分、および第1フォルマント周波数の時間変化の分析結果。

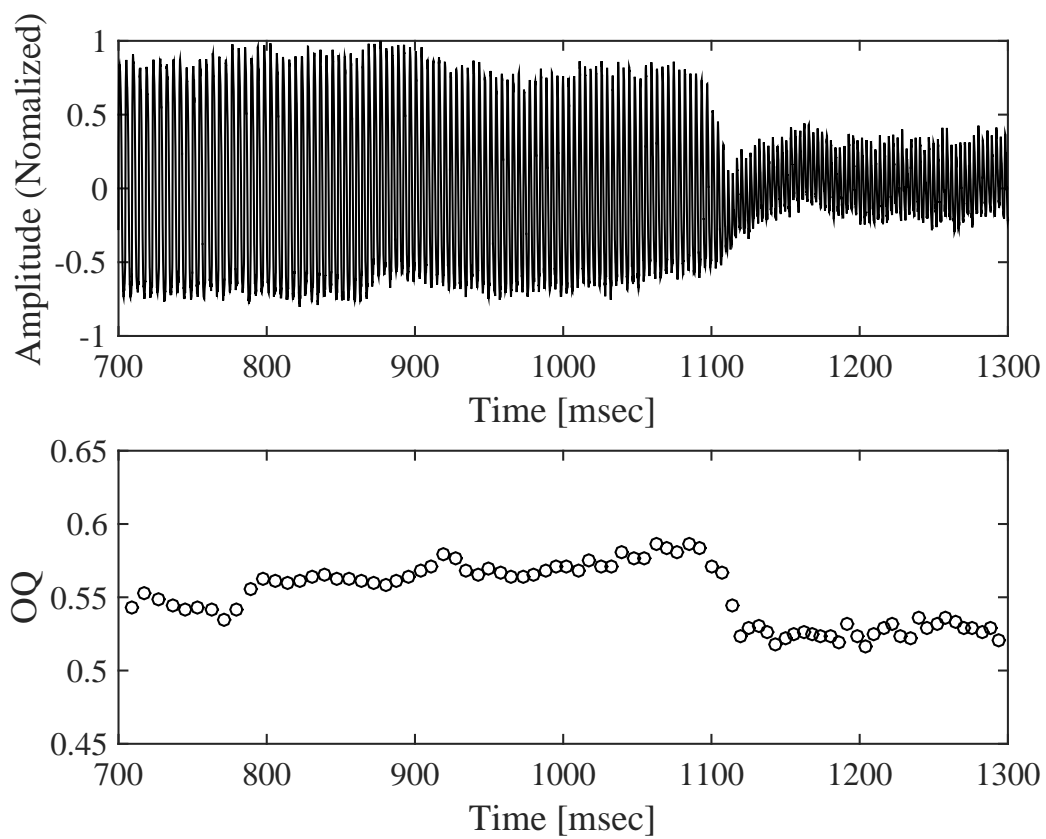


図 4.7: 母音/a/で声区転換を生じる際における、EGG 信号の時間波形と得られた声門開放率の時間変化の分析結果。

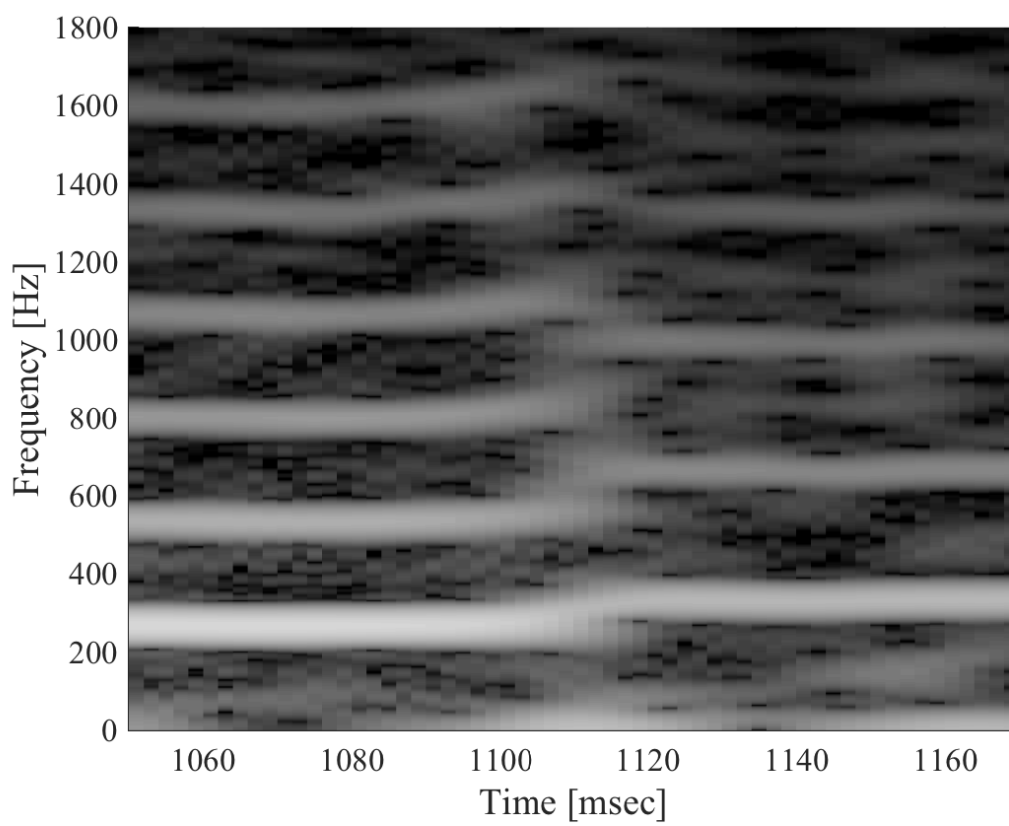


図 4.8: 母音/a/で声区転換を生じる際における、EGG 信号のスペクトログラムの声区転換部を拡大した分析結果。

4.3.2 母音/i/の分析結果

図 4.9 ~ 4.11 に、母音/i/の上昇グリッサンド発声時における分析結果の例を示す。

図 4.9 は声区転換を生じた際における基本周波数と低次の倍音成分、および第 1 フォルマント周波数の時間変化を示す。これより、基本周波数の周波数ジャンプがおよそ 970 msec 付近で生じていることがわかる。この周波数ジャンプは、母音/a/での場合のものと比べて、より不連続になっているように見える。また、周波数ジャンプを生じる直前、基本周波数が第 1 フォルマントに明らかに近接していることがわかる。

図 4.10 は声区転換を生じた際における EGG 信号の時間波形と、得られた声門開放率の時間変化を示す。図 4.9 で周波数ジャンプを生じていたおよそ 970 msec 付近において、EGG 波形の振幅が鋭く低下していること、声門開放率の時間変化もより急激に変化しており、瞬間的に声門の開閉がほとんど生じていない状態になっていることがわかる。この“ブレイク”の後、EGG 信号の振幅は回復しているが、ブレイク前と比較すると振幅が小さくなっていることがわかる。以上より、この分析結果の例ではおよそ 970 msec 付近において、母音/a/での場合とは質の異なる声区転換を生じているといえる。

図 4.11 は EGG 信号のスペクトログラムの分析結果である。声区転換を生じた 970 msec 付近を拡大して示している。これより、基本周波数とその倍音成分が急な周波数ジャンプを生じていることがわかる。加えて、周波数ジャンプの際に、母音/a/での場合では見られなかった周波数成分の明瞭な分岐とサブハーモニクスが発生が見られる。このような特徴を持つ声区転換は、他の実験参加者による母音/i/の声区転換の分析結果においても同様に見られた。

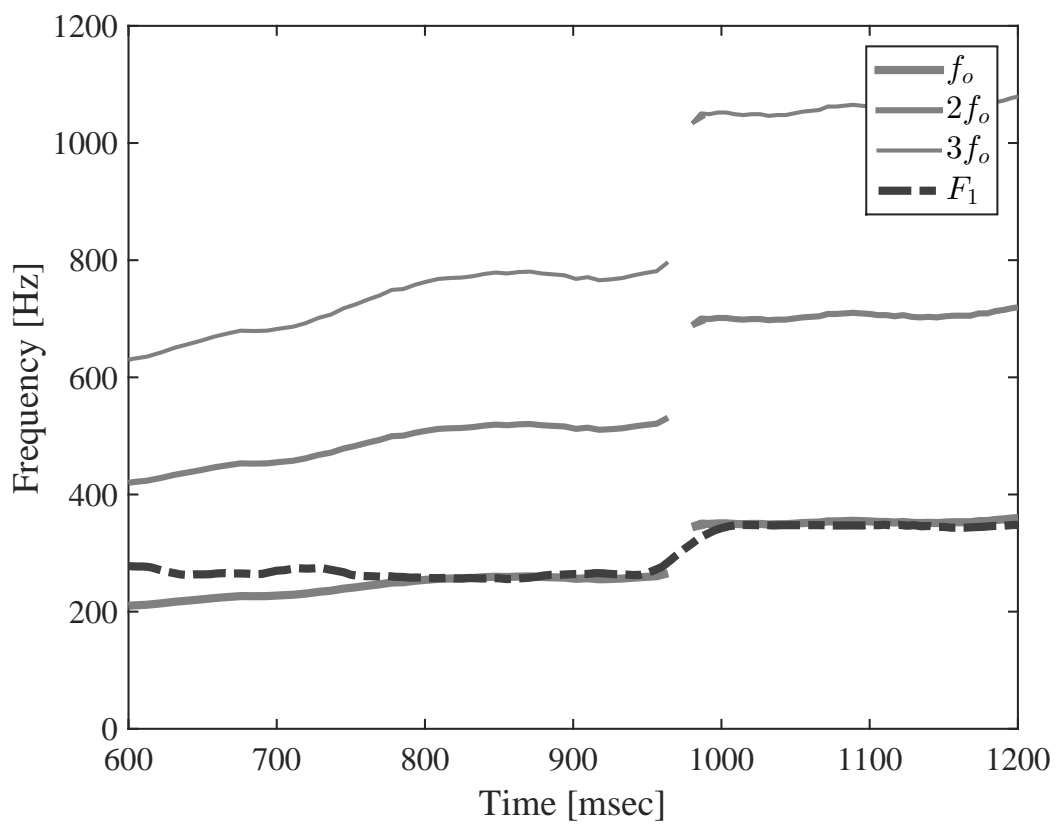


図 4.9: 母音/i/で声区転換を生じる際における、基本周波数と低次の倍音成分、および第1フォルマント周波数の時間変化の分析結果。

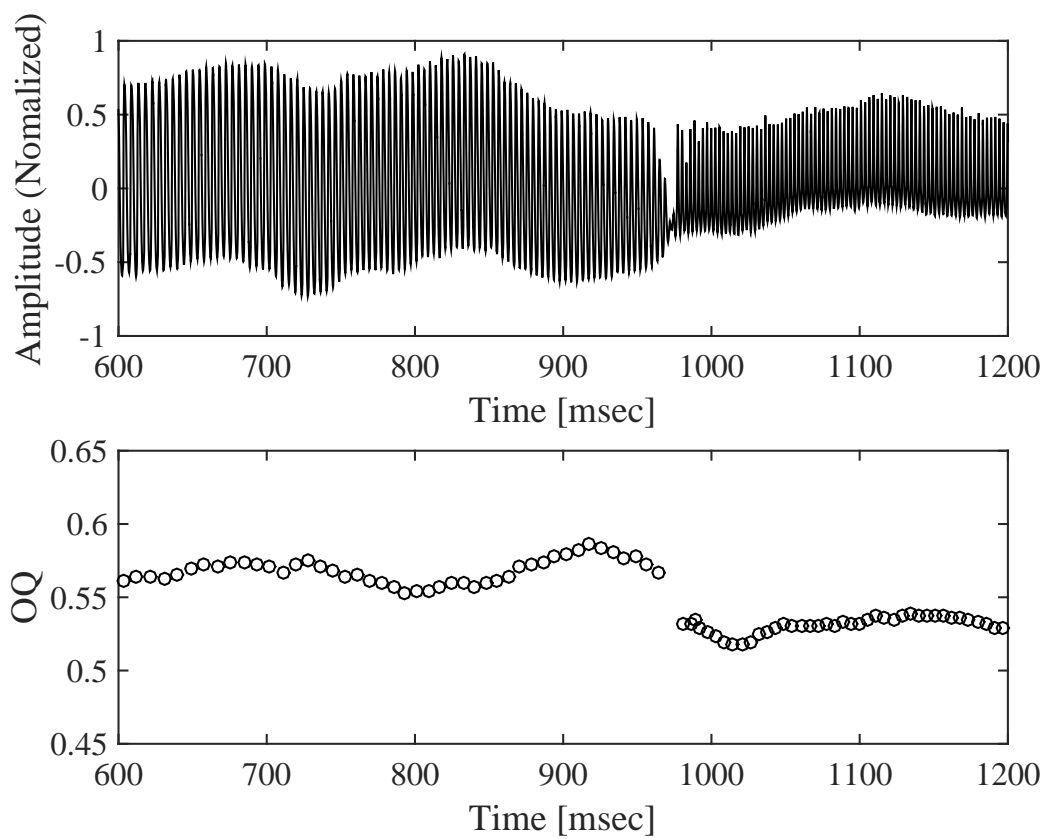


図 4.10: 母音/i/で声区転換を生じる際における、EGG 信号の時間波形と得られた声門開放率の時間変化の分析結果。

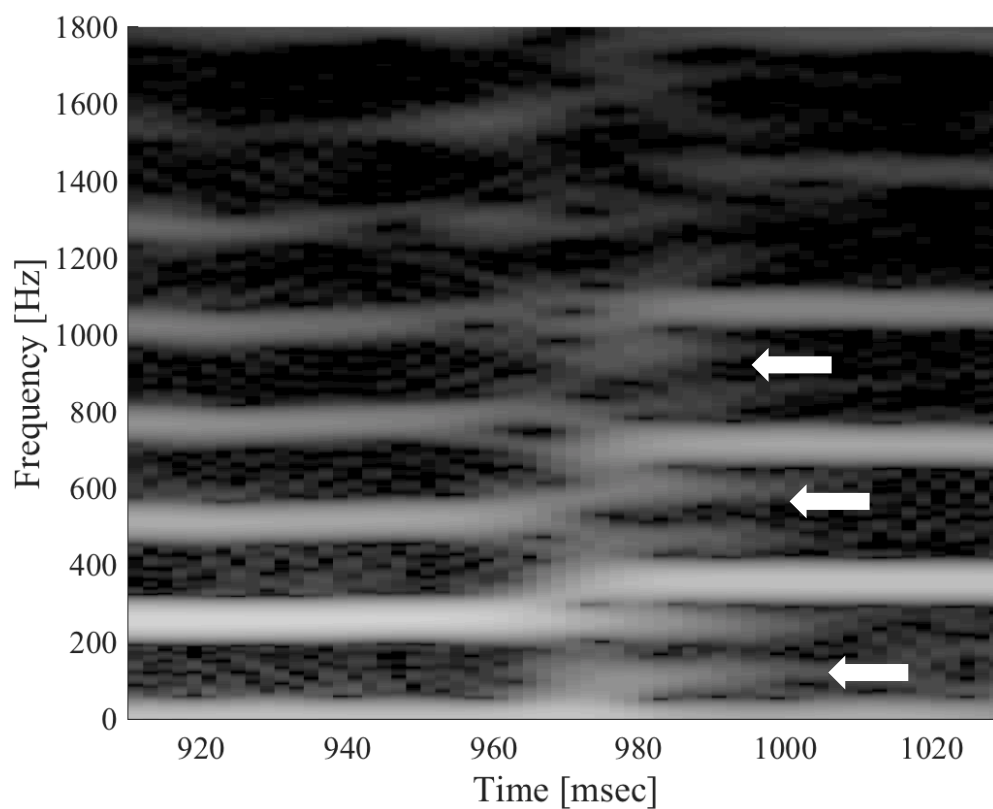


図 4.11: 母音/i/で声区転換を生じる際における、EGG 信号のスペクトログラムの声区転換部を拡大した分析結果。

4.3.3 全試行の分析結果

図 4.12 に、各実験参加者ごとの母音/a/と/i/での上昇グリッサンド発声時において、声区転換を生じる際における基本周波数とその倍音成分、および第1フォルマン트의周波数の関係性を示す。図中の各プロットは、外部音響励振法によって測定した全データのうち、声道の励振を十分に行えた場合のデータを用いている。各図の横軸は基本周波数および倍音成分の周波数(Hz)を、縦軸は第1フォルメント周波数(Hz)をそれぞれ示している。これらの値は、地声声区から裏声声区への声区転換を生じる際に見られる周波数ジャンプが起こるときの、声門の開閉パターンと声道の音響特性のデータを分析したものである。各図中に見られる傾きを持つ線分は、基本周波数もしくは倍音成分の周波数と、第1フォルマン트의周波数が一致している場合を示す。この線分の近傍に位置する濃色のマーカーは、基本周波数もしくは倍音成分が、第1フォルメント周波数に近接している場合を示す。なお、ここでは、両者の周波数差が40 Hz以下である場合において、両者の周波数が近接関係にあるとみなす。

図 4.12 に示す母音/i/の場合の結果より、すべての実験参加者のほとんどのデータにおいて、声区転換を生じる際に基本周波数が第1フォルメント周波数に近接していることがわかる。加えて、濃色のマーカーのほとんどが、周波数の一致を示す線分の右側に位置していることがわかる。このことは、周波数ジャンプを生じる際の基本周波数が、第1フォルメント周波数よりも僅かに高いことを示しており、シミュレーション実験による先行研究([15,23])による示唆と一致している。

一方、図 4.12 に示す母音/a/の場合の結果より、すべての実験参加者のデータにおいて、声区転換を生じる際に、基本周波数が第1フォルメント周波数に近接している場合は見られなかったが、第2倍音や第3倍音が第1フォルメント周波数に近接している場合がいくつか見られた。図 4.13 ~ 図 4.15 は、母音/a/での声区転換を生じる際に、第2倍音が第1フォルメント周波数に近接している場合の分析結果の例を示す。また、図 4.16 ~ 図 4.18 は、母音/a/での声区転換を生じる際に、第3倍音が第1フォルメント周波数に近接している場合の分析結果の例を示す。ここで、それぞれの結果における EGG 信号の分析結果には、母音/i/の場合に見られるような特徴が見られることがわかる。すなわち、周波数ジャンプが起こる際に、EGG 信号の振幅が鋭く低下すること、声門開放率が急激に変化すること、そして周波数の分岐やサブハーモニクスを生じることなどである。

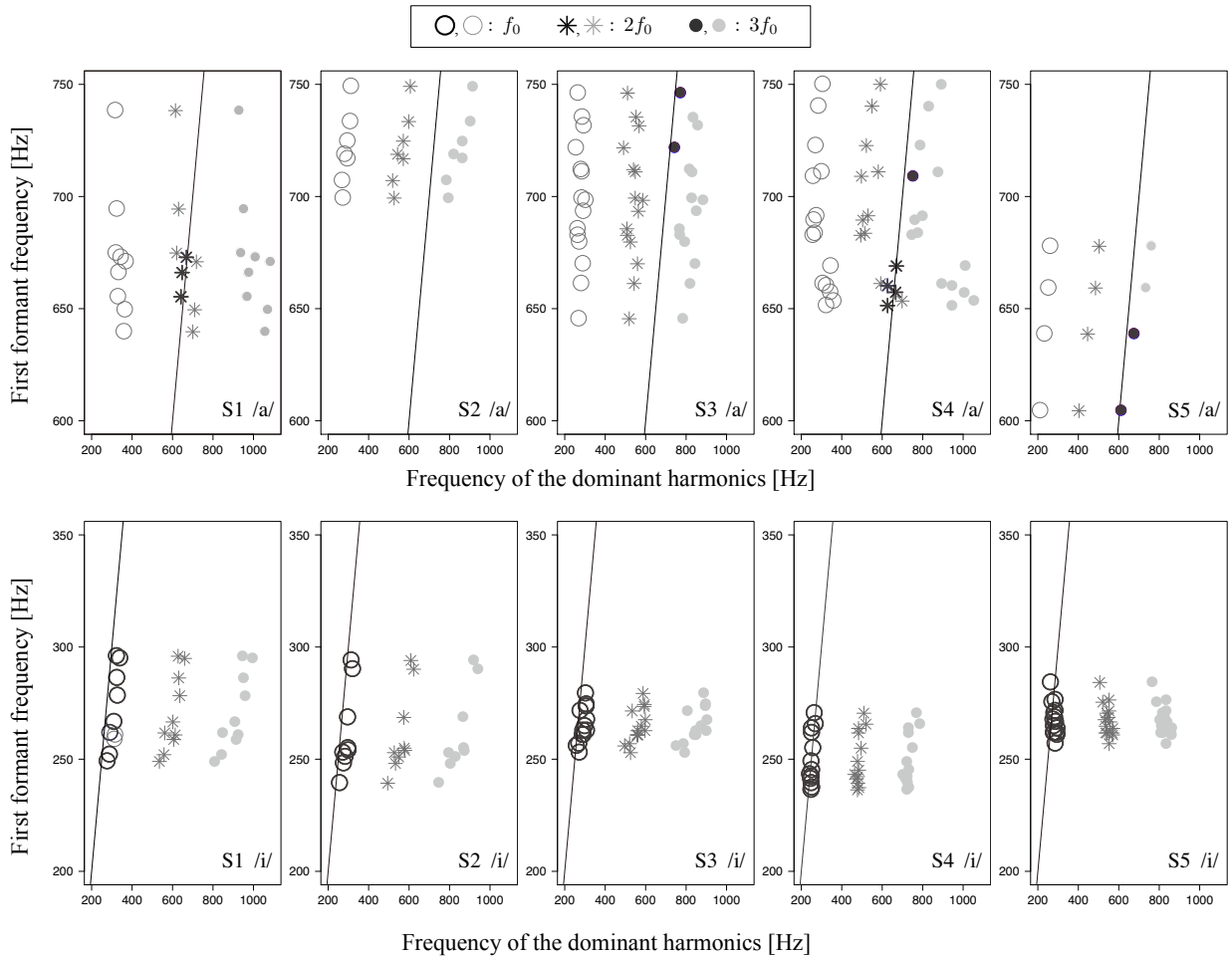


図 4.12: すべての実験参加者での母音/a/および母音/i/での声区転換を生じる際における、基本周波数とその倍音成分、および第1フォルマントの周波数の関係性。横軸は基本周波数もしくは倍音成分の周波数 (Hz)、縦軸は第1フォルマントの周波数 (Hz) を示す。各図中に見られる直線は、基本周波数もしくは倍音成分の周波数と、第1フォルマント周波数が一致する場合を示す。この直線の近傍に見られる濃色のマーカーは、両者の周波数が近接関係にあることを示す。

