

電子科学研究科

GD

K

0002512937

R

75

静岡大学附属図書館

静岡大学 博士論文

軟口蓋と声道壁の振動を考慮した音声生成機構の  
音響モデルに関する研究



平成4年5月

大学院電子科学研究科

電子応用工学専攻

党 建武

## 内容梗概

本論文は、軟口蓋と声道壁の振動を考慮した音声生成機構の音響モデルに関する研究を論ずるものである。この研究では、発声器官各部位の音響放射特性と声道壁の振動特性の測定結果に基づいて、軟口蓋の動きと声道壁インピーダンスを推定して、軟口蓋と声道壁の振動を考慮した音響モデルを構築した。

最初に、発声器官各部位からの放射音を分離して測定する音響測定システムと、その導入により音響測定に与える影響について述べた。測定装置に関する予備実験の結果より、本研究で用いた測定システムは音声の音響測定に十分有効であることがわかった。

本研究の内容は、大別して、軟口蓋の動きの推定と声道壁インピーダンスの測定の2つの部分となる。まず、軟口蓋の動きの推定とモデル化について述べる。

従来の音響モデルにおいては、軟口蓋がONとOFFの二状態のスイッチで切替えるものとされているが、実際の調音過程は単なる切替えではない。軟口蓋は有限の機械インピーダンスを持つ粘弾性体なので、口腔内圧により振動する。非鼻音化音声においても軟口蓋の振動により鼻腔と口腔とが音響的に結合していると考えられる。鼻音化音声では軟口蓋の開閉により開口面積が連続的に変化する。そのゆえ、本研究では、音声発声時の軟口蓋の動きを非鼻音化母音、非鼻音化子音と鼻音化音声の3つの側面から考察して現実に近い軟口蓋の音響モデルを構築した。

非鼻音化母音の考察では、発声器官各部位からの放射音を口唇放射、鼻孔放射と咽頭壁放射に分離して測定した。測定結果により、非鼻音化母音においては、口唇開口部以外の鼻孔や咽頭壁などからも音が放射されていることを明らかにした。口唇放射では狭母音が広母音より小さいが、鼻孔放射では狭母音が広母音より大きい。このような音響現象を口腔内音圧につれて軟口蓋の振動による口腔と鼻腔の音響結合であると考えて、粘弾性体からできている軟口蓋をばねとダッシュポットでつながる二枚の振動板としてモデル化した。シミュレーションの結果から、軟口蓋の振動板モデルを用いて非鼻音化母音における音響放射をよく説明できることがわかった。

非鼻音化子音、特に有声破裂子音では、口腔内圧が大幅に変動している。この変動による軟口蓋の状態がどのように変化しているかを、鼻孔放射と口腔内圧の関係をを用いて推定した。推定結果では、有声子音区間で軟口蓋の軟らかさが変化しており、この変化には受動的な成分と能動的な成分があることを示した。軟口蓋の受動的な変化

を定式化して、上述の振動板モデルに導入した。さらに、従来有声子音バズ区間における咽頭腔容積の測定データに基づいて、咽頭腔容積の能動的な増大を定式化して、有声破裂子音の音響モデルを作成した。有声破裂子音を発声する時鼻腔の音響的な結合の効果について聴取実験を行ったところ、このモデルで合成した音声は実音声とほぼ同じ結果を示した。

鼻音化音声の考察では、口腔内音圧と放射音の関係を用いて、軟口蓋開口面積の等価インピーダンスと軟口蓋振動の等価インピーダンスの関係を考察した。さらに、それらの関係を用いて軟口蓋の等価開口面積率を推定した。

上述の3つの側面での考察結果をまとめて、連続音声における軟口蓋の音響モデルを提案した。このモデルでは、非鼻音化音声の場合は軟口蓋が有限の機械インピーダンスを持つ振動板として働き、鼻音化音声の場合は鼻音化の度合により軟口蓋の開口面積が連続的に変化している。鼻音化の度合が大きい場合は、主に軟口蓋の開口面積による結合が働くが、鼻音化の度合が小さい場合は、軟口蓋の開口面積による結合が軟口蓋の振動による結合と同時に働く。

次に、声道壁インピーダンスの推定について述べる。この推定では、2つの方法を用いた。一つの方法は直接測定方法である。本論文では、音声発声時の声道壁の振動と声道壁からの音響放射は声道内音圧によって引き起こされたものであることに基づいて、従来とは異なる方法を提案した。実測した頬のインピーダンスでは、単位面積あたりの機械抵抗は高周波領域でかなり大きくなり、単位面積あたりの質量とステイフネスは周波数と共に若干変動するがほぼ一定の値である。従来提案した値に比べて、本測定方法で得られた値は十分に信頼できる。

もう一つの方法は合成音のパラメータ整合法である。この方法では、従来は母音を音声資料として使っているが、声道壁の母音に対する効果が必ずしも大きくない。それに対して、有声破裂子音のバズ区間の音響特性はほとんど声道壁の特性に決められる。それゆえ、本論文では、母音の代わりに有声破裂子音のバズ区間の音響特性を用いて、声道壁インピーダンスを推定した。求めた単位面積あたりの質量とステイフネスは上述の測定結果とほぼ一致している。機械抵抗については、推定値と高周波領域での実測値はほぼ一致した。

# 目次

第1章 序論	
1. 1 まえがき	1
1. 2 音声生成と音声合成に関する研究の流れ	1
1. 3 本研究の目的と内容	3
1. 4 音声器官の機構	5
1. 5 言語音の分類	7
第2章 測定システムの構成	
2. 1 まえがき	9
2. 2 測定装置	9
2. 3 防音箱と遮蔽板の音響特性	12
2. 3. 1 防音箱の共振周波数	12
2. 3. 2 防音箱の透過損失	13
2. 3. 3 遮蔽板の透過損失	14
2. 4 測定装置に対する予備実験	15
2. 4. 1 防音箱の残響時間の推定	15
2. 4. 2 口唇放射音に対する防音箱内の影響	16
2. 4. 3 鼻孔などの放射音に対する防音箱外壁の影響	17
2. 5 マイクロホン出力の校正	18
2. 5. 1 マイクロホンの特性	18
2. 5. 2 マイクロホン出力の校正	19
2. 6 加速度ピックアップについて	21
2. 7 むすび	22
第3章 非鼻音化母音における音響特性と軟口蓋の振動板モデル	
3. 1 まえがき	23
3. 2 放射音圧と振動量の測定システム	23
3. 3 発声器官各部位の放射音と振動加速度の比較	24
3. 3. 1 放射音と振動加速度の波形	24
3. 3. 2 放射音と振動加速度のスペクトル	27
3. 3. 3 壁の振動量と放射音圧	28

3. 4	母音における各部位の放射音圧レベル	29
3. 5	音声生成の声道モデル	37
3. 6	軟口蓋の振動板モデル	40
3. 6. 1	軟口蓋振動の等価回路	41
3. 6. 2	振動板モデルのパラメータの推定	42
3. 6. 3	シミュレーションの結果	43
3. 7	考察	45
3. 8	むすび	47
第4章 子音区間での音響測定と軟口蓋の動きの推定		
4. 1	まえがき	48
4. 2	測定装置と音声資料	48
4. 2. 1	測定装置	48
4. 2. 2	音声資料	49
4. 3	子音区間における放射音の音圧レベル	49
4. 3. 1	音圧レベルの求め方	49
4. 3. 2	単音圧の放射音圧レベル	50
4. 3. 3	連続音声におけるバズ区間の放射音圧	58
4. 4	口腔内圧と放射音	59
4. 4. 1	口腔内圧と放射音の波形	59
4. 4. 2	鼻孔放射音と鼻孔空気流	59
4. 4. 3	無声子音における音響放射	61
4. 4. 4	有声破裂子音の音響放射と鼻子音の音響放射	63
4. 5	軟口蓋の動きの推定	65
4. 5. 1	軟口蓋の受動的な変化	65
4. 5. 2	軟口蓋の能動的な変化	67
4. 5. 3	子音区間における軟口蓋の動きと音響放射	68
4. 6	有声破裂子音における音響放射のシミュレーション	68
4. 6. 1	有声子音区間における軟口蓋の音響モデル	68
4. 6. 2	シミュレーションの結果	70
4. 6. 3	聴取実験の結果	73
4. 7	むすび	75
第5章 鼻音化音声における音響放射の測定および軟口蓋の等価開口面積率		

の推定		
5. 1	まえがき	76
5. 2	測定装置と音声資料	76
5. 2. 1	測定装置	76
5. 2. 2	音声資料	77
5. 3	言語音声の文脈による放射音の変化	77
5. 3. 1	単音節発声について	78
5. 3. 2	連続音声について	81
5. 4	口唇放射と鼻孔放射の周波数特性	82
5. 4. 1	鼻孔放射と口唇放射の比較	82
5. 4. 2	口唇放射と鼻孔放射の極零	84
5. 4. 3	ホルマントに対する鼻孔放射の影響	86
5. 5	軟口蓋の等価インピーダンス	88
5. 5. 1	軟口蓋の等価インピーダンスと放射音圧	89
5. 5. 2	軟口蓋開口面積の等価インピーダンスと軟口蓋 振動の等価インピーダンス	91
5. 6	軟口蓋の等価開口面積率	92
5. 6. 1	軟口蓋の等価開口面積率の定義	92
5. 6. 2	軟口蓋の等価開口面積率の推定結果	93
5. 7	連続音声における軟口蓋の音響モデル	96
5. 8	むすび	97
第6章	声道内音圧及び頬の振動加速度による頬インピーダンスの測定	
—	および合成音のパラメータ整合法で声道壁インピーダンスの推定	
6. 1	まえがき	98
6. 2	測定装置	99
6. 3	測定等価回路とインピーダンスの推定方法	101
6. 3. 1	測定等価回路	101
6. 3. 2	インピーダンスの推定方法	102
6. 4	予備実験	103
6. 4. 1	機械工学的な方法での測定	104
6. 4. 2	内圧と壁の振動による測定	105
6. 5	測定結果と頬インピーダンスの推定	106

6. 5. 1	測定した音圧と振動量	106
6. 5. 2	頬インピーダンスの推定	110
6. 6	測定結果の考察	111
6. 7	合成音のパラメータ整合法による声道壁インピーダンスの推定	115
6. 7. 1	実音声の音響特性	115
6. 7. 2	声道壁インピーダンスの推定	115
6. 8	むすび	125
第7章	結論	126
付録 A	引張り試験	130
付録 B	等価質量と等価ステイフネスから静的な値への変換	131
付録 C	データベース「MCSDB」について	134
参考文献		143
論文目録		148
謝辞		150

# 第1章 序論

## 1.1 まえがき

音声は人間の間のもっとも基本的な情報伝達の手段である。文字による情報伝達は記録性において音声より優れているが人間相互間の情報の交換において音声は他の手段にはない特徴を持っている。音声の特徴は、他の手段より速いことや、他の仕事と並列的に処理できることや豊富な情報が含まれることなどである。音声の情報には大別して2種の情報内容が含まれている。一つは意味内容に関する情報で、話し手の意志や、指示を伝えるための基本的な情報である。もう一つは発声者に関する情報で、発声者の性別、年齢、名前や、発声時の感情、気分など種々の情報が含まれる。このような音声の特徴を十分に利用するため、音声は古くから生成・知覚・合成・認識等の諸分野で研究されている。

## 1.2 音声生成と音声合成に関する研究の流れ

音声生成と音声合成に関する研究は長い歴史を持っている。今日記録に残っている音声合成器は、1779年のKratzensteinの母音合成器と、1791年のVon Kempelenの音声合成器であった<sup>1)</sup>。それらの構造は機械的であるが、いずれも人間の音声器官の構造や動作を模擬したものであった。

1876年電話が発明されて以後、通信媒体としての音声の特性を解明して、その特性に整合の取れた通信システムを実現することを目標とする新しい音声音響学の展開が始まった<sup>2)</sup>。1939年には、H. DudleyによってVocoderが発表されて、音声音響学における音声情報処理の研究の基礎が開かれることになった<sup>2)</sup>。また、千葉ら(1941)は母音に関して音響・生理・心理の各側面から卓越した研究を行った<sup>3)</sup>。G. Fant(1960)はこれを継承して調音と音声の関連をさらに明確にし、音声生成の音響理論を確立した<sup>4)</sup>。これらの古典的な業績は、現在の音声研究の基礎を築いたものとして、高く評価されている。

1960年代にはコンピュータの実用化に伴って、音声の信号処理にデジタル技術が広く用いられるようになり、今日のデジタル信号処理の基礎ができあがった。音声

の分析と合成に線形予測を直接応用した研究は、斉藤と板倉（1966）およびAtalとSchroeder（1967）によって行われた<sup>5)</sup>。1970年代以後、板倉らはPARCOR（Partial Auto-Correlation）方式<sup>6)</sup>、LSP（Line Spectrum Pair）方式<sup>7)</sup>等実用的なアルゴリズムを開発した。その後、これらのアルゴリズムはLSI化され、たくさんの音声応答システムに用いられている。しかしながら、音声合成技術という点から考えてみると、現在実用の段階に入っているのは主として自然音声を素材にして、そのまま録音したもの、またはそれらの編集ないし分析合成したものである。

それに対して、人間の音声の録音によらずに人間の発声器官のうち共鳴系を形作っている声道を音響管として作り、声帯音源に代わる適当な駆動源でこの音響管を音響的に励振して音声を生成する方法に関する研究も数多く行われてきた。千葉ら（1941）<sup>8)</sup>、平松ら（1964）<sup>9)</sup>はX線写真撮影法を用いて発声時の声道パターンを測定して声道断面積を推定した。Fant（1960）<sup>4)</sup>は鼻子音を発声する時の声道（口腔と鼻腔）断面積をX線写真で推定した。竹内ら（1977）<sup>10)</sup>は頭蓋標本により各副鼻腔の形状や容積を推定した。また、駆動源としての声帯に対しても色々な研究を行ってきた。Sawashimaら（1968）<sup>11)</sup>はファイバーコープを鼻腔を通じて咽頭腔に挿入して撮影することによって声帯の振動を直接観測した。その後、類似的な方法を用いて、高速撮影での観測がいくつか行われてきた。石坂ら（1972）<sup>12)</sup>は以前の研究に基づいて声帯の二質量振動モデルを提案した。

一方、音声波形から声道形を推定する研究では、Wakita（1973）<sup>13)</sup>および白井ら（1978）<sup>14)</sup>は調音器官の構造に基づいて、声道系を表現する調音モデルを設定し、音声波からモデルマッチングの手法によって、調音状態を推定した。

実用化の程度からみると、音声の生成過程の研究はデジタル信号処理方式の音声合成に比べかなり遅れている。しかし、音声信号を確率過程と見なし、そのパラメータを推定し、あるいはパラメータから信号を再合成する方法や、波形を統計的に処理する方法では、個人差のあらゆる状態を再現する適切なモデルとなりうるかどうか疑問であり、適応可能の領域が限られた範囲にあると考えざるをえない。音声生成系を物理的実体に対応した形でモデル化できるならば、人間の発声メカニズムや会話の科学的理解に役立つばかりなく、音声合成のためのシミュレーションとして、合成技術の発展にも寄与できる。それゆえ、音声生成の音響モデルをもう一度新たに検討するというモチベーションも強まった。この数年間、超音波を用いて舌の調音運動の観測<sup>15)</sup>や、高速映画撮影法によって調音時の口唇運動の特徴の解析<sup>16)</sup>、韻律制御における喉頭筋の協調的活動の測定<sup>17)</sup>と声道壁インピーダンスの推定<sup>18)</sup>・<sup>19)</sup>など音声生

成過程の調音器官に関する研究は行われてきている。しかし、これまでの研究による多くの観測結果は音声生成モデルには具体化されていない。特に、音声生成過程の静的な側面と比べて、その動的な特性の観測と定式化には、多くの問題が残されている。

### 1. 3 本研究の目的と内容

音声を発声する時は音が一つの出口だけでなく発声器官のいくつかの出口から同時に放射される。音の主としての出口は、母音では口唇、鼻子音では鼻孔、有声子音のバズバーでは声道壁、鼻音化音声では口唇と鼻孔の双方である。このように、発声器官各部位からの放射音は言語音声に種々の音色を与えている。しかし、それらの放射音の音響特性と割合はまだ明らかにされていない。自然度が高く個人情報が豊富な音声を生成するために、発声器官の放射機構と各放射音の音響特性を研究する必要がある。本研究では、発声時の放射音を口唇放射、鼻孔放射と声道壁放射に分離して測定する。これらの放射音の音響特性により発声器官の動きを推定する。さらに、軟口蓋と声道壁の振動を考慮した音声生成機構の音響モデルを作成する。

従来の音声生成モデルでは、軟口蓋は鼻音化音声と非鼻音化音声の場合、ONとOFFの二状態のスイッチで切替えられるものとされている。実際の調音過程は単なる切替えではない。鼻音化音声においては、鼻音化の度合により軟口蓋の開口面積が連続的に変化している。非鼻音化音声の場合、口腔と鼻腔の間の隔壁の一部である軟口蓋は有限の機械インピーダンスを持つ粘弾性体なので閉鎖していてもそのインピーダンスにより口腔と鼻腔が結合すると考えられる。現実に近い軟口蓋の音響モデルを構築するため、非鼻音化母音、有声破裂子音および鼻音化音声の3つの側面から軟口蓋の働きを考察する。また、音声生成モデルには声道壁の振動を考慮するため、本論文では、従来の測定方法と異なる新しい測定方法を提案して、音声発声に近い状態で声道壁インピーダンスを推定する。

本論文は次のような構成である。

第2章「測定システムの構成」では、発声器官各部位の放射音を分離するために用いた防音箱と遮蔽板などの測定装置を紹介する。防音箱と遮蔽板を導入することによる発声器官各部位からの放射音の分離の度合、音響測定に与える影響およびこれらの装置の音響測定に対する有効性について、予備実験で確認する。

第3章「非鼻音化母音における音響特性と軟口蓋の振動板モデル」では、非鼻音化音声を発声する時の口唇放射、鼻孔放射と咽頭壁放射を分離して収録し、同時に各放

射部位付近の声道壁の振動加速度を観測する。これらの放射音と振動量の関係を調べる。非鼻音化音声における軟口蓋の働きを推定するために、口唇放射と鼻孔放射を母音の調音特徴と結び付け、この場合の鼻孔放射を軟口蓋の振動による口腔と鼻腔間の音響結合として考察する。軟口蓋の解剖学的な特徴を考慮した上で軟口蓋をばねとダッシュポットでつながる二枚の板としてモデル化し、口唇放射と鼻孔放射の関係をを用いて軟口蓋の振動板モデルのパラメータを推定する。

第4章「子音区間での音響測定と軟口蓋の動きの推定」では、有声破裂子音の音響特性に基づいて、バズ区間で口腔内圧の大幅な変化による軟口蓋の軟らかさの変化を考察する。軟口蓋の受動的な変化を、口腔内圧と鼻孔放射で定式化する。また、軟口蓋の軟らかさの変化を軟口蓋の振動板モデルに入れ、同時に咽頭腔容積が能動的に増大する機構を導入して有声破裂子音の音響モデルを作る。さらに、有声破裂子音の音響特性の表現に対するこのモデルの有効性を聴取試験で確認する。

第5章「鼻音化音声における音響放射の測定および軟口蓋の等価開口面積率の推定」では、まず単音節と連続音声における言語文脈による放射音の変化を考察する。次に、分離して収録した口唇放射と鼻孔放射の周波数特性を分析する。さらに、上述の観測結果に基づいて軟口蓋開口面積の等価インピーダンスと軟口蓋振動の等価インピーダンスの関係を考察する。軟口蓋の等価開口面積率を放射音圧と結び付けて、鼻音化音声における軟口蓋の等価開口面積率を推定する。

第6章「声道内音圧及び頬の振動加速度による頬インピーダンスの測定 — および合成音のパラメータ整合法で声道壁インピーダンスの推定」では、音響放射モデルに必要な声道壁のインピーダンスを測定または推定する。音声発声時、声道壁の振動と声道壁からの放射音は声道内音圧によって引き起こされたものである。このことに基づいて新しい測定方法を提案する。この測定方法を用いて声道壁インピーダンスを測定する。また、実音声の有声破裂子音バズ区間の口腔内音圧、鼻孔放射と声道壁放射を用いて合成音パラメータ整合法で声道壁インピーダンスを推定する方法と推定した結果についても述べる。

第7章「結論」は本研究のまとめである。本研究で用いた測定システムの有効性と、いくつかの測定結果、およびそれらの結果の定式化またはモデル化について述べる。また、今後の課題についても述べる。

#### 1. 4 音声器官の機構<sup>20)</sup>

音声の生成過程の研究は人間の発声器官と調音過程を着目するものである。本節では、音声生成に関与する諸器官、特に本研究に関する発声器官の構造および機能を音響的な面から概説する。

言語音声の生成過程には、呼気、発声と調音という三つの過程がその重要な生理的要素として含まれている<sup>20)</sup>。呼吸運動とそれにしばしば伴う発声とは、音声生成のエネルギー源となり、または調音系への入力信号を与えるものであり、器官としては主に喉頭以下の呼吸器官である。発声の場合には、肺からの空気圧（声門下圧）が高くかつ声帯の緊張が大きいと、声帯の開閉周期、すなわち振動周期が短くなって、音源の音の高さが高くなり、逆のときは低くなる。これが声の高さ（ピッチ）に対応する。声帯の振動周期のことを基本周期、この逆数を基本周波数と呼ぶ。このように、肺から生じた呼気流が声門を通過する際、準周期的な気流に変化され音源信号として調音系に入る。調音というのは、さまざまな言語音を発声するために声道の形を調整することであり、主に喉頭から上の調音器官によって構成制御される音響系において実現する。

人間の音声器官としては、喉頭、口蓋、顎、舌、唇などがあり、その解剖学的位置の関係を図1. 1に示す。それらは全体として一つの連続した管となっている。喉頭より上の部分は声道と呼ばれ、顎、舌、口唇などを動かすことより種々の形に変化する。鼻腔は共鳴腔の一部であり、口腔の上側にあつて、口腔とは口蓋で境されている。鼻腔は軟口蓋を持ち上げることにより咽頭や口の奥から遮断される。

本研究では、非鼻音化音声と鼻音化音声の口唇と鼻孔及び咽頭壁の音響放射に関する測定を行って、各部位の放射音により、軟口蓋の動きと声道壁インピーダンスを推定する。ここで、軟口蓋と鼻腔の生理学的な特徴を述べる。