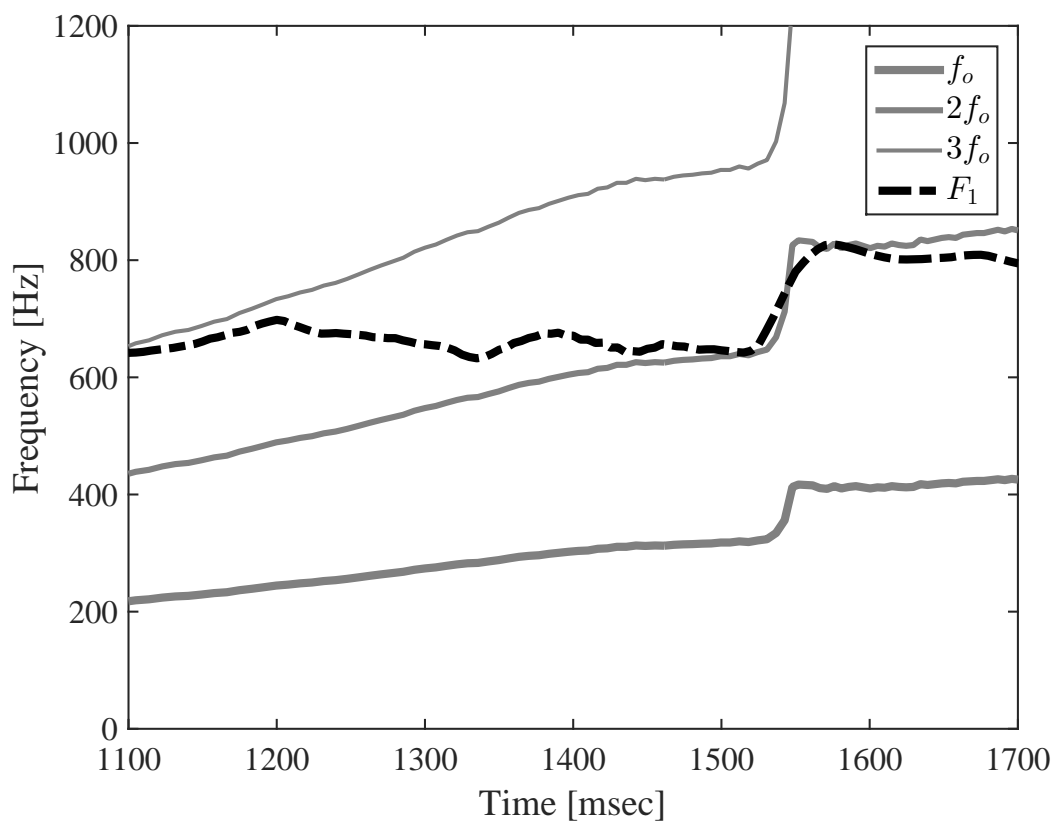


図 4.13: 母音/a/で声区転換を生じる際に、第2倍音が第1フォルマント周波数に近接している場合における、基本周波数と低次の倍音成分、および第1フォルマント周波数の時間変化の分析結果。

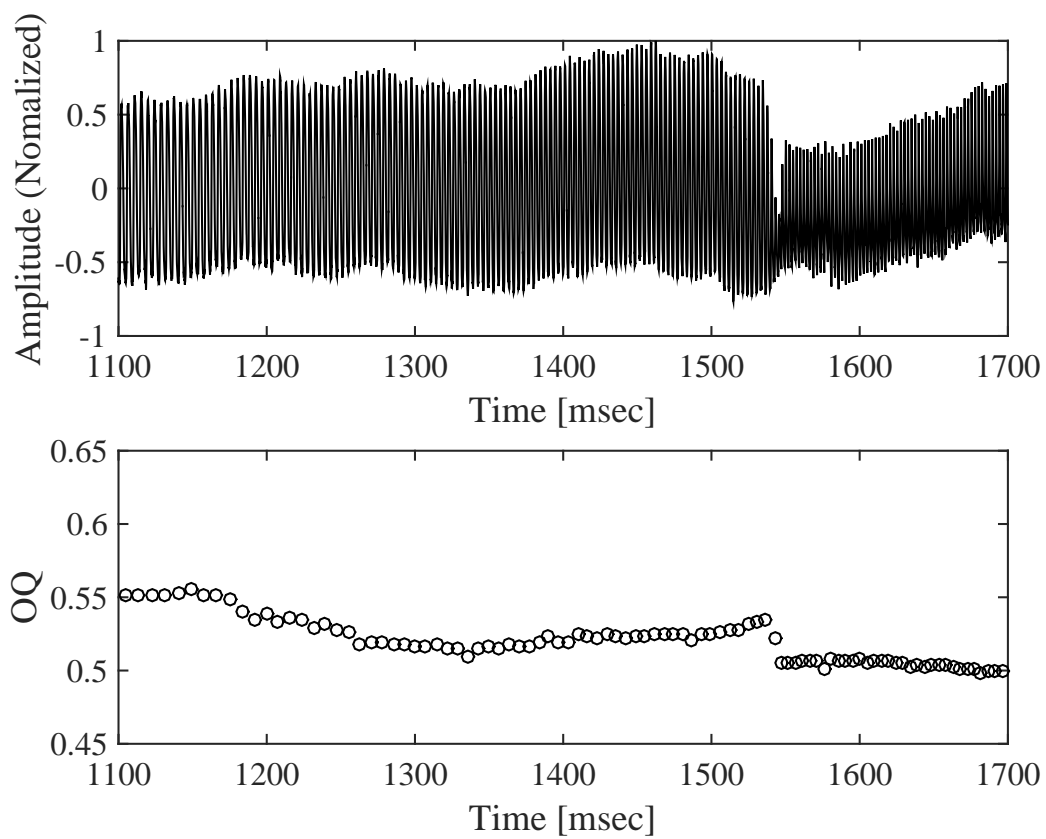


図 4.14: 母音/a/で声区転換を生じる際に、第2倍音が第1フォルマント周波数に近接している場合における、EGG信号の時間波形と得られた声門開放率の時間変化の分析結果。

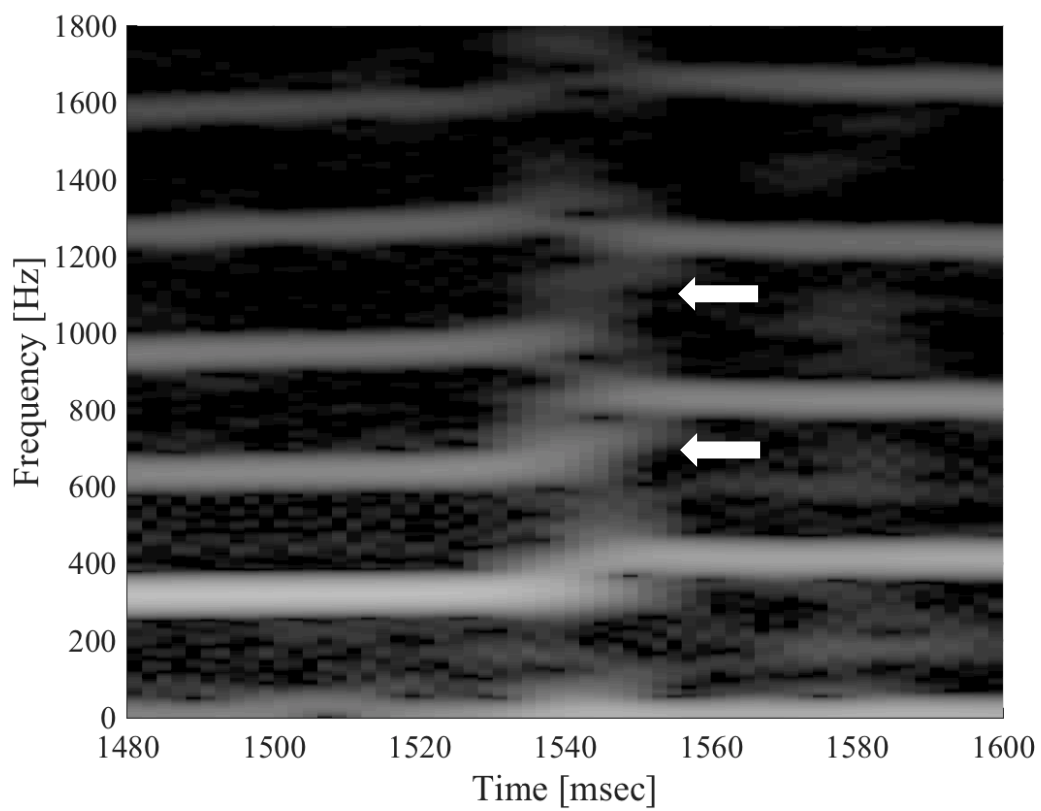


図 4.15: 母音/a/で声区転換を生じる際に、第2倍音が第1フォルマント周波数に近接している場合における、EGG信号のスペクトログラムの声区転換部を拡大した分析結果。

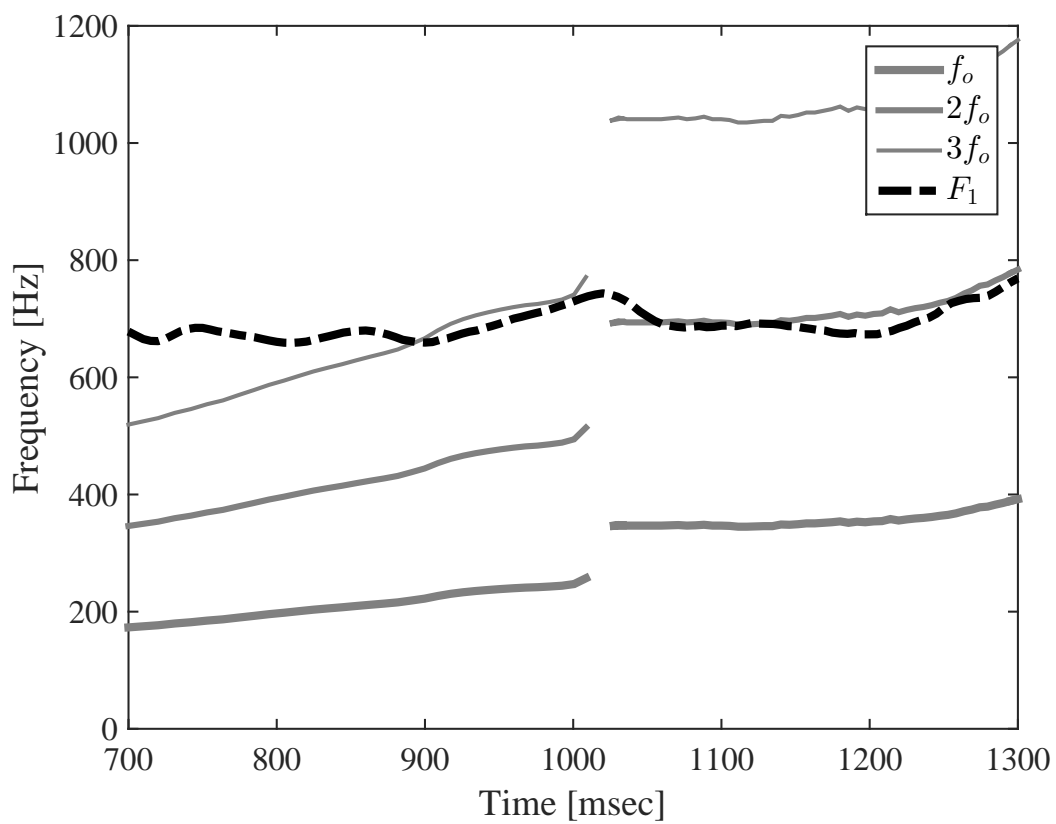


図 4.16: 母音/a/で声区転換を生じる際に、第3倍音が第1フォルマント周波数に近接している場合における、基本周波数と低次の倍音成分、および第1フォルマント周波数の時間変化の分析結果。

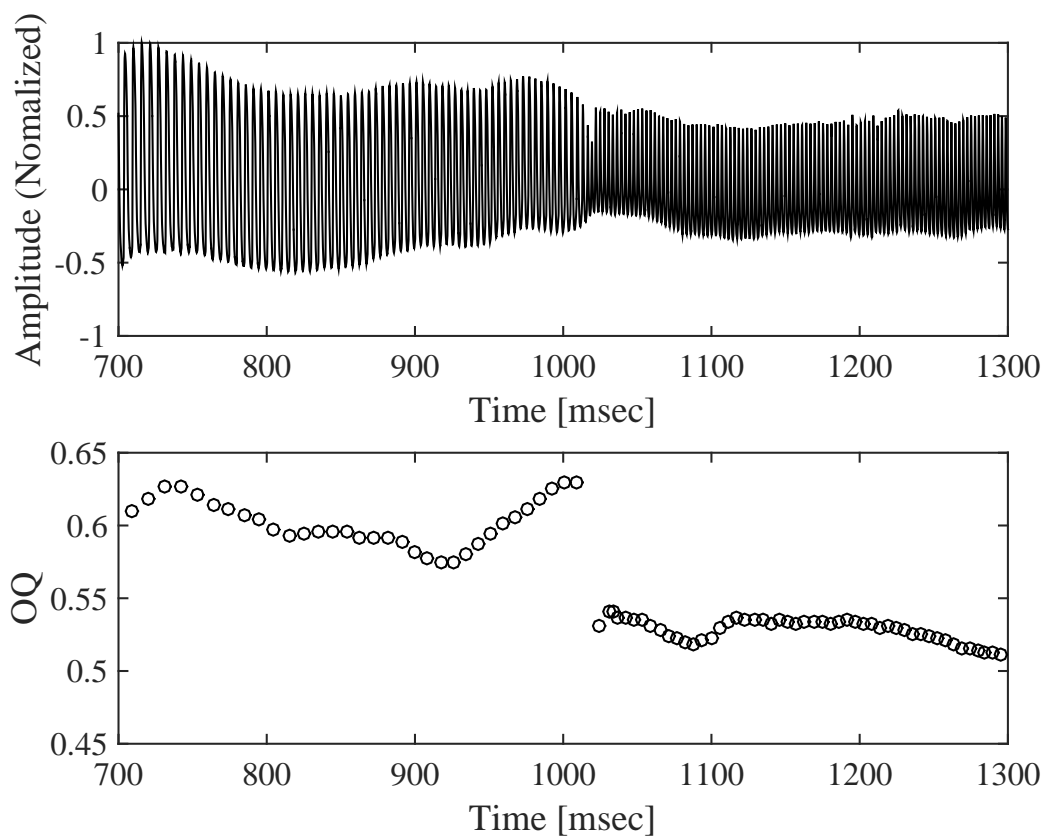


図 4.17: 母音/a/で声区転換を生じる際に、第3倍音が第1フォルマント周波数に近接している場合における、EGG 信号の時間波形と得られた声門開放率の時間変化の分析結果。

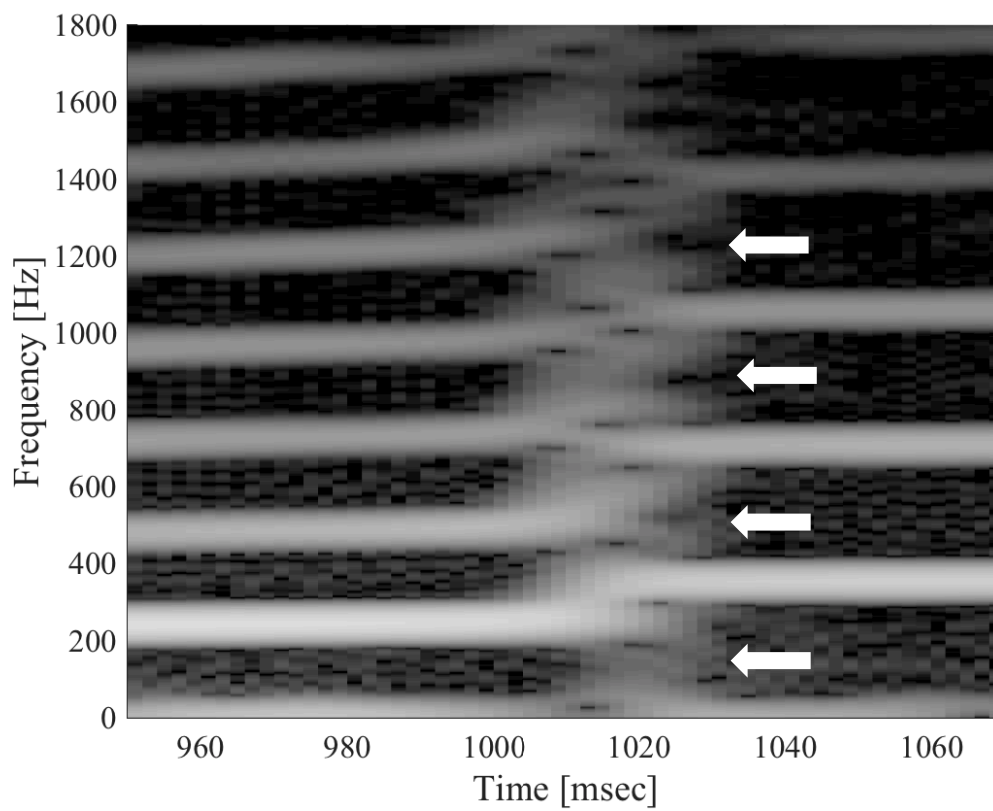


図 4.18: 母音/a/で声区転換を生じる際に、第3倍音が第1フォルマント周波数に近接している場合における、EGG信号のスペクトログラムの声区転換部を拡大した分析結果。

表 4.2: すべての被験者と母音の組み合わせについて、声区転換を生じる直前の第1フォルマント周波数 F_1 (Hz) と基本周波数 f_o^{pre} (Hz)、周波数ジャンプ後の基本周波数 f_o^{post} (Hz)、および周波数ジャンプの幅 Δf_o (Cent) の分析結果における、平均と標準偏差 (S.D.)。また、 F_1 と f_o^{pre} の周波数の関係から、声区転換の要因として考えられるメカニズムを区分している。SI は音源由来、AI は音響由来の声区転換をそれぞれ示す。

要因	被験者と母音	試行回数	F_1 (Hz) 平均 (S.D.)	f_o^{pre} (Hz) 平均 (S.D.)	f_o^{post} (Hz) 平均 (S.D.)	Δf_o (Cent) 平均 (S.D.)
SI	S1 /a/	6	677.9 (32.3)	332.7 (22.1)	415.4 (12.1)	387.3 (82.5)
	S2 /a/	7	721.3 (15.3)	281.1 (15.7)	333.8 (16.7)	297.5 (52.8)
	S3 /a/	13	692.7 (26.1)	271.9 (11.8)	334.7 (12.7)	360.1 (57.6)
	S4 /a/	11	699.5 (30.4)	276.0 (30.4)	364.1 (24.7)	485.1 (96.1)
	S5 /a/	2	668.5 (13.2)	247.0 (6.4)	293.3 (2.9)	297.7 (28.0)
AI	S1 /a/	3	664.8 (8.9)	334.2 (7.1)	414.3 (6.2)	414.0 (12.1)
	S2 /a/	-	- (-)	- (-)	- (-)	- (-)
	S3 /a/	2	733.9 (17.3)	250.3 (6.9)	313.2 (5.1)	388.3 (19.8)
	S4 /a/	4	659.5 (7.3)	323.7 (12.2)	405.5 (5.4)	390.8 (81.1)
	S5 /a/	2	621.6 (24.2)	212.4 (15.1)	265.9 (6.6)	391.2 (79.6)
AI	S1 /i/	10	270.4 (16.4)	301.4 (18.8)	378.0 (21.0)	392.3 (27.6)
	S2 /i/	9	261.4 (17.9)	281.3 (19.3)	340.8 (19.7)	333.7 (97.2)
	S3 /i/	14	264.6 (7.7)	280.9 (16.2)	362.0 (14.0)	440.8 (58.5)
	S4 /i/	13	250.0 (11.9)	242.7 (7.5)	320.6 (12.1)	481.4 (41.3)
	S5 /i/	15	267.5 (7.0)	272.7 (8.7)	329.5 (19.3)	325.7 (75.0)

4.4 考察

得られた分析結果、図 4.1 に示すような基本周波数と倍音成分、フォルマント周波数の関係性と声区転換を引き起こす要因の仮定から、以下のことが考えられる。(1) ほとんどの母音/a/の場合において、声区転換の際には近接関係が見られなかった。このことから、この場合では音源由来のメカニズムが声区転換の主要な要因である。(2) ほぼ全ての母音/i/の場合と、一部の母音/a/の場合において、声区転換の際に基本周波数もしくは低次の倍音成分が、第1フォルマント周波数に近接していた。このことから、この場合では音響由来のメカニズムが声区転換の主要な要因である。この一部の母音/a/における実験結果は、音源-フィルタ相互作用の影響が、基本周波数が第1フォルマント周波数に近接する場合だけでなく、低次の倍音成分が第1フォルマント周波数に近接する場合でも強まること、そして声区転換を引き起こす可能性を

持っていることを示している。このことは、先行研究による示唆にも一致している [3]。

表 4.2 は、図 4.12 に示すデータについて、声区転換の開始点における第 1 フォルマント周波数 (Hz) と基本周波数 (Hz)、周波数ジャンプ後の基本周波数 (Hz)、周波数ジャンプの幅 (Cent) の平均値と標準偏差を各実験参加者と母音の種類ごとにまとめたものである。また、上述した声区転換の要因ごとに分類している。ここで、周波数ジャンプの値は Cent 値で算出した。これはある周波数の基準の周波数に対する対数比であり、式 (4.1) によって求められる。なお、ここでは周波数ジャンプ前とジャンプ後の基本周波数の値を用いている。

$$1200 \log_2(f_o^{post} / f_o^{pre}) \quad (4.1)$$

第 1 フォルマント周波数の平均値は、母音/a/でおよそ 620 Hz から 740 Hz、母音/i/で 250 Hz から 270 Hz であった。声区転換を生じる直前の基本周波数の平均値は、すべての母音で 210 Hz から 340 Hz であった。これらの値は実験参加者により若干の個人差が見られた。母音/i/において、声区転換前の基本周波数は、第 1 フォルマント周波数よりも僅かに高い値を示した。また興味深いことに、周波数ジャンプ幅に着目すると、母音/a/の音源由来のメカニズムによる声区転換の場合と比べて、音響由来のメカニズムによる声区転換のほうが、より高い値を示す傾向が見られた。そこで、周波数ジャンプ幅について、音源由来のメカニズムと音響由来の場合とでの比較検討を行った。Titze はシミュレーション実験により、epilarynx tube が狭くなると、音源-フィルタ相互作用がより大きな周波数ジャンプを引き起こすことを示唆した [3]。しかしながら、周波数ジャンプの幅という観点において、音源由来のメカニズムと音響由来のメカニズムとでは差異があるかどうかは明らかにされていない。ここでは、母音/i/の音響由来のメカニズムによる周波数ジャンプの幅が、母音/a/の音源由来のメカニズムによるものと異なるかどうかを調べた。周波数ジャンプの幅について、独立したサンプルの t 検定を行ったところ、母音/i/の音響由来のメカニズムによる周波数ジャンプの幅 (平均：412.56 Cent、標準偏差：117.16 Cent) と母音/a/の音源由来のメカニズムによる周波数ジャンプの幅 (平均：380.83 Cent、標準偏差：97.40 Cent) とで有意差は見られなかった。なお、t(100) は -1.41、p 値は 0.163 であった。

声区転換は音源由来のメカニズムだけでなく、音響由来のメカニズムによっても生じることを発声実測実験でも確かめることができた。音響由来のメカニズムによる声区転換は、声区転

換を生じる直前に基本周波数が第1フォルマント周波数に近接した結果、音源-フィルタ相互作用の影響が強まったことによるものであると考えられ、これは計算機によるシミュレーション実験の結果と一致した。また、声区転換時に伴う周波数ジャンプの幅については、メカニズムの違いによる影響はあまり存在しないことがわかった。

4.5 まとめ

本研究では、声区転換を引き起こす要因として考えられる2つのメカニズムについて、発声測定実験を行い検討した。声門の開閉パターンと声道の音響特性を同時測定し、声区転換時における発声の基本周波数と低次の倍音成分、第1フォルマント周波数の関係性を調べた。声門の開閉パターンは電気喉頭計によってEGG信号として計測し、基本周波数と声門開放率はEGG信号にDECOM法を適用することで分析した。一方、声区転換のような基本周波数が比較的高い場合においても、一般的な信号処理法と比べてより精確にフォルマント周波数を測定するために、声道共鳴は、外部音響励振法を適用することで計測した。また、励振信号による応答から実験参加者の自発発声成分を取り除くために、ケプストラム分析を適用することでフォルマントの抽出を行った。得られた基本周波数と倍音成分、および第1フォルマント周波数の関係性を調べることで、声区転換と周波数ジャンプを生じる要因を決定した。

測定と分析を通して、異なる2種類の声区転換が存在することを確認した。一つは母音/a/のほとんどの場合のように、声区転換を生じる際に基本周波数と第1フォルマント周波数があまり近接していない場合である。この場合では、喉頭調節による声帯振動の不安定さ、すなわち音源由来のメカニズムが主な要因として声区転換を引き起こしていると考えられる。もう一つは母音/i/の場合のように、声区転換を生じる際に基本周波数と第1フォルマント周波数がしばしば近接している場合である。この場合では、音源-フィルタ相互作用の影響、すなわち音響由来のメカニズムが主な要因として声区転換を引き起こしていると考えられる。このような声区転換は一部の母音/a/の場合でも見られ、声区転換を生じる際に第2倍音や第3倍音が第1フォルマント周波数に近接した場合でも、音響由来のメカニズムによる声区転換を生じることがあることを示した。声区転換の要因と周波数ジャンプ幅の大きさに関しては、メカニズムによる大きな違いは確認されなかった。

第5章 発声条件がフォルマントに与える影響の検討

5.1 背景と目的

母音などの有声音は、声帯振動により生じた音源信号に、声道の音響特性が付与されることで生成される。一般に、声道音響特性は声道の形状によって決定されるが、MRIによる声道形状の観測によって、地声声区と裏声声区での声道形状の変化はわずかに存在することが示されている [24]。一方で、地声声区と裏声声区では発声時における声門の開口面積が異なることが知られており、一般に地声声区の方が大きいとされる。Hoppe [25] による内視鏡を用いた声帯振動の観測では、発声時における最大声門面積は、地声声区では 32mm^2 、裏声声区では 20mm^2 であった。また、声門開放率は、地声声区よりも裏声声区の方が大きいことが知られている [26]。このように声区の違いは、声道のわずかな形状の違いや、声門側の境界条件の違いをもたらす、結果として声道音響特性に差異をもたらすと考えられる。そこで本研究では、発声条件の違いが声道音響特性に与える影響について検討する。異なる発声条件による持続発声を行った際の声道音響特性を測定し、フォルマント周波数を分析する。発声条件によって、フォルマント周波数にどのような違いが生じるかを調べた上で、そのような違いをもたらす要因について検討を行う。

5.2 実験

発声測定実験は防音室で行い、被験者として成人男性3名を用いた。本実験を行う前に、地声声区と裏声声区の両声区で発声することが可能な音域であるオーバーラップレンジを、各被験者ごとに測定した。はじめに、地声声区で発声しやすい典型的な音高で母音/a/を持続発声してもらい、音高を確認した。その音高を開始点として、半音だけ上昇させた音を被験者に呈示し、母音/a/で地声発声できるか確認した。これを繰り返し、地声声区で発声可能な上限

の音高を確認し、オーバーラップレンジの最高音とした。続いて、裏声声区で発声しやすい典型的な音高で母音/a/を持続発声してもらい、音高を確認した。その音高を開始点として、半音だけ下降させた音を被験者に呈示し、母音/a/で裏声発声できるか確認した。これを繰り返して、裏声声区で発声可能な下限の音高を確認し、オーバーラップレンジの最低音とした。測定したオーバーラップ域において、地声声区と裏声声区の両声区で最も発声しやすい音高をターゲット音とした。各被験者の声域データを表 5.1 に示す。

本実験では、各被験者に異なる 4 つの発声条件で母音を持続発声してもらった。発声時の声道音響特性は、外部音響励振法 [17-20] を用いて測定した。本手法を用いることで、基本周波数によらず、また、非発声時でも声道音響特性を測定できる。同時に、EGG を用いて声帯振動パターンを計測し、DECOM 法を用いることで、EGG 信号から基本周波数と声門開放率を得た。実験で測定した母音は、日本語 5 母音/a/,/i/,/u/,/e/,/o/とした。各母音を、以下に示す 4 つの条件で発声してもらい、その際の声道音響特性を測定した。

- 地声声区：ターゲット音の音高で、指定の母音を地声声区で持続発声する
- 裏声声区：ターゲット音の音高で、指定の母音を裏声声区で持続発声する
- 声門閉鎖：ターゲット音の音高で、指定の母音を持続発声した直後に、続けて呼吸を止めて声帯を完全に閉鎖する
- 声門解放：ターゲット音の音高で、指定の母音を持続発声した直後に、続けてささやき声で声帯を開放する

これらの発声条件に対して想定される声門の状態を、表 5.2 に示す。このうち、声門閉鎖の条件では、声門が完全に閉鎖していると考えられることから、4 つの発声条件の中では声門

表 5.1: 表各被験者の声域データ。地声声区と裏声声区で発声可能な声域を特定した後、オーバーラップレンジを求めた。これより、本実験で発声を行ってもらう音高をターゲット音として決定した。

	被験者 1	被験者 2	被験者 3
地声の声域	E1-G3	E1-F3	D1-A3
裏声の声域	G2-G4	G2-E4	A2-A4
オーバーラップレンジ	G2-G3	G2-F3	A2-A3
ターゲット音	E3	D3	E3

面積が最小となり、声門開放率は最小値である 0 を示すと想定される。一方で、声門開放の条件では、声門が完全に開放していると考えられることから、4 つの発声条件の中では声門面積が最大となり、声門開放率は最大値である 1 を示すと想定される。

5.3 結果と考察

各母音について、それぞれの条件における第 1 フォルマント周波数を比較した。図 5.1 に、得られた地声声区と裏声声区における第 1 フォルマント周波数の関係を示す。これより、地声声区と裏声声区とでは、地声声区のほうが第 1 フォルマントが高い傾向を示した。また、声門開放と声門閉鎖の条件で比較すると、声門開放のほうが第 1 フォルマント周波数が高い傾向を示した。表 5.3 に、第 1 フォルマント周波数および第 2 フォルマント周波数の差の平均値を、「地声声区-裏声声区」と「声門開放-声門閉鎖」とで求めた結果を示す。いずれの差においても、 $p=0.05$ の t 検定において有意差が認められた。

表 5.3 より、第 1 フォルマント周波数、第 2 フォルマント周波数いずれにおいても、「地声声区-裏声声区」よりも「声門開放-声門閉鎖」の方が周波数差が大きい。これらのことより、発声条件の差異が低次のフォルマント周波数に影響を与えられと考えられる。北村らは、MRI データを用いて声道断面積関数を抽出し、声門閉鎖時と声門開放時の声道伝達特性を求め、低次フォルマントが声門開放時に高周波数方向にシフトする傾向があることを示した [27]。本実験でも同様の結果が得られ、発声時の最大声門面積が大きいほど低次のフォルマント周波数が高くなる可能性が示唆された。

表 5.2: 本実験で行った発声測定実験における 4 つの発声条件。具体的な発声方法と、発声条件によって想定される声門の状態 (声門面積と声門開放率)。

発声条件	発声方法	声門面積	声門開放率
地声声区	ターゲット音の音高で指定母音を持続発声	大	小
裏声声区	ターゲット音の音高で指定母音を裏声声区で持続発声	小	大
声門閉鎖	ターゲット音で指定母音を持続発声した直後、呼吸を止める	最小	最小 (=0)
声門開放	ターゲット音で指定母音を持続発声した直後、ささやき声で発声	最大	最大 (=1)

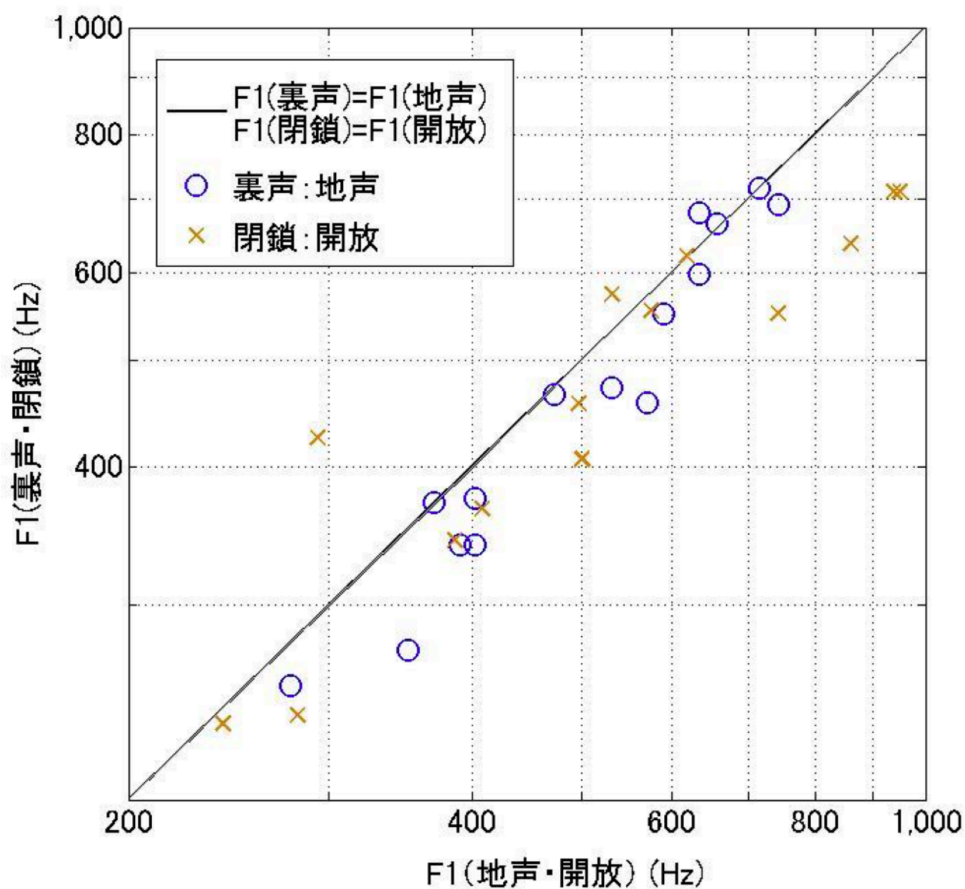


図 5.1: 日本語 5 母音について、4つの発声条件における第 1 フォルマントの分布図。横軸は地声声区と声門開放の発声条件における第 1 フォルマント周波数 (Hz) を、縦軸は裏声声区と声門閉鎖の発声条件における第 1 フォルマント周波数 (Hz) をそれぞれ示す。図中の○は、地声声区 - 裏声声区が発声条件における第 1 フォルマント周波数を、× は、声門開放 - 声門閉鎖の発声条件における第 1 フォルマント周波数をそれぞれ示す。

表 5.3: 第 1 フォルマント周波数および第 2 フォルマント周波数の差の平均値を、「地声声区-裏声声区」と「声門開放-声門閉鎖」とで求めた結果。いずれの差においても、 $p=0.05$ の t 検定において有意差が認められた。

	F1 の差の平均 (Hz)	F2 の差の平均 (Hz)
地声声区 - 裏声声区	32.8	76.7
声門開放 - 声門閉鎖	73.6	117.2

次に、声門閉鎖、地声声区、裏声声区における第1フォルマント周波数、第2フォルマント周波数のQ値を図5.2に示す。低次のフォルマントに関して、声門閉鎖、地声声区では鋭いピークが形成されているのに対し、裏声声区ではQ値が減少し、ピークがなだらかなる傾向が見られた。特に、母音/a/と/o/に関しては、声門開放時に第2フォルマント周波数が消失することもあった。今回の実験では、声門開放率の平均値は地声声区で0.38、裏声声区で0.71であり、裏声声区の方が声門開放率が高い。声門が開いているときは、声門下に流出する音響エネルギーが多くなり、結果としてピークが減衰すると考えられる。したがって、声門開放率が高く、声帯の振動周期に対して声門が開いている時間の割合が高い発声法ほど、ピークがなだらかなる傾向がみられたと考えられる。

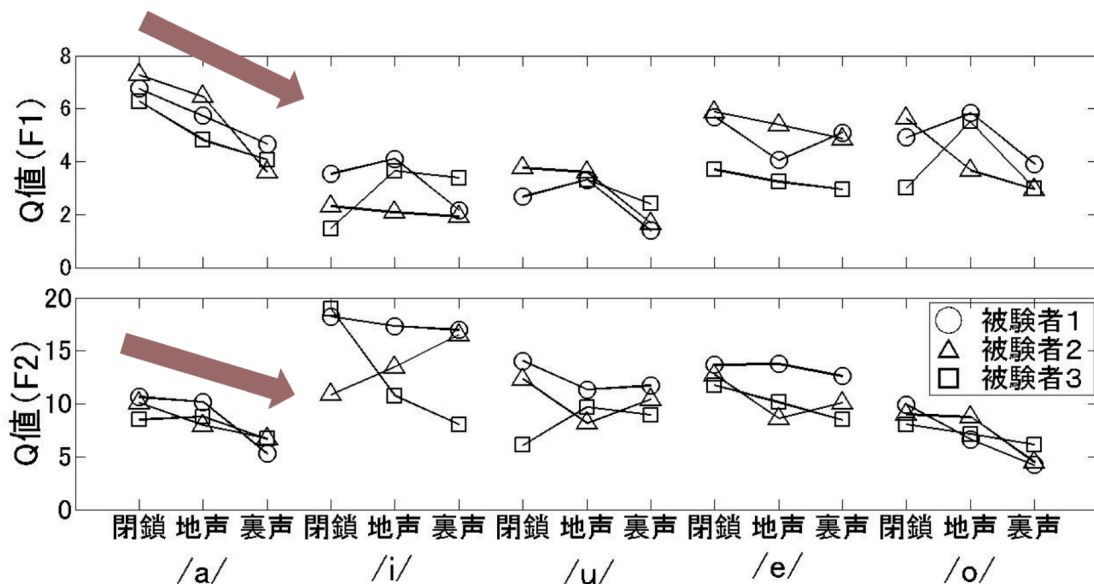


図 5.2: 日本語5母音について、各発声条件における第1フォルマント周波数と第2フォルマント周波数のQ値。縦軸は各フォルマント周波数のQ値を示す。○は実験参加者1、△は実験参加者2、□は実験参加者3の結果をそれぞれ示す。

発声法の違いにより、高次フォルマントでは特に第4フォルマントに特徴的な差異がみられた。特定の母音を除く第4フォルマントでは、声門閉鎖、裏声声区、地声声区の順に、周波数が上昇しピークが減衰した。その様子を図5.3に示す。この現象は、被験者1～3における母音/a/, /u/, /o/, 被験者1, 2における母音/e/で観察された。第4フォルマントの生成には、

喉頭腔の共鳴が密接に関係していることが知られている。竹本らはMRIデータを元に、声門インピーダンスを考慮した縦続音響管モデルを用いて声道伝達関数を算出し、声門面積が喉頭腔共鳴に与える影響を検討した [28]。その結果、声門面積が大きくなるにつれ喉頭腔の共鳴周波数が上昇し共鳴が弱くなり、喉頭腔共鳴が誘導する第4フォルマントは周波数を上昇させながらピークが減衰することが示された。

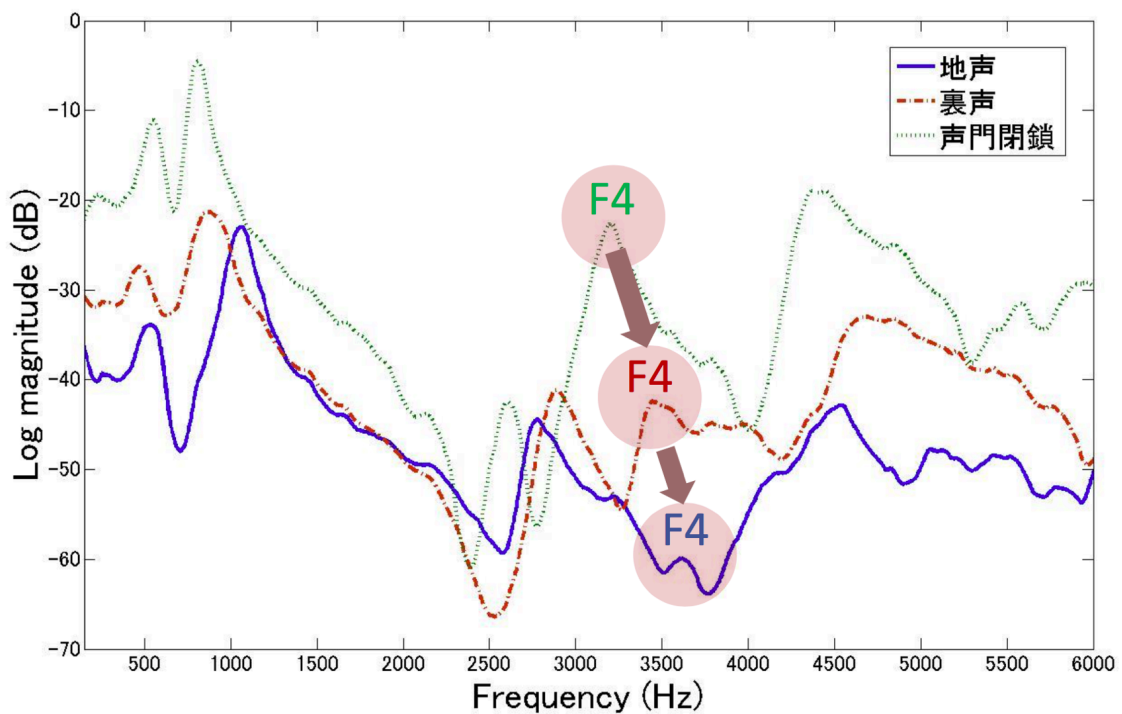


図 5.3: 実験参加者 3 による発声条件の異なる母音/o/ 発声時における声道音響特性の分析結果。横軸は周波数を、縦軸は対数振幅をそれぞれ示す。第4フォルマントに着目すると、声門閉鎖、裏声声区、地声声区が発声条件につれて、フォルマント周波数が上昇するとともに、ピークが減衰するような傾向が見られる。

本実験においても、声門閉鎖、裏声声区、地声声区と想定される声門面積が大きくなるにつれ、第4フォルマントが上昇しながら減衰する傾向がみられた。各被験者の各母音における第4フォルマントを、表 5.4 に示す。上記の現象が観測された被験者 1～3 の母音/a/, /u/, /o/, 被験者 1,2 の母音/e/における第4フォルマントは、声門閉鎖時におよそ 3300～3700 Hz 辺りに明確なピークを形成していた。これらのピークは、喉頭腔共鳴によって誘導された第4フォルマントであると考えられる。一方、上記の現象がみられなかった被験者 1～3 の母音/i/, 被

表 5.4: 各被験者の各母音における第4フォルマント周波数。想定される声門面積が次第に大きくなるにつれて、第4フォルマントが高域に上昇しつつ、ピークが減衰するような傾向を示していたものに、*を付してある。

	被験者 1		被験者 2		被験者 3	
	周波数 (Hz)	S.D. (Hz)	周波数 (Hz)	S.D. (Hz)	周波数 (Hz)	S.D. (Hz)
/a/	* 3293	22.1	* 3645	24.5	* 3359	-
/i/	3707	-	3910	-	4297	12.2
/u/	* 3355	32.0	* 3426	24.3	* 3273	20.5
/e/	* 3547	25.9	* 3715	21.2	4008	6.4
/o/	* 3313	30.4	* 3414	31.0	* 3203	24.1

験者3の母音/u/では、第4フォルマントの周波数が他の第4フォルマントよりも高く、Q値が小さい。したがって、これらの第4フォルマントは喉頭腔共鳴によって誘導されたのではなく、別の要因によって形成された可能性が考えられる。

5.4 まとめ

外部音響励振法を用いて、発声条件の違いが声道音響特性に与える影響について検討した。第1フォルマント周波数、第2フォルマント周波数について、想定される声門面積が大きいほどピークの周波数は上昇し、また声門開放率が大きいほどピークはなだらかになる傾向がみられた。また、高次フォルマントについて、喉頭腔共鳴によって誘導されたと考えられる第4フォルマントは、想定される声門面積が上昇するにつれて周波数が上昇しながら減衰した。したがって、声区の違いに伴う声道音響特性の変化は、声道形状の違いだけではなく、声門面積や声門開放率といった発声条件の違いによっても生じる可能性が示された。

第6章 声門の境界条件がフォルマントに与える影響 の検討

6.1 背景と目的

音声生成の音源-フィルタ理論では、音源機構と声道フィルタが独立であると仮定されている。また、声道の末端である声門の境界条件として、声門の完全閉鎖が仮定されている。しかし実際の発声では、呼気流による声帯の準周期的な自励振動に伴って声門が開閉するため、声門の境界条件も時間とともに変化する。また、地声や裏声などの発声条件によって、声門の開閉パターンも様々に変化しうる。この声門の境界条件の変化は、声道と声門下部のカップリングの程度にも影響を与える。このように実際の発声では、厳密には声門閉鎖の仮定は成立せず、声門や声門下部の音響的な特徴が、声道のフォルマントに影響を与えたと考えられる。

声門の境界条件とフォルマントの関係を調べた先行研究として、Barney ら [30] は、周期的に開閉する声門模型とアクリル製の一樣矩形管を組み合わせた模型実験を行った。声門の開閉面積を固定した場合では、声門の開閉面積が大きくなるに従って、第1フォルマントと第2フォルマントが上昇することがわかった。また、最大声門面積と声門開放率が大きくなると、第1フォルマントが上昇することがわかった。

また、Barney らの実験では、声門の境界条件は動的に変化させることができるが、音源信号は定常的に与えられ、声道模型は常に音響駆動された。しかし実際の発声では、声道を駆動する音圧源は声門流量の時間変化によって生じる。特に地声の発声において、声門流量は声門閉小期において急激に減少するため、声道は声門閉鎖の直前の時点において、瞬時的に駆動されると考えられる。このように、Barney らの実験は、実際の声道の駆動状態を必ずしも正しく模擬したものとはなっておらず、声道の境界条件として声門が開くことの影響が過大に評価された可能性がある。

本研究では、周期的な声門の開閉を考慮した音声生成モデルによる音声合成シミュレーショ

ンと、合成音の分析を通して、声門の境界条件がフォルマントに与える影響について検討する。はじめに、声道・声門・声門下部を連結した音響管モデルと、声門面積の時間変化を表す多項式モデルを組み合わせた音声生成モデルについて述べる。次に、本モデルを用いて、声門開放率や最大声門面積などの発声パラメータを様々に変化させた音声を合成する。これらの合成音声からフォルマントを分析し比較することで、声門の境界条件とフォルマントの関係について考察する。

6.2 音声生成の物理モデル

6.2.1 声門面積の多項式モデル

本研究では、音源部を簡易に表すため、声門面積の時間波形を表す多項式モデルを用いる。声門面積の時間波形 G_A と、声門体積流の時間波形 U_g は比例関係にあることから、Rosenberg による声門体積流の多項式モデル [31] を用いて、声門面積の時間波形 G_A と声門体積流の時間波形 U_g は式 (6.1) および式 (6.2) のようにそれぞれ表される。

$$G_A = \alpha(t^2 - t^3)G_M \quad (0 \leq t \leq 1) \quad (6.1)$$

$$U_g = G_A \sqrt{2P_0/\rho} \quad (6.2)$$

ここで、 G_M は最大声門面積、 $\alpha(= 27/4)$ は正規化のための係数、 P_0, ρ は肺圧および空気の密度である。なお、声門面積の時間波形 G_A の時間軸は相対的なものであり、実際にはこれをスケーリングして合成に用いる。得られる声門面積の時間波形 G_A の例を図 6.1 に示す。音声合成シミュレーションの際には、最大声門面積 G_M と声門開放率を発声パラメータとして与える。声門開放率を与えることで、声門開閉の基本周期に対する開放期の割合を決定することができる。