口蓋は、鼻腔、上咽頭と口腔の境を作っている。口蓋の前の約3分の2の部分は硬口蓋である。硬口蓋の粘膜の下は頭蓋に固定されている骨の板である。口蓋の後3分の1は軟口蓋であり、粘膜の下は筋肉である。粘膜に覆われた筋肉板を口蓋帆と呼ぶ。口蓋帆挙筋は口蓋帆を後上方に引き上げ、中咽頭から上咽頭、すなわち口腔から鼻腔を遮断する。安静呼吸の状態では、軟口蓋は下がっているので口腔と鼻腔の間には空気通路がある。

鼻腔は左右の外鼻孔によって外界と連絡している腔室であり、鼻中隔といわれる骨と軟骨の薄板によって左右に分かれている。鼻腔の側壁にはいくつかの凸凹があり、

その形は複雑で、個人差も大きい。また、鼻腔の壁全体を覆う粘膜の状態によって内腔の形や容積はかなり変化する。鼻の解剖学的な構造で固有鼻道以外に数多く容積の小さい副鼻腔が存在する。副鼻腔のうち上顎洞や、前頭洞や、蝶形洞と篩骨洞がある。前の3つは容積が大きい。篩骨洞は数は多いがそれぞれの容積は小さい。しかし、調音運動としては、上咽頭後部が軟口蓋の上下運動によって形を変える以外には変化しない。音声発声時、軟口蓋が垂れ下がると鼻腔と口腔は一つの共鳴形になり、特有な共鳴音を生じる。また、鼻孔は唇のように自由な運動ができなく、鼻腔の中には舌のような運動をする器官がなく、鼻腔自身も外から力を加えない限りそれ自身では形が変わらないので、音声生成の観点から見ると鼻腔部分の研究は比較的単純である。然し、鼻腔の作用を操縦する軟口蓋の運動は他の調音器官の運動よりかなり遅い<sup>20)</sup>ので、鼻腔は鼻子音のみでなく、鼻子音の前後にある音素に直接影響する。そのため、鼻音化音声は言語音声にはさまざまな形で現れている。

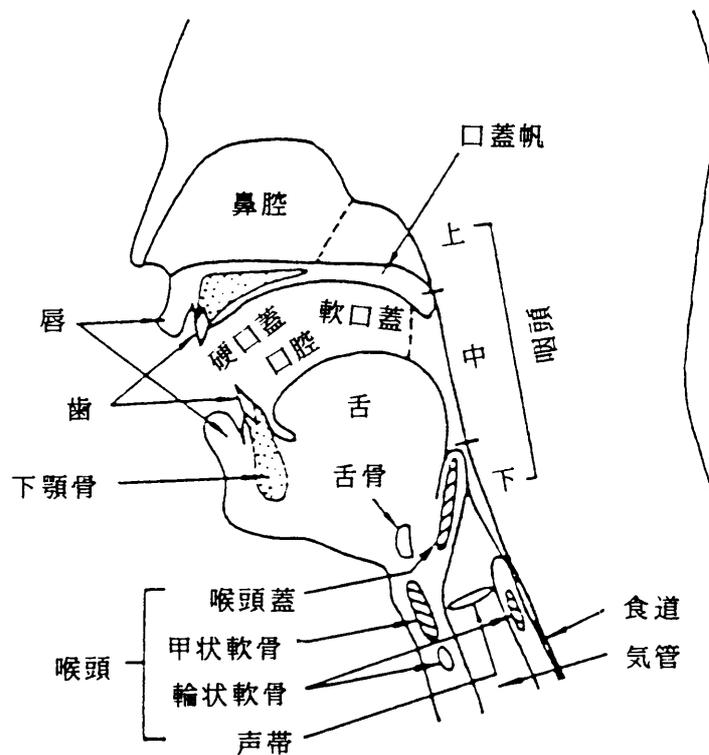


図1. 1 音声器官正中断面図 (文献(20)による)

## 1. 5 言語音の分類<sup>22)</sup>

言語音を大別すると、母音と子音がある。口唇・口腔・咽頭などの声道の途中に閉鎖を持たず、狭めによる雑音性音源を伴わないで、口唇から放射される音は母音という。母音以外の音は子音である。半母音は、母音と子音の中間的な存在であるが、母音と子音に2大別した場合は子音に分類される<sup>21)</sup>。

母音は調音位置で分類すると、前舌母音、中舌母音と後舌母音に分けられる。前舌母音とは、舌による狭めが声道の前（口唇に近い方）にある母音である。同様に、中舌母音と後舌母音では舌による狭めがそれぞれ声道の中部と後部にある。顎の開き程度によって分類すると、母音は狭母音や、半狭母音や、半広母音および広母音に分けられる。母音の分類<sup>22)</sup>を図1. 2に示す。上述した二つの条件と口唇の狭めの長さや狭さを組み合わせると、母音の調音の基本的な性質をほぼ表現することができる。図中、一对の母音が左右に隣接しているのは口唇の調音のみ異なる一組の母音であり、左側が口唇の狭めのないもの、右側が口唇の突き出しを伴うものである。

子音の生成は、声帯の振動の有無とは独立して行われる。声帯の振動を伴う子音を有声子音、伴わない子音を無声子音と呼ぶ。発声の方式によると子音は、摩擦音や、

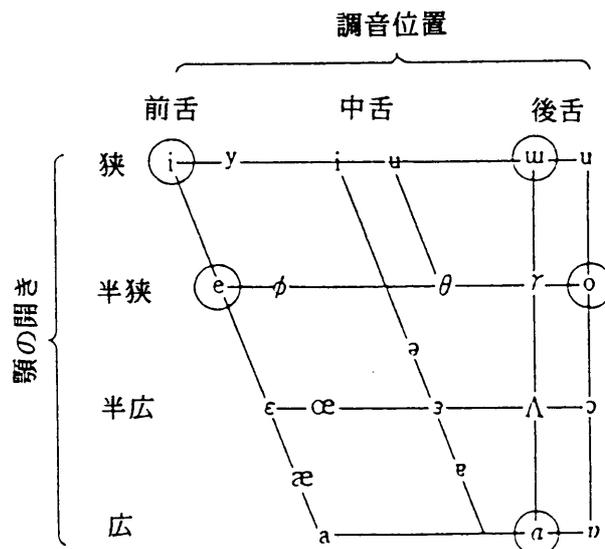


図1. 2 母音の分類 (○：日本語（東京方言）で主として用いられる母音。西欧の言語を対象とした分類であるので、厳密にはいずれも日本語の母音とはやや異なる。)

破裂音や、破擦音、半母音及び鼻音となる。発声の特徴においては、摩擦音は舌によって声道のある部分に狭め場所を作り、そこを空気流が通り抜けるときに、乱流を生じて雑音的な音を生成するものである。破裂音は舌や口唇で声道を遮断することによって空気流が一時的に止め、空気その後方にたまって圧力が十分高まったところで、これを急に開放してインパルス的な音を生成するものである。破擦音は破裂後と同じ場所での強い狭めを定常的に保って摩擦音に移動することにより生成するものである。これは二つの調音が密接に連結したものとみなすことができる。

半母音は、その生成の様式は母音とほとんど同じであるが持続されることなく移動的に変化する特徴がある。鼻音の場合、軟口蓋が下がると共に、口腔のいずれかの位置で気流を遮断することによって、鼻腔にも空気流が供給され、声道に分岐が生じた形になる。母音発声時にも軟口蓋が下がって、鼻腔と口腔と並んで声道の一部を形成することがあり、このときの音を鼻音化音声と称する。日本語の子音の分類を表1.1に示す。ただし、日本語の音節をローマ字で表示するにはヘボン式、訓令式と日本式があるが、本論文以下の部分で子音の符号の表示ではヘボン式を用いる。

表1.1 日本語子音の分類\*

調音位置		口唇		歯、歯茎		口蓋		声門
		有声	無声	有声	無声	有声	無声	無声
調音方式	摩擦音		f <sup>1)</sup> ッ	z ッ <sup>2)</sup>	s ッ	ʃ シ	ʃ シ	h <sup>1)</sup> ハ
	破擦音			dz ッ <sup>2)</sup>	ts ッ	dʒ シャ	tʃ チャ	
	破裂音	b ハ <sup>2)</sup>	p ハ <sup>2)</sup>	d ッ <sup>2)</sup>	t ッ	g カ <sup>2)</sup>	k カ	
	半母音	w ッ		r <sup>2)</sup> ッ		j ッ		
	鼻音 <sup>3)</sup>	m ッ		n ッ		ŋ ッ		

\* 子音の記号は同欄の音節の子音部分を表示するものである。

- 1) 日本語のハ行音は特殊な構造を持つ。
- 2) 日本語のラ行音は弾音として分類される。
- 3) 日本語のンの音は撥音と呼ばれ、環境によって種々に変化する。

## 第2章 測定システムの構成

### 2.1 まえがき

音声発声時の調音器官の働きと調音過程に関する研究には、従来の観測手法として、X線、光、超音波、電気、磁気等を介する間接的な方法と、筋電図、機械的変位、呼気圧、気流等を実測する直接的な方法とがある。これらの方法のほとんどは発声器官の音響放射に直接結び付けていない。人間の発声過程では、各発声器官が調和して動いており、結果としては音響放射の形式で現われる。それゆえ、放射音から調音器官の働きを推定するのは重要な手法の一つである。発声された音素に応じて、主としての発声器官が異なり音響の放射部位も異なるので、放射音波形から調音器官の働きを正確に推定するには各部位の放射音を分離する必要がある。これまでの研究では、竹内ら<sup>1)</sup>は鼻孔放射と口唇放射を分離して測定し、その鼻孔音を用いて副鼻腔の影響を考察した。神内ら<sup>2)</sup>は分離された口唇放射と鼻孔放射を音声認識に使用することを考察した。鼻孔放射と口唇放射を分離するには両研究とも一枚の遮蔽板を用いた。遮蔽板の透過損失は両者とも20dB程度である。一般に、口唇放射は鼻孔放射より大きく、特に低周波領域では遮蔽板の周りから音の回折による影響が大きいので十分に分離することができない。当研究室<sup>3), 4)</sup>では、以前発声時における声道開口部と壁部からの放射音に関する研究として、口唇と口唇以外の放射を分離して測定したが、その測定では声道壁からの放射と鼻孔からの放射は分離しなかった。

音声発声時、音が主に口唇、鼻孔と声道壁から放射されている。本研究では、上述の3つの放射音源を分離して測定するため、防音箱と遮蔽板など装置を導入した。本章では、防音箱と遮蔽板を導入することにより発声器官各部位からの放射音の分離の度合、音響測定に与える影響、およびこれらの装置の音響測定に対する有効性について考察する。

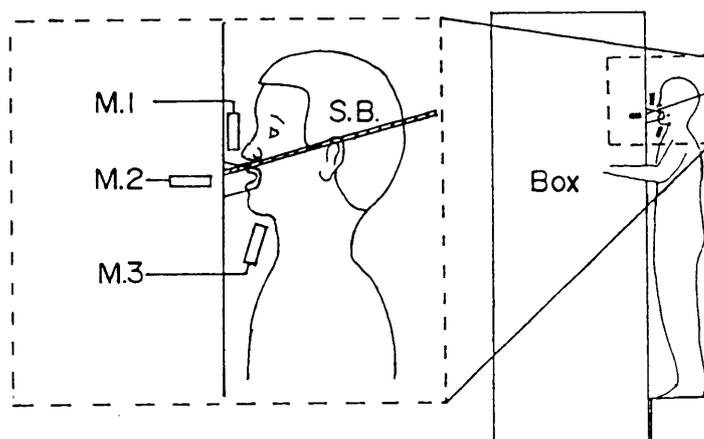
### 2.2 測定装置

口唇、鼻孔と声道壁からの放射音を分離して測定するために、図2.1のように防音箱と遮蔽板を用いて被験者に発声させた。外部からの騒音を避けるためこれら全体

を無響室内に置いた。

防音箱と遮蔽板の形とサイズを図2.2に示す。防音箱は厚さ2.1cmの合板で、大きさがH190×W64×L45cmである。箱にはマウスピース(口の方の直径6cm、箱に結合する部分の直径8.7cm)を取り付けた。箱の共振を防ぐため箱の内壁に、幅4cm厚さ2cm長さ50~100cmの棒を十字形に取り付け、その上に吸音材を貼付した。また、箱の外壁にも吸音材を貼付した。吸音材は厚さ2cmの粗毛フェルトで、吸音率を図2.3に示す。実際には、防音箱内壁に取り付けてある棒のため、吸音材と箱壁の距離は0~2cmの間で変化している。その上、吸音材は多箇所重なっているため、低周波領域で吸音効果は一枚の吸音材だけの場合よりよいと予想できる。この効果は2.4節で考察する。

鼻孔からの放射音を声道壁からの放射音と分離するために遮蔽板を用いた。遮蔽板の中央部の縁は顔面を鼻の下から両頬を横断する形で顔に密着し、前縁は箱と接合している。遮蔽板は幅が94cm、長さが40cm、厚さ6mmの合板と厚さ2mmのアルミニウム板を重ね合わせ、表裏に吸音材を貼付した。それらの装置を用いて発声器官からの放射音を口唇放射、鼻孔放射と声道壁放射の3つに分離することができる。



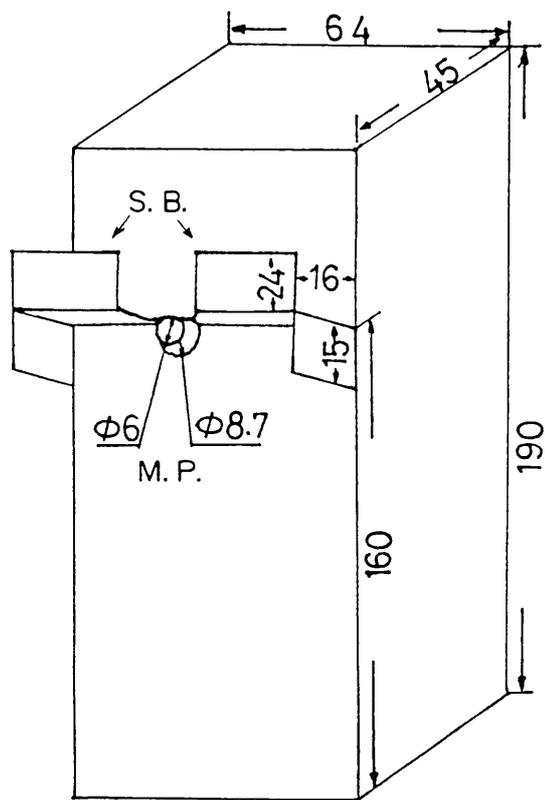
M.:Microphone S.B.:Shield board

図2.1 測定するシステムの略図

M.1:口唇部のマイクロホン

M.2:鼻孔部のマイクロホン

M.3:咽頭壁部のマイクロホン



S.B.:Shield board M.P.:Mouth piece

図 2 . 2 防音箱と遮蔽板の略図(cm)

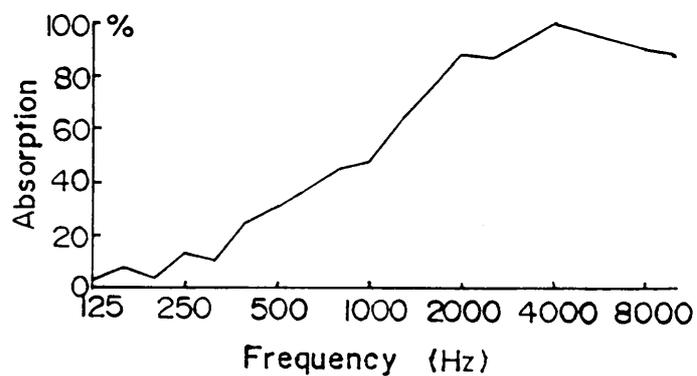


図 2 . 3 粗毛フェルト吸音材の吸音率

## 2.3 防音箱と遮蔽板の音響特性

### 2.3.1 防音箱の共振周波数

防音箱は直方体形である。直方体の囲いの中の音場では固有周波数は次式で表わせる<sup>5)</sup>。

$$f_{nms} = \frac{c}{2} \sqrt{(n/lx)^2 + (m/ly)^2 + (s/lz)^2} \quad (2.1)$$

ここで、 $c$ は音速で、 $340\text{m/s}$ と仮定する。 $lx$ 、 $ly$ 、 $lz$ は壁間の距離で、それぞれ $190\text{cm}$ 、 $64\text{cm}$ と $45\text{cm}$ である。 $n$ 、 $m$ 、 $s$ はモード次数で、それぞれ $n=0,1,2,3,\dots$ 、 $m=0,1,2,3,\dots$ 、 $s=0,1,2,3,\dots$ である。 $f_{nms}$ は第 $nms$ 次基準振動姿態の固有値である。 $m=s=0$ の場合、音場の粒子の運動方向は $x$ 軸方向のみとなるのでこの場合を $x$ 軸平行波動という。同様に、 $s=n=0$ の場合は $y$ 軸平行波動、 $n=m=0$ の場合は $z$ 軸平行波動と呼ぶ。また、 $n=0$ の場合、粒子速度は $y$ 成分と $z$ 成分を有するようになるが $x$ 成分はないので、これを $y-z$ 面平行波動と呼ぶ。同様に $m=0$ の場合を $z-x$ 面平行波動、 $s=0$ の場合を $x-y$ 面平行波動と呼ぶ。ただし、防音箱内に吸音材があるので二次以上のモードの固有周波数での影響は小さくなる。式(2.1)で求めた二次モードまでの防音箱の共振周波数を表2.1に示す。箱の共振現象は箱内の音場と固有周波数付近の透過損失にある程度

表2.1 推定した防音箱の共振周波数

$n$	$m$	$s$	$f(\text{Hz})$
1	0	0	89
0	1	0	266
0	0	1	378
1	1	0	280
1	0	1	388
0	1	1	462
2	0	0	179
0	2	0	531
0	0	2	756

影響を与える。

### 2.3.2 防音箱の透過損失

放射音の測定では、口唇放射音を防音箱に入れることによって他の放射音と分離する。実際には、防音箱の透過損失が無限大ではないので、防音箱の内側のマイクロホンで収録した音圧は口唇放射音圧と箱の壁から透過された外側音圧の和である。同様に、外側のマイクロホンで収録した音圧には箱の壁から透過された内側音圧の分も含めている。その影響を量的に考察するため、防音箱の透過損失を測定した。

防音箱の透過損失の測定では、Node社のType7030発振器とFostex社のGS100のスピーカを用いて1オクターブ幅の帯域雑音を発生した。スピーカを防音箱の真ん中に吊し、箱の周りの4つの面に向けてそれぞれ5回測定した。この20回の平均値を表2.2に示す。63Hzから8kHzまでの周波数範囲においては防音箱の内側から外側への透過損失は31dB以上で、外側から内側への透過損失は30dB以上である。

もし、透過音圧を $P_t$ (dB)、測定した対象音圧を $P_m$ (dB)、透過音圧のない場合の対象音圧を $P_o$ (dB)とすると、3つの音圧の関係は次式で表わせる。

$$P_o = 20 \log_{10}(10^{P_m/20} - 10^{P_t/20}) \quad (2.2)$$

表 2.2 防音箱の透過損失(dB)

周波数(Hz)	内→外(dB)	外→内(dB)*
63	41.8	30.4
125	31.4	31.7
250	35.6	32.0
500	42.8	37.2
1k	41.5	33.9
2k	42.1	47.0
4k	51.5	38.5
8k	56.3	54.5

\* 外側から内側への測定データは文献4)より

防音箱を用いた場合、透過音圧  $P_t$  が対象音圧  $P_o$  に与える影響は式(2.2)で推定できる。 $P_t$  と  $P_m$  とのdB差を求めれば、 $P_o$  を次式で推定できる。

$$P_o = P_m + \beta \quad (2.3)$$

ここで  $\beta$  (dB) は  $P_o$  と  $P_m$  の間の補正值である。 $P_m - P_t$  と対応している補正值  $\beta^{(6)}$  を表2.3に示す。

箱の内側と外側の音圧の大きさが大体同じ場合、防音箱の透過損失は30dB以上 ( $P_m - P_t > 30\text{dB}$ ) なので、透過音圧の影響は無視できる。箱の内側と外側の音圧差が20dB以内 ( $P_m - P_t > 10\text{dB}$ ) の場合、表2.3により、透過音圧の影響は1dBのオーダーで無視できる。内側と外側の音圧差が防音箱の透過損失と同じ程度なら、小さい方が大きい方の透過音圧の影響を受けて、無視できない。

### 2.3.3 遮蔽板の透過損失

鼻孔放射と声道壁放射を十分に分離するために、大きな透過損失を持つ遮蔽板が必要である。遮蔽板の音響漏洩は主に透過音圧と低周波の回折音圧である。低周波の回折を減少するために遮蔽板の面積をできるだけ増大する必要があるので、本研究では遮蔽板の前部を防音箱に密着して、その部分での回折による巻き込みを無くすようにした。図2.2に示したように遮蔽板の両側を下に曲げ、後ろの部分を上を曲げた。このようにした場合、遮蔽板の有限な面積で効率的に低周波の回折を減少させることができる。

透過音圧については、遮音の質量則によると遮蔽板の質量を2倍にすると透過損失は約5dB増大する。遮蔽板の材質が決まった場合厚さを増加することによって、透過

表2.3 デシベル差に対応する補正值<sup>(6)</sup>

$P_m - P_t$	$\beta$						
1.0	-6.9	3.5	-2.6	7	-1.0	12	-0.3
1.5	-5.3	4.0	-2.2	8	-0.7	13 ~ 14	-0.2
2.0	-4.3	4.5	-1.9	9	-0.6	15 ~ 19	-0.1
2.5	-3.6	5.0	-1.7	10	-0.5	20 以上	0
3.0	-3.0	6	-1.3	11	-0.4		

損失を増大させることができる。しかし、遮蔽板の厚さは口唇と鼻孔との間の距離で制限されている。

一方、完全に独立した遮蔽板を二枚で重ね合わせた場合、一枚の板で遮音された音が次の板で更に遮音される。結果的には、全体の透過損失は二枚の板の透過損失の和となる。二枚で重ね合わせた遮蔽板は一枚の遮蔽板と同じ厚さでもより大きな遮音効果が得られる。そのため、本研究では遮蔽板を2枚の板の重ね合わせとした。一枚は厚さ6mmの合板で、もう一枚は厚さ2mmのアルミニウム板である。遮蔽板の透過損失と回折による音の損失との和の測定値を表2.4に示す。

## 2.4 測定装置に対する予備実験

音響測定に対する測定装置の影響は主に、口唇放射音に対する防音箱内の音場の影響と、鼻孔放射と咽頭壁放射に対する防音箱の外壁による反射の影響である。ここでは、本実験で用いた装置の音響測定に対する有効性を確認するための予備実験を行い、その結果を述べる。

### 2.4.1 防音箱の残響時間の推定

吸音材の吸音効果を調べるために、吸音材を貼付してある場合と貼付していない場合の防音箱の残響時間を、無響室の残響時間と比較する測定を行った。

残響時間を推定する方法としては、1オクターブ幅の帯域雑音(Node社のType7030発振器使用)をスピーカ(Fostex社のGS100)から約3秒出した後遮断し、減衰振動波形を記録観測して60dB減衰するまでの時間を、繰り返し求める方法を用いた。この方法では、残響時間にスピーカの過渡特性による減衰特性も含まれた値が求められるが、箱と無響室との比較や吸音材の有無による相対比較は可能である。帯域雑音の中心周波数を63Hzから8kHzまで8点とり、各周波数で14回ずつ測定して平均値を求めた。

表2.4 遮蔽板の透過損失と回折による音の損失の和

周波数(Hz)	63	125	250	500	1000	2000	4000	8000
実測値(dB)	17.1	21.3	19.3	26.8	28.8	27.6	36.0	31.2

測定結果を図2.4に示す。測定結果では、吸音材が貼付していない場合は、箱の共振の影響のため500Hzで残響時間がかなり長くなっているが、吸音材が貼付してある場合は箱の共振の影響はほとんど見られず、125Hz以上の周波数範囲では無響室に近い残響時間となっている。

#### 2.4.2 口唇放射音に対して防音箱内の影響

口唇音を限られた容積の防音箱に取り込む場合防音箱による低域の周波数特性に及ぼす影響を調べるため、スピーカの音を箱の中に放射した場合と無響室内に放射した場合との比較をした。

図2.5(a)のように直径8cmの小型スピーカを箱のマウスピースに取り付け、箱内と箱外に置いたマイクロホンで同時に音圧を測定した。箱内のマイクロホンとスピーカの振動膜との距離を音声測定時と同じく7.5cmにし、外のマイクロホンとスピーカの振動膜との距離をも7.5cmにした。音源は1オクターブバンド幅の帯域雑音で、中心周波数63Hz~8kHzでオクターブ間隔で測定した。内側の音圧から外側の音圧を減算した結果を図2.5(b)に示す。この結果から、防音箱を使用して得られた音声信号に対しては図2.5(b)の逆特性でスペクトルを補正した。

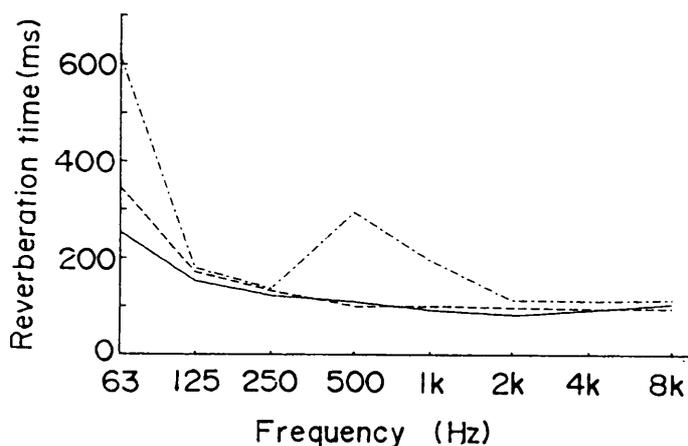


図2.4 無響室と防音箱の残響時間の比較

実線：無響室の残響時間

破線：吸音材を貼っている防音箱の残響時間

一点鎖線：吸音材を貼っていない防音箱の残響時間

### 2.4.3 鼻孔などの放射音に対する防音箱外壁の影響

鼻孔と咽頭壁からの放射音に対する防音箱の外壁による反射の影響について述べる。実音声を測定する時の条件を模擬するために、図2.6(a)のようにマイクロホンを防音箱の外壁より2cm、床面より160cmの所に置き、スピーカをマイクロホンから3cm、防音箱の壁から5cmのところに置く。スピーカはFostexのGS100で、エンクロージャのサイズはW14.5×H14.5×D11.5cmである。測定結果を図2.6(b)に示す。その結果では、500Hz以下では3~4dB高くなっている。この特性についても防音箱内の場合と同様に、鼻孔放射と咽頭壁放射の音声信号に補正を加えた。

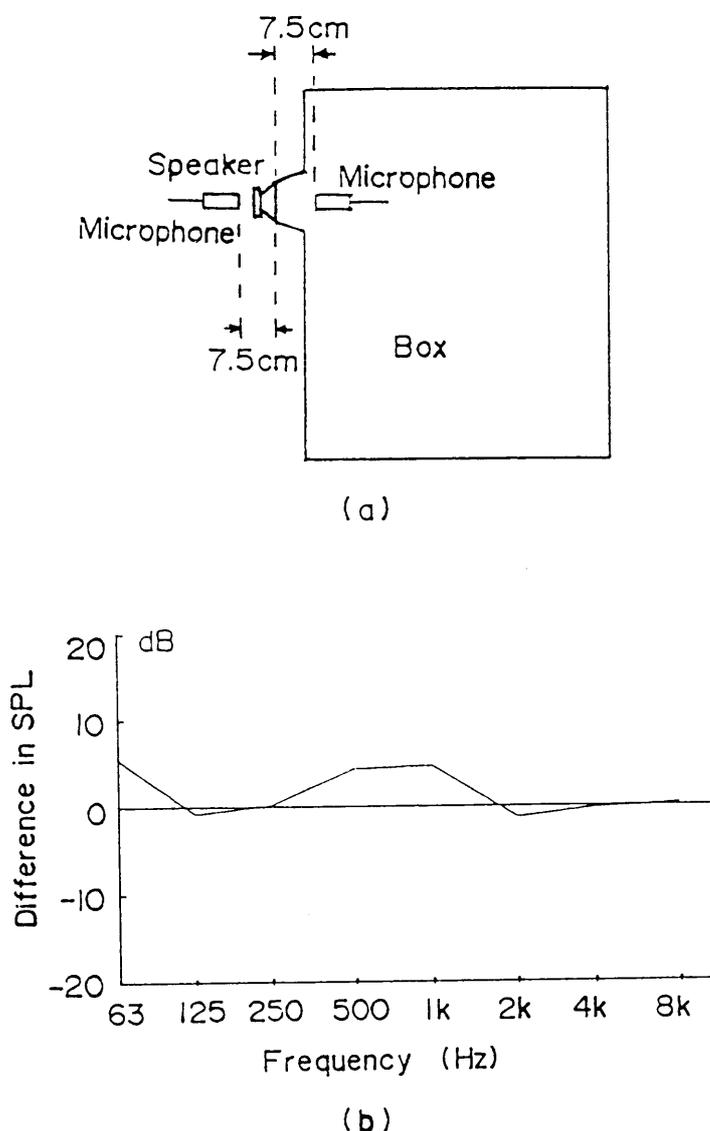


図2.5 防音箱内の周波数特性  
(a) 測定方法  
(b) 測定結果

## 2.5 マイクロホンの出力の校正

### 2.5.1 マイクロホンの特性

本研究では、4個(3種類)のマイクロホンを使用した。それらは、Rion-UC26のマイクロホンが2個と、B&K-4133のマイクロホンおよびB&K-4182のプロープマイクロホンが1個ずつである。Rion-UC26とB&K-4133マイクロホンの周波数特性を図2.7(a)(b)に示す。本測定では主に音声信号の5kHz以下の周波数特性を分析するので、上述のマイクロホンの特性は十分に満足できる。プロープマイクロホン4182の周波数特性

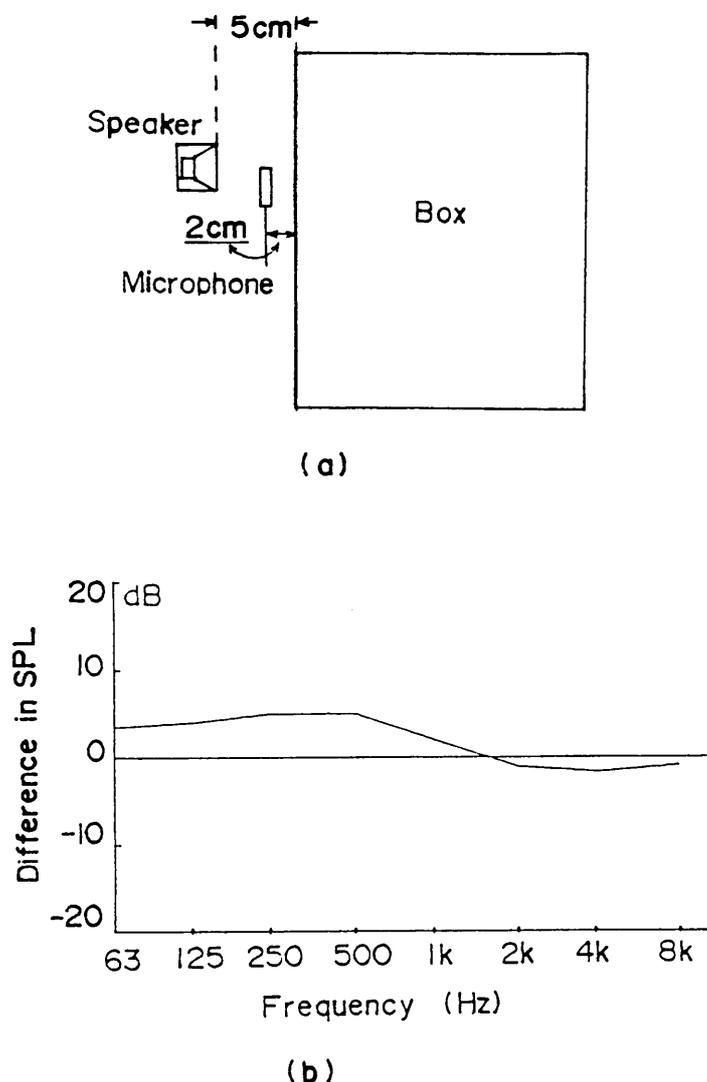


図2.6 防音箱外壁が放射音に対する影響  
(a) 測定方法  
(b) 測定結果

を図2.7(c)に示す。プローブマイクロホン4182は5kHz以下の広い帯域で優れた周波数特性を持っているので口腔内圧の測定では直流成分と交流成分とも測定できる。図での3つの曲線はそれぞれプローブチューブの長さ25mm、50mmと100mmに対応している。第4章と第6章の口腔内圧の測定では100mmのプローブチューブを用い、第5章の鼻孔放射音の測定では50mmのプローブチューブを用いた。

## 2.5.2 マイクロホンの出力の校正

防音箱と遮蔽板によって分離された鼻孔、口唇と咽頭壁の放射音をマイクロホンM1、

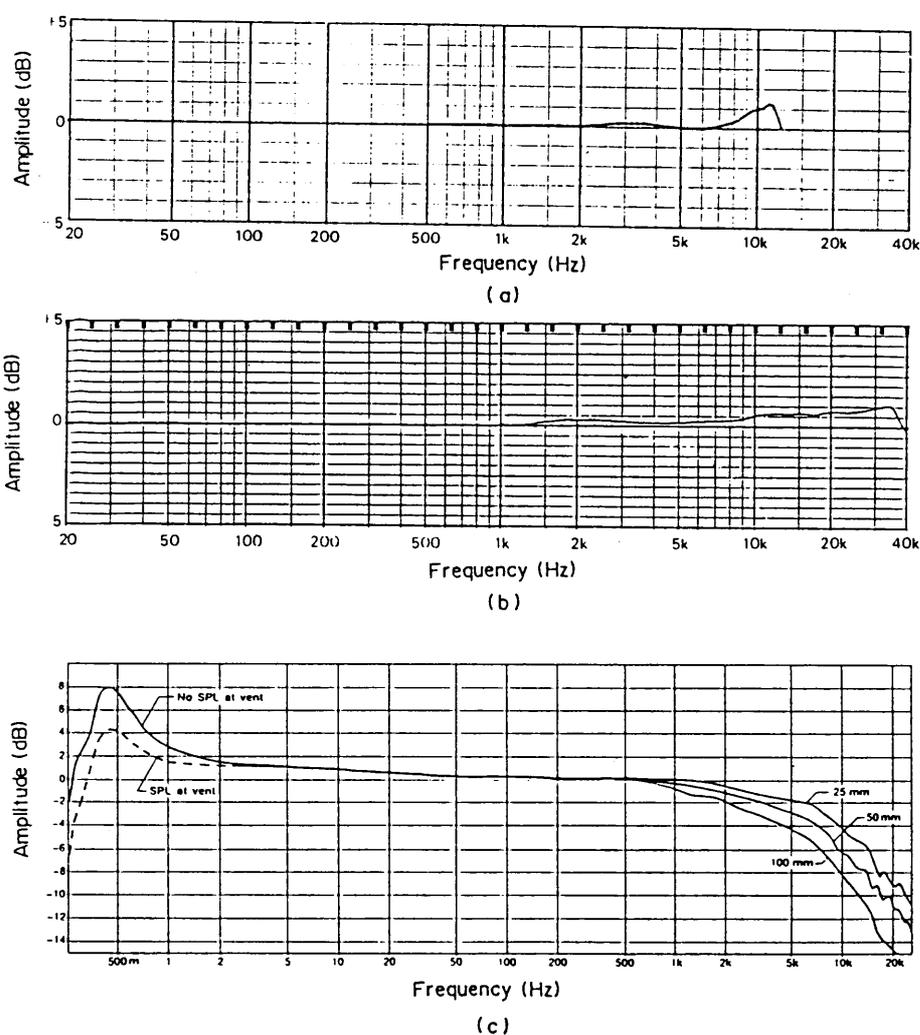


図2.7 本実験で用いたマイクロホンの周波数特性 (説明書による)  
 (a) Rion-UC26型  
 (b) B&K-4133型  
 (c) B&K4182型プローブマイクロホン

M2とM3でそれぞれ測定した。各マイクロホンで収録した音声信号は相互に比較するため、マイクロホンの出力を校正する必要がある。

口唇部のマイクロホンと鼻孔部のマイクロホンの出力の校正では鼻子音付きの音声資料（例えば／ma／、／na／、／oN／など）を用いて行った。防音箱を用いて分離された口唇放射と鼻孔放射の和の鼻子音区間の音声波形と母音区間の音声波形との振幅比を、箱を用いず一個のマイクロホンで収録した音声の対応区間の振幅比に等しくなるようにM1とM2の回路の利得を調整した。調整の一例として、／oN／の防音箱を用いた場合の口唇放射、鼻孔放射及びそれらの和と、防音箱を用いず一個のマイクロホンで収録した音声波形を図2. 8に示す。図2. 8(c)の母音区間と鼻子音区間の振幅比は(d)のそれと同じになっている。

防音箱を用いた場合口唇放射音と鼻孔放射音の和のスペクトルと防音箱を用いない場合の音声のスペクトル（512サンプル長さのハミング窓をかけ、FFTで分析したものを2. 4で述べた箱の特性で補正したもの）を図2. 9に示す。母音区間では、防音箱を用いない場合で求めた4つのホルマンントの周波数は、それぞれ544Hz、840Hz、2421Hzと3334Hzとなっている。それに比べて、防音箱を用いた場合の各ホルマンント周波数の相対誤差は10%以下であり、第一ホルマンントに対する各ホルマンントの相対的大きさでは、両方の差は3dB以内である。鼻子音区間では、防音箱を用いない場合で求めた3kHz以下の範囲での4つのピークの周波数（ケプストラムで求めた）は、それぞれ177Hz、821Hz、1500Hz、2210Hzとなっている。1500Hzのピークを除けば、各ピークの周波数の誤差は10%である。177Hzのピークに対する各ピークの相対大きさについて、両場合の差は、820Hzでは3.1dB、2210Hzでは4.7dBである。1500Hzのピークでの違いは二回の発声の差とみられる。防音箱を用いたと用いてない場合得られた音声信号のスペクトルの相関係数（ケプストラムで求めた）は、母音では0.86、鼻子音では0.93で、両方がかなり一致することを示している。上述の考察結果により、出力の校正と周波数領域での補正は有効であることを確認した。

一方、咽頭壁部のマイクロホン(M3)と鼻孔部のマイクロホン(M1)の間の校正では、上述の場合のように咽頭壁放射を他の放射とはっきり区別できる音素はないので直接的な調整はできない。咽頭壁部のマイクロホンと咽頭壁の面（音源）との距離を、鼻孔部のマイクロホンと鼻孔（音源）からの距離と等しく、3cmにすることによって、それぞれの音圧を同じ感度で収録するようにした。

収録した音声信号を4.8kHzのローパスフィルタを通しA-D変換してワークステーション（H.P. アポロDN3500とDN4000）に取り込んだ。サンプリング周波数は10kHz、語長

は12bitである。

## 2.6 加速度ピックアップについて

音声を発声する時、口腔内音圧によって声道壁が振動する。声道壁の振動量を測定するために、加速度ピックアップを用いた。加速度ピックアップはENDEVCO社の2250A-10であり、直径5.84mm、重さ0.4gである。感度は100Hzでは10mV/g (gは重力加速度

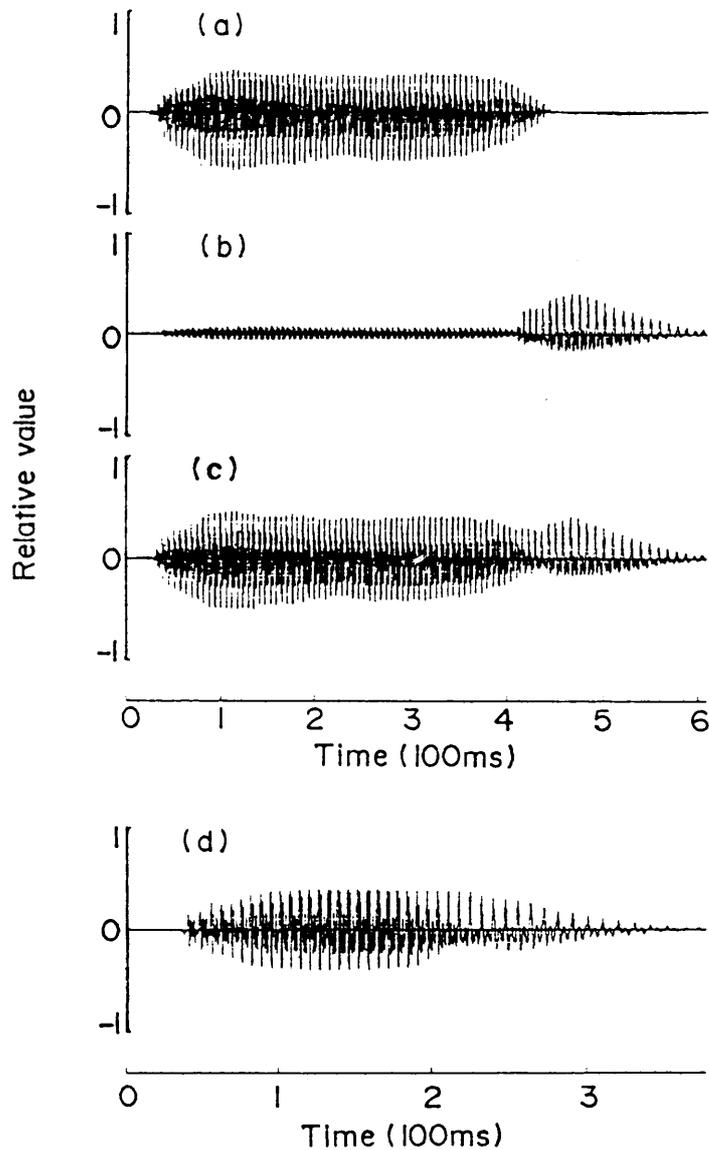


図2.8 口唇部マイクロホンと鼻孔部マイクロホンの校正の例  
(a)(b) 防音箱と遮蔽板を用いて収録した/oN/の口唇と鼻孔放射  
(c) 口唇放射と鼻孔放射の和  
(d) 防音箱と遮蔽板を用いていない場合の放射音

で $980\text{cm/s}^2$ )である。仕様書によれば、加速度ピックアップは $20\text{Hz}\sim 10\text{kHz}$ で偏差は5%である。

## 2.7 むすび

本研究で用いた防音箱と遮蔽板は、発声時の口唇放射、鼻孔放射と咽頭壁放射の分離に対して十分な遮音特性を持っている。防音箱の放射音に対する影響は図2.5(b)と図2.6(b)の逆特性で補正できる。マイクロホンと加速度ピックアップの周波数特性は測定周波数領域で満足できる。それゆえ、この測定システムは本研究での音響測定に十分有効である。