声門面積 G_A の時間波形

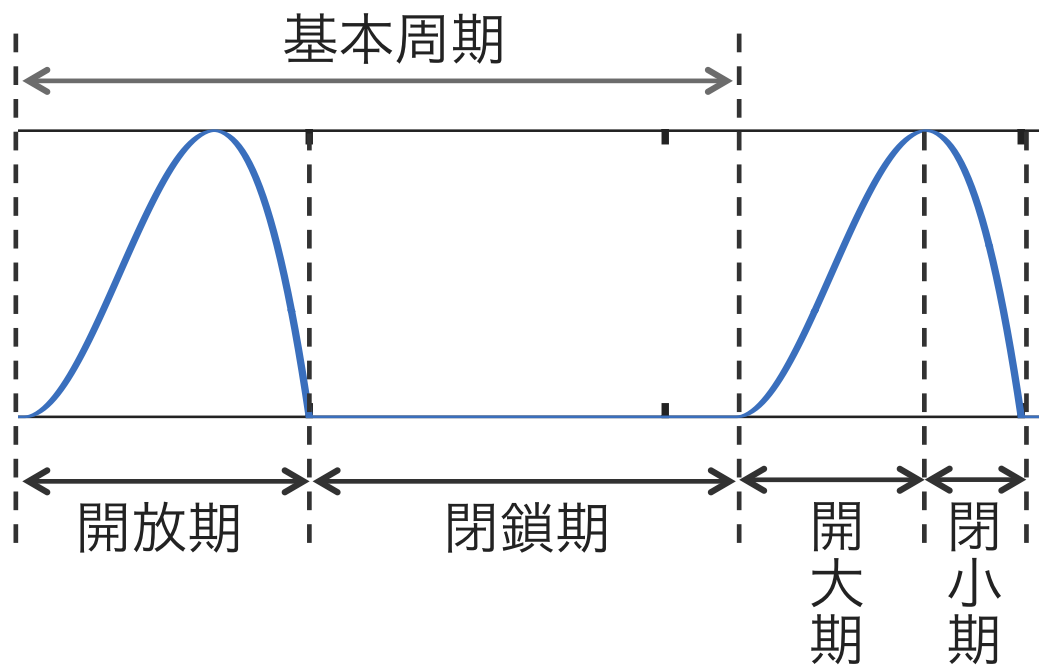


図 6.1: 声門面積の多項式モデルによって計算した声門面積の時間波形 G_A の例。なお、 G_A の時間軸は相対的なものであり、実際にはこれをスケールリングして合成に用いる。

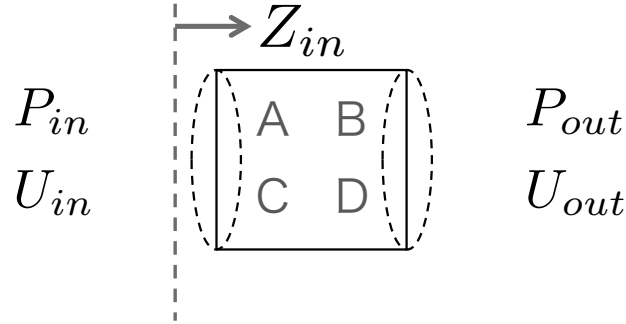


図 6.2: 単一円筒管の音響管モデルの概略図。 P_{in} と P_{out} はそれぞれ入力音圧と出力音圧を、 U_{in} と U_{out} はそれぞれ入力流量と出力流量を示す。また、 Z_{in} は入力インピーダンスを示す。

6.2.2 声道・声門・声門下部の連結音響管モデル

声道の物理モデルとして、Sondhi らの音響管モデル [32] を導入する。はじめに、単一の円筒管における音響管モデルについて説明する。図 6.2 に、単一円筒管の音響管モデルの概略図を示す。 P_{in} と P_{out} は音圧の入出力を、 U_{in} と U_{out} は体積速度の入出力をそれぞれ示す。ここで、円筒管の音圧と体積速度の入出力の関係は、式 (6.3) のように伝搬行列を用いて表すことができる。またその伝搬係数は、式 (6.4) によって表される。 L_g, S_g は円筒管の長さと同面積、 ρ, c は空気の密度と音速、 σ, γ は周波数依存の係数 [32] である。さらに、円筒管の入力インピーダンス Z_{in} および伝達関数 H は、式 (6.5) および式 (6.6) によって求められる。

$$\begin{pmatrix} P_{out} \\ U_{out} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} P_{in} \\ U_{in} \end{pmatrix} \quad (6.3)$$

$$\begin{aligned} A &= \cosh(\sigma L_g/c) \\ B &= -(\rho c/S_g)\gamma(\sinh(\sigma L_g/c)) \\ C &= -(S_g/\rho c)(\sinh(\sigma L_g/c))/\gamma \\ D &= \cosh(\sigma L_g/c) \end{aligned} \quad (6.4)$$

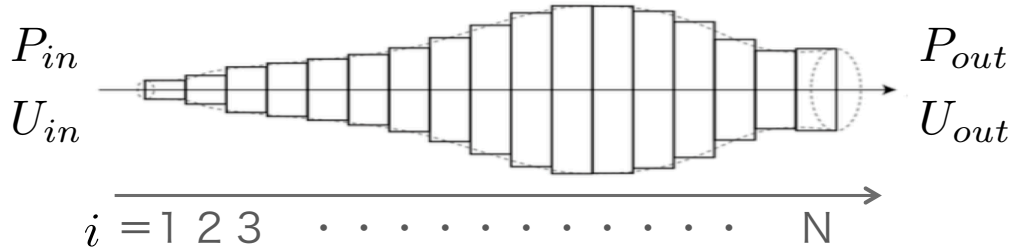


図 6.3: 声道音響管モデルの概略図。断面積の異なる単一円筒管を連結させてモデル化する。インデックス i は声門側からの円筒管のセクション数 (1, 2, ... N) を示す。 P_{in} と P_{out} はそれぞれ入力音圧と出力音圧を、 U_{in} と U_{out} はそれぞれ入力流量と出力流量を示す。また、 Z_{in} は入力インピーダンスを示す。

$$Z_{in} = \frac{P_{in}}{U_{in}} \quad (6.5)$$

$$H = \frac{U_{out}}{U_{in}} \quad (6.6)$$

音響管モデルにおいて、声道は断面積の異なる複数の円筒管を連結した多管近似によって表される (図 6.3)。このとき、声道における音圧と体積速度の入出力の関係は式 (6.3) と同様に表されるが、その伝搬行列および伝搬係数は式 (6.7) と式 (6.8) によって表される。

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \prod_{i=1}^N \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix} \quad (6.7)$$

$$\begin{aligned} A_i &= \cosh(\sigma L_{gi}/c) \\ B_i &= -(\rho c/S_{gi})\gamma(\sinh(\sigma L_{gi}/c)) \\ C_i &= -(S_{gi}/\rho c)(\sinh(\sigma L_{gi}/c))/\gamma \\ D_i &= \cosh(\sigma L_{gi}/c) \end{aligned} \quad (6.8)$$

ここで、 i は声道のセクションを表すインデックスである。音圧と体積速度の入出力の関係と、式 (6.5) および式 (6.6) から、声道入力インピーダンス Z_{in} と伝達関数 H が求められる。

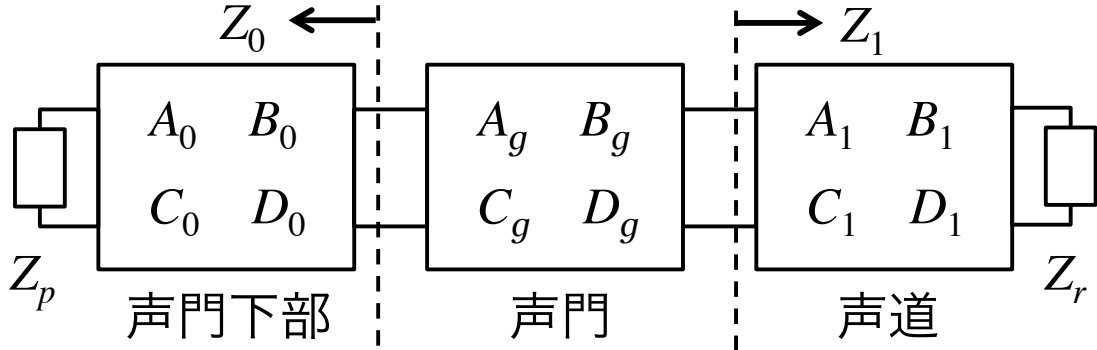


図 6.4: 声道・声門・声門下部を連結した音響管モデル [12] の概略図。 A, B, C, D は各区間における伝搬係数、 Z_1, Z_0 は声道、声門下部の入力インピーダンス、 Z_r は口唇の放射インピーダンス、 Z_p は声門下部の終端インピーダンスをそれぞれ示す。

本研究では、このような声道音響管モデルに基づいた、声道・声門・声門下部の連結音響管モデル [12] を用いる。以下に、本モデルについて説明する。図 6.4 に、連結音響管モデルの概略図を示す。ここで、 A_0, B_0, C_0, D_0 は声門下部の、 A_1, B_1, C_1, D_1 は声道の、 A_g, B_g, C_g, D_g は声門の伝搬係数を示す。また、 Z_0 は声門下部の、 Z_1 は声道の入力インピーダンスを示す。 Z_r は口唇の放射インピーダンス [5]、 Z_p は声門下部の終端インピーダンス [34] である。各入力インピーダンス Z_1, Z_0, Z_g および声道の伝達関数 H_1 はそれぞれ、式 (6.9)、式 (6.10)、式 (6.11)、および式 (6.12) のように求められる。

$$Z_0 = -\frac{A_0 Z_p + B_0}{C_0 Z_p + D_0} \quad (6.9)$$

$$Z_1 = \frac{D_1 Z_r - B_1}{A_1 - C_1 Z_r} \quad (6.10)$$

$$Z_g = -\frac{B_g - A_g Z_0}{D_g - C_g Z_0} \quad (6.11)$$

$$H_1 = \frac{1}{A_1 - C_1 Z_r} \quad (6.12)$$

最終的に、声道・声門・声門下部を連結した音響管モデルによる伝達関数 H_{vt} は、式 (6.13) に従って計算することができる。

$$H_{vt} = \frac{H_1 Z_g}{Z_1 + Z_g} \quad (6.13)$$

6.2.3 音声合成シミュレーション

ここでは、2.1 節および 2.2 節で述べたモデルを組み合わせた、音声合成のシミュレーション手順について述べる。図 6.5 に、そのフローチャートを示す。

まず、声門面積の時間波形 G_A 、声門下部の入力インピーダンス Z_0 、声道の入力インピーダンス Z_1 、声道の伝達関数 H_1 を計算する。ここで G_A は、発声パラメータ (ピッチ f_0 、声門開放率 OQ 、最大声門面積 G_M) を与えて求める。また Z_1 および H_1 は、声道断面積関数より算出する。

次に、以下の計算を時刻を更新して繰り返し行う。 G_A から各時刻 t における声門面積を計算し、声門体積流 U_g と声門の入力インピーダンス Z_g を求める。次に、 Z_g と事前に求めた Z_0, Z_1, H_1 から、声道・声門・声門下部を連結した音響管モデルの伝達関数 H_{vt} を求める。さらに、 H_{vt} のインパルス応答と U_g をたたみこむことで、口唇の体積流を求める。

最後に、上記の繰り返し計算で求めた口唇体積流に、伝達インピーダンスとして微分フィルタをかけることで、合成音声を得る。この音声合成シミュレーションにおいて、声道の断面積関数と、発声パラメータとして基本周波数、声門開放率、最大声門面積を与えることで、任意の母音で多様な声門境界条件をもつ音声を合成することができる。

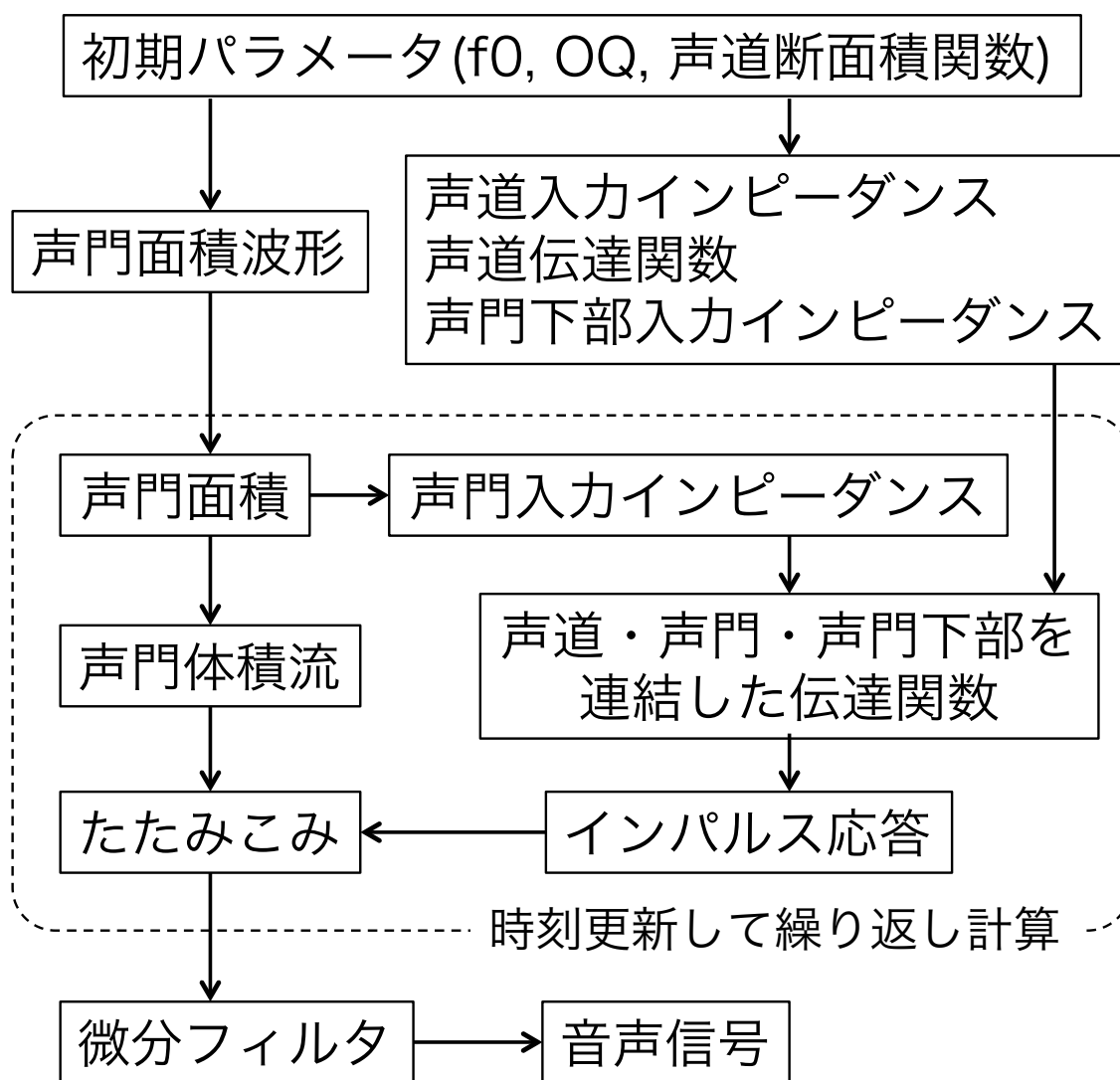


図 6.5: 音声生成モデルによる音声合成シミュレーションのフローチャート。

6.3 合成音声の作成と分析

2章で述べた音声生成モデルを用いて、様々な声門境界条件で音声を合成した。音声合成に用いた声道断面積は、日本人成人男性1名を被験者として、3次元声道MRI撮像によって得られた、日本語5母音/a/, /i/, /u/, /e/, /o/の声道断面積データである。また、声門下部の断面積関数として、Weibel [35] の気管支と肺の断面積データを用いた。各母音の声道断面積データを、図6.6～6.10に示す。

本シミュレーションに用いた声道・声門・声門下部の連結音響管モデル(図6.4)において、声門を表す音響管の長さは3 mmとした。発声パラメータとして、基本周波数は100 Hz、声門開放率は0.2間隔で、0.4から0.8までとした。また、最大声門幅は0.25 mm間隔で、0.25から3.00 mmまでとし、最大声門面積 = 最大声門幅 × 17 mm²として求めた。これにより、実際に用いた最大声門面積は4.25～51 mm²である。これらの発声パラメータの組み合わせにより、1つの母音に対して、60通りの音声を合成した。その他のパラメータは、肺圧 $P_0 = 8 \text{ cmH}_2\text{O}$ 、空気の密度 $\rho = 1.184 \times 10^{-3} \text{ g/cm}^3$ 、音速 $c = 34630 \text{ cm/sec}$ 、合成音の時間長は400 msec、サンプリング周波数は20 kHz、DFT点数は 2^{14} とした。

得られたすべての合成音声から、線形予測分析によって声道共鳴特性を求めた。さらに、ピークピッキングによってフォルマント周波数を分析した。線形予測分析の次数は12とし、分析対象のフォルマントは F_1, F_2 とした。なお、線形予測分析の際には、合成音声の100～350 msecの区間を切り出した上で、10 kHzへのダウンサンプリング、プリアンファシス、ハミング窓による窓掛けを行った。

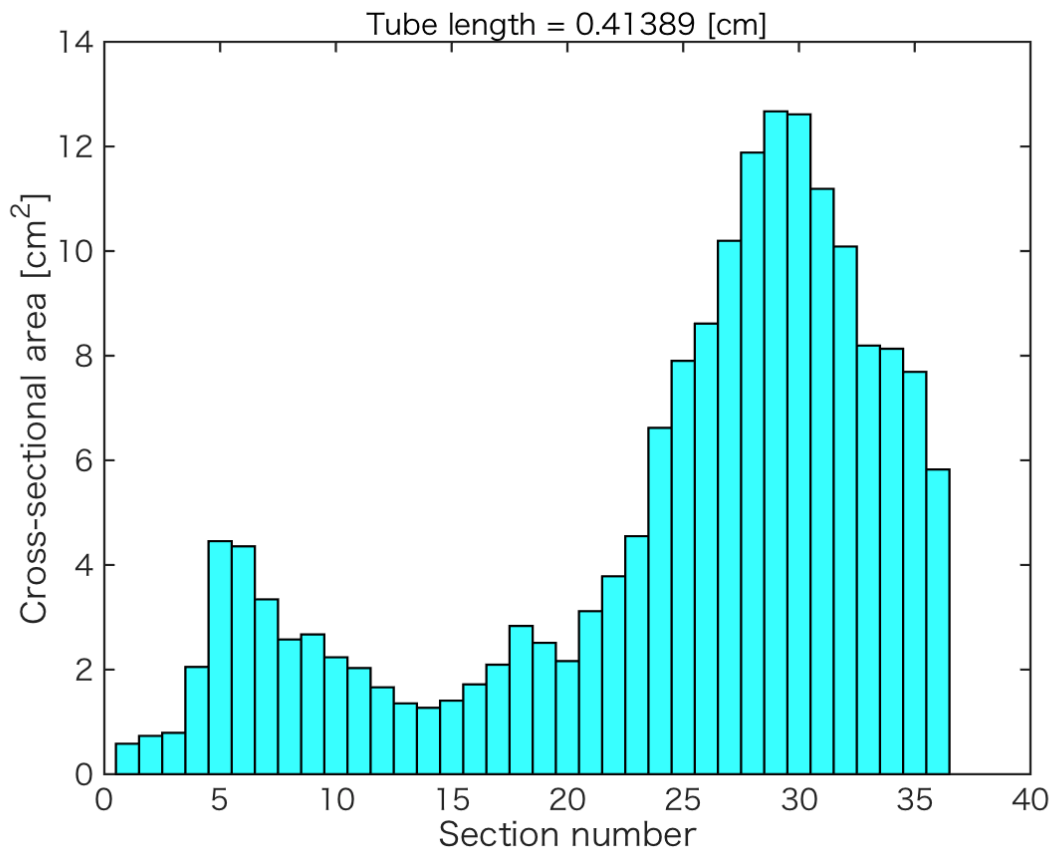


図 6.6: 音声合成に用いた、日本語 5 母音 /a/ の声道断面積データ。Section number は声門から口唇の順に対応している。各セクションの長さは一定値として、図の上に表示している。

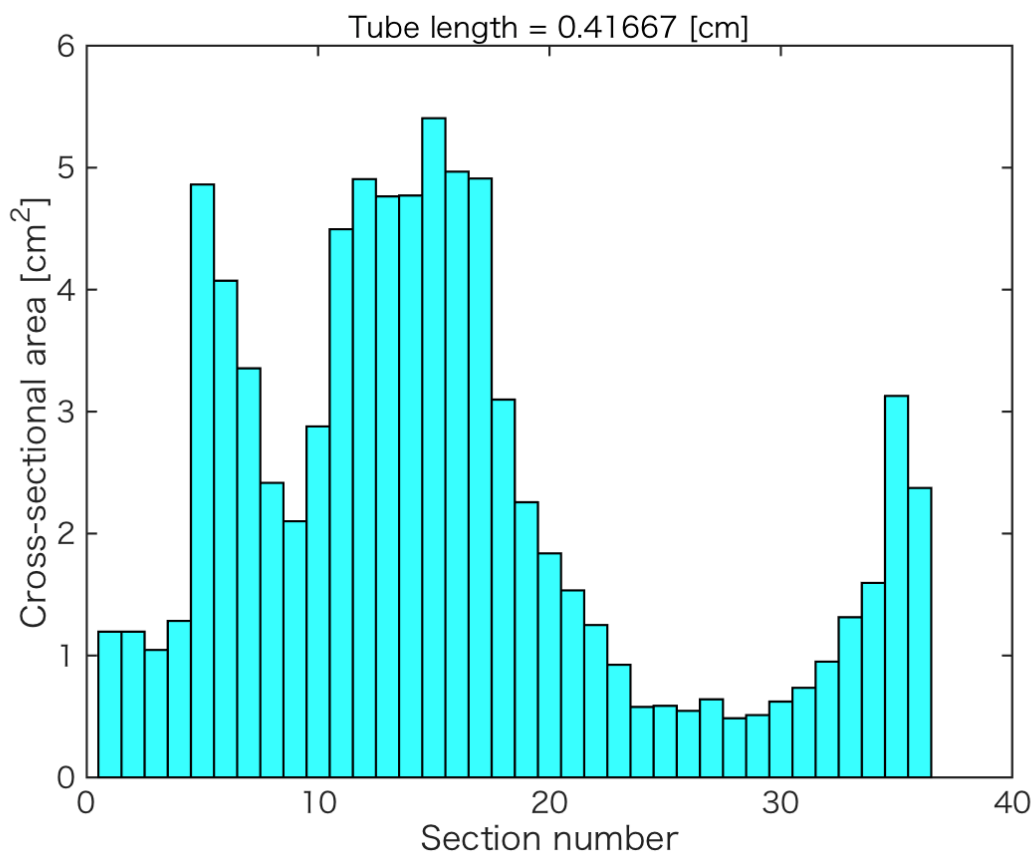


図 6.7: 音声合成に用いた、日本語 5 母音 /i/ の声道断面積データ。Section number は声門から口唇の順に対応している。各セクションの長さは一定値として、図の上に表示している。