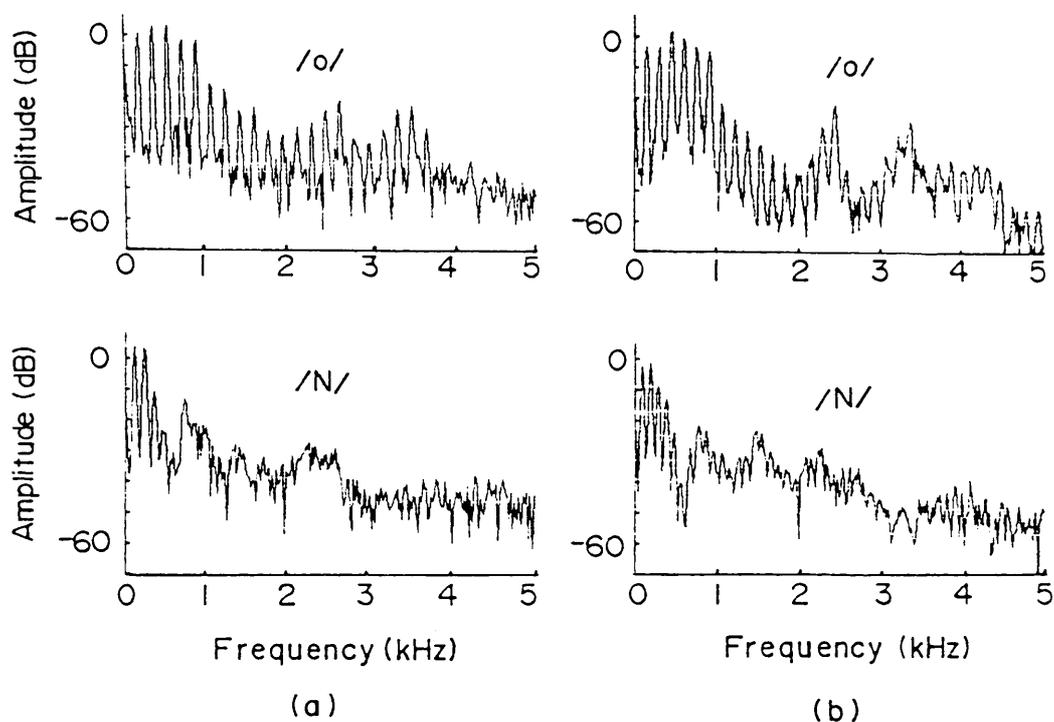


図2.9 防音箱を用いた場合と用いていない場合

収録した/oN/のスペクトル

(a) 防音箱を用いた場合口唇放射と鼻孔放射の和のスペクトル

(b) 防音箱を用いていない場合のスペクトル

## 第3章 非鼻音化母音における音響特性と軟口蓋の振動板モデル

### 3.1 まえがき

軟口蓋は、鼻腔と口腔の結合の度合を制御しているため、言語音声特に鼻音化音声に重要な役割がある。それゆえ、軟口蓋に関する研究が数多く行われてきた。たとえば、Mollら<sup>1)</sup>は鼻音と母音及び非鼻音子音を色々に組み合わせた短文を発声する時の軟口蓋の調音過程を高速映画フィルムで撮影して調べた。Björkら<sup>2)~4)</sup>は発声時の鼻咽腔と軟口蓋の動きを測定した。しかし、それらの測定は定性的な測定であって、音声生成モデルにほとんど具体化されていない。しかも、軟口蓋に関する研究は鼻音化音声の場合に限られている。

鼻音化音声に比べて、非鼻音化母音における鼻腔と口腔の結合に関する研究はほとんどなく、鼻孔放射音と口唇放射音の関係は明らかにされていない。従来の音響モデルでは、非鼻音化音声における鼻腔と口腔の結合を無視して、軟口蓋はOFFとONの二状態のスイッチとされている。しかし、口腔と鼻腔の間の隔壁の一部である軟口蓋は有限の機械インピーダンスを持つ粘弾性体なので、閉鎖していてもそのインピーダンスにより口腔と鼻腔が結合すると考えられる。さらに、言語音声における非鼻音化音声から鼻音化音声へのわたりでは、軟口蓋は閉鎖状態から開放状態へ変化する。この変化を音声生成モデルで表わすには、まず非鼻音化音声における軟口蓋の状態を考察する必要がある。本章では、非鼻音化母音における発声器官各部位からの放射音を分離して考察することにより、閉鎖している軟口蓋をばねとタッチポットでつながる二枚の振動板としてモデル化し、口唇放射と鼻孔放射を用いて軟口蓋の振動板モデルのパラメータを推定する。

### 3.2 放射音圧と振動量の測定システム

非鼻音化母音における音響放射の測定では、2.2で述べた測定装置を用いて、口唇放射、鼻孔放射と咽頭壁放射を分離して測定し、同時に加速度ピックアップを用いて、咽頭壁や、鼻孔壁と頬の加速度振動量を測定した。測定システムを図3.1に示

す。

鼻孔、口唇、咽頭壁の放射音はそれぞれコンデンサマイクロホン M1(B&K-4133)、M2(Rion-UC26)、M3(Rion-UC26)で抽出した。口唇部のマイク M2には口唇の開口出力と、口唇周りのマウスピースで囲まれた部分の放射音を取り込まれる。マイクロホンはそれぞれ精密騒音計B&K-2209、Rion-NA60及びRion-NA40に接続し、そのAC出力を取り出した。口唇放射は図2.5(b)に示した箱の逆特性で補正した。鼻孔放射と咽頭壁放射は図2.6(b)に示した箱の外壁の反射特性の逆特性で補正した。加速度ピックアップENDEVCO-2250Aは3個使用し、それぞれ鼻孔壁(A3)、頬(A2)、咽頭壁(A1)に密着し、振動量を抽出した。

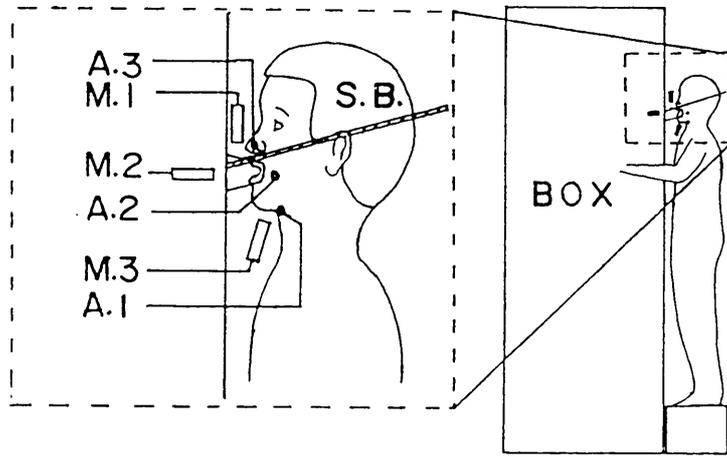
測定した6チャンネルの信号は遮断周波数4.8kHzの低域ろ波フィルタを通しA-D変換して、直接にワークステーション(H.P.アポロDN4000)に取り込んだ。サンプリング周波数は10kHzで、語長は12bitである。音声資料は表3.1に示した日本語100音節である。発声者は成年男性5名である。

### 3.3 発声器官各部位の放射音と振動加速度の比較

#### 3.3.1 放射音と振動加速度の波形

3個のマイクロホンと3個の加速度ピックアップで収録した母音/a/と/i/の放射音と振動加速度の波形をそれぞれ図3.2と図3.3に示す。母音/a/の放射音では、口唇放射以外に、鼻孔と咽頭壁からも音が放射されており、咽頭壁放射は鼻孔放射よりやや大きくなっている。振動加速度では、咽頭壁が一番強く頬より2~3倍大きくなっており、鼻孔壁の振動がほとんど見られない。

母音/i/の場合、鼻孔放射が咽頭壁放射より2倍程度大きくなっている。振動加速度では、鼻孔壁と頬の振動量は同じ程度であるが、咽頭壁の方が前两者より5倍程度大きくなっている。口唇放射では/a/が/i/より4倍以上大きいが、鼻孔放射では/i/の方がはるかに大きくなっている。鼻孔壁の振動量にも/i/が/a/よりはるかに大きくなっている。非鼻音化音声といっても、母音/i/の放射音には鼻孔放射の割合がかなり大きい。/a/と/i/の放射量と振動量の増加の割合においては、鼻孔部では放射より振動の増加量が少ないが、咽頭壁部では放射より振動の増加量が多くなっている。これは放射と振動の因果関係が両者では逆になっているからである。すなわち、鼻孔壁の振動は鼻孔部の放射によるものであり、咽頭壁の放射は咽頭壁の振動によるものである。



A.:Accelerometer M.:Microphone S.B.:Shield board

図3. 1 音声を測定するシステムの略図

M.1、M2、M3: 鼻孔、口唇、咽頭壁のマイクロホン

A.1、A2、A3: 咽頭壁、頬、鼻孔壁の加速度ピックアップ

表3. 1 日本語の単音節音声資料

ア	a	イ	i	ウ	u	エ	e	オ	o
カ	ka	キ	ki	ク	ku	ケ	ke	コ	ko
サ	sa	シ	shi	ス	su	セ	se	ソ	so
タ	ta	チ	chi	ツ	tsu	テ	te	ト	to
ナ	na	ニ	ni	ヌ	nu	ネ	ne	ノ	no
ハ	ha	ヒ	hi	フ	fu	ヘ	he	ホ	ho
マ	ma	ミ	mi	ム	mu	メ	me	モ	mo
ヤ	ya			ユ	yu			ヨ	yo
ラ	ra	リ	ri	ル	ru	レ	re	ロ	ro
ワ	wa								
ガ	ga	ギ	gi	グ	gu	ゲ	ge	ゴ	go
ザ	za	ジ	zi	ズ	zu	ゼ	ze	ゾ	zo
ダ	da	ヂ	zi	ヅ	zu	デ	de	ド	do
バ	ba	ビ	bi	ブ	bu	ベ	be	ボ	bo
パ	pa	ピ	pi	プ	pu	ペ	pe	ポ	po
キャ	kya			キュ	kyu			キョ	kyo
シャ	sha			シュ	shu			ショ	sho
チャ	cha			チュ	chu			チョ	cho
ニャ	nya			ニュ	nyu			ニョ	nyo
ヒャ	hya			ヒュ	hyu			ヒョ	hyo
ミャ	mya			ミュ	myu			ミョ	myo
リャ	rya			リュ	ryu			リョ	ryo
ギャ	gya			ギュ	gyu			ギョ	gyo
ジャ	zya			ジュ	zyu			ジョ	zyo
ビャ	bya			ビュ	byu			ビョ	byo
ピャ	pya			ピュ	pyu			ピョ	pyo

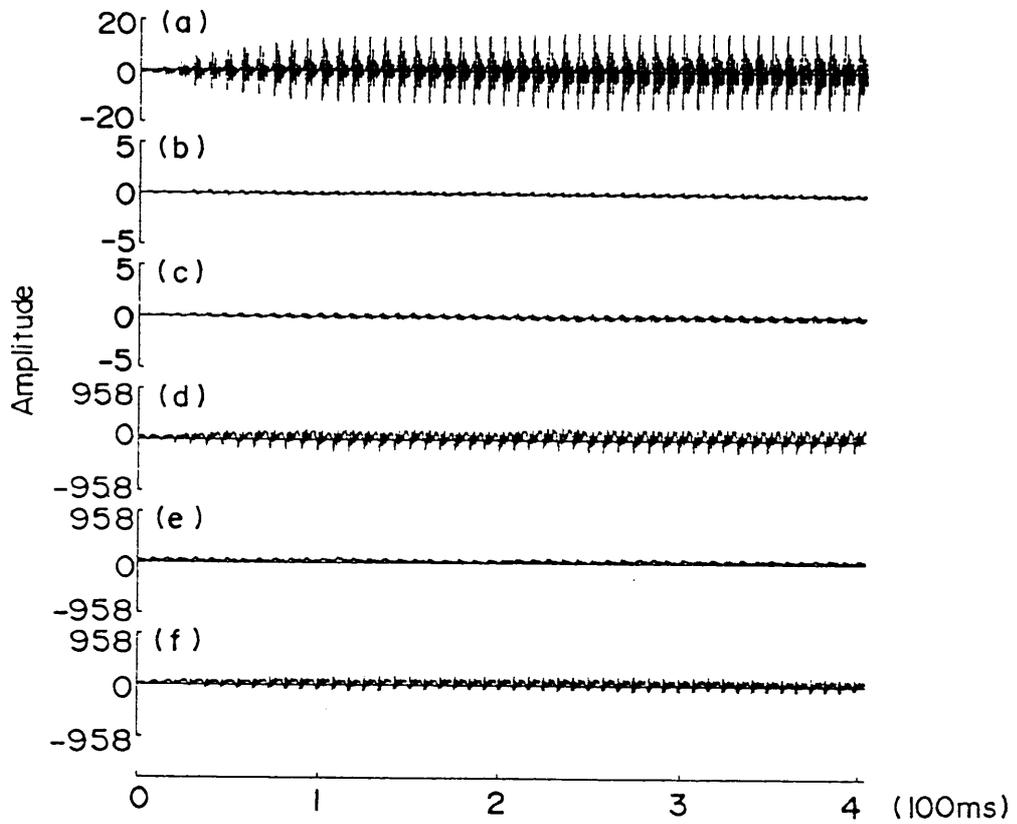


図3. 2 /a/の各部位の放射音と振動加速度の波形  
 (a)、(b)、(c) 口唇、鼻孔及び咽頭壁からの放射音圧(Pa)  
 (d)、(e)、(f) 咽頭壁、鼻孔壁及び頬の加速度( $\text{cm/s}^2$ )

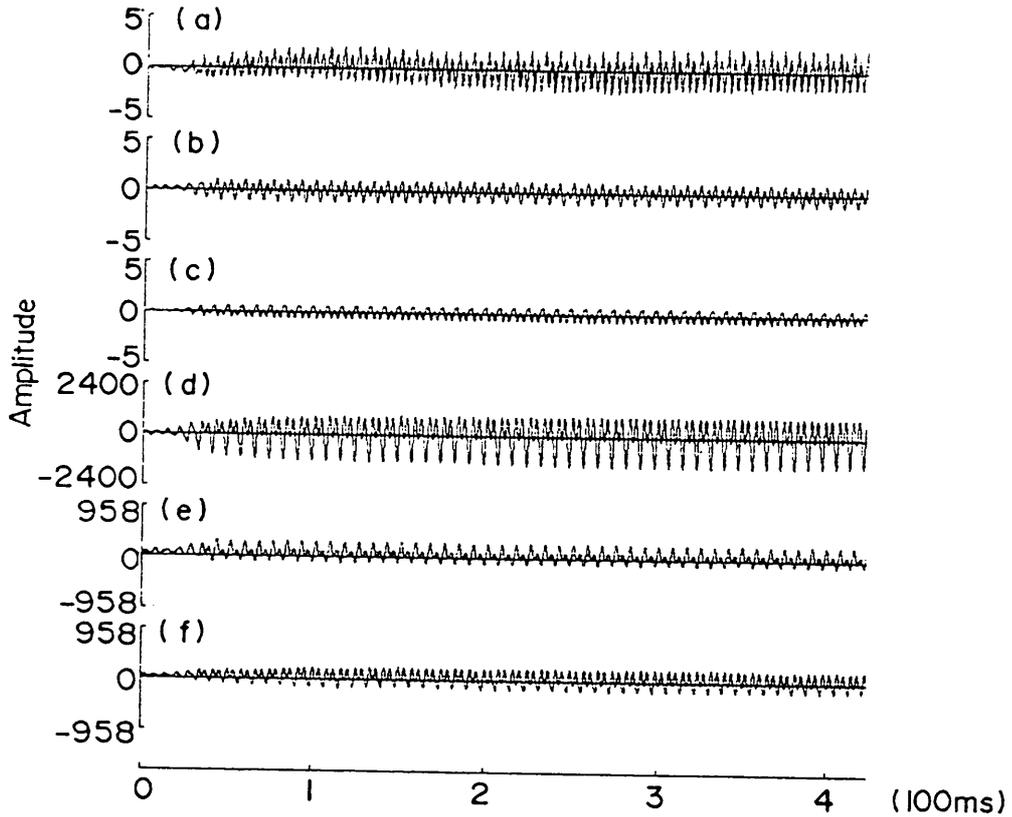


図3. 3 /i/の各部位の放射音と振動加速度の波形  
 (a)、(b)、(c) 口唇、鼻孔及び咽頭壁からの放射音圧(Pa)  
 (d)、(e)、(f) 咽頭壁、鼻孔壁及び頬の加速度( $\text{cm/s}^2$ )

### 3.3.2 放射音と振動加速度のスペクトル

各部位の放射音と振動加速度の周波数特性を調べるには、各チャンネルの信号が高い周波数までの周波数成分を持つ音声資料が必要である。そのため、例として /ma/ の6チャンネルの波形を図3.4に示し、それらのスペクトルを図3.5に示す。ここで、音声データは512サンプル長のハミング窓をかけ、FFTで分析した。分析区間は図3.4の縦点線で示した位置からの512サンプルである。A-D変換器のダイナミックレンジの制約から、図3.5に示した口唇放射では4kHz以上の周波数成分はノイズと区別がつかなくなった。同様な原因で、鼻孔放射では3kHz以上、咽頭壁放射では1.5kHz以上の周波数成分はノイズと区別がつかない。

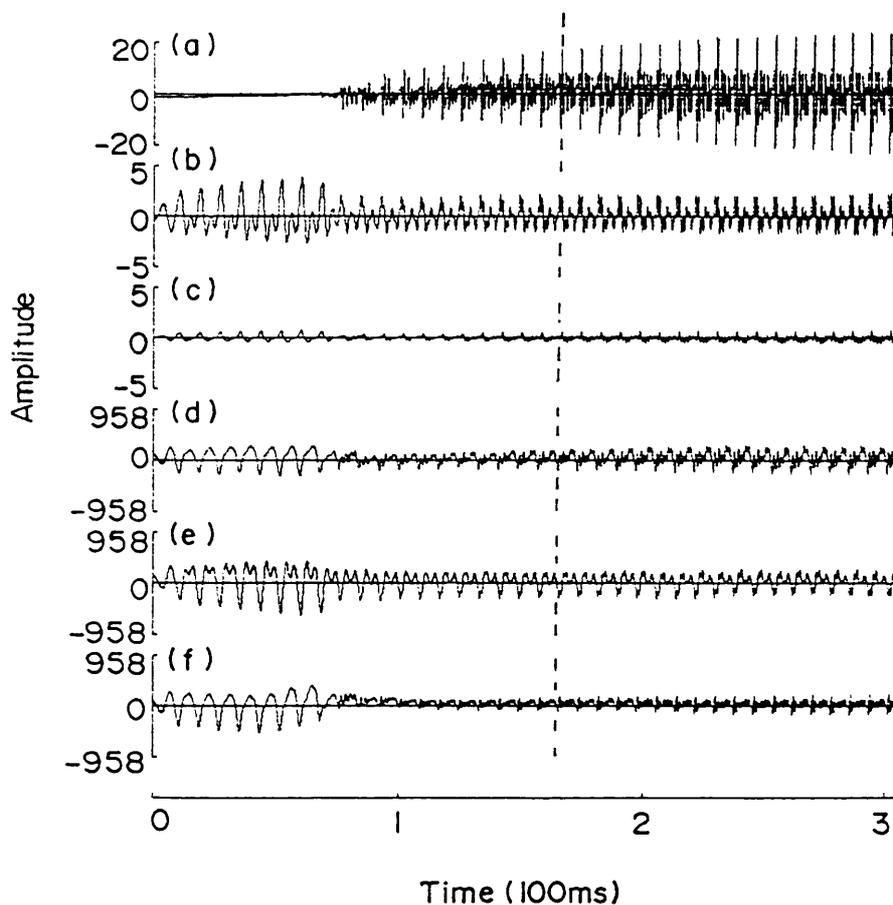


図3.4 /ma/の各部位の放射音と振動加速度の波形  
(a)、(b)、(c) 口唇、鼻孔及び咽頭壁からの放射音圧(Pa)  
(d)、(e)、(f) 咽頭壁、鼻孔壁及び頬の加速度( $\text{cm/s}^2$ )

放射と振動加速度の相関係数を求めたところ、鼻孔部では0.91、咽頭壁部では0.92となっている。また、鼻孔放射での250Hz、700Hz、1300Hz、2300Hz付近のピークと500Hz、1200Hzの零点が鼻孔壁の振動加速度に現れ、咽頭壁放射での200Hzと600Hz付近のピークおよび450Hzの谷が咽頭壁の振動加速度に現れていることから、それぞれの放射音と振動加速度のスペクトルはほぼ同じであることがわかる。

### 3. 3. 3 壁の振動量と放射音圧

放射音圧と声道壁の振動量との関係を明らかにするため、母音区間及び鼻音区間での鼻孔放射量と鼻孔壁の振動加速度、咽頭壁放射量と咽頭壁の振動加速度の関係を図3. 6に示す。その結果では、鼻孔部と咽頭壁部とも、放射音圧と壁の振動加速度のレベルはほぼ比例している。

図3. 6では、鼻孔のデータは値の大きいグループ（鼻音）と小さいグループ（母音）に分かれているのに対して、咽頭壁のデータは鼻音化の有無に関係なく一つのグループとなっている。鼻音区間（黒マークつき）で鼻孔放射量は大きくなっても咽頭壁の放射量は必ずしも大きくならない。鼻孔放射量と咽頭壁放射量の間にはほと

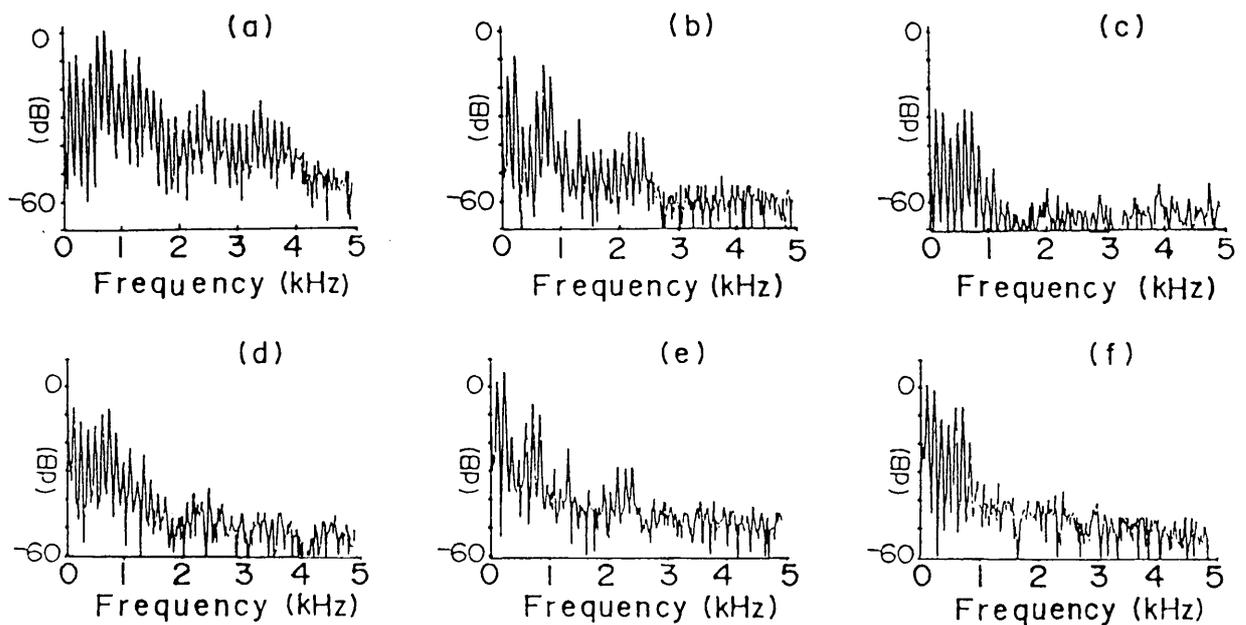


図3. 5 /ma/の/a/の定常部のスペクトル  
 (a)、(b)、(c) 口唇、鼻孔及び咽頭壁からの放射音圧  
 (d)、(e)、(f) 頬、鼻孔壁及び咽頭壁の加速度