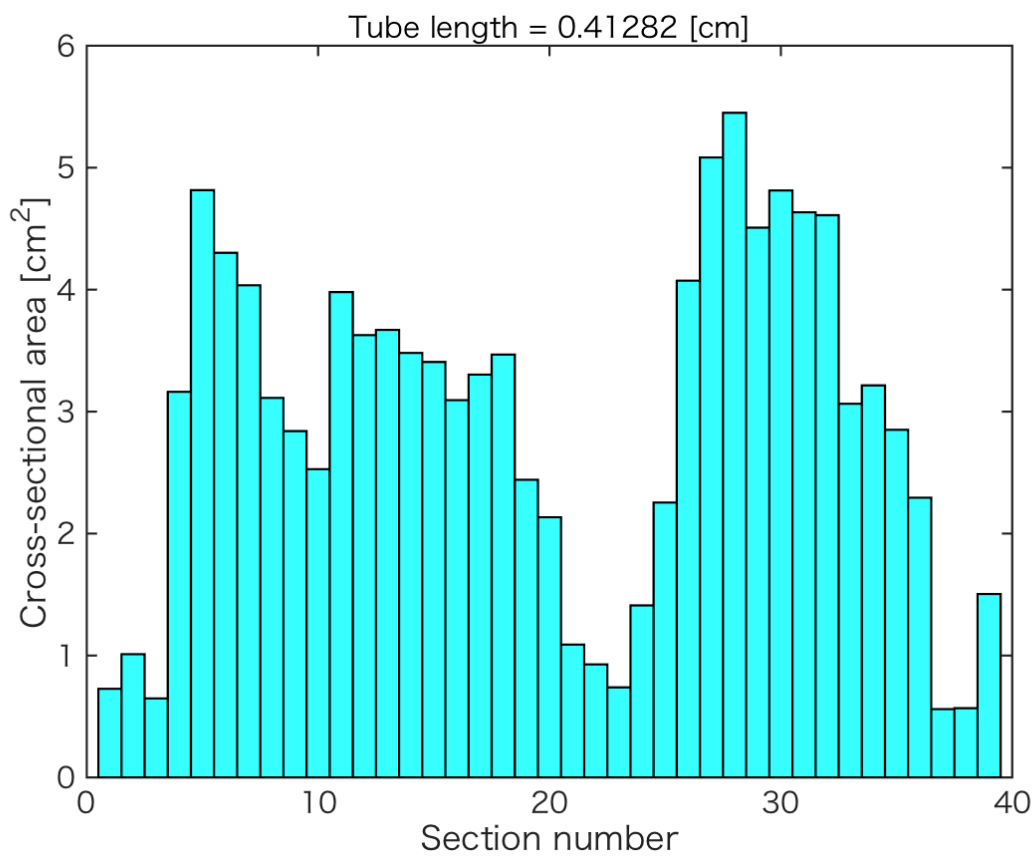


図 6.8: 音声合成に用いた、日本語 5 母音 /u/ の声道断面積データ。Section number は声門から口唇の順に対応している。各セクションの長さは一定値として、図の上に表示している。

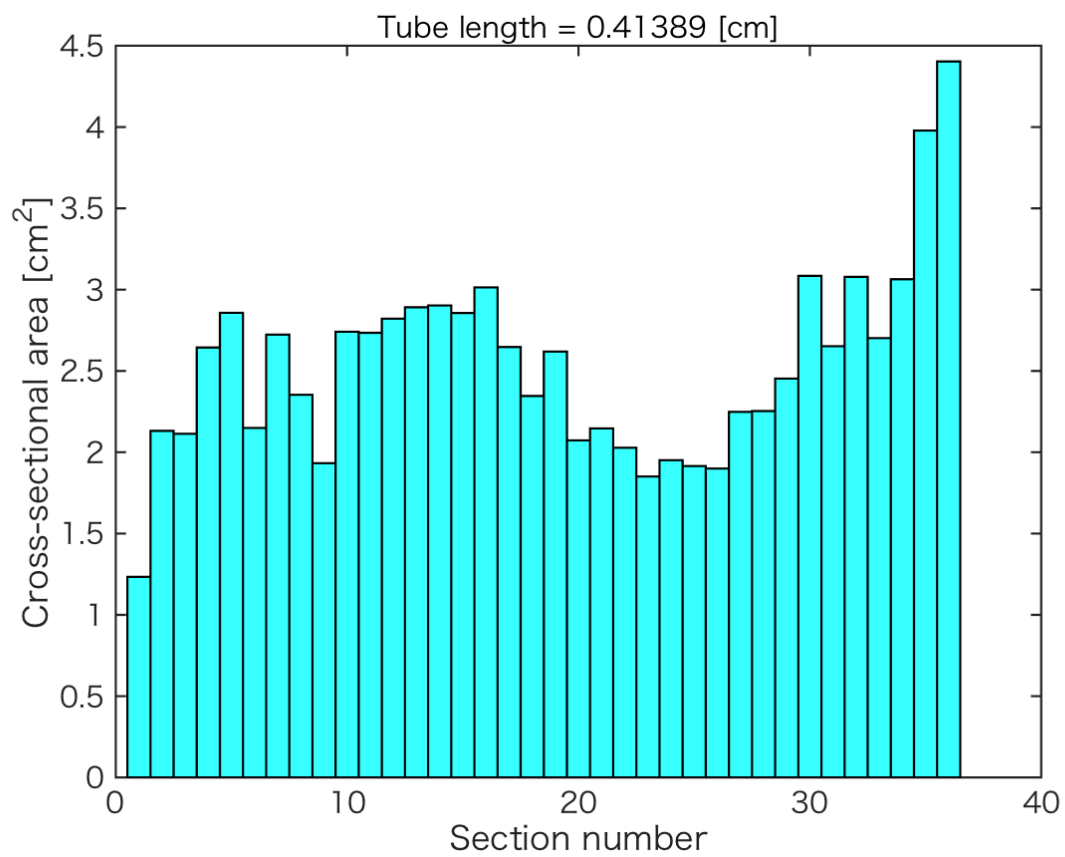


図 6.9: 音声合成に用いた、日本語 5 母音 /e/ の声道断面積データ。Section number は声門から口唇の順に対応している。各セクションの長さは一定値として、図の上に表示している。

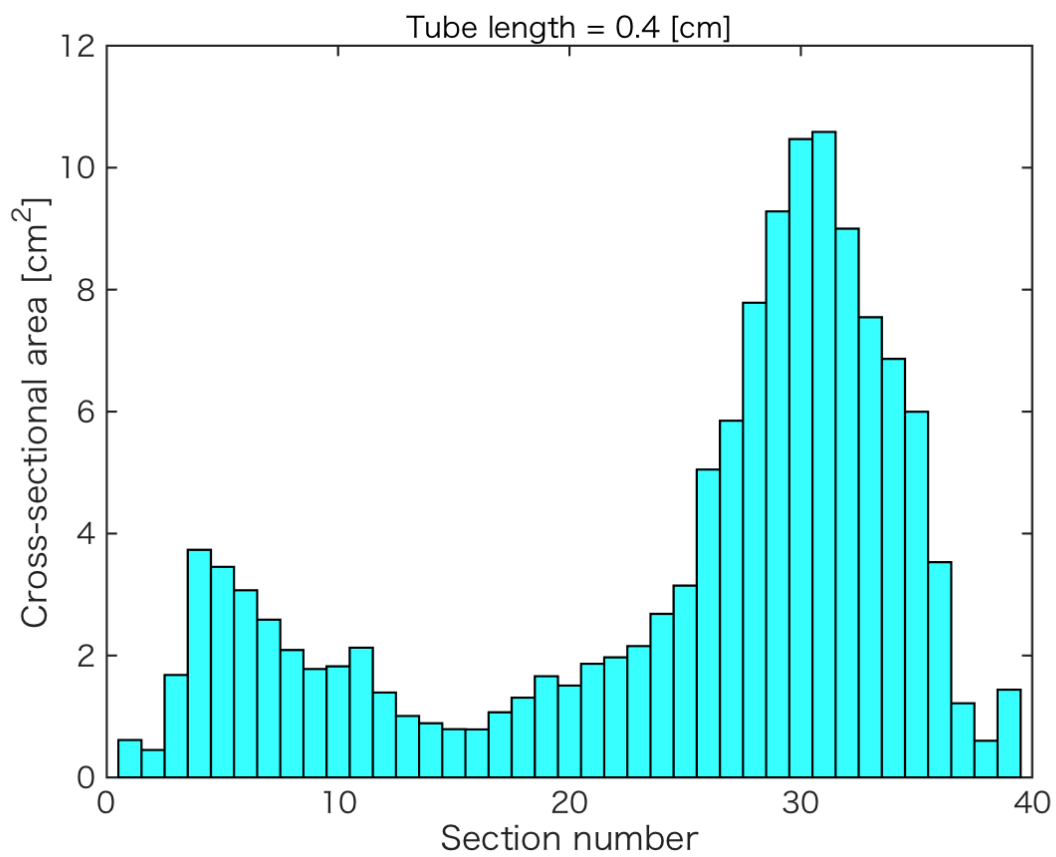


図 6.10: 音声合成に用いた、日本語 5 母音 /o/ の声道断面積データ。Section number は声門から口唇の順に対応している。各セクションの長さは一定値として、図の上に表示している。

6.4 結果と考察

6.4.1 第1フォルマント周波数

図6.11～図6.15において、日本語5母音の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値を示す。各図において、赤色の実線は声道断面積関数の音響管モデルを用いて計算された主声道の第1フォルマント周波数 F_1' を示している。また、赤色の破線は $F_1' \pm 10\%$ の範囲を、赤色の点線は $F_1' \pm 20\%$ の範囲を示している。

第1フォルマント周波数の結果に着目すると、いずれの母音の場合でも、最大声門幅や声門開放率にかかわらず、主声道の第1フォルマント周波数 F_1' に比べて高くなる傾向を示した。特に母音/i/の場合では、いずれの場合においても、主声道の第1フォルマント周波数より20%以上の上昇が見られた。また、母音/o/の場合では、ほぼすべての場合において、主声道の第1フォルマント周波数より10%以上の上昇が見られた。これらの傾向は、先行研究 [28,30] と一致している。すなわち、声門の境界条件が完全閉鎖でなくなったため、声道が開管共鳴に近づいた結果、全体として第1フォルマント周波数が上昇したと考えられる。

各母音について着目すると、母音/a/, /e/ では、最大声門幅が大きくなってもさほど第1フォルマント周波数に影響していないのに対し、母音/i/, /u/, /o/ では、最大声門幅が大きくなるにつれて、第1フォルマント周波数も上昇する傾向が見られた。特に母音/i/, /o/は顕著にフォルマント周波数が上昇した。また、声門開放率が上昇するにつれて、第1フォルマント周波数も上昇する傾向が見られたが、声門開放率の上昇と第1フォルマント周波数の上昇傾向は必ずしも一致しなかった。

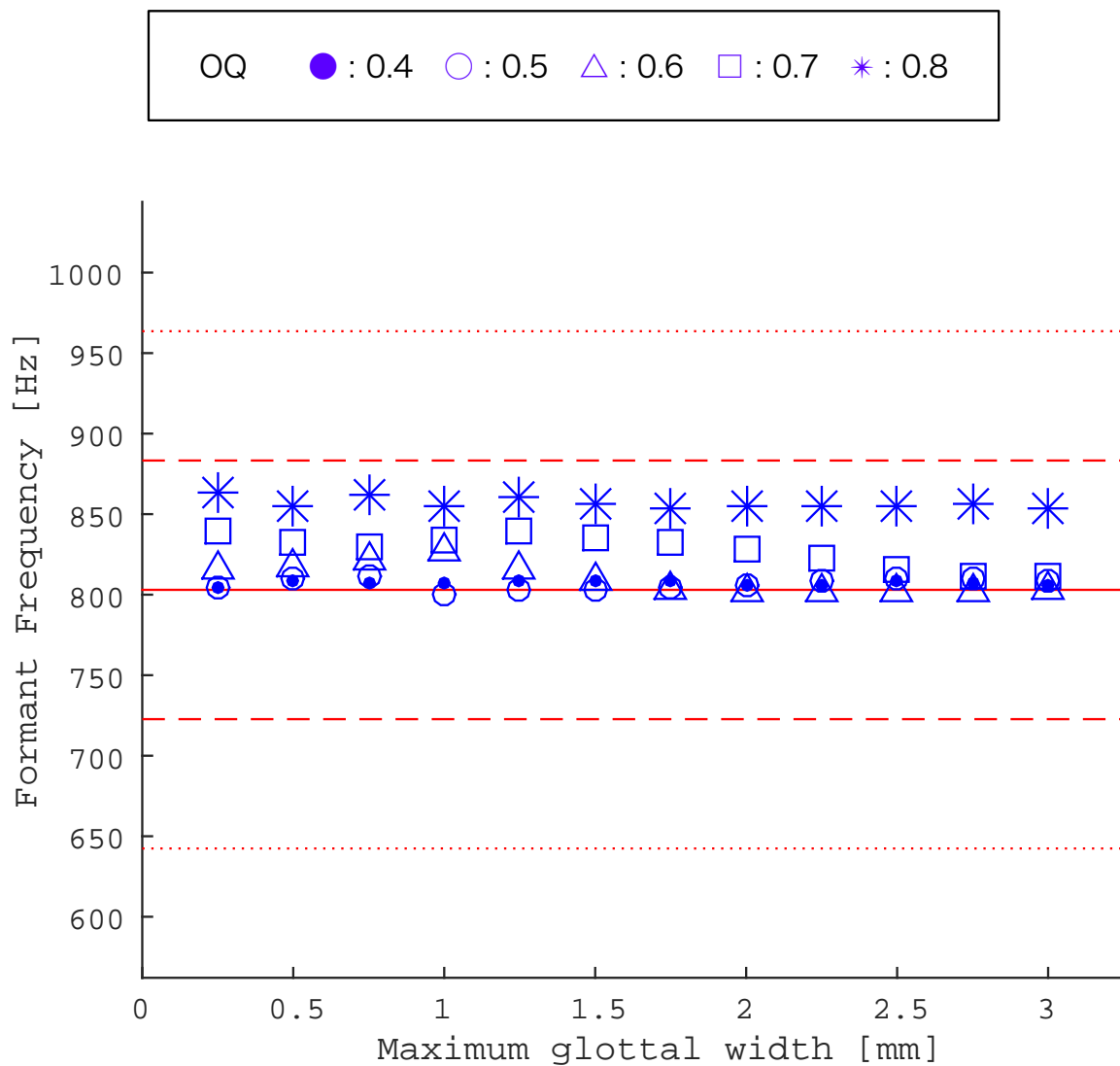


図 6.11: 母音/a/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値。赤色の実線は主声道の第1フォルマント周波数 F_1' (= 803 Hz)、赤色の破線は $F_1' \pm 10\%$ 、赤色の点線は $F_1' \pm 20\%$ を示す。

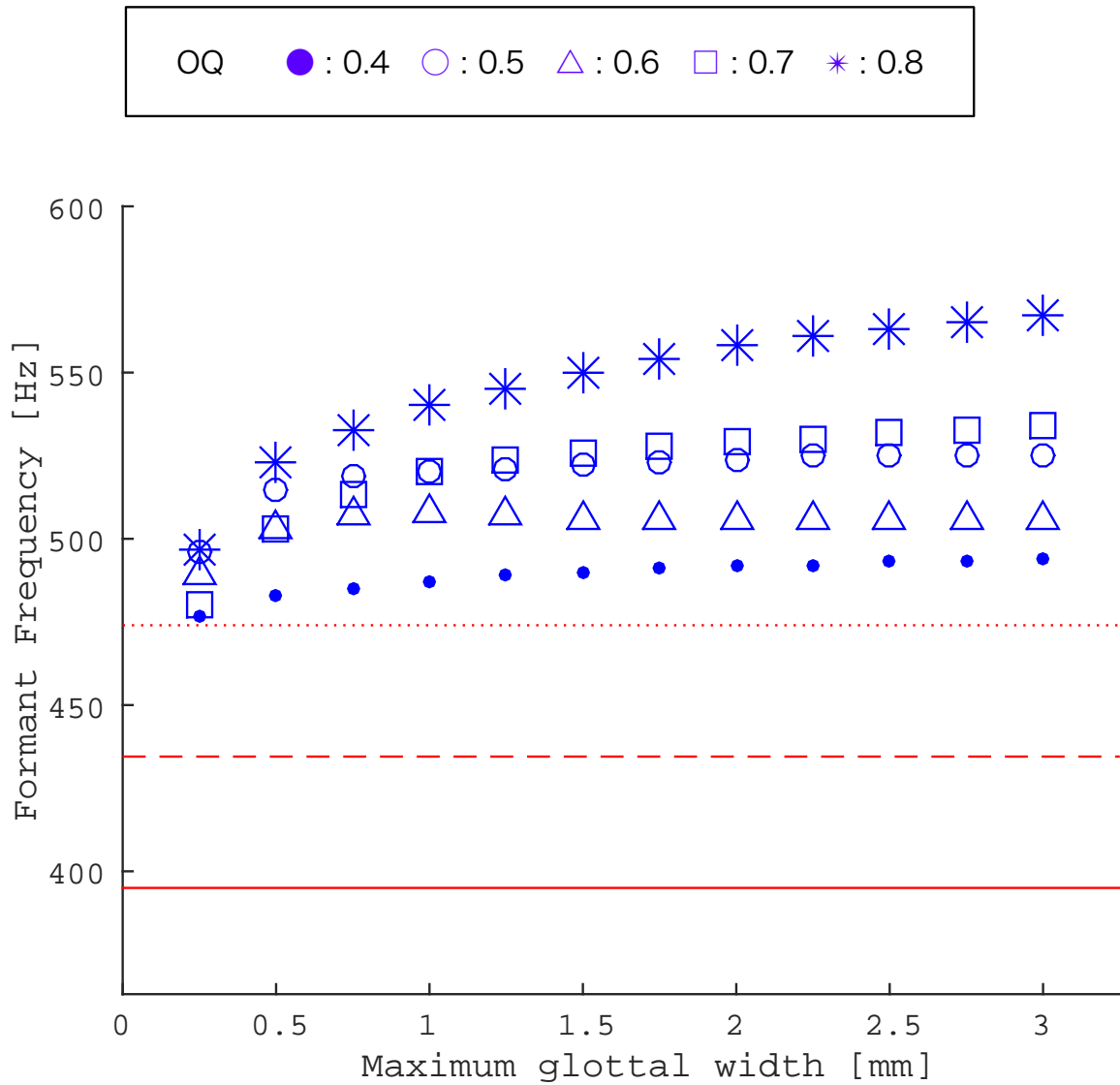


図 6.12: 母音/i/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値。赤色の実線は主声道の第1フォルマント周波数 F_1' (= 395 Hz)、赤色の破線は $F_1' \pm 10\%$ 、赤色の点線は $F_1' \pm 20\%$ を示す。

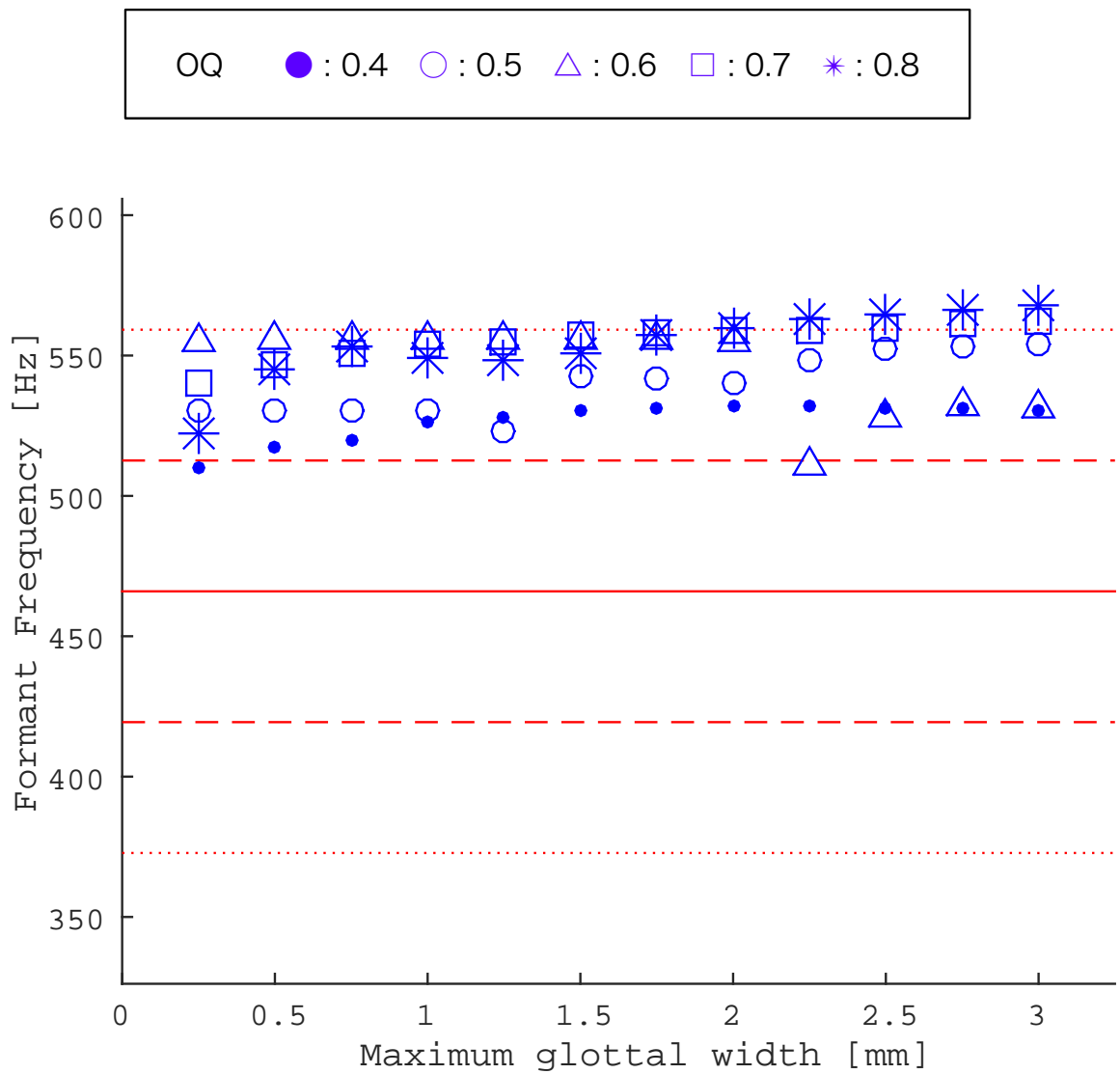


図 6.13: 母音/u/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値。赤色の実線は主声道の第1フォルマント周波数 F_1' (= 466 Hz)、赤色の破線は $F_1' \pm 10\%$ 、赤色の点線は $F_1' \pm 20\%$ を示す。