んど相関はない。咽頭壁放射音が咽頭壁振動に比例しているのは、その放射音が咽頭壁の振動によるものであることを示している。

上述した結果では、声道壁の振動加速度と放射音については、量的に比例しており、周波数特性においては鼻孔部では3kHz以下、咽頭壁部では1.5kHz以下のスペクトルがほぼ同じである。

### 3. 4 母音における各部位の放射音圧レベル

非鼻音化母音における各部位の放射音に対しては、式(3.1)にしたがって母音定常部のサンプル値の平均値からのずれの2乗平均のデシベル値P(以下音圧レベルと呼ぶ)を求めた。

$$P = 10 \log_{10} \left( \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - X)^2 \right) \quad (3.1)$$

ここで、Nは切り出し区間のサンプル数で、x(n)はn番目のサンプルの値で、Xは切り出し区間の平均値である。話者5名の母音における口唇放射、鼻孔放射及び咽頭壁放射の音圧レベルを表3.2～表3.6に示す。5名のうち、話者Nの鼻孔放射の音

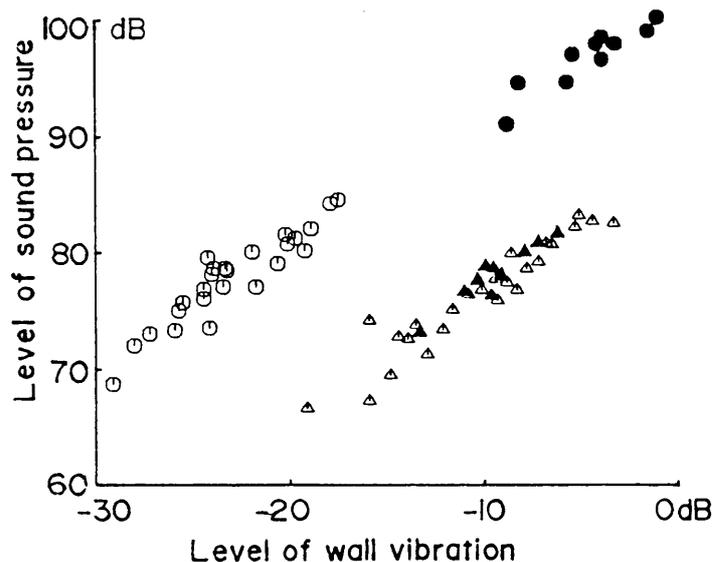


図3.6 壁の振動量と放射音圧の関係図

- : 母音における鼻孔のデータ
- △: 母音における咽頭壁のデータ
- : 鼻音における鼻孔のデータ
- ▲: 鼻音における咽頭壁のデータ

表3.2 母音区間での放射音の音圧 (話者N)  
口唇放射音の音圧(dB)

	a	i	u	e	o
a~o	104.6	91.1	93.2	101.9	103.9
k	106	92.3	96.3	100.7	102.8
s	104.4	92.4	98	100.1	103.5
t	105.7	91.8	96.8	107.4	106.6
h	107.5	93.1	95.5	103.8	107.8
r	107.8	96.3	99.7	104.5	108.5
g	107.4	93.4	100	105.7	108.9
z	107.4	95.8	97.8	105.9	108
d	107	95.4	95	107.2	108
b	108.5	95.8	96.8	105.5	107.6
p	107.2	98.4	99.3	108.9	109
標準偏差	1.3	2.3	2.1	2.8	2.3
平均	106.7	94.2	97.1	104.7	106.8

鼻孔放射音の音圧(dB)

	a	i	u	e	o
a~o	82.5	85.5	83.5	78.2	79.9
k	80.8	86.5	84.3	80.4	82
s	79.6	85.9	87.3	79.4	82.4
t	86.1	83.8	84.2	77.8	79.4
h	88.5	96.4	92.5	87.3	88.9
r	83.8	84.8	86	82.6	87.1
g	84.3	87.4	86.9	83.1	88.3
z	85.1	87.2	83.1	83.1	86.5
d	81.5	86.4	84.3	85.4	87.9
b	78.4	88.6	84.2	80	81.8
p	88.9	91.6	86.2	81.2	82.8
標準偏差	3.4	3.6	2.7	3	3.5
平均	83.6	87.6	85.7	81.7	84.3

咽頭壁放射音の音圧(dB)

	a	i	u	e	o
a~o	75.8	83.1	80.6	74.9	77.2
k	77.2	82.6	82.6	74.1	75.8
s	75.7	80.9	81.1	74	76.4
t	76.9	80.6	79.4	77.9	79.4
h	78.9	83.2	81.9	77.3	81
r	77.2	80.4	82.2	75.7	78.1
g	78.1	84.2	83.5	77.7	82.1
z	77.9	83.2	79.2	78.5	78.9
d	77.4	82.7	81.2	79.5	79.6
b	78	84.6	81.9	77.5	82.4
p	80.8	87.8	81.6	78.8	80.4
標準偏差	1.4	2.1	1.3	1.9	2.2
平均	77.6	83	81.4	76.9	79.2

表 3. 3 母音区間での放射音の音圧 (話者 C)  
口唇放射音の音圧 (dB)

	a	i	u	e	o
a~o	95.5	80.4	82.9	88.8	94.6
k	94	79.3	83.7	89.4	91.6
s	92.4	79	84.2	86.1	91.1
t	91.4	82.3	85.4	84.1	90.8
h	91.2	82	82.8	85.3	92.6
r	91.1	83.1	87.8	87.1	92.6
g	93.2	80.4	85.9	85.4	91.5
z	88.9	83.5	84.4	86.4	91.2
d	90.2	83	85.6	87	92
b	90.6	84.4	86.3	84.1	92.1
p	91.7	83.5	83.9	88.5	92.7
標準偏差	1.9	1.8	1.5	1.8	1.1
平均	91.8	81.9	84.8	86.6	92.1

鼻孔放射音の音圧 (dB)

	a	i	u	e	o
a~o	78.9	78.4	78.7	78	82.3
k	80.3	79.9	78.5	79.7	79.7
s	81.9	78.4	78.2	77	81.6
t	81.2	78	80.3	76.5	79.4
h	79	79.9	77.6	76.6	78.7
r	79.6	77.2	80.2	79.1	82.4
g	79.3	77.4	77.6	78.3	78.8
z	79	79.9	79.5	78.6	81.2
d	80	79	79.4	76.9	81.1
b	79.2	77.3	77.6	77.4	78.2
p	81.1	77.5	75.9	80.1	76.4
標準偏差	1	1.1	1.3	1.3	1.9
平均	80	78.4	78.5	78	80

咽頭壁放射音の音圧 (dB)

	a	i	u	e	o
a~o	66.3	74.3	71	63.1	70
k	66.2	73.8	70.4	64.7	67.2
s	69.7	73.3	69.9	66.1	71.8
t	66.9	76.1	73.6	63.5	66.8
h	64.5	74.5	71.2	61.6	67.3
r	68.5	76.7	73.6	67.4	67.8
g	70.7	77.9	73.4	66.4	69.2
z	66.7	78	73.5	66.6	71.3
d	66.5	77.3	75.3	67.3	74.2
b	70.4	77.3	76.7	68.9	70.1
p	70.7	77.5	75.2	69.2	70.2
標準偏差	2.2	1.8	2.2	2.4	2.3
平均	67.9	76.1	73.1	65.9	69.6

表3.4 母音区間での放射音の音圧(話者S)  
口唇放射音の音圧(dB)

	a	i	u	e	o
a~o	100.6	94.7	96.1	96.6	99.9
k	101.9	91.1	92.7	94.8	99.8
s	100.5	94.2	97.3	98.6	99.7
t	101	91.6	93.3	93.5	96.5
h	101.5	91.5	94	96.5	99.9
r	100.3	89.8	93	95.9	97.5
g	101	89.7	90.5	95.4	96.3
z	95.1	89.9	88.9	92.7	96.6
d	99.4	87.2	89.6	95.5	95.1
b	100.2	90.7	92.5	92.4	100.1
p	102.3	89.8	92.7	92.9	97.3
標準偏差	1.9	2.1	2.5	1.9	1.8
平均	100.3	90.9	92.8	95	98.1

鼻孔放射音の音圧(dB)

	a	i	u	e	o
a~o	80.3	81.7	79.6	78	80.1
k	80.6	81.4	80.7	75.8	76.8
s	79.7	82.9	84.6	80.4	76.7
t	78.4	82.9	82.5	79.6	79.6
h	77.9	82.1	81.4	78.6	78.6
r	81.3	82.7	81	75.7	76.3
g	79.6	82.5	80.3	82.1	80.4
z	83	81.1	80.1	82.8	81.7
d	78.2	80.6	79.9	82.2	80.8
b	77.7	82.6	80.8	77.9	78.6
p	80.3	81.9	80.5	77.5	74.3
標準偏差	1.6	0.8	1.4	2.5	2.3
平均	79.7	82	81	79.1	78.5

咽頭壁放射音の音圧(dB)

	a	i	u	e	o
a~o	72.2	79.2	75.3	71.1	74.3
k	72	80.7	76.7	70.3	75.1
s	72.6	80.6	77.3	73.3	73
t	70.8	80.3	75.8	71	73.2
h	70.7	82.6	76.5	72.1	73.8
r	72.4	83	78.1	73.4	73.1
g	74.8	82.8	77.9	76.8	75.5
z	73.1	78	78.1	73.5	75.6
d	71.3	78.1	77.4	75.5	74.6
b	72.8	82.9	77.5	69.8	76.4
p	72.4	82	75.6	71.1	72.6
標準偏差	1.1	1.9	1	2.2	1.3
平均	72.3	80.9	76.9	72.5	74.3

表3.5 母音区間での放射音の音圧(話者Y)  
口唇放射音の音圧(dB)

	a	i	u	e	o
a~o	109.2	92.9	101.1	100.7	104.1
k	106.3	90.6	100.7	101	107
s	108.5	92.6	102.1	103.1	107.8
t	108.5	93.3	103.2	99.6	108.3
h	108.6	93.8	103.3	101	107.8
r	109.5	94.3	107.1	101.2	109.5
g	109.1	95.5	105.7	101.3	110.7
z	109.1	96.5	107.7	102.7	111.1
d	109.7	96.7	105	102.3	111.4
b	109.3	96	108.8	102.5	110.6
p	106.9	95.6	106.8	102.3	110.9
標準偏差	1.1	1.9	2.8	1.1	2.3
平均	108.6	94.3	104.7	101.6	109

鼻孔放射音の音圧(dB)

	a	i	u	e	o
a~o	80.2	86	82.9	80.4	79.1
k	80.1	85.7	81.8	80.1	79.4
s	80.8	82.6	81.8	79.4	80.1
t	81.1	82.8	82.2	79.9	81.7
h	82.9	85.1	83.2	80.5	79.6
r	81.8	86.1	83	79.2	80.7
g	81	87.9	88.2	79.6	80.1
z	81	87.5	88.1	79.6	81.1
d	81.8	88.1	85.7	79.6	82.2
b	81.2	87.8	85.3	79.7	80.2
p	79.7	85.3	84	79.1	79.7
標準偏差	0.9	1.9	2.3	0.5	1
平均	81.1	85.9	84.2	79.7	80.4

咽頭壁放射音の音圧(dB)

	a	i	u	e	o
a~o	80.2	83	83.8	78.3	79
k	78.9	83.8	84.5	78.4	79.4
s	79	82.1	84.7	78.3	79.6
t	79.1	82.5	83.8	78.6	80.2
h	81.7	83.8	85.5	79.9	80.3
r	79.7	84.1	86	78.4	81.6
g	79.9	85.9	86.1	78.7	82.2
z	80.1	85.6	85.8	78.6	82.4
d	80.5	86	86.6	78.9	81.3
b	80.4	85.4	86.7	78.7	81.7
p	79.8	83.5	84.8	78.5	81.9
標準偏差	0.8	1.4	1	0.5	1.2
平均	79.9	84.2	85.3	78.7	80.9

表3.6 母音区間での放射音の音圧(話者H)  
口唇放射音の音圧(dB)

	a	i	u	e	o
a~o	100.4	86.9	89.3	101.2	102.2
k	104.4	89.2	93.9	99.8	102.6
s	105	89.8	92.3	98.7	101.7
t	104.1	90.3	96.6	100.3	101.5
h	104.5	89.5	93.3	99.3	103.6
r	104.4	90.5	95.5	99.2	103.8
g	106.4	91.4	95.9	102.8	104.5
z	104.4	91.7	92	100.2	102.4
d	106.5	91.1	93.2	101.6	104.3
b	104.4	89.6	94	100.2	103.3
p	106	90.8	92.5	100.8	105.1
標準偏差	1.6	1.3	2.1	1.2	1.2
平均	104.6	90.1	93.5	100.4	103.2

鼻孔放射音の音圧(dB)

	a	i	u	e	o
a~o	78.5	84.8	80.3	78.6	78.6
k	79.1	84.8	79.9	79.2	79
s	79.3	82.2	79.7	78.6	78.8
t	79	81.5	79.1	78.6	78.7
h	79.3	82.8	80.2	78.8	80.3
r	79	84.4	79.4	78.7	80.1
g	80.7	85.8	81.1	79.7	80
z	79.5	84.6	81.8	79.5	80.2
d	80.6	84.1	81	79.5	79.4
b	80.1	86.4	81.1	79.5	79.1
p	80.8	84.2	80.7	79.4	80.3
標準偏差	0.8	1.5	0.8	0.4	0.7
平均	79.6	84.1	80.4	79.1	79.5

咽頭壁放射音の音圧(dB)

	a	i	u	e	o
a~o	78.8	81.7	79.7	79.2	79.8
k	80.3	82.3	80.8	79.5	80.5
s	80.4	81	79.5	78.9	79.5
t	80	80.7	79.1	79.1	79.6
h	80.3	81.7	81	78.9	81.1
r	79.6	81.8	79.8	79	80.8
g	81.3	83.7	81.8	80.1	81.2
z	80.2	82.5	81.2	79.7	80.9
d	80.6	82.4	81	80.1	79.8
b	81	81.6	80.1	79.8	79.7
p	81.7	81	79.4	80.1	81.8
標準偏差	0.8	0.9	0.9	0.5	0.8
平均	80.4	81.9	80.3	79.5	80.4

圧レベルの標準偏差が3dB程度である以外に、他の放射音圧レベルの標準偏差は2dB程度である。咽頭壁放射量と鼻孔放射量の相対的大きさによって、話者5名はAとB 2つのグループに分けられる。Aグループは3名で、話者N、話者Sと話者Cである。Bグループは2名で、話者Yと話者Hである。

代表的なデータ（話者Nと話者Y）について、単母音と鼻音以外の12種類のCV型音節における母音の各部位の音圧レベルの平均値を母音ごとに求めた。それらの平均値を図3. 7に示す。Aグループは鼻孔の方が咽頭壁より明らかに強く、Bグループ

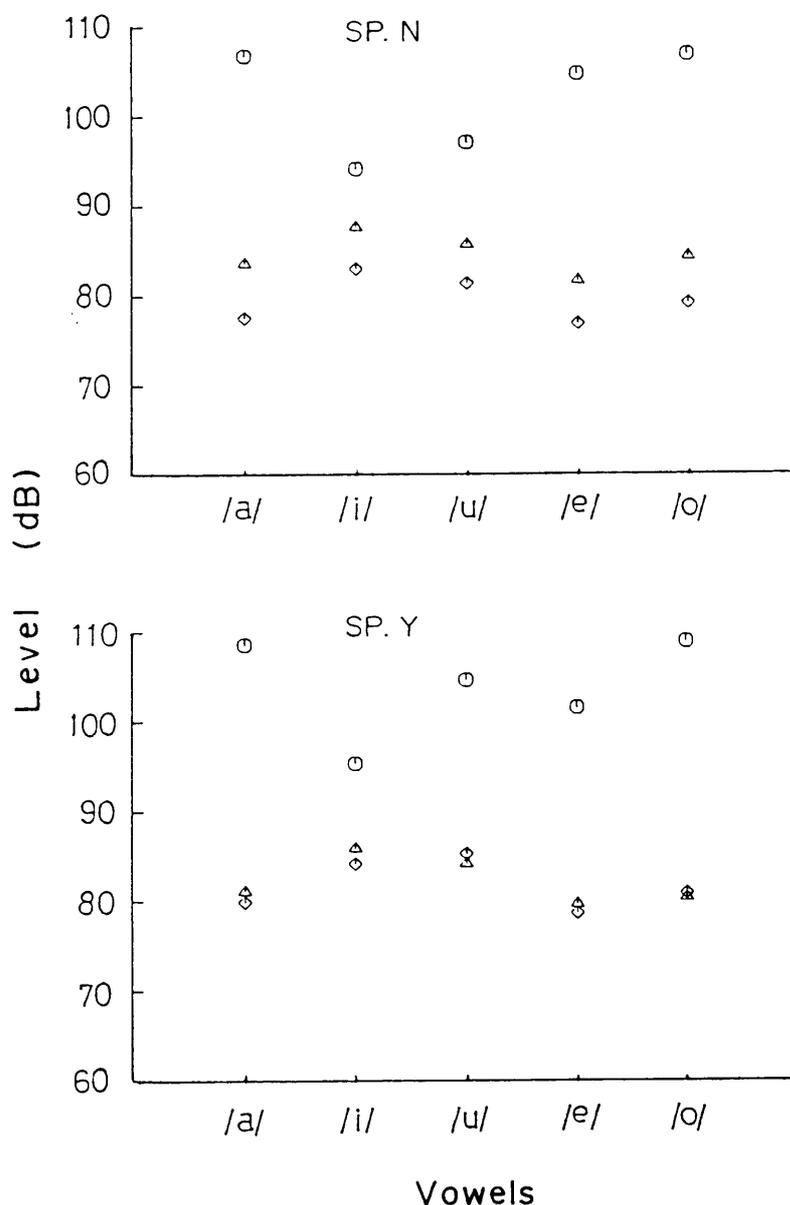


図3. 7 非鼻音化母音の各部位からの放射音圧  
○：口唇放射 △：鼻孔放射 ◇：咽頭壁放射

プはその差が少ない。

母音*/i/*の場合、口唇放射音圧は他の母音より3~14dB小さいが、鼻孔放射音圧は他の母音より2~6dB大きくなっている。*/i/*の口唇放射と鼻孔放射の音圧差は6~8 dBであるが、他の母音では、その差は12~27dBとなっている。放射量と調音特徴に結び付けると、口唇放射では広母音が狭母音より大きく、鼻孔放射では広母音が狭母音より小さくなっている。話者Cの場合では、母音による口唇放射と鼻孔放射の音圧差の変化は他の話者と同じであるが、母音による鼻孔放射の大きさの変化は他の話者と一致しない。咽頭壁の放射を参照して総合的にみると、それは口唇音の漏れでなく個人差であるとみられる。

母音*/a/*と*/i/*の口唇放射と鼻孔放射のスペクトルを図3. 8に示す。*/a/*の場合、鼻孔放射のスペクトルに第1、第2ホルマントが現れているが、口唇放射に比べると30dB程度小さくほとんど影響はない。その周波数の近傍で防音箱の透過損失が40dB以上になっているので、口唇放射から鼻孔放射への透過音圧の影響は無視できる程度である。このことより、鼻孔音は口腔内音圧が閉鎖している軟口蓋のインピーダンスを通し鼻腔を経て放射した音声であることを示している。

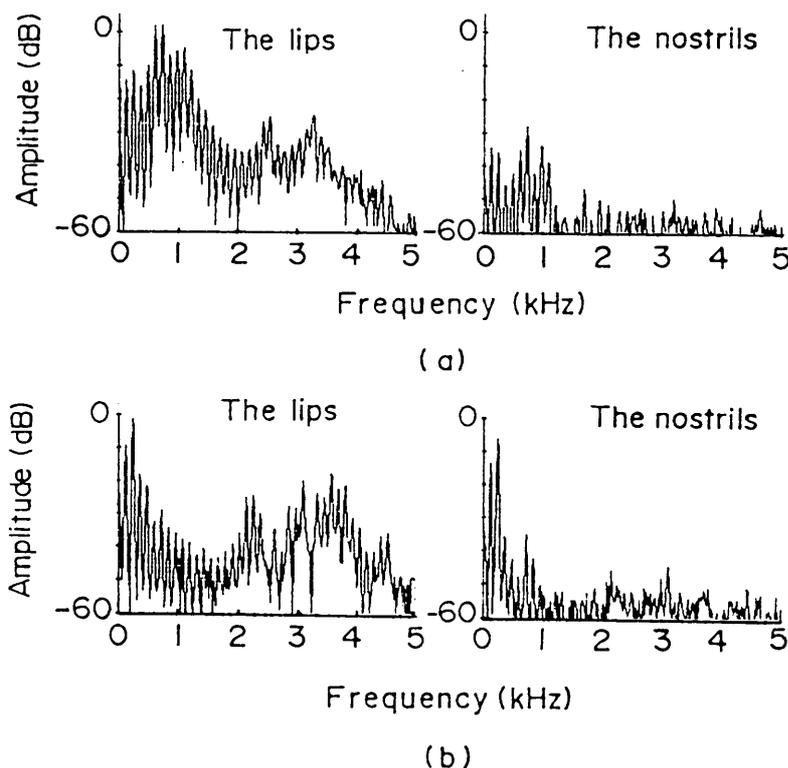


図3. 8 実音声の口唇放射音と鼻孔放射音のスペクトル  
(a) 母音*/a/* (b) 母音*/i/*

母音 /i/ では、第1ホルマントにあたる部分に鼻孔放射のエネルギーの大部分があり、口唇放射に比べても4dB低いだけなので第1ホルマントに影響があると思われる。/i/の鼻孔放射音では、強いピークは母音の第1ホルマントの付近にあり、それは口唇に比べて多少小さい程度である。

### 3.5 音声生成の声道モデル

非鼻音化音声を発声する時、軟口蓋は閉鎖している。然し、上述の測定結果より、非鼻音化母音でも音が鼻孔と咽頭壁から放射されている。このような音響特性を音声生成モデルで説明するため、図3.9に示したFlanagan-Ishizakaらの音声合成システム<sup>6)</sup>を用いた。声帯モデルは2質量モデルで、口腔断面積の値は文献[7]によるデータを用い、鼻腔の断面積はFantのデータ<sup>8)</sup>を用いた。このモデルを口腔と鼻腔の音