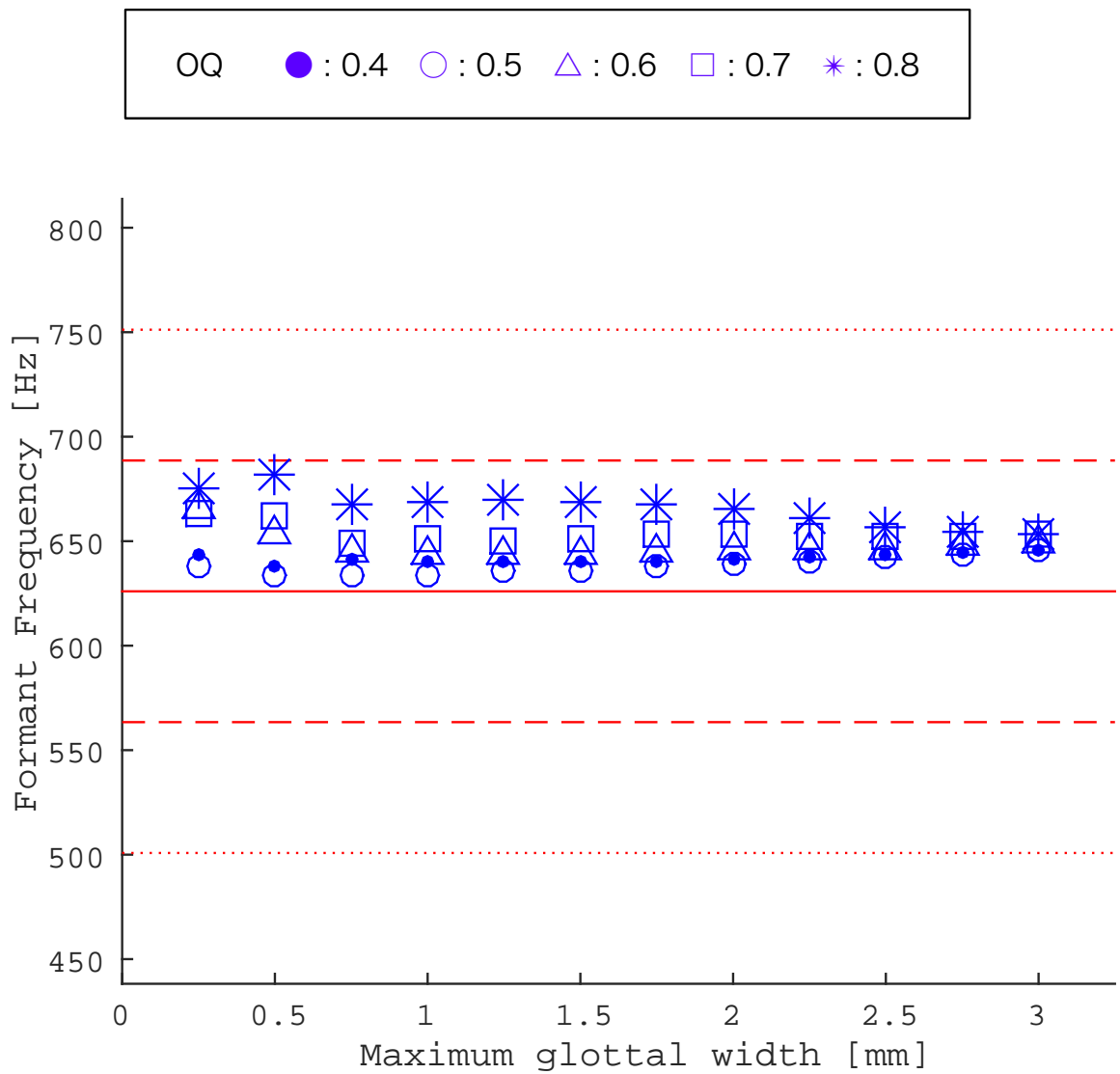


図 6.14: 母音/e/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値。赤色の実線は主声道の第1フォルマント周波数 F_1' (= 626 Hz)、赤色の破線は $F_1' \pm 10\%$ 、赤色の点線は $F_1' \pm 20\%$ を示す。

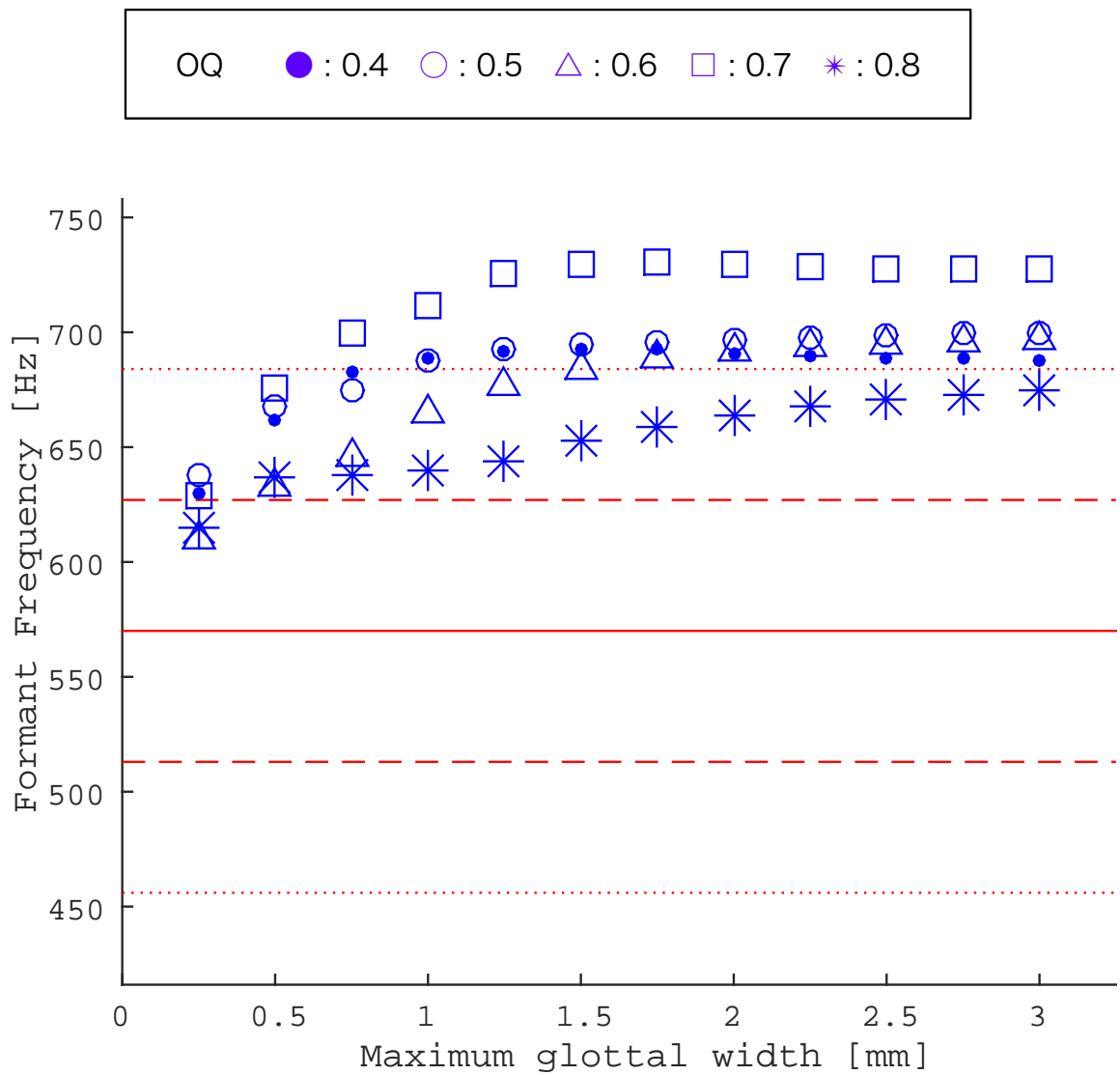


図 6.15: 母音/o/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第1フォルマント周波数の値。赤色の実線は主声道の第1フォルマント周波数 F_1' (= 570 Hz)、赤色の破線は $F_1' \pm 10\%$ 、赤色の点線は $F_1' \pm 20\%$ を示す。

6.4.2 第2フォルマント周波数の結果

図6.16～図6.20において、日本語5母音の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値を示す。各図において、赤色の実線は声道断面積関数の音響管モデルを用いて計算された主声道の第2フォルマント周波数 F_2 を示している。また、赤色の破線は $F_2 \pm 10\%$ の範囲を、赤色の点線は $F_2 \pm 20\%$ の範囲を示している。

第2フォルマント周波数は、第1フォルマント周波数の結果で見られた傾向とは異なる傾向を示した。第1フォルマント周波数の場合では、主声道のフォルマント周波数と比較して、全体的に周波数が上昇していたのに対し、第2フォルマント周波数では、主声道のフォルマント周波数よりも低下している場合が見られた。特に母音/a/において、その傾向は顕著であった。また、最大声門幅の大きさに対する第2フォルマント周波数の変化は、母音によって異なる様子を示した。母音/a/, /o/では、最大声門幅の変化に対して、第2フォルマント周波数も大きく変化している。これに対し、母音/i/, /u/, /e/においては、最大声門幅の変化に対する第2フォルマント周波数の変化は比較的小さいことがわかる。声門開放率の変化に対する第2フォルマント周波数は、声門開放率の値の大きさと、フォルマント周波数の変化の程度に、必ずしも一致した傾向は見られなかった。また最大声門幅の値によって、声門開放率の変化に対するフォルマント周波数の変化の程度が異なる傾向を示した。このことから、声門開放率が第2フォルマント周波数に与える影響は、最大声門幅の値に依存することが示唆された。

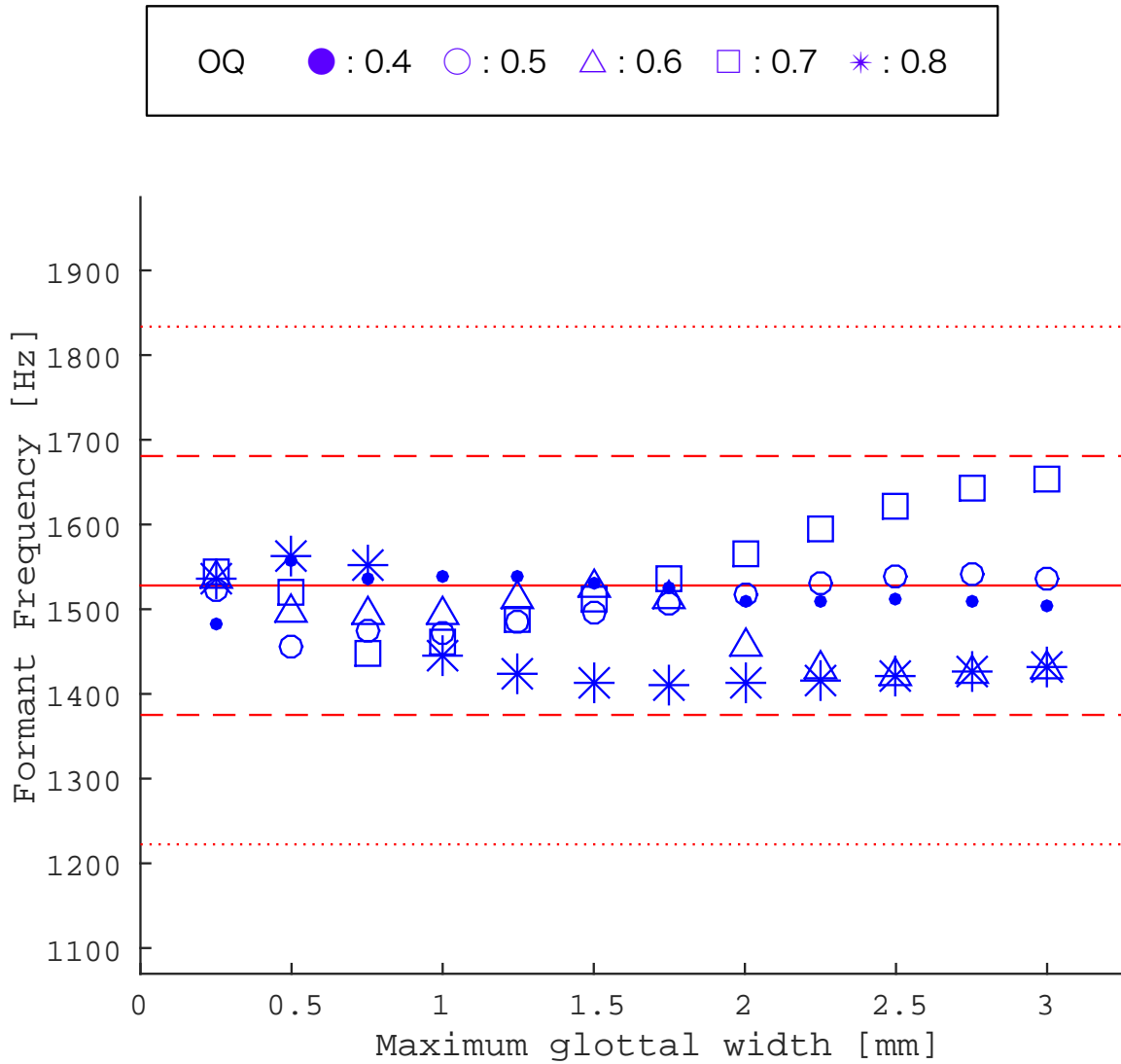


図 6.16: 母音/a/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値。赤色の実線は主声道の第2フォルマント周波数 F_2' ($= 1528$ Hz)、赤色の破線は $F_2' \pm 10\%$ 、赤色の点線は $F_2' \pm 20\%$ を示す。

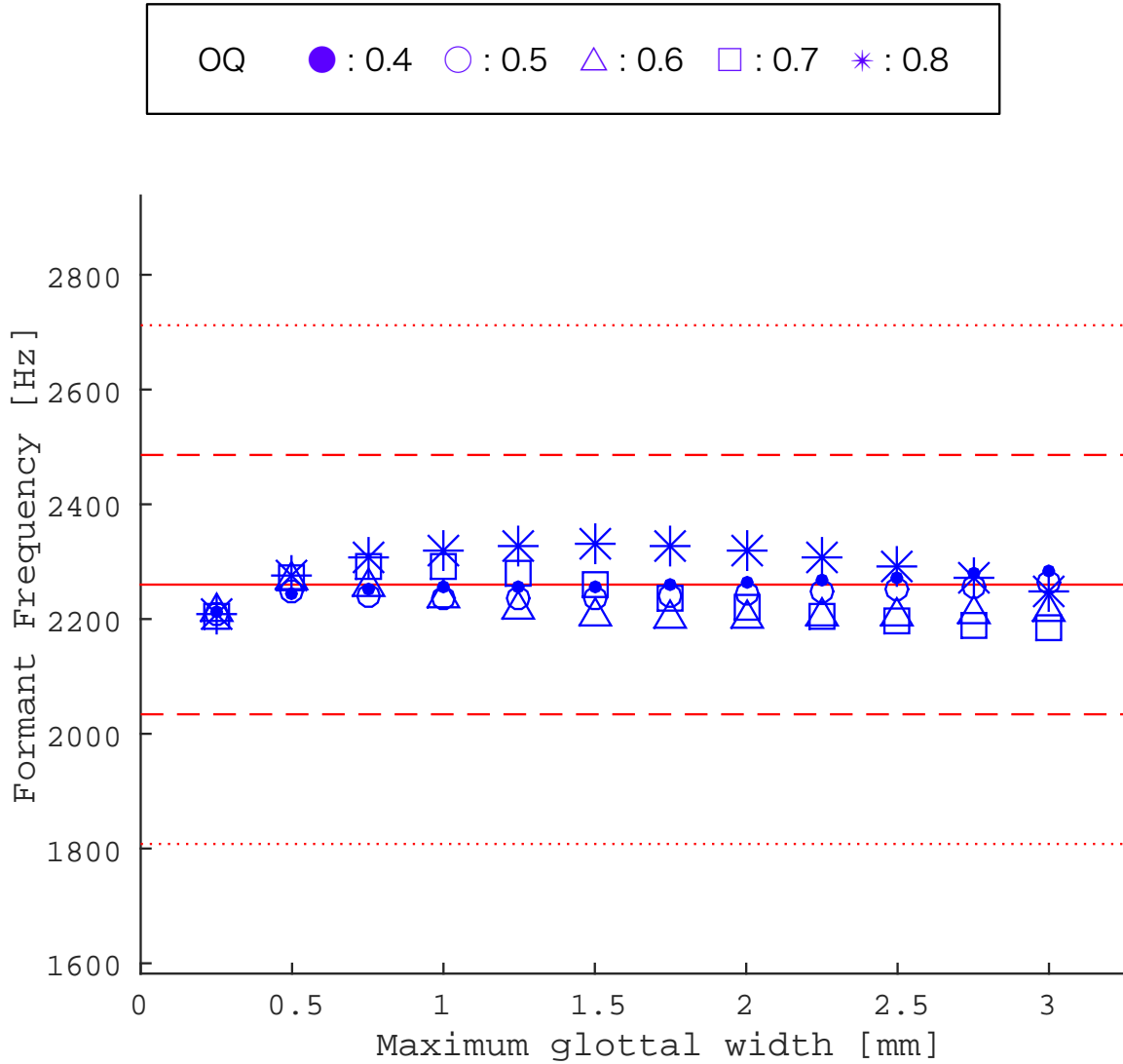


図 6.17: 母音/i/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値。赤色の実線は主声道の第2フォルマント周波数 F_2' ($= 2260$ Hz)、赤色の破線は $F_2' \pm 10\%$ 、赤色の点線は $F_2' \pm 20\%$ を示す。

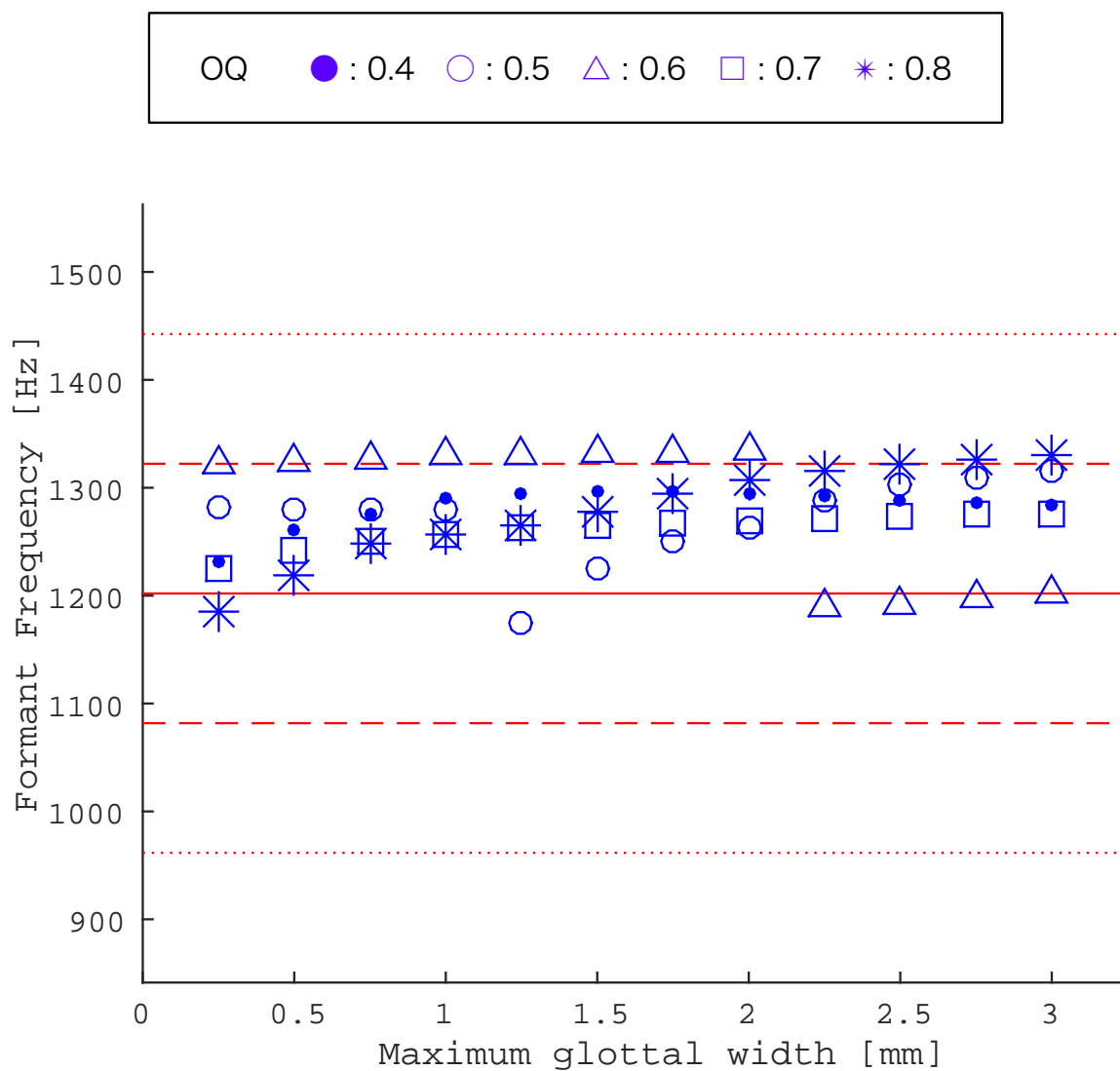


図 6.18: 母音/u/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値。赤色の実線は主声道の第2フォルマント周波数 F_2' (= 1202 Hz)、赤色の破線は $F_2' \pm 10\%$ 、赤色の点線は $F_2' \pm 20\%$ を示す。

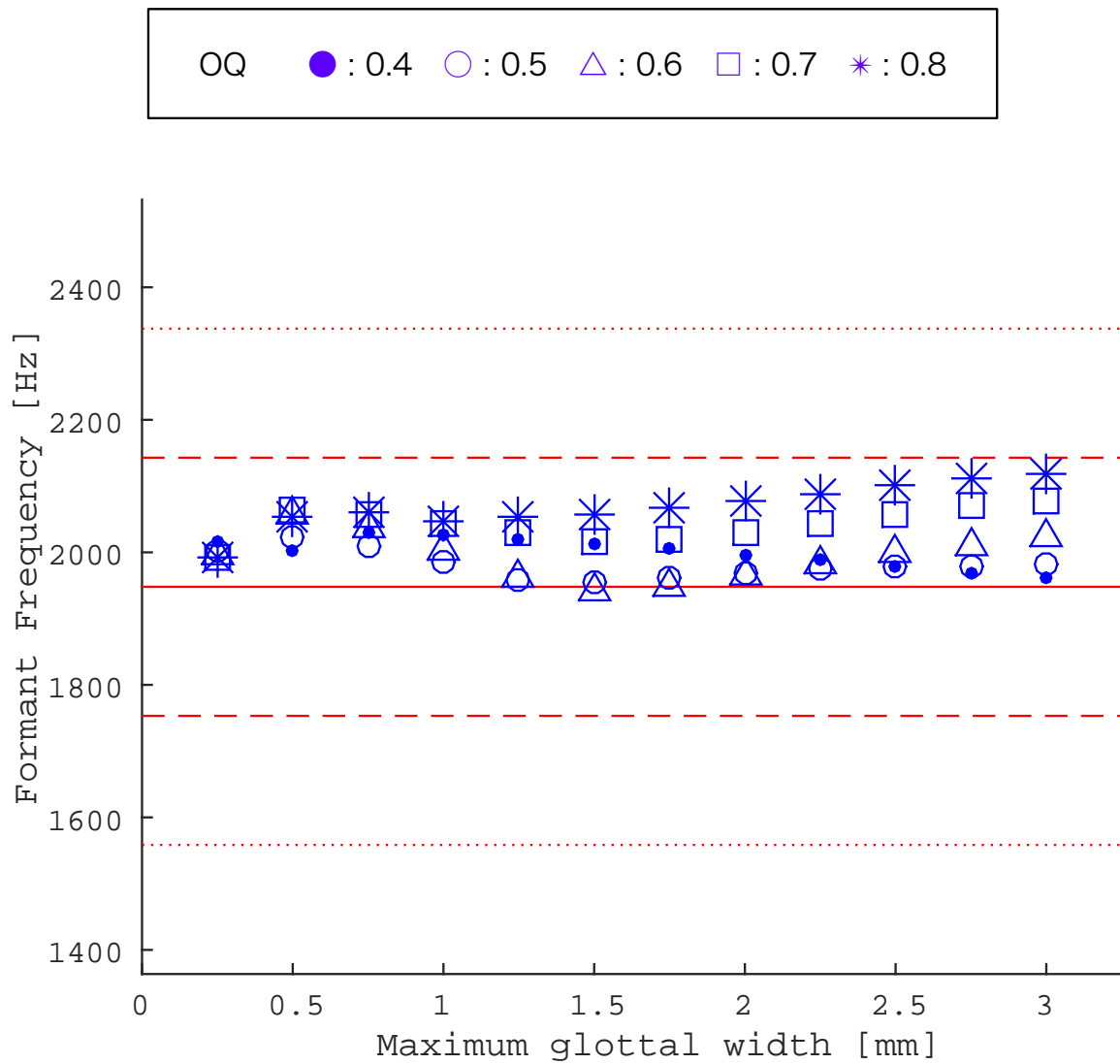


図 6.19: 母音/e/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値。赤色の実線は主声道の第2フォルマント周波数 F_2' (= 1948 Hz)、赤色の破線は $F_2' \pm 10\%$ 、赤色の点線は $F_2' \pm 20\%$ を示す。