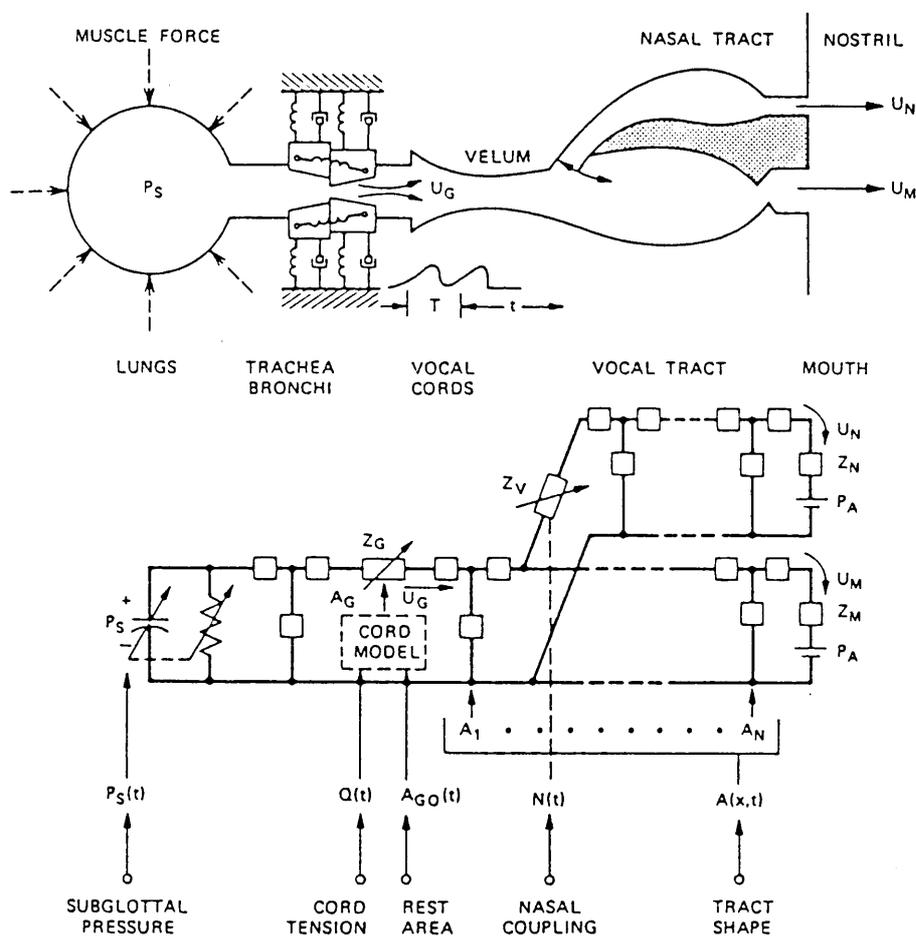


図3.9 音声生成モデル (文献[6]による)

響結合の研究に対して有効にさせるため、まず、実音声に合う鼻腔モデルを導入する。

鼻腔モデルの特性を実際の特性に合うようにするため、そのモデルを用いて合成した /ma/ の /m/ の特性を話者 5 名の平均値に近い話者 N の鼻子音付き音節 /ma/ と /mi/ の /m/ に近づけるようにした。話者 N の /ma/ と /mi/ の /m/ のスペクトルを図 3. 10 に示す。鼻の解剖学的な構造では固有鼻道以外に数多く容積の小さい副鼻腔が存在する。ヘルムホルツ共鳴の原理により、これらの副鼻腔は周波数領域でいくつかの零点を導入する。副鼻腔のうち上顎洞や、前頭洞や、蝶形洞と篩骨洞がある。まず、容積の大きい上顎洞、前頭洞と蝶形洞の影響を考慮に入れた。この 3 つの副鼻腔は文献[9]のデータを参考にして、アンチホルマント周波数が 500Hz、1000Hz、1600 Hz になるように導入した。このようにしても、高周波成分がまだ大き過ぎる。ここで、さらに篩骨洞を 3 つの副鼻腔として導入した。鼻腔と副鼻腔の断面積を図 3. 11 に示す。また、副鼻腔のない場合とある場合に合成した /ma/ の /m/ を図 3. 12 に示す。副鼻腔がある場合の /m/ を実音声に比べて 3kHz 以下のホルマントの周波数(実音声では 260Hz、800Hz と 2200Hz)においては誤差は 20% 以下である。第一ホルマントに

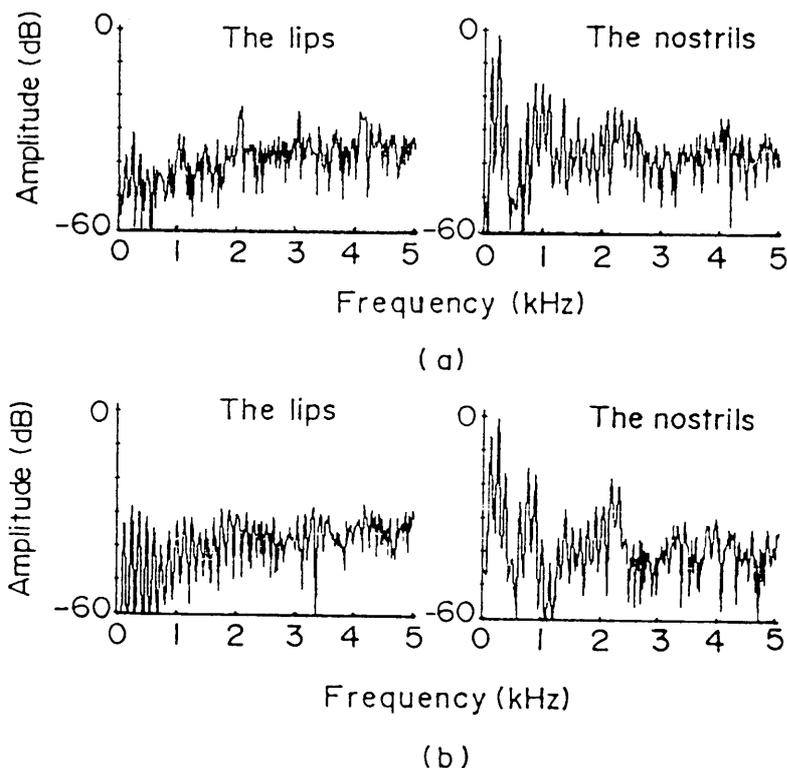


図 3. 10 実測した /m/ のスペクトル  
(a) /ma/ の /m/ の口唇放射と鼻孔放射  
(b) /mi/ の /m/ の口唇放射と鼻孔放射

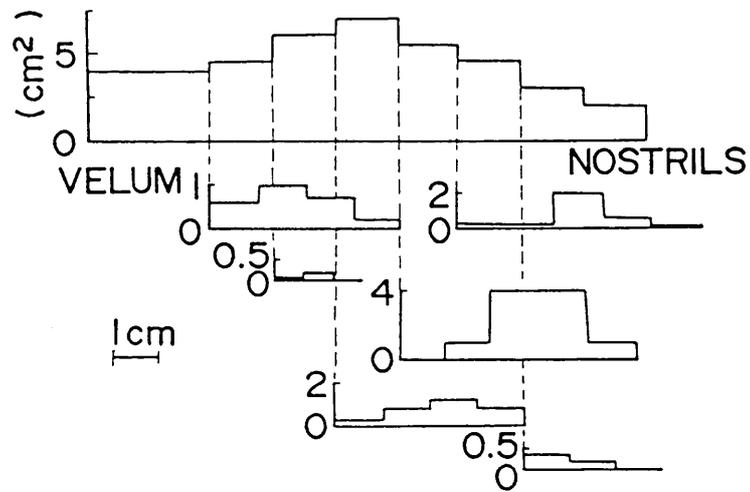


図 3. 1 1 鼻腔と副鼻腔の断面積

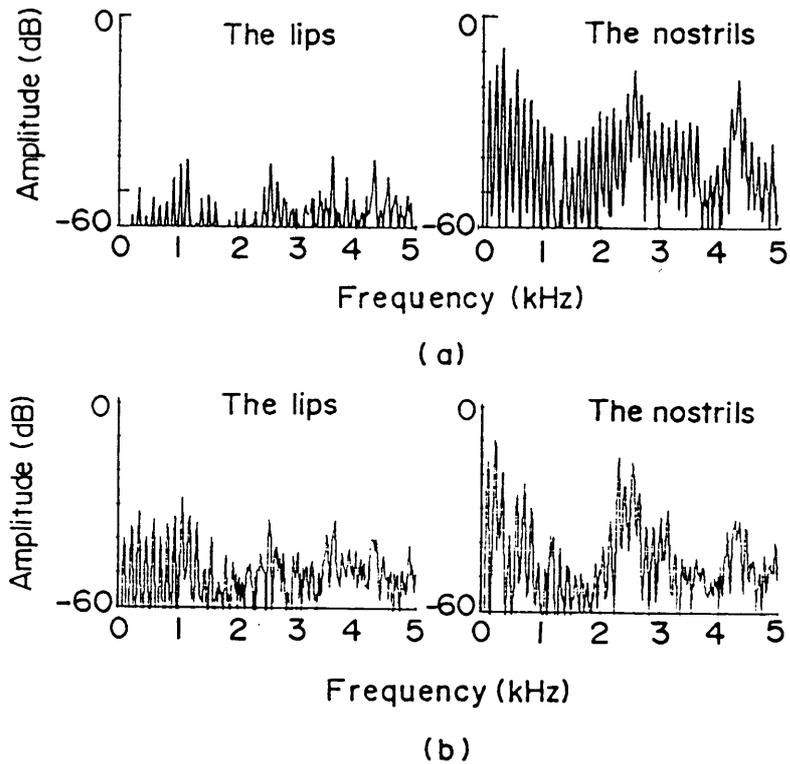


図 3. 1 2 副鼻腔がない場合とある場合に合成した /ma/ の /m/ (a) 副鼻腔がない場合 (b) 副鼻腔がある場合

対する相対的な大きさについては、2200Hz付近の差が大きいが、第二ホルマントでの差は3dB程度である。鼻孔放射の場合実音声との相関係数は副鼻腔のない場合0.75、副鼻腔のある場合0.80となっているので、後者は実音声に近い。

副鼻腔を考慮したモデルで軟口蓋を単純な音響抵抗にした場合、合成した母音は実測データに合わない。そのため、軟口蓋の機構を詳しく考察する必要がある。

### 3. 6 軟口蓋の振動板モデル

発声時軟口蓋の状態をあらためて考察するために、日本語の母音 (/i, u, a/) を持続発声した時の正中矢状断面のX線写真(文献[13]より)を図3. 13に示す。

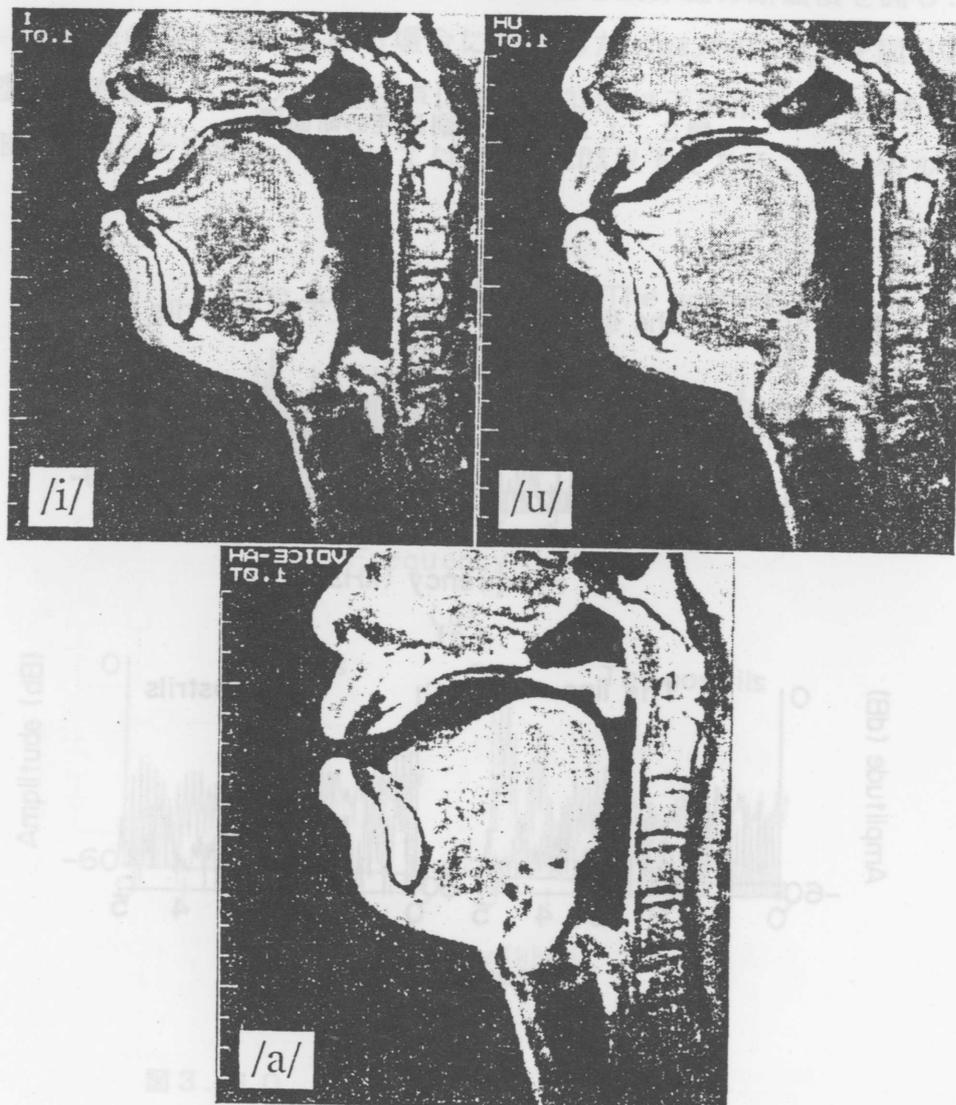


図3. 13 母音/i, a, u/発声時の調音器官の正中断面MRI画像  
(文献[13]より)

3つの発声とも、軟口蓋はしっかり閉じた状態で、口腔と鼻腔との間に空気の通路はない。しかし、観測した結果では鼻孔からも音が放射されている。このような音響放射は軟口蓋の筋肉板の上下振動による口腔と鼻腔の音響的な結合によるものと考えられる。本節では、軟口蓋の機構を考察して軟口蓋の振動に関する音響モデルを構築する。さらに、このモデルを用いて母音における音響放射をシミュレーションする。

### 3. 6. 1 軟口蓋振動の等価回路

軟口蓋が閉鎖した場合、前の部分は硬口蓋とつながり、後の部分は咽頭壁に密着している。この場合の軟口蓋の筋肉板が円形に近い。軟口蓋の面積は咽頭壁の横方向の振動によって微動するが、簡略化のため円板と仮定する。さらに、この円板は剛壁にはまり、周辺はばねで支えられ、機械抵抗が存在し、上下に振動すると仮定する。

軟口蓋がある厚さを持つ粘弾性体なので、口腔内音圧が軟口蓋に力を加えるとき、鼻腔側の面は口腔側の面と全く同じには動かない。そこで、軟口蓋を図3. 14(a)のような二枚の板の複合でモデル化する。一枚毎の単位面積あたりの質量、機械抵抗及びステイフネスを軟口蓋の単位面積あたりの質量 $M$ 、機械抵抗 $R_m$ 及びステイフネス $K_m$ の半分にして、それぞれ $m$ 、 $r_m$ 、 $k_m$ で表わす。二枚の板の間に挟まれている粘弾性体は、オークト模型<sup>10)</sup>で考えると、弾性ステイフネス $k_{m0}$ と、粘性抵抗 $r_{m0}$ の並列接続として表わされる。 $F$ は軟口蓋に口腔側から加わった力で、 $v_1$ と $v_2$ はそれぞれ軟口蓋の口腔側の面と鼻腔側の面の速度である。この機械モデルの運動方程式は

$$\begin{aligned} m \frac{dv_1}{dt} + r_m v_1 + k_m \int v_1 dt &= F - r_{m0}(v_1 - v_2) + k_{m0} \int (v_1 - v_2) dt \\ m \frac{dv_2}{dt} + r_m v_2 + k_m \int v_2 dt &= r_{m0}(v_1 - v_2) + k_{m0} \int (v_1 - v_2) dt \end{aligned} \quad (3.2)$$

となる。

音圧を $P$ 、軟口蓋の面積を $S$ 、体積速度を $U_i$  ( $i=1,2$ )とすれば、音響パラメータと機械パラメータの関係は

$$\begin{aligned} P &= F/S, \quad U_i = S v_i \quad (i=1,2) \\ R &= r_m/S, \quad C = S/k_m, \quad L = m/S, \\ R_0 &= r_{m0}/S, \quad C_0 = S/k_{m0} \end{aligned} \quad (3.3)$$

となる。式(3.2)と(3.3)を用いて変換すると、軟口蓋の音響等価回路は図3. 14(b)のようになる。

### 3. 6. 2 振動板モデルのパラメータの推定

図3. 14のような振動板モデルの機械インピーダンスを推定するには、声道の音響管モデルによる音声合成法で、上述の機械インピーダンスの値を現実になりそうな範囲で変化させ、合成した音声の放射量とスペクトルが実音声のそれと合うような値を求める。手がかりとする音声サンプルとしては、口唇放射が小さく鼻孔放射が大きい/i/と口唇放射が大きく鼻孔放射が小さい/a/を選んだ。

軟口蓋の面積は $4\text{cm}^2$ とし、単位面積あたりのステイフネス $k_m (=2k_m)$ を声道壁と同じ $8.45 \times 10^4 \text{dyne/cm}^2$ と仮定する。声道壁のステイフネスの測定値はいくつかあるがここでは値の小さい方(すなわち、軟らかい方)<sup>12)</sup>を選んだ。軟口蓋の密度は人体の平均密度とほぼ同じ $1\text{g/cm}^3$ であると仮定する。実音声のデータとしては、3.

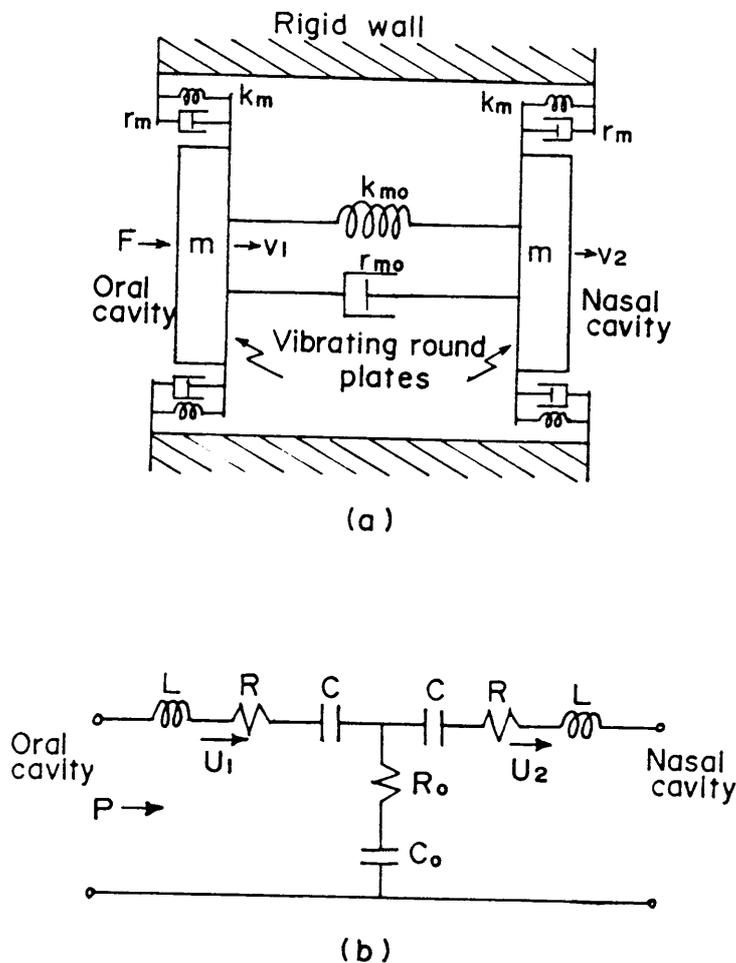


図3. 14 軟口蓋の機械モデルと音響モデル  
 (a) 軟口蓋の機械モデル  
 (b) 軟口蓋の音響モデル

3の測定値から5人のデータの平均値に近い一人(話者N)のデータ(図3.8参照)を選んだ。

まず、500Hz以下の帯域で、合成音声の鼻孔放射音の大きさを実音声に合わせるよう、軟口蓋の単位面積あたりの質量 $M(=2m)$ を $0\sim 1\text{g/cm}^2$ の間で変えてみたところ、 $0.15\text{g/cm}^2$ が最適であった。つぎに300Hz付近の鼻孔放射と口唇放射の相対的な大きさについて、合成音声/i/を実音声と同じく鼻孔放射量が口唇放射量より4dB程度小さくなるようにすると、単位面積あたりの機械抵抗 $R_m(=2r_m)$ は $50\text{g/s}\cdot\text{cm}^2$ となった。さらに、/a/と/i/の2000Hz付近の音響特性を用いて、合成音を実音声と同じく鼻孔放射が口唇放射より約-25dB小さくなるようにすると、弾性ステイフネス $k_{m0}$ は $2.01\times 10^5\text{dyne/cm}^2$ 、粘性抵抗 $r_{m0}$ は $400\text{g/s/cm}^2$ となった。

上述の機械インピーダンスを式(3.3)に代入すると、軟口蓋の音響パラメータは次のようになる。

$$\begin{aligned}L &= 1.875 \times 10^{-2} \text{g/cm}^4 \\ C &= 9.47 \times 10^{-5} \text{cm}^4 \text{s}^2/\text{g} \\ R &= 6.25 \text{g/s/cm}^4 \\ C_0 &= 1.99 \times 10^{-5} \text{cm}^4 \text{s}^2/\text{g} \\ R_0 &= 100 \text{g/s/cm}^4\end{aligned}\tag{3.4}$$