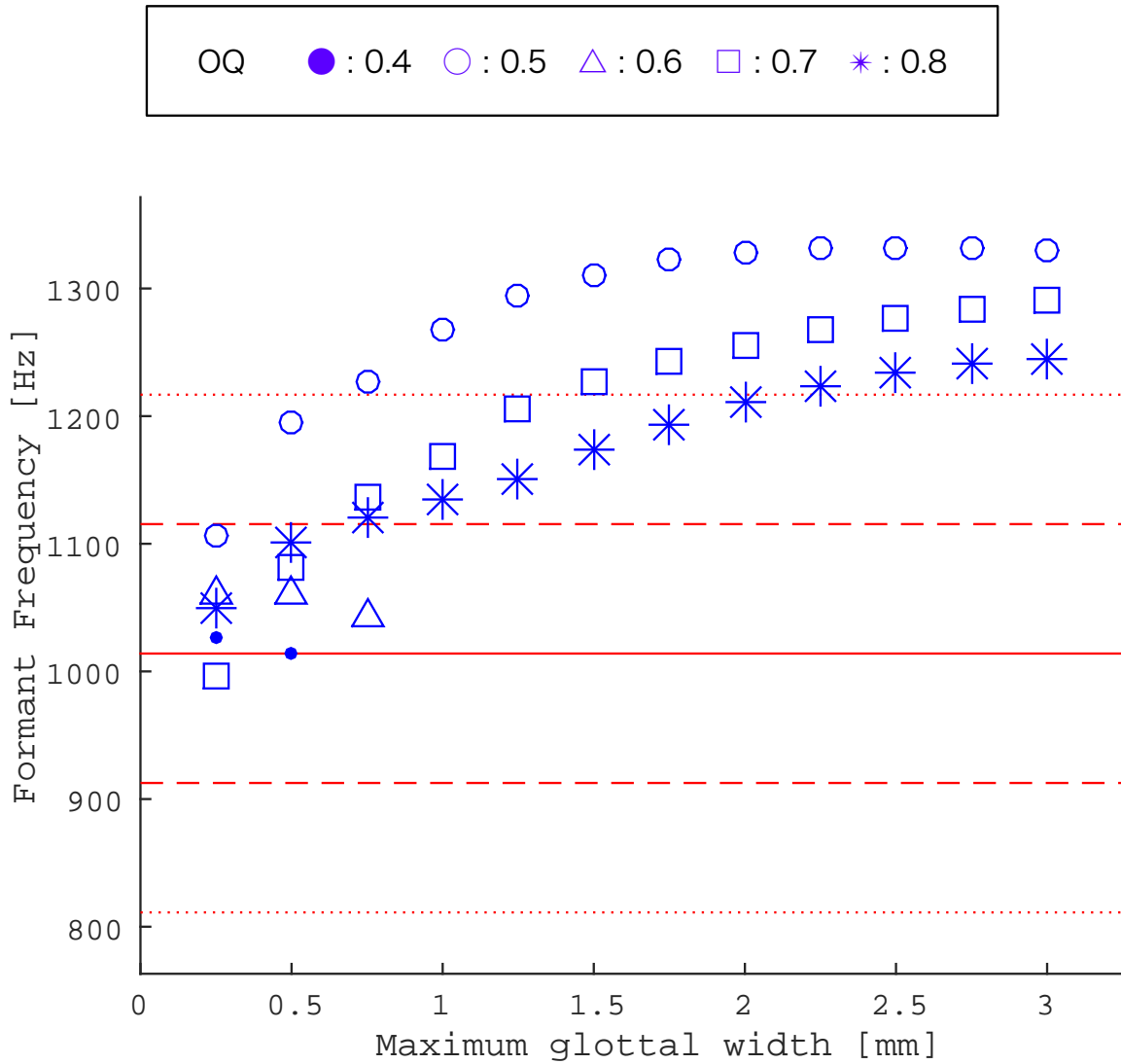


図 6.20: 母音/o/の声道断面積を用いて、最大声門幅と声門開放率を様々に変化させて合成した音声から得られた、第2フォルマント周波数の値。赤色の実線は主声道の第2フォルマント周波数 F_2' (= 1014 Hz)、赤色の破線は $F_2' \pm 10\%$ 、赤色の点線は $F_2' \pm 20\%$ を示す。

6.4.3 声門下部インピーダンスと声道伝達関数

最大声門幅や声門開放率に対して、第1および第2フォルマントの周波数の変化の程度は、母音によって大きく異なった。これは、各母音の第1および第2フォルマント周波数の値と、声門下部インピーダンスのピーク周波数の値の関係性による影響も考えられる。図6.21に、本研究の音声合成に用いた声門下部インピーダンスを示す。これより、およそ600 Hzに第1ピーク、およそ1200 Hzに第2ピークが生じていることがわかる。

図6.11～6.15に示した第1フォルマント周波数の結果より、母音/i/, /u/, /o/は最大声門幅や声門開放率の影響を比較的大きく受けていることがわかった。これらの第1フォルマントはそれぞれ約400 Hz～500 Hz程度であり、声門下部インピーダンスの第1ピークの近傍にある。そのため、声門の境界条件が変化することによって、声門下部インピーダンスの第1ピークの影響をより強く受けたために、第1フォルマント周波数が大きく変化したものだと考えられる。一方で母音/a/, /e/は、第1フォルマント周波数と声門下部インピーダンスの第1ピークが離れている（母音/e/の場合はほぼ同程度の値）であるため、声門の境界条件の変化に対して、さほど声門下部の影響を受けなかったと考えられる。

図6.16～6.20に示した第2フォルマント周波数の結果より、母音/a/, /o/は最大声門幅や声門開放率の影響を大きく受けていることがわかったが、これらの第2フォルマントはそれぞれ約1500 Hzおよび約1000 Hzであり、声門下部インピーダンスの第2ピークの近傍にある。そのため、声門の境界条件が変化することによって、声門下部インピーダンスの第2ピークの影響をより強く受けたために、第2フォルマント周波数が大きく変化したものだと考えられる。一方で母音/i/, /u/, /e/は、第2フォルマント周波数と声門下部インピーダンスの第2ピークが離れている（母音/u/の場合はほぼ同程度の値）であるため、声門の境界条件の変化に対して、さほど声門下部の影響を受けなかったと考えられる。

これらの声門下部インピーダンスのピークによるフォルマントへの影響は、各母音の声道伝達関数の計算結果においても確認することができる。図6.22～図6.26は、各母音の声道断面積を用いた場合における、ある最大声門面積の値に対する声道伝達関数の計算結果を示す。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた最大声門面積の値は0.25mm間隔で、 $0.25 \sim 3.00 \times 17$ [mm²]として、求められた伝達関数の結果を最大声門面積の小さい順に

下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。第1フォルマントと第2フォルマントに注目すると、最大声門面積が大きくなるにつれて、声門下部インピーダンスの第1ピークおよび第2ピークの周波数の方向に、フォルマント周波数が近づくように次第に変化していることがわかる。また、エネルギーの損失も大きくなるため、フォルマントのQ値も次第に低下している。

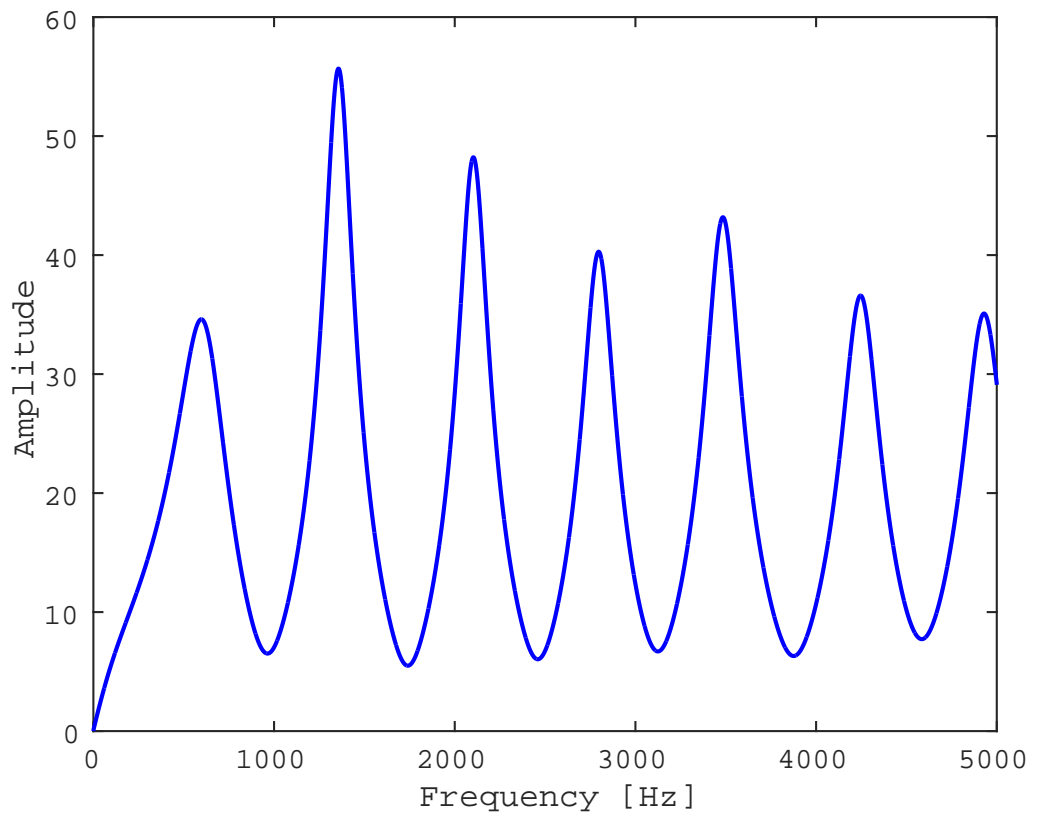


図 6.21: 音声合成に用いた、声門下部組織のインピーダンス。約 600Hz, 1200Hz にピークをもつ。

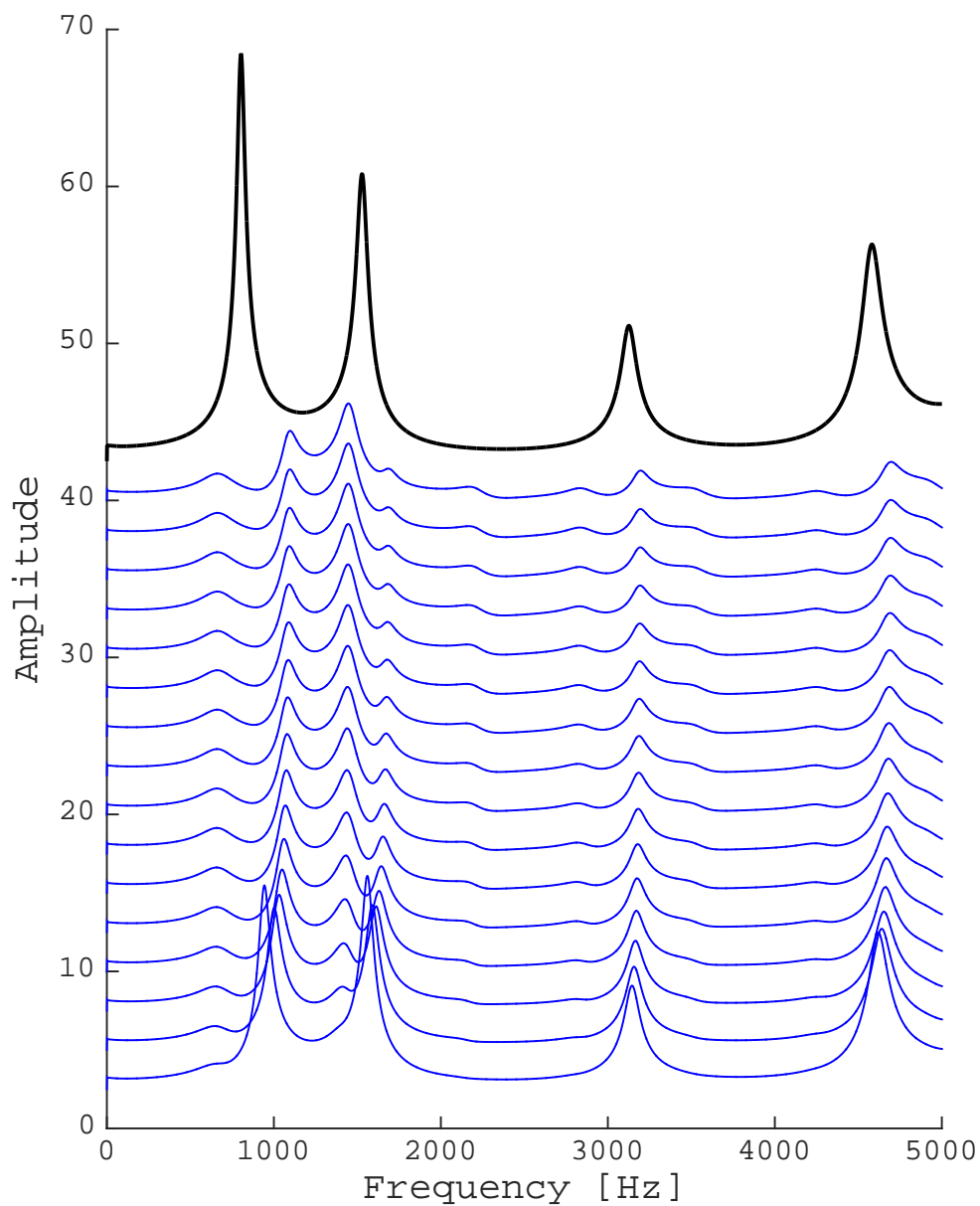


図 6.22: 母音/a/の声道断面積を用いた場合における、ある声門面積の値に対する声道伝達関数の計算結果。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた声門面積の値は $(0.25, 0.50, \dots, 3.00 \times 17 \text{ [mm}^2\text{)})$ として、得られた伝達関数の結果は声門面積の小さい順に下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。

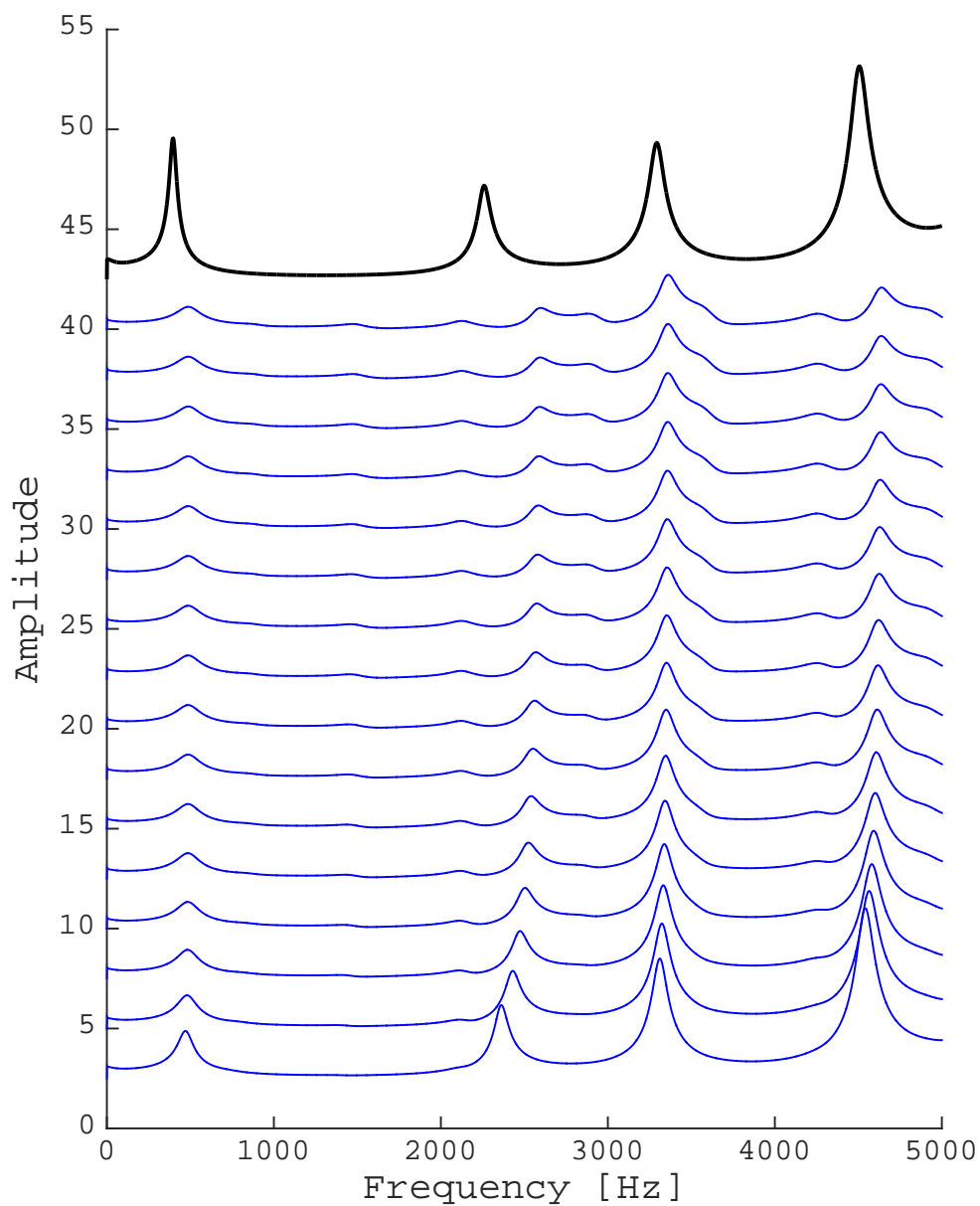


図 6.23: 母音/i/の声道断面積を用いた場合における、ある声門面積の値に対する声道伝達関数の計算結果。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた声門面積の値は $(0.25, 0.50, \dots, 3.00 \times 17 \text{ [mm}^2])$ として、得られた伝達関数の結果は声門面積の小さい順に下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。

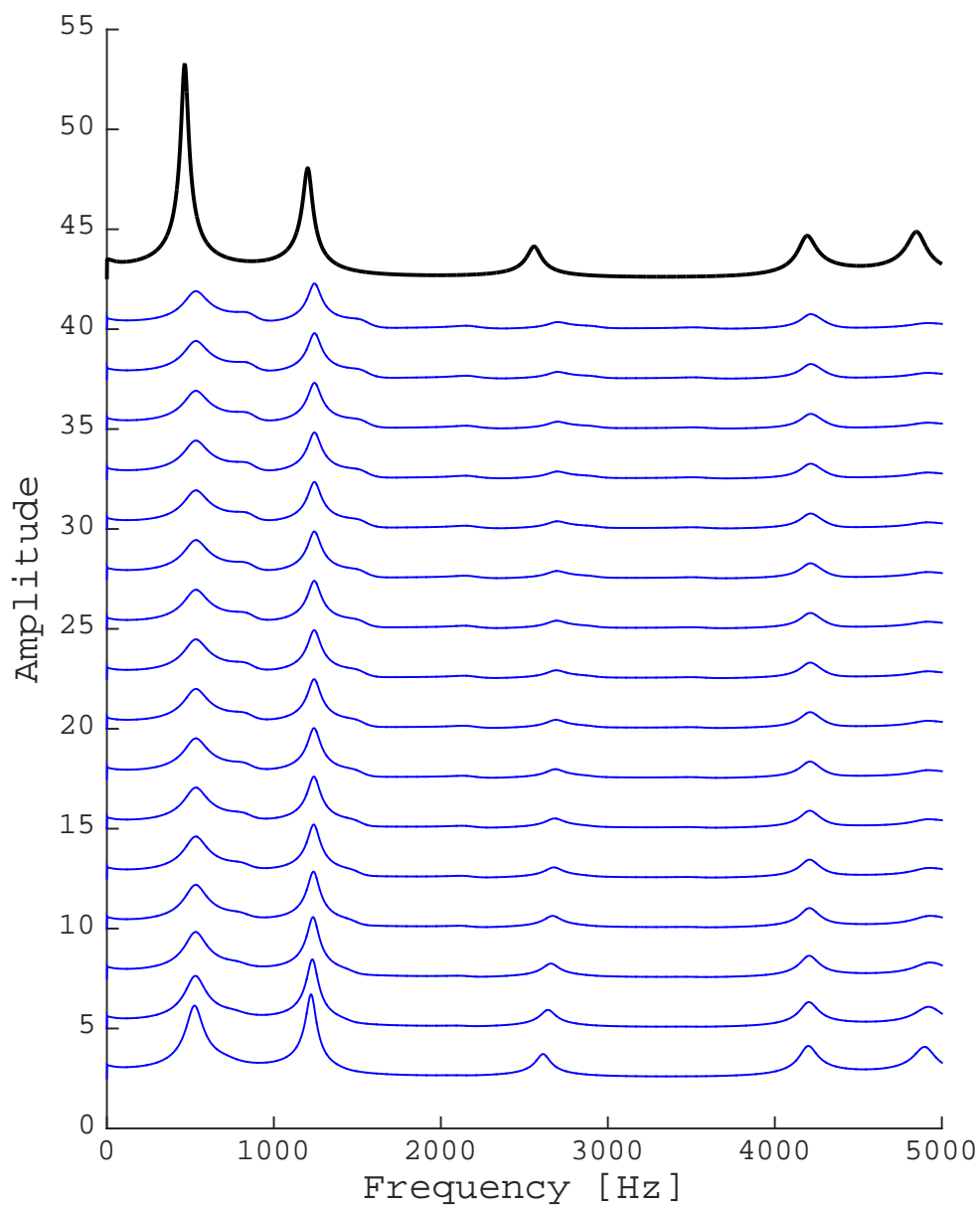


図 6.24: 母音/u/の声道断面積を用いた場合における、ある声門面積の値に対する声道伝達関数の計算結果。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた声門面積の値は $(0.25, 0.50, \dots, 3.00 \times 17 \text{ [mm}^2\text{)})$ として、得られた伝達関数の結果は声門面積の小さい順に下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。

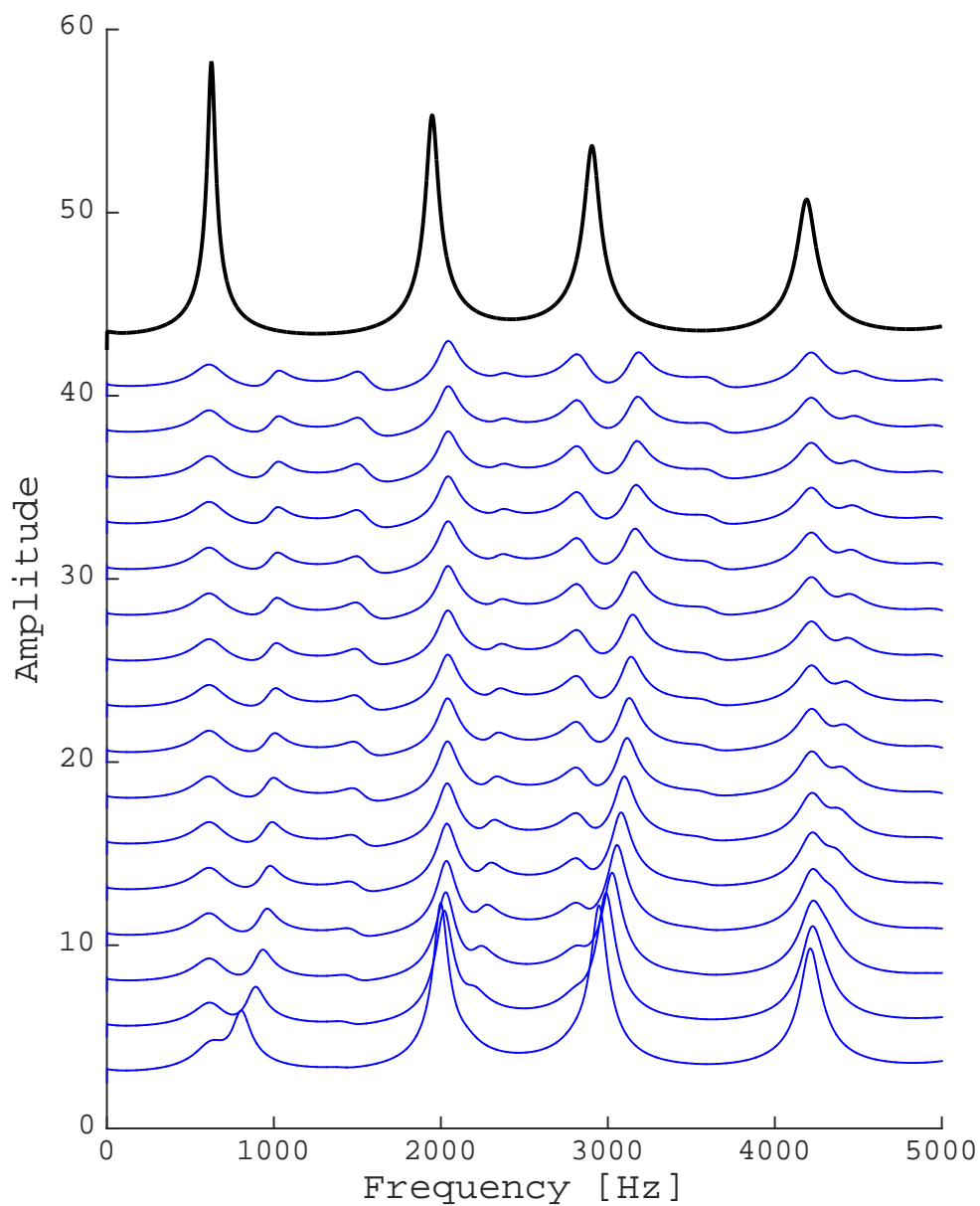


図 6.25: 母音/e/の声道断面積を用いた場合における、ある声門面積の値に対する声道伝達関数の計算結果。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた声門面積の値は $(0.25, 0.50, \dots, 3.00 \times 17 \text{ [mm}^2\text{)})$ として、得られた伝達関数の結果は声門面積の小さい順に下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。