非鼻音化母音の場合、口腔と鼻腔とは、上述のパラメータを持ち、図3.14のような軟口蓋の振動板モデルで結合している。

### 3.6.3 シミュレーションの結果

図3.9の音声生成モデルに上述の軟口蓋振動モデルを入れた音響回路を用いて、5つの母音を合成した。例として、/a/と/i/の口唇放射と鼻孔放射のスペクトルを図3.15に示す。/a/の場合、目標通り鼻孔放射音圧の最大値は口唇放射音圧の対応値より25dB程度低くなっている。鼻孔から出る音は500Hz付近の零点の影響で第3、4高調波成分が小さくなっている。口唇の第一ホルマントにあたるものは鼻孔からの放射音にも現れた。1kHz以上の領域で意味のあるピークと谷はない。合成音と実音声との相関係数は口唇放射では0.81、鼻孔放射では0.90となっている。母音/i/の場合もほぼ目標通り鼻孔放射の最大音圧は口唇放射音圧の対応値より4dB程度低く、/a/と同じ500Hz付近に零点が現れる。2kHz付近には第二ホルマントに対応するピークが現れた。合成音と実音声との相関係数は口唇放射では0.71、鼻孔放射では0.88となっている。鼻孔放射の第一ホルマントの周波数は口唇放射に比べて、実音声

と合成音とも /i/ では多少下に移り、 /a/ では多少上に移る現象がある。これは中心周波数が500Hzにある副鼻腔の影響の為である。また、軟口蓋の振動板モデルで合成した他の母音のスペクトルも実音声の測定値の偏差範囲に入っている。

合成した5つの母音に対して、口唇放射と鼻孔放射の音圧レベルをそれぞれ求めた。それと3.4で測定した5人の実音声の音圧レベルを図3.16に示す。黒マークは実測データの平均値で、縦棒は測定データの偏差範囲を示している。口唇放射と鼻孔放射の音圧差については、 /e/ では合成音声の実音声の平均値より2.7dB小さくなっているが、他の母音では両者は同じである。シミュレーションの結果より、本章で構築した軟口蓋の振動板モデルは、非鼻音化母音区間の軟口蓋の働きをよく表わし、非鼻音化母音における鼻孔放射を軟口蓋の振動による口腔と鼻腔の音響的結合の結果として説明できることがわかる。

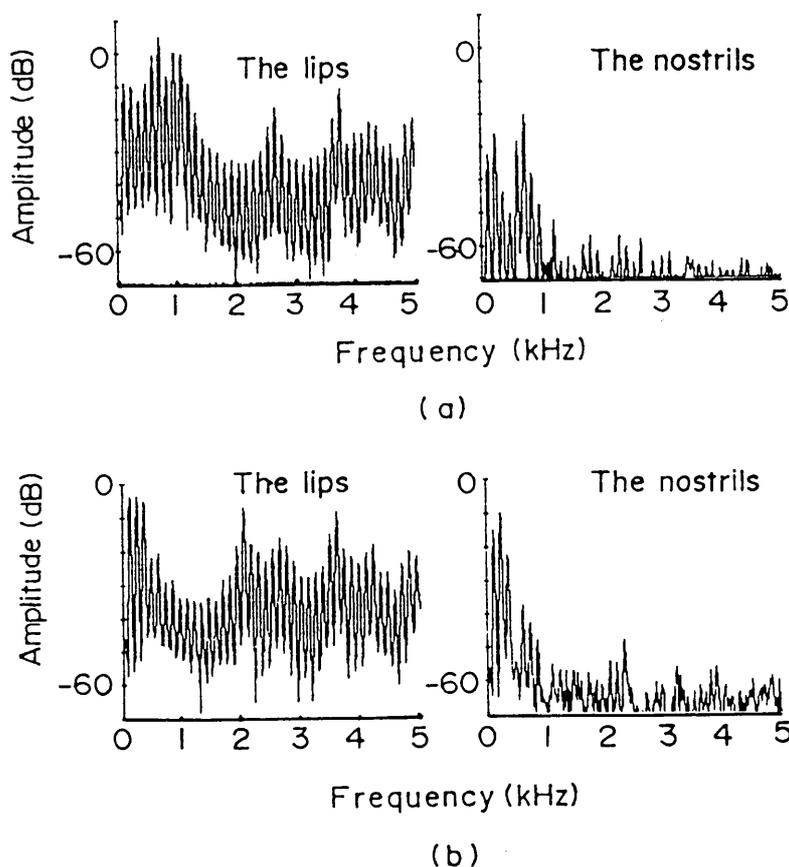


図3.15 合成音の口唇放射音と鼻孔放射音のスペクトル  
(a) 母音/a/ (b) 母音/i/

### 3. 7 考察

本節では、3. 4の測定結果と3. 6のシミュレーション結果について考察する。

非鼻音化母音の測定結果では、広母音/a, e, o/は口唇放射が大きくて、それに比べて鼻孔放射と咽頭壁放射が無視できる程度小さい。広母音に反して狭母音/i, u/では口唇放射は小さくなるが鼻孔放射と咽頭壁放射は大きくなった。周波数領域の分析結果より、非鼻音化音声においても狭母音の場合鼻孔放射は必ずしも無視できないことがわかった。

母音を発声する時の咽頭壁及び軟口蓋の動きは筋電図や口腔内圧などの方法で測定されている<sup>2)~5), 11)</sup>。それらの測定によれば、母音発声時の咽頭壁と軟口蓋の動きは咽頭腔内圧に応じて生ずる受動的な運動であり、狭母音の咽頭腔内圧は広母音のそれより高いので咽頭壁と軟口蓋の動きは狭母音の方が強い。本研究で測定した母音発声時の口腔内圧を表3. 7に示す。そのデータは口唇から6cm奥のところの口腔内音圧であり、上述した咽頭腔の内音圧と若干差がある。総合的にみると口腔内音圧は狭母音の方が広母音より高い傾向がある。鼻孔放射と咽頭壁放射の相対的な大きさが話者によって変わるが、母音に対しては両放射音の変化傾向は同じである。すなわち、鼻孔放射と咽頭壁放射とも狭母音（口腔内音圧の大きい）の方が大きく広母音の方が小さい。このように咽頭壁放射と鼻孔放射の大きさが口腔内音圧に応じて変化するの

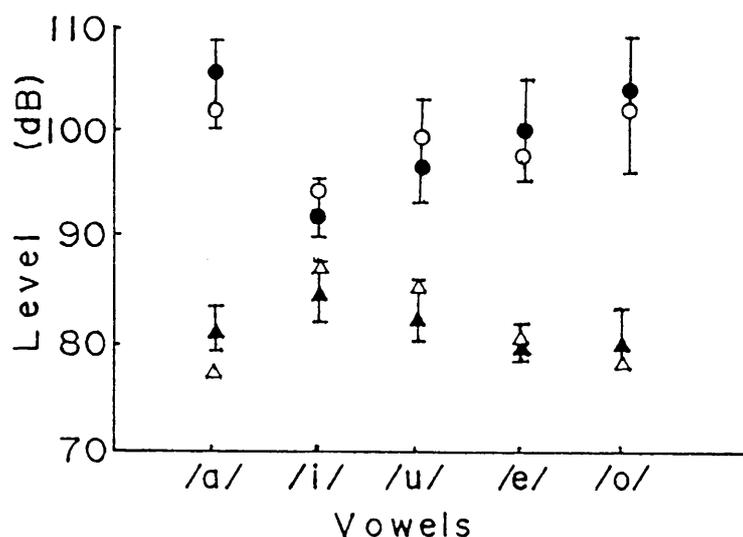


図3. 16 実測音と合成音の非鼻音化母音の音圧

●▲：実音声の口唇放射と鼻孔放射

○△：合成音の口唇放射と鼻孔放射

は、咽頭壁放射が咽頭壁の振動によるもので、鼻孔放射が軟口蓋の振動による口腔と鼻腔の結合であることを示している。3. 6の音声生成モデルの作成では、軟口蓋を二枚の複合振動板で構成して、そのパラメータは一部の母音サンプルから推定したが、それで合成した音声は推定に用いた母音ばかりでなく他の母音でも音圧レベルとスペクトルにおいても実音声とよく一致した。このことより、軟口蓋の振動板モデルは非鼻音化母音の鼻腔と口腔との音響結合の働きをよく表現できる。

また、閉鎖した場合の軟口蓋のモデルをばねとダッシュポットでつながる二枚板にした理由について説明する。軟口蓋を一枚の円板と仮定すると、軟口蓋モデルは図3. 14の $R_0$ と $C_0$ の分岐をはずしたRLCの直列回路となる。一枚の振動板モデルと3. 6で述べた二枚の振動板モデルを用いて合成した/i/のスペクトルを図3. 17に示す。一枚の振動板モデルの場合、鼻孔放射では2000Hz近くに強いピークは現れているが実音声にはほとんどみられない。また、そのピークは口唇放射のピークに比べて15dBだけ低い、実測値にはそのピークにあたる周波数での差は25dB程度となっている。実音声の特性により、軟口蓋をフィルタとしてみると2次以上のローパスフィルタの特性が必要である。一方、2次ローパスフィルタに相当する二枚の振動板モデルを用いて合成した音声は実音声とほぼ一致している。それゆえ、本論文では軟口蓋を二枚の振動板に粘弾性体をはさむ振動板モデルにした。

表 3. 7 実音声の口腔内音圧 (dB)

	/a/	/i/	/u/	/e/	/o/
Speaker N	122.4	132.1	137.2	123.0	130.2
Speaker K	120.2	128.2	130.3	126.6	130.9
Speaker M	113.1	129.2	131.7	122.9	130.0
Speaker H	103.1	123.4	125.0	107.3	117.0

### 3.8 むすび

本章では、発声時の口唇、鼻孔及び咽頭壁の放射音と、頬、鼻孔壁及び咽頭上部壁の振動を同時に測定した。各部位からの放射音は同部位の声道壁の振動で確認した。放射音の測定結果より、音は非鼻音化母音でも鼻や咽頭壁部などから放射され、鼻孔放射は狭母音の方が広母音より大きく、口唇放射と鼻孔放射の音圧差は狭母音の方が小さいことがわかった。このことに基づいて、閉鎖している軟口蓋を機械抵抗とステイフネスでつながる二枚の振動板としてモデル化し、口唇放射と鼻孔放射の関係を用いてモデルのパラメータを推定した。シミュレーションの結果では、この振動板モデルは非鼻音化母音の口腔と鼻腔の音響的結合をよく説明することができる。

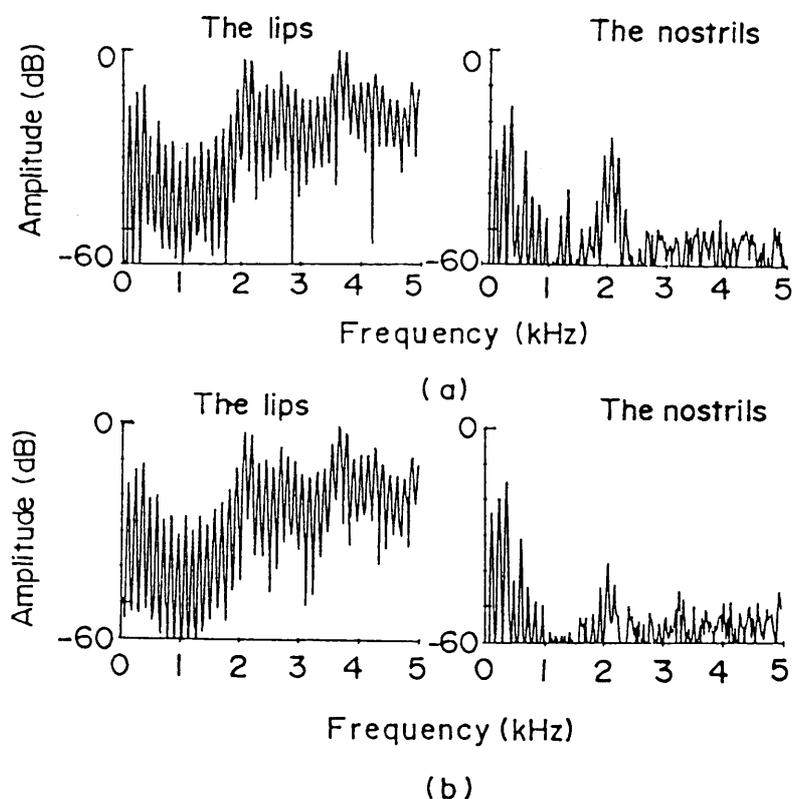


図3.17 合成音/i/の口唇放射と鼻孔放射のスペクトル  
(a) 軟口蓋を一枚振動板とした場合  
(b) 図3.14のモデルを用いた場合

## 第4章 子音区間での音響測定と軟口蓋の動きの推定

### 4.1 まえがき

非鼻音化母音区間では、口腔内音圧により軟口蓋が振動しており、その振動によって鼻腔と口腔とが結合している。母音と比べて、非鼻音化子音、特に破裂子音区間では、口腔内圧が激しく変化している。この場合、軟口蓋はどのように変化するか、またその変化は各部位の音圧にどのように関連しているかなどの問題がある。これらの問題は軟口蓋の音響モデルの作成では避けられない。本章では、有声破裂子音の音響特徴に着目して、口腔内圧と放射音圧、鼻孔放射と軟口蓋の動きの関係を考察する。そのため、有声破裂子音をはじめ他の子音を含む音声資料を発声する時の口腔内圧や、口唇放射、鼻孔放射と咽頭壁放射をそれぞれ分離して測定する。単音節と連続音声における発声器官各部位の音圧の関係を考察する。さらに、それらの観測結果に基づいて軟口蓋の音響モデルを作り、このモデルを用いて有声破裂子音の音響放射についてシミュレーションを行う。

### 4.2 測定装置と音声資料

#### 4.2.1 測定装置

発声時の口腔内圧や、口唇放射、鼻孔放射と声道壁放射を分離して観測するために、図4.1のように防音箱と遮蔽板を用いた。分離された鼻孔放射、口唇放射と咽頭壁放射をそれぞれマイクロホンM1、M2とM3で収録した。マイクロホンは鼻孔部にはRion-UC26、口唇部にはRion-UC26、咽頭壁部にはB&K-4133を用い、それぞれ精密騒音計Rion-NA60、Rion-NA40とB&K-2209に接続して、そのAC出力を取り出した。3個のマイクロホンの出力レベルの校正は2.5で述べた方法で行った。また、プローブマイクロホンB&K-4182を口腔内(口唇から6cm奥)に挿入して口腔内音圧を測定した。これら4チャンネルの信号はデータレコーダ(Sony社DFR-3415)に収録し、4.8kHzのローパスフィルタを通しA-D変換して、ワークステーション(H.P.アポロDN3500)に取り込んだ。サンプリング周波数は10kHz、語長は12bitである。

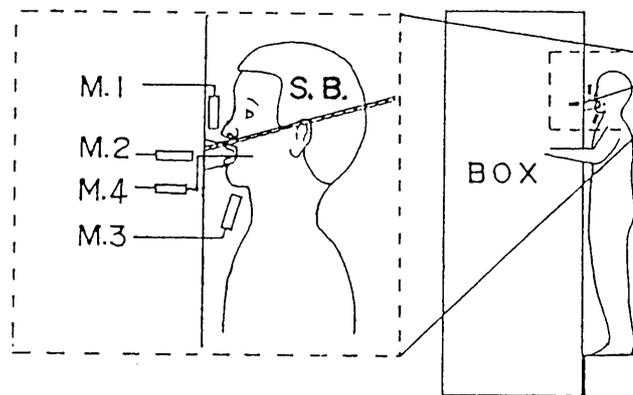
#### 4. 2. 2 音声資料

音声資料は表3. 1で示した日本語の100音節と、表4. 1に示している単語あるいは意味のない音素系列、有声破裂子音、破擦子音と無声破裂子音を含むものである。発声時、口腔内にプローブマイクロホンが挿入されているので、語中の音素の選択では発声器官の動きを妨げない音素を選んだ。また、被験者には慣れるまで練習させた。表3. 1の音声資料を成人男性5人、表4. 1の音声資料を成年男性7人に発声させた。

#### 4. 3 子音区間における放射音の音圧レベル

##### 4. 3. 1 音圧レベルの求め方

防音箱を用いて発声する時、箱内の音圧には呼気の流入によって音声の開始期に信



M.:Microphone S.B.:Shield board

図4. 1 音声を測定するシステムの略図

M.1、M2、M3: 鼻孔、口唇、咽頭壁のマイクロホン

M.4: 口腔内に挿入しているプローブマイクロホン

表4. 1 音声資料

bunbetsu(分別)	binjyo(便乗)
bunbutsu(文物)	daibubun(大部分)
zaba	dobo banba

号の揺れが生ずる事があるので、それに対して非巡回形フィルタ（FIR）による高域通過デジタルフィルタを用いて信号を濾波した。このフィルタの伝達関数は

$$H(z) = \sum_{n=0}^{N-1} h(n)z^{-n} \quad (4.1)$$

である。ここで、 $h(n)$ はフィルタの係数で、 $z$ は $Z$ -変換の演算子である。フィルタの次数 $N$ は329にした。急峻な遮断特性を得るためチェビシェフ・フィルタを用いた。このフィルタについては、遮断周波数は75Hzで、通過域の許容偏差は0.5dB以内で、阻止域の減衰は25dB程度である。

有声破裂子音における音響放射の特徴を明らかにするため、有声破裂子音以外の子音における発声器官各部位からの放射音も収録して比較しながら分析した。4チャンネルの信号波形をディスプレイ画面に並列表示し、目視で子音区間を切り出し、式(4.2)に従ってその区間のサンプル値の平均値からのずれの2乗平均のデシベル値 $P$ （以下音圧レベルと呼ぶ）を求めた。

$$P = 10 \log_{10} \left( \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - \bar{X})^2 \right) \quad (4.2)$$

ここで、 $N$ は切り出し区間のサンプル数で、 $x(n)$ は $n$ 番目のサンプルで、 $\bar{X}$ は切り出し区間の平均値である。

#### 4.3.2 単音節の放射音圧レベル

子音区間としては、無声子音は破裂区間または摩擦区間、有声破裂子音はバズ区間、鼻子音は nasal murmur、有声摩擦音は有声摩擦区間、他の有声子音は母音の定常部になるまでの過渡区間を切り出し、式(4.2)を用いて音圧レベルを求めた。第3章と同じ話者5名のデータを表4.2～表4.6に示す。表の数値がない項目はバズバーがないかあるいは子音区間の判断ができない音節である。子音ごとに求めた音圧レベルの標準偏差は5dB程度で、母音の場合より3dB程度大きくなっている。咽頭壁放射と鼻孔放射の相対的な大きさによって話者5名を二つのグループに分けることができる。グループ1は話者3名（話者N、話者C、話者S）であり、鼻孔放射の方が咽頭壁より大きくなっている。グループ2は話者2名（話者Y、話者H）であり、鼻孔放射は咽頭壁放射と同じ程度である。代表的な話者Nと話者Yの子音放射音圧を子音ごとにまとめた平均音圧（後続母音によって1～5個の発声音を平均したもの）を図4.2に示す。そのうち、/f/は発声「フ」の子音で、/j/は発声「ジ」と「チ」の子音である。鼻子音以外の有声子音の場合、鼻孔放射音圧と咽頭壁放射音圧については、

表4.2 子音区間での放射音の音圧 (話者N)  
口唇放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	80.4	77.9	75.7	79.8	82.7	2.6	79.3
s	71	70.5	68.9	65.9	69.2	2	69.1
t	74.1	68.9	69.4	73.6	72.9	2.4	71.8
n	56.8	57	52.6	51.5	54	2.5	54.4
h	70.7	65.3	61.5	64.2	66.8	3.4	65.7
m	60.7	63	61.4	60.6	61.2	1	61.4
r	83.3	59.5	74.3	69.8	64.1	9.2	70.2
g	75.5	68.2	72.2	68.4	68.1	3.3	70.5
z	66.2	62.3	68	66.5	67.6	2.3	66.1
d	56.4	62.1	69.7	65.6	63.1	4.9	63.4
b	60.8	57.8	61.6	59.9	63.9	2.2	60.8
p	78.9	70.7	76.6	79.3	82.6	4.4	77.6

鼻孔放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	66.6	66.9	62.2	69.9	67.1	2.8	66.5
s	43.8	41.9	42.1	43.6	46	1.6	43.5
t	65.7	43.4	54.8	62.9	55.8	8.7	56.5
n	91.9	92.5	93.2	91	92.8	0.9	92.3
h	52.2	49.6	47	46	45.6	2.8	48.1
m	91	93	92.5	91	91.8	0.9	91.9
r	75.2	90	89.4	91.9	91.5	7	87.6
g	88.5	80.5	87.4	79.5	88.9	4.6	85
z	80.1	87.4	88.4	64.6	85	9.8	81.1
d	85.5	87.3	86.2	84.6	89	1.7	86.5
b	88.1	91.8	90.8	86	89.5	2.3	89.2
p	69.2	66.6	65.6	67.3	70.7	2.1	67.9

咽頭壁放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	60.6	63.9	59	67.2	61.2	3.2	62.4
s	51	48.1	49.3	48.7	48.2	1.2	49.1
t	60.2	50.4	53.7	60.3	55.6	4.3	56
n	76.4	77.1	78	75.2	76.7	1	76.7
h	48.6	49.1	47.9	47.8	48	0.6	48.3
m	75.8	77.6	78.1	76	76.7	1	76.8
r	69.5	78.8	78.1	78.3	76.7	3.9	76.3
g	77.7	74.4	78.5	70.2	76.1	3.3	75.4
z	76	78.2	78.6	56.1	77.4	9.6	73.3
d	77.3	78.9	78.5	75.8	79.3	1.4	78
b	73.9	76.5	75.5	71.8	74.3	1.8	74.4
p	63	60.2	59.7	61.9	65.4	2.3	62

表4.3 子音区間での放射音の音圧 (話者C)  
口唇放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	78.9	66.7	74.3	73.2	78.1	4.9	74.2
s	57	65.5	64.1	61.2	62.8	3.3	62.1
t	76.6	67.9	64.3	74.2	72.1	4.9	71
n	67.9	69.7	68.7	66.5		1.4	68.2
h	64.5	59.2	57.1	62.9	60.7	2.9	60.9
m	64	66.2	69.1	62.5	64.1	2.6	65.2
r	71.8	63.6	65	63.4	66.8	3.5	66.1
g	62.2	63.4	63			0.6	62.9
z	63.7	60.7	64.5	60.5		2	62.4
d	61	56.7	63	61.1	48.2	5.9	58
b	59	55.6	55.5	51.3	54.1	2.8	55.1
p		59.9	64.8	67	59.9	3.6	62.9