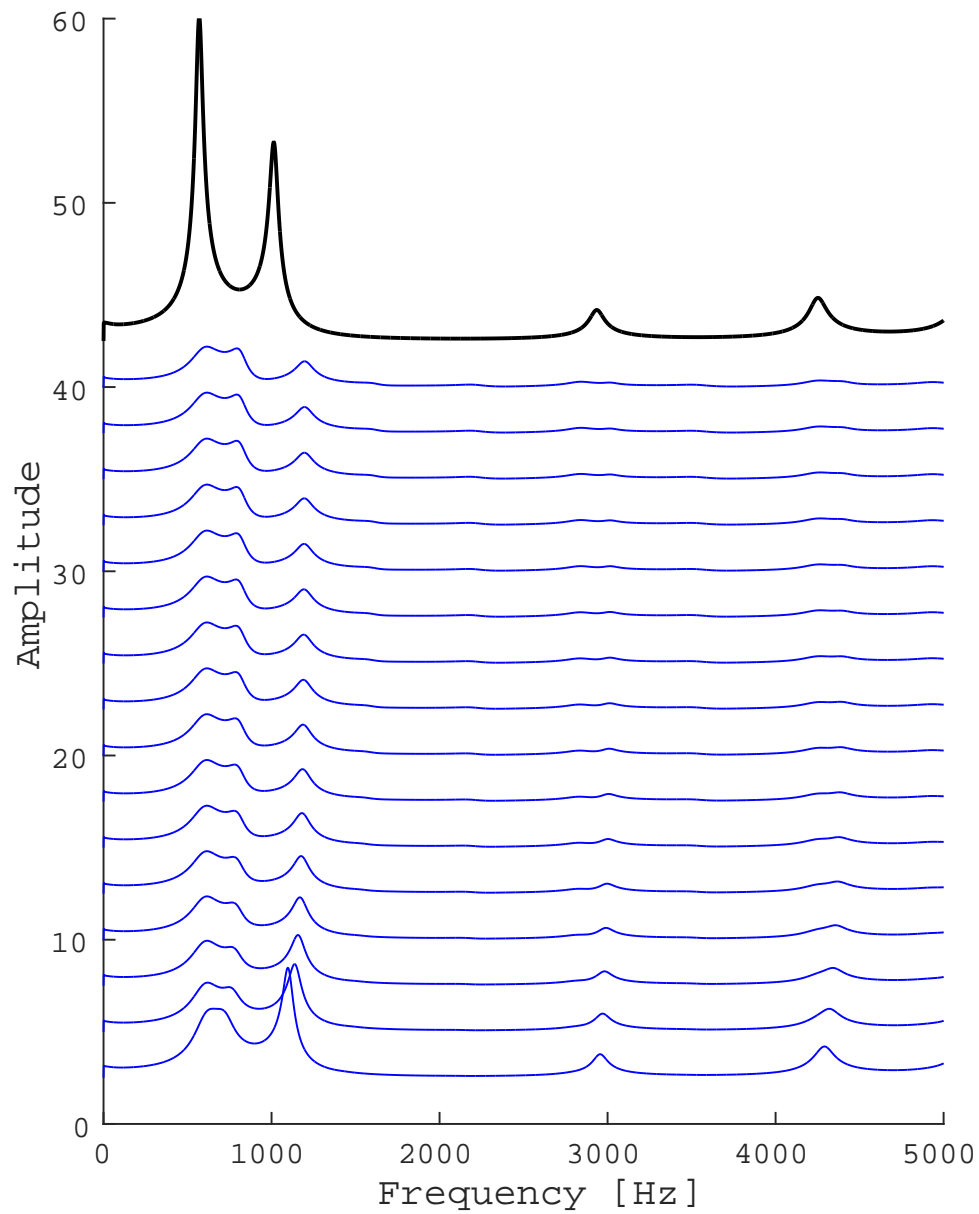


図 6.26: 母音/o/の声道断面積を用いた場合における、ある声門面積の値に対する声道伝達関数の計算結果。黒の実線は、主声道のみを用いて求めた声道伝達関数を示す。青の実線は、主声道・声門・声門下部を連結した経路全体の伝達関数を示す。なお、与えた声門面積の値は $(0.25, 0.50, \dots, 3.00 \times 17 \text{ [mm}^2\text{)})$ として、得られた伝達関数の結果は声門面積の小さい順に下から上に並べている。また見やすさのために、各伝達関数には一定の間隔を設けている。

6.5 まとめ

本稿では、音声生成シミュレーションによる合成音を用いて、声門の境界条件がフォルマントに与える影響について検討した。声道・声門・声門下部を連結した音響管モデルと、声門面積の多項式モデルを組み合わせた音声生成モデルによって音声を合成した。合成時の発声パラメータとして、声門開放率と最大声門幅を様々に変化させて与えた。声門の境界条件がフォルマントに与える影響は、声道断面積関数によって異なること、第1フォルマント周波数はすべての母音で周波数が上昇すること、声門の境界条件では、最大声門幅のほうが声門開放率よりもフォルマント周波数の変化により支配的であることが見られた。これらのフォルマント周波数の変化は、声道が開管共鳴に近づいたために生じていることや、声門下部のインピーダンスのピークの影響をより強く受けるために生じていることが示唆された。

第7章 総括

本章では、本研究のまとめを述べる。また、音源-フィルタ相互作用の影響に関する研究について、今後の展望と課題について述べる。

7.1 まとめ

本研究では、従来の音声生成モデルである音源-フィルタ理論において、互いに独立であり影響を及ぼし合わないと仮定された音源と声道が、実際の音声生成メカニズムのように独立ではない場合では、互いにどのように影響を及ぼし合うのか、すなわち音源-フィルタ相互作用の影響について検証を行った。

声道が音源に対して影響を及ぼすような音源-フィルタ相互作用の影響については、従来より様々な先行研究による検討がなされてきた。特に、近年の音源-フィルタ相互作用の影響を考慮した音声生成シミュレーション実験において、声道の第1フォルマント周波数に音源の基本周波数が近接すると、音源-フィルタ相互作用の影響が強まることで音源生成機構が不安定になり、声区転換を生じることが示唆されていた。このような声道が音源に対して影響を及ぼすような音源-フィルタ相互作用の影響に関して、本研究では第4章において検証した。一方で、音源が声道に対して影響を及ぼすような音源-フィルタ相互作用の影響についても、これまでいくつかの先行研究による検討がなされてきた。特に、音源の生成機構である声門の開きの程度の違いが、音源-フィルタ理論では完全閉鎖とみなされた声道の境界条件を変化させることによって、声道のフォルマントに影響を与えることが示唆されていた。このような音源が声道に対して影響を及ぼすような音源-フィルタ相互作用の影響に関して、本研究では第5章と第6章にて検証した。

7.1.1 声道が音源に対して影響を及ぼす影響について

第4章では、声区転換の発声測定実験を行い、声区転換時における第1フォルマント周波数と基本周波数の関係性を調べることで、声道が音源にもたらす音源-フィルタ相互作用の影響を検討した。声区転換の発声測定実験では、声門の開閉パターンと声道音響特性を同時測定した。声門の開閉パターンは電気喉頭計によってEGG信号として計測し、得られたEGG信号にDECOM法を適用することで、基本周波数と声門開放率を分析した。一方で、声区転換のように基本周波数が比較的高い発声の場合でも、より精確にフォルマント周波数を測定するために、外部音響励振法を用いることで声道音響特性を計測した。得られた励振信号による応答から被験者の自発発声成分を取り除くために、ケプストラム分析を適用して、フォルマント周波数の抽出を行った。このようにして得られた基本周波数と倍音成分、および第1フォルマント周波数の関係性を調べ、さらには声区転換時の発声データのスペクトログラム分析などを行った。その結果、母音/i/の場合では、声区転換時における第1フォルマント周波数と基本周波数の近接が見られたこと、また声区転換時の発声がより不安定であったことなどから、音源-フィルタ相互作用の影響、すなわち音響由来のメカニズムが主要な要因として声区転換を引き起こしていることがわかった。このような声区転換は一部の母音/a/の場合でも見られ、声区転換を生じる際に第2倍音や第3倍音が第1フォルマント周波数に近接した場合でも、音響由来のメカニズムによる声区転換を生じることがあることを示した。

7.1.2 音源が声道に対して影響を及ぼす影響について

第5章では、発声測定実験を通して、発声条件の違いが声道のフォルマントに与える影響について検討した。発声時において、声門開放率と最大声門面積が異なると想定される4つの異なる発声条件のもとで、外部音響励振法を用いて声道のフォルマントを測定し、分析した。その結果、第1フォルマント周波数と第2フォルマント周波数については、想定される声門面積が大きいほどフォルマント周波数は上昇し、また声門開放率が大きいほどフォルマントのピークはなだらかになる傾向がみられた。一方で、喉頭腔共鳴によって誘導されたと考えられる第4フォルマントは、想定される声門面積が上昇するにつれて周波数が上昇しながら減衰した。したがって、声道形状の違いだけでなく、最大声門面積や声門開放率といった発声条件の違い

によっても、声道音響特性に変化を生じる可能性が示された。

第6章では、音源-フィルタ相互作用の影響を考慮した音声生成シミュレーションを通して、声門の境界条件が声道のフォルマントに与える影響について検討した。声道・声門・声門下部を連結した音響管モデルと、声門面積の多項式モデルを組み合わせた音声生成モデルによって、日本語5母音の持続音声を合成した。音声合成時の発声パラメータとして、声門開放率と最大声門幅を様々に変化させて与え、これらの声門の境界条件が、声道のフォルマントに対してどのように影響を及ぼすのかを調べた。結果より、声門の境界条件がフォルマントに与える影響は、声道断面積関数によって異なること、第1フォルマント周波数はすべての母音で周波数が上昇すること、声門の境界条件では、最大声門幅のほうが声門開放率よりもフォルマント周波数の変化により支配的であることが見られた。これらのフォルマント周波数の変化は、声道が開管共鳴に近づいたために生じていることや、声門下部のインピーダンスのピークの影響をより強く受けるために生じていることが示唆された。

7.2 今後の展望と課題

7.2.1 歌唱音声の音声生成メカニズムの解明と応用

第1章でも述べたように、多くの歌唱音声は特有の発声法によって生成され、特有の音響的特徴を持つ。その中には、楽器の音量に負けないような非常に大きな声量を伴うもの、歌声の基本周波数が1 kHzを超え、あたかもホイッスルを吹いているように聴こえるもの、意図的にサブハーモニクスを多分に含ませることで、歪んだように聴こえるものなど、従来の音源-フィルタ理論では十分に説明できないような歌唱音声が存在する。実際に、Titzeら [13] は伸びやかな高音を大きな声量で歌う“ベルティング唱法”について、音源-フィルタ相互作用の影響を考慮した音声生成モデルによるシミュレーション実験を行っており、多くの歌唱音声の生成メカニズムにも、同様に音源-フィルタ相互作用の影響が寄与していると考えられる。また、これらの歌唱音声の発声法は完全に確立されていないものも多いため、もしこれらの歌唱音声の生成メカニズムを解明することができれば、その歌唱技法の体系化にも寄与できるだろう。更に、近年流行している歌唱音声合成技術に対して応用することによって、よりリアルで豊かな表現の歌唱音声合成を行えるようになるだろう。

7.2.2 非線形現象としての音源-フィルタ相互作用

声区転換ではピッチが周波数ジャンプを示すことをこれまでに述べたが、このジャンプを生じる周波数は、ピッチを上昇させる系列と下降させる系列とでは必ずしも一致せず、一種のヒステリシスを示す。また、音源-フィルタ相互作用の影響が強まった場合には、周波数ジャンプに加えてピッチの分岐やサブハーモニクスなども生じることを示した。特にサブハーモニクスは、上述したように歌唱音声の技法としても用いられることが多い。このようなヒステリシスや分岐現象、サブハーモニクスは、非線形現象において特有のものであり、線形モデルである音源-フィルタ理論では説明することができない。これらの音声生成に見られる様々な非線形現象と、音源-フィルタ相互作用を結びつけて検討することができれば、音声生成のメカニズムについて新たな観点から検討することができるだろう。

謝辞

本研究は、筆者が九州大学大学院芸術工学府芸術工学専攻に在学中に、九州大学大学院芸術工学府・音声情報研究室において、2014年より約3年間に渡って行ったものです。

本研究を遂行するにあたり、博士後期課程の3年間にわたりご指導いただきました、九州大学大学院芸術工学研究院の鏑木時彦教授には、深く感謝し、厚く御礼を申し上げます。ご多忙の中にも関わらず、私の研究活動全般において大変熱心にご指導・ご助言いただき、また学内外での測定実験や、国内外での学会における研究発表、学会誌論文の作成に至るまで、学部時代から数えて6年間もの長きにわたり、研究者として非常に恵まれた機会を数多く与えていただきました。

九州大学大学院芸術工学研究院の岩宮眞一郎教授には、本論文の副査を快く引き受けていただき、専門的な観点から本論文の論点や考察など、非常に有益なご指摘・ご助言をいただきました。深く感謝いたします。

九州大学大学院芸術工学研究院の高田正幸准教授には、本論文の副査を快く引き受けていただき、本論文を作成するにあたって、細部にわたる非常に的確なご指摘・ご助言をいただきました。深く感謝いたします。

九州大学大学院芸術工学研究院の若宮幸平助教には、同じ音声情報研究室に所属している教員として、日頃の研究活動やゼミなどにおいて、これまでに様々なご指導・ご助言を数多くいただき、大変お世話になりました。深く感謝いたします。

私が学部の頃から6年間にわたり所属した音声情報研究室において、同じ月日を過ごした学生の皆様には、日頃の学生生活から研究活動に至るまで、大変お世話になりました。特に、松田祥弘さん、古川貴博さん、安藤百代さんには、歌唱音声という私の研究テーマに共通で携

わった者として、日々活発な議論をさせていただきました。また、儀保伸吾さん、岡田昌大さんには、音声とは異なる分野を専門としながらも、私の研究に対して非常に興味を持って、様々な議論をさせていただきました。さらに、清水拓磨さん、大黒貴弓さんには、私が行った測定実験のデータをもとにして、素晴らしい論文を作成していただきました。そして、ご多忙の中にもかかわらず、私が行ってきた様々な測定実験の被験者を快く協力してくださった皆様にも、深く感謝いたします。

最後に、学部から博士後期課程に至るまでの9年間もの私の大学生活を、はるか遠方より金銭面・精神面から日々支え続けていただき、また、常日頃から私の研究活動を熱心に応援してくださった、私の両親と姉弟、祖父母、親戚一同に、心から深く感謝し、厚く御礼を申し上げます。

参考文献

- [1] G. Fant, *Acoustic theory of speech production with calculations based on x-ray studies of Russian articulations*, Second printing, Mouton, The Hague (1970).
- [2] J. Sundberg, *The science of the singing voice*, Northern Illinois University Press, DeKalb, (1987).
- [3] I. R. Titze, “Nonlinear source-filter coupling in phonation: Theory,” *The Journal of Acoustical Society of America*, **123**, 2733–2749 (2008).
- [4] K. Ishizaka and J.L. Flanagan, “Synthesis of voiced sounds from a two-mass model of the vocal cords,” *Bell Syst. Tech. J.*, 51(6), 1233-1268 (1972).
- [5] J. L. Flanagan, *Speech Analysis Synthesis and Perception*, 2nd ed. Springer Verlag, New York, 36–38 (1972).
- [6] A. Löfqvist, B. Carlborg and P. Kitzing, “Initial validation of an indirect measure of subglottal pressure during vowels,” *The Journal of the Acoustical Society of America*, **72**, 633–635 (1982)
- [7] Jw. van den Berg, W. Vennard, D. Burger and C. C. Shervanian, *Voice Production: The Vibrating Larynx*, Instructional film, University of Groningen, The Netherlands (1960).
- [8] J. G. Švec, H. K. Schutte and D. G. Miller, “On pitch jumps between modal and falsetto registers in voice: Data from living and excised human larynges,” *The Journal of Acoustical Society of America*, **106**, 1523–1531 (1999).
- [9] D. G. Miller, J. G. Švec and H. K. Schutte, “Measurement of characteristic leap interval between modal and falsetto registers,” *Journal of Voice*, **16**, 8–19 (2002).

- [10] I. Tokuda, J. Horacek, J.G. Švec and H. Herzel, “Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments,” *The Journal of Acoustical Society of America*, **122**, 519-531 (2008).
- [11] I. T. Tokuda, M. Zemke, M. Kob and H. Herzel, “Biomechanical modeling of register transitions and the role of vocal tract resonators,” *The Journal of Acoustical Society of America*, **127**, 1528–1536 (2010).
- [12] T. Kaburagi, “Voice production model integrating boundary-layer analysis of glottal flow and source-filter coupling,” *The Journal of Acoustical Society of America*, **129**, 1554–1567 (2011).
- [13] I. R. Titze, and A. S. Worley, “Modeling source-filter interaction in belting and high-pitched operatic male singing” *The Journal of Acoustical Society of America*, **126**, 1530–1540 (2011).
- [14] D. A. Berry, H. Herzel, I. R. Titze and B. H. Story, “Bifurcations in excised larynx experiments,” *Journal of Voice*, **10**, 129–138 (1996).
- [15] I. R. Titze, T. Riede and P. Popolo, “Nonlinear source-filter coupling in phonation: Vocal exercises,” *The Journal of Acoustical Society of America*, **123**, 1902–1915 (2008).
- [16] M. Zaňartu, D. D. Mehta, J. C. Ho, G. R. Wodicka and R. E. Hillman, “Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study,” *The Journal of Acoustical Society of America*, **129**, 326–339 (2011).
- [17] J. Epps, J.R Smith and J. Wolfe, “A novel instrument to measure acoustic resonances of the vocal tract during phonation,” *Measurement Science and Technology*, **8**, 1112–1121 (1997).
- [18] E. Joliveau, J. R. Smith and J. Wolfe, “Vocal tract resonances in singing: The soprano voice,” *The Journal of Acoustical Society of America*, **116**, 2434–2439 (2004).

- [19] Y. Swerdlin, J. R. Smith and J. Wolfe, “The effect of whisper and creak vocal mechanisms on vocal tract resonances,” *The Journal of Acoustical Society of America*, **127**, 2590–2598 (2010).
- [20] J. M. Chen, J. R. Smith and J. Wolfe, “Do trumpet players tune resonances of the vocal tract?,” *The Journal of Acoustical Society of America*, **131**, 722–727 (2012).
- [21] N. Henrich, J. R. Smith and J. Wolfe, “Vocal tract resonances in singing: Variation with laryngeal mechanism for male operatic singers in modal and falsetto registers,” *The Journal of Acoustical Society of America*, **135**, 491–501 (2014).
- [22] N. Henrich, C. d’Alessandro, B. Doval and M. Castellengo, “On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation,” *The Journal of Acoustical Society of America*, **115**, 1321–1332 (2004).
- [23] J.C. Lucero, K. G. Lourenço, N. Hermant, A. Van Hirtum and X. Pelorson, “Effect of source–tract acoustical coupling on the oscillation onset of the vocal folds,” *The Journal of Acoustical Society of America*, **132**, 403–411 (2012).
- [24] M. Echternach, J. Sundberg, S. Arndt, T. Breyer, M. Markl, M. Schumacher and B. Richter, “Vocal tract and register changes analysed by real-time MRI in male professional singers - A pilot study,” *Logopedics Phoniatrics Vocology*, **33**, 67–73 (2008).
- [25] U. Hoppe, F. Rosanowski, M. Döllinger, J. Lohscheller, M. Schuster and U. Eysholdt, “Glissando: laryngeal motorics and acoustics,” *Journal of Voice*, **17**, 370–376 (2003).
- [26] N. Henrich, C. d’Alessandro, B. Doval and M. Castellengo, “Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity and fundamental frequency,” *The Journal of Acoustical Society of America*, **117**, 1417–1430 (2005).
- [27] T. Kitamura, H. Takemoto, S. Adachi, P. Mokhtari and K. Honda, “Cyclicality of laryngeal cavity resonance due to vocal fold vibration,” *The Journal of Acoustical Society of America*, **120**, 2239–2249 (2006).

- [28] 竹本浩典, 北村達也, パーハム・モクタリ, 足立整治, 本多清志, “声門開放時の声道共鳴特性,” 電子情報通信学会技術研究報告, SP2006-165, 31-36 (2007).
- [29] H. Takemoto, S. Adachi, T. Kitamura, P. Mokhtari and K. Honda, “Acoustic roles of the laryngeal cavity in vocal tract resonance,” *The Journal of Acoustical Society of America*, **120**, 2228-2238 (2006).
- [30] A. Barney, A. De Stefano and N. Henrich, “The Effect of Glottal Opening on the Acoustic Response of the Vocal Tract,” *Acta Acustica united with Acustica*, **93**, 1046-1056 (2007).
- [31] A. Rosenberg, “Effect of the glottal pulse shape on the quality of natural vowels,” *The Journal of Acoustical Society of America*, **49**, 583-590 (1971).
- [32] M. Sondhi and J. Schroeter, “A hybrid time-frequency domain articulatory speech synthesizer,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, **35**, 955-967 (1987).
- [33] B. H. Story, I. R. Titze and E. A. Hoffman, “Vocal tract area functions from magnetic resonance imaging,” *The Journal of Acoustical Society of America*, **100**, 537-554 (1996).
- [34] J. van den Berg, “An electrical analogue of the trachea, lungs and tissues,” *Acta physiologica et pharmacologica neerlandica*, **9**, 361-385 (1960).
- [35] E. R. Weibel, *Morphometry of the human lung*. Springer-Verlag, New York (1965).
- [36] 籙木時彦, 音声生成の計算モデルと可視化, (社)日本音響学会(編), コロナ社, 東京 (2010).
- [37] V. W. Kahle, H. Leonhardt and W. Platzer 著, 越智淳三訳, 解剖学アトラス (第3版), 文光堂, 東京 (1998).