鼻孔放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	63.7	56.2	62.6	61	66.3	3.8	62
s	47.4	58.1	54.4	62.1	59.4	5.7	56.3
t	68.5	57.3	54.4	64.4	60.2	5.6	61
n	83.8	83.1	82.5	80.8		1.3	82.6
h	56.1	55.8	56.5	58.3	52.2	2.2	55.8
m	82.6	82	84.7	81.7	83.6	1.2	82.9
r	76.6	78.8	80.8	78.4	78.9	1.5	78.7
g	81.3	72.6	79.2			4.5	77.7
z	61.7	63.7	68.4	76.4		6.5	67.6
d	76.5	70.9	75.2	77.9	62.1	6.4	72.5
b	77.3	72.6	75.7	66.9	74.9	4	73.5
p		58.4	60.7	55.9	58	2	58.3

咽頭壁放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	59.1	51.4	57.6	57.2	59.6	3.3	57
s	44.1	46.6	45.7	49.5	45.2	2	46.2
t	63.8	45.8	45.6	60.5	58.6	8.6	54.9
n	70.4	71.4	70.9	68.4		1.3	70.3
h	48.3	47.7	46.7	45.8	44.8	1.4	46.7
m	68.7	69.4	72	68	70	1.5	69.6
r	65.3	73.4	72.5	68.9	68.1	3.3	69.6
g	73.4	68.1	71.7			2.7	71.1
z	51.3	50	62.6	68.7		9	58.2
d	70.1	65.2	70.6	70	54.9	6.7	66.2
b	70.1	69	69.4	63.4	65.5	2.9	67.5
p		57.6	56.7	60.4	53.1	3	57

表4. 4 子音区間での放射音の音圧 (話者S)  
口唇放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	87.1	70.3	80.9	70	85.3	7.3	78.7
s	60	67.5	60.4	62.9	57.3	3.8	61.6
t	73.2	71.2	66.8	71.8	69.8	2.4	70.6
n	67.6		68.7	66.7	67.1	0.9	67.5
h	66.2	56.2	76.3	67.6	63.5	7.3	66
m	62.9	67.7	62		63.2	2.6	64
r	64.9	48.2	63.2	60.8	53.5	7	58.1
g	52.4	71.4	67.2			10	63.7
z	61	58.1	60.6		66.8	3.7	61.6
d	64.1	63.8	63.1	64.1		0.5	63.8
b	64.5	65	55.2	63.4	59.4	4.2	61.5
p	67.3	58.8	58	64.1	58.7	4.1	61.4

鼻孔放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	66.9	54.6	65.6	56.3	68.2	6.4	62.3
s	45.8	55.1	52.4	54.2	49.6	3.8	51.4
t	60	55.9	54.6	58.2	54.8	2.3	56.7
n	88.2		86.1	87.1	87.2	0.9	87.2
h	52.7	49.3	53.9	44.7	51.6	3.6	50.4
m	88.8	87.7	81.7		88.2	3.3	86.6
r	76.8	59.9	81.4	85.4	74.4	9.7	75.6
g	72.2	81.8	85.5			6.9	79.8
z	76.5	75.5	78.1		79.5	1.8	77.4
d	81.9	80.9	79.2	77.8		1.8	80
b	85.3	81	68.6	86.5	75.5	7.4	79.4
p	55.2	39.6	37.9	50.3	42.9	7.3	45.2

咽頭壁放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	62.5	53.4	64	54.6	63.4	5.1	59.6
s	45.8	45.7	45.6	51.8	45.1	2.8	46.8
t	59.7	51.3	51	57	52.9	3.8	54.4
n	74.6		72.3	72.7	72.9	1	73.1
h	44.3	45.1	44.6	44.1	44.2	0.4	44.5
m	72.7	73.4	67.5		73.2	2.8	71.7
r	72	61.1	74.7	73	65.8	5.7	69.3
g	69	78.8	76.8			5.2	74.9
z	69.6	73.1	73.4		74	2	72.5
d	73.9	77.2	75.2	74.6		1.4	75.2
b	76.6	77.2	67.6	75.3	72.9	3.9	73.9
p	58	45.3	46.7	46.5	46.3	5.3	48.6

表4.5 子音区間での放射音の音圧 (話者Y)

口唇放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	86.4	78.4	81.5	79.7	83.5	3.2	81.9
s	72.3	68.6	61.9	69.9	65.4	4.1	67.6
t	82.8	75.2	70.4	77.9	79.8	4.7	77.2
n	63.4	54.4	56.8	57.5	60.3	3.5	58.5
h	64.2	71.9	68.7	77.9	69.6	5	70.5
m	69.3	67.8	71.5	66.7	64.5	2.6	68
r	94.5	68.3	68.5	89	94	13.4	82.9
g	66.6	64.8	66.4		73.7	4	67.9
z	63.1	72.9	70.9	75.2	75.2	5	71.5
d	59.6	65.6	67.7	67.4	77.1	6.3	67.5
b	70.4	69.7	71.4	61.5	67.9	3.9	68.2
p	71.5	73.6	74.3	69.9	63.3	4.4	70.5

鼻孔放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	65.9	63.5	61.9	63.6	66.1	1.8	64.2
s	59.3	40.5	49.8	53.8	43.8	7.6	49.4
t	67.6	53.1	58	63.1	68.4	6.5	62
n	95.5	95.1	93.8	92.1	96.3	1.6	94.6
h	52.7	52.5	59.5	43.8	38.7	8.2	49.4
m	96.8	98.9	99.3	95.7	92.5	2.7	96.6
r	79.8	82.7	83.5	75.3	75	4	79.3
g	80.3	83.1	78.4		73.7	4	78.9
z	77.5	82	80.5	75.7	78.3	2.5	78.8
d	80.2	84.6	83.4	80.6	81.4	1.9	82
b	79.7	82.2	83.5	73.5	79	3.9	79.6
p	68.1	65.1	66.2	65.2	63.2	1.8	65.6

咽頭壁放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	62.1	60.7	58.6	60.2	61.9	1.4	60.7
s	55.4	49.2	49	51.3	49	2.8	50.8
t	65.3	55.1	52.3	60.3	65.7	6	59.7
n	76.1	75.8	76.8	71.8	76.3	2	75.4
h	48.2	50.9	50	49	48.8	1.1	49.4
m	77.6	80	82	75.3	75.7	2.9	78.1
r	80.1	80.5	82	77.1	77.9	2	79.5
g	80.4	82.6	79		74.5	3.4	79.1
z	77.1	81	80.1	75.1	77.7	2.4	78.2
d	79.2	84.6	82.2	79.4	80.8	2.2	81.2
b	79.2	82.2	82.5	72.6	78.1	4	78.9
p	66.8	65.2	64.3	62.5	58.6	3.1	63.5

表4.6 子音区間での放射音の音圧 (話者H)  
口唇放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	72.8	76.1	79.1	75.2	74.4	2.3	75.5
s	71.5	69.1	65.6	63.2	63.7	3.6	66.6
t	78	70.8	68.3	72.1	72	3.6	72.2
n	72.4	66.4	72.2		75.2	3.7	71.6
h	63.5	56.7	86.3	62.1	63.3	11.5	66.4
m	76		70.5	66.3	75	4.5	72
r	85.3	86.3	75.8		86.3	5.1	83.4
g	76.5	67.4	77	73.9	76.8	4.1	74.3
z		64.9	62.7	67.7	61.4	2.8	64.2
d	64.3	64.4	56.6		66.3	4.3	62.9
b	68.4	62.5	68.5	66.2	64.7	2.5	66.1
p	76.5	69.3	69.8	70.6	81.9	5.5	73.6

鼻孔放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	60.8	64.2	61.7	63.5	58.7	2.2	61.8
s	60.4	58.3	60.8	59.1	56.8	1.6	59.1
t	63.8	54.6	56.1	59	56.2	3.6	57.9
n	87.3	90.8	94		91.9	2.8	91
h	56.2	54.7	56	55.7		0.7	55.7
m	93		90.5	86.6	91.9	2.8	90.5
r	80.4	81.7	84.9		84.7	2.2	82.9
g	89.1	85	89.7	87.2	78.1	4.7	85.8
z		78.8	75.5	84.1	75.4	4.1	78.5
d	80.2	78	65.9		75.1	6.3	74.8
b	74.4	74.6	74.2	77.6	74	1.5	75
p	70	64.3	61.85	67.6	68.5	3.3	66.5

咽頭壁放射音の音圧(dB)

	a	i	u	e	o	標準偏差	平均
k	61.2	64.8	63.2	64.6	59.3	2.3	62.6
s	53.3	46.3	48	48.6	49.4	2.6	49.1
t	67.2	54.1	52.8	60.1	57.8	5.7	58.4
n	72.4	76.1	80		76.6	3.1	76.3
h	44.8	45.6	45.8	43.8	44.6	0.8	44.9
m	80		76.6	72.3	79.4	3.5	77.1
r	72.7	79.7	77.5		75.5	3	76.4
g	80.2	80.3	79.8	81.9	79.6	0.9	80.4
z		78.8	75.8	78.5	76.3	1.5	77.4
d	80.9	78.8	67.2		76.1	6	75.8
b	75.3	75	74.6	78	75.3	1.4	75.6
p	70	64.3	61.8	67.6	68.5	3.3	66.4

話者Nは鼻孔の方が大きく、話者Yは双方が同じ程度である。

有声破裂子音のバズ区間や鼻子音区間では口唇は閉じている筈だがマイクM2からある程度の出力がある。図4. 2(a)の/n/の口唇放射には、母音部とnasal murmurとのレベル差が大きくて、使用したA-D変換器のダイナミックレンジの制約からnasal murmurに対応する口唇音がノイズと区別がつかないものがあったが、それ以外の鼻子音区間や有声破裂音のバズ区間ではたしかに口唇からの出力がある。

鼻子音区間での口唇部の出力を考察するため、第3章の測定方法で収録した/ma/と/na/の発声器官各部位の放射音と振動加速度の波形をそれぞれ図4. 3と4. 4

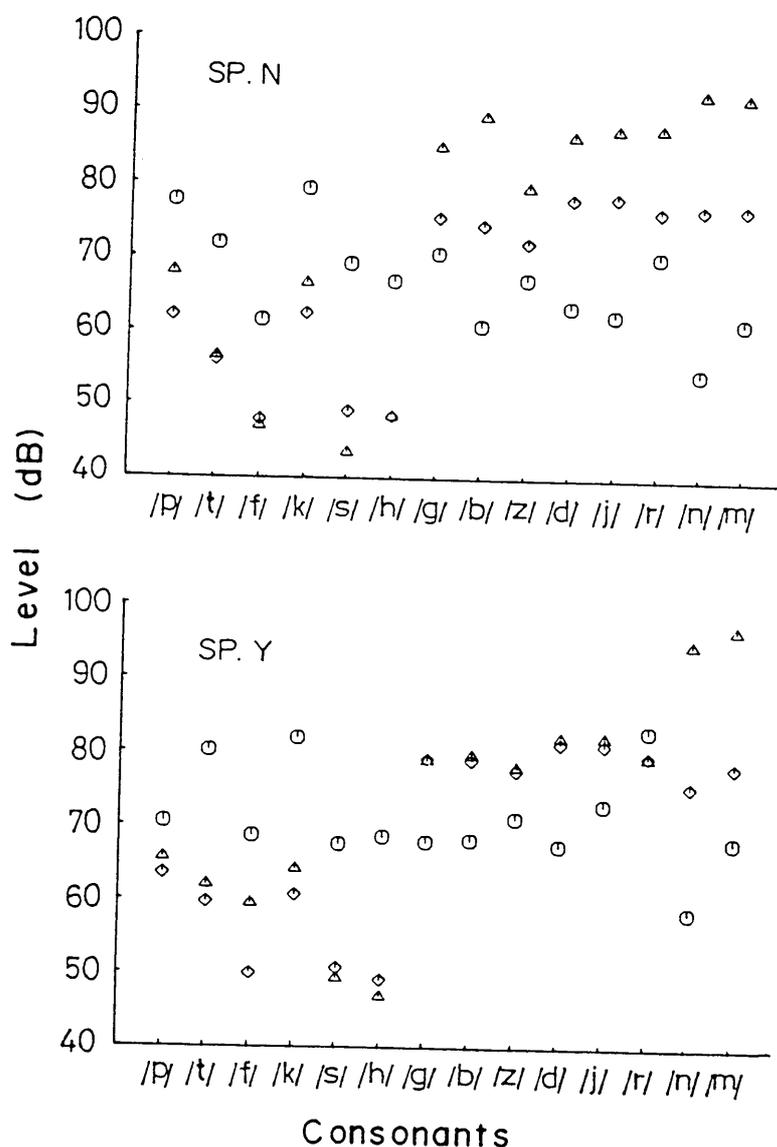


図4. 2 子音における放射音圧

○：口唇放射 △：鼻孔放射 ◇：咽頭壁放射

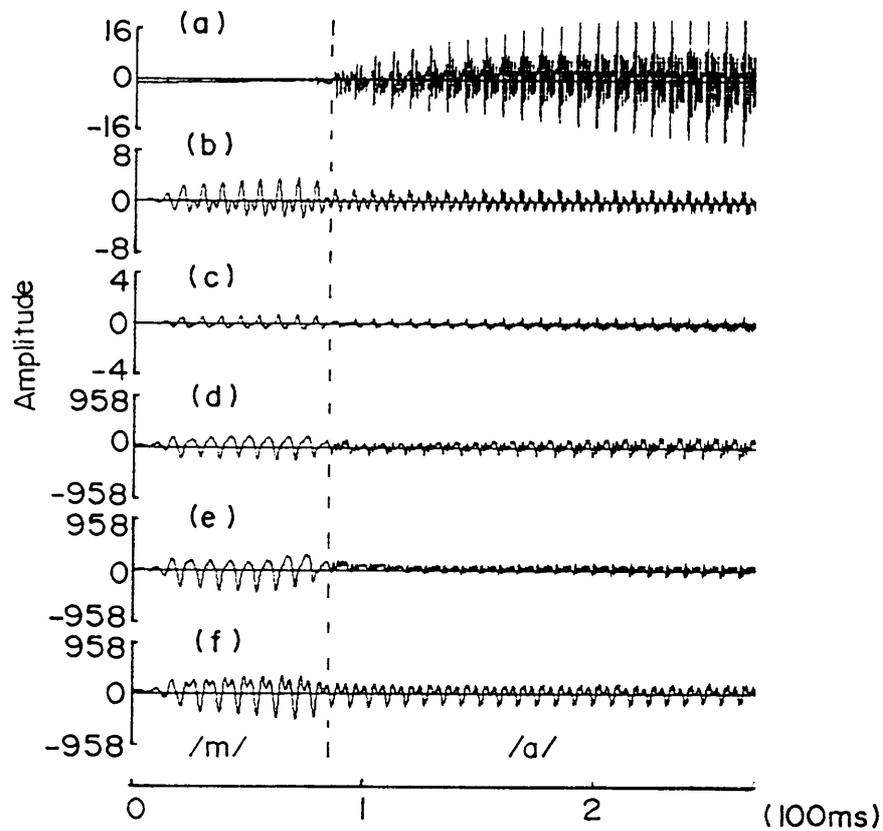


図4.3 実測音/ma/の各部からの放射音圧と加速度波形  
 (a)(b)(c) 口唇、鼻孔及び咽頭壁からの放射音圧 (Pa)  
 (d)(e)(f) 咽頭壁、頬及び鼻孔壁の加速度 ( $\text{cm/s}^2$ )

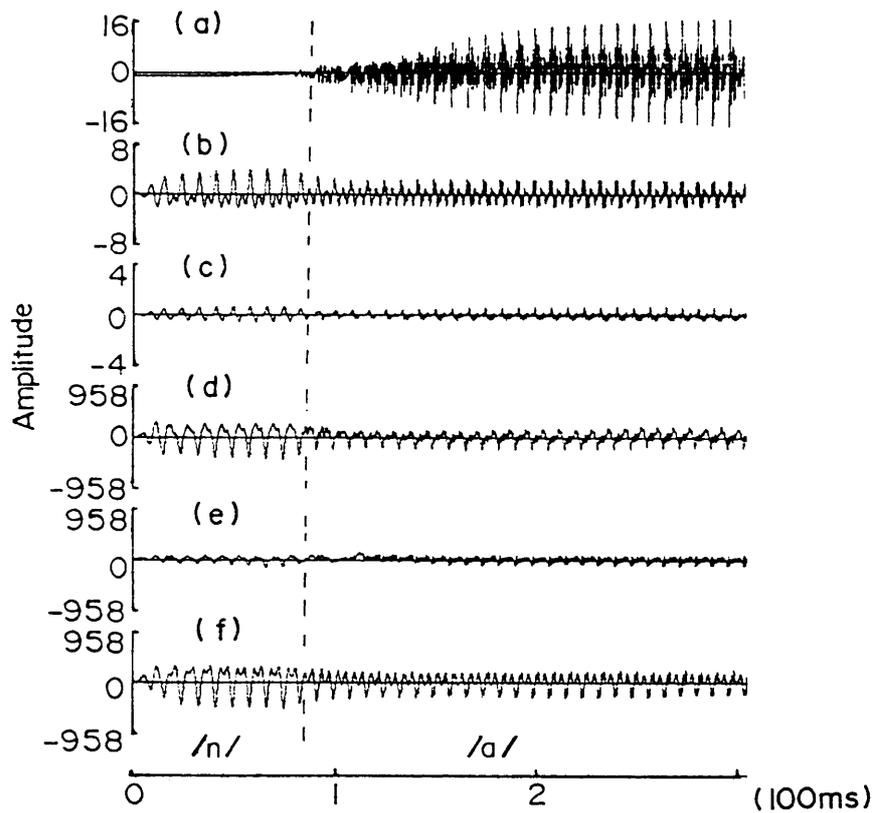


図4.4 実測音/na/の各部からの放射音圧と加速度波形  
 (a)(b)(c) 口唇、鼻孔及び咽頭壁からの放射音圧 (Pa)  
 (d)(e)(f) 咽頭壁、頬及び鼻孔壁の加速度 ( $\text{cm/s}^2$ )

に示す。鼻子音区間では、双方の口唇以外の放射音圧と、咽頭壁および鼻孔壁の振動ともほぼ同じであるが、/m/の頬（口唇付近）の振動量は/n/のその2倍以上となっている。鼻子音/m/と/n/において調音位置の違いで、/m/の方が/n/より口唇付近の振動が大きいので、マイクM2で収録した音圧が大きくなっていることが容易にわかる。図4.2にもその特徴が表れている。総合的にみると鼻子音区間で口唇からの出力は口唇とその周辺部の振動によるものと考えられる。

鼻子音(/m,n/)の鼻孔放射は口唇放射より25~30dB大きく、他の子音の鼻孔放射より話者Nで2~35dB、話者Yで10~39dB大きい。鼻子音以外の有声子音の場合、口唇放射は、鼻孔放射と咽頭壁放射より小さい。無声子音の場合、口唇放射は鼻孔放射より話者Nでは8~17dB、話者Yでは4~15dB大きくなっている。

#### 4.3.3 連続音声におけるバズ区間の放射音圧

連続音声における有声子音のバズ区間の音響放射の特徴は単音節の場合と比べて、鼻孔放射と咽頭壁放射との相対的な強さとバズバーの出現頻度が多少異なることである。声道閉鎖区間における音響放射部位について単音節の音声資料150個と単語の音声資料120個を調べた結果を表4.7にまとめる。この結果によると、バズバーの出現頻度は単音節の場合では89%で、連続音声の場合では100%である。相対的な大きさについては、単音節バズバーの85%と語頭のバズバーの65%は鼻孔放射の方が大きく、連続発声の語中にあるバズバーの69%は咽頭壁放射の方が大きくなっている。

また、有声子音が単音節または連続発声の語頭にある時、バズバーのないものが11%程度観測されたが、この場合、口唇からの破裂性の放射の後に声帯の振動が始ま

表4.7 声道閉鎖区間における非鼻音有声子音の鼻孔放射(Rn)と咽頭壁放射(Rp)の頻度の統計

voiced consonants and their context		sound radiating %	Rn > Rp %	Rn < Rp %
/g,b,z,d,j/ in mono-syllable		89	85	5
/g,b,z,d,j/ in a word	at begining	89	65	25
	other case	100	11	69
/r/		70	86	7

り、破裂以前には鼻孔と咽頭壁からの放射は無かった。口腔内圧の直流成分が声道閉鎖の開始から急に増大すると口腔内圧の交流成分は現れず、鼻孔と咽頭壁から音も放射されない。

上述の測定では、有声破裂子音（／b, d, g／）だけでなく、／j, r, z／など有声子音の声道閉鎖区間で鼻孔と咽頭壁から音が放射されている。つまり、鼻子音を除いて日本語有声子音においては声道の閉鎖区間では声帯が振動する限り、鼻孔と咽頭壁の双方から音が放射されている。

#### 4.4 口腔内圧と放射音

##### 4.4.1 口腔内圧と放射音の波形

有声破裂子音の例として／ba／の口腔内圧と発声器官各部位からの放射音の波形を図4.5に示す。有声子音のバズ区間で鼻孔と咽頭壁から放射した音声波形の包絡線は紡錘形をしている。それは、バズ区間で声帯の振動が立ち上がってから徐々に強くなり、その後徐々に弱くなることを示している。口腔内圧から見ると、発声が始まる時、口腔内圧の直流成分（基本周波数より低くゆっくり変化する成分を指す）は上昇し始め、最大値の200～500Pa（19.6～49cmH<sub>2</sub>O）まで上昇し、その後下がる。口腔内圧の声帯振動に伴う交流成分は直流成分に重畳して現れる。有声子音のバズバーはその交流成分に対応して、口腔内圧の直流成分の上昇部分、下降部分あるいはその全区間に現れる。

単音節の場合の／b／と連続音声の語頭にある／d／は、図4.5と図4.6に示しているように、咽頭壁放射の極大値は口腔内圧の交流成分の極大値と時間的に一致しているが、鼻孔放射の極大値は口腔内圧の直流成分と交流成分が極大値に達する前に現れている。一般に、鼻孔放射の極大値は口腔内圧の交流成分の極大値の5～20ms前、直流成分の極大値の40～110ms前にある。鼻孔放射音と口腔内圧の関係は次のようになる。(1)口腔内圧の直流成分が増大すると共に鼻孔放射の振幅は減少する。(2)鼻孔放射音のピークは口腔内圧の交流成分と直流成分のピークの前に現れる。(3)口腔内圧の直流成分に強いピークが現れる場合、そのピークの出現前に鼻孔放射はなくなるかまたは小さくなる。

##### 4.4.2 鼻孔放射音と鼻孔空気流

破裂子音の声道閉鎖区間においては、軟口蓋の上昇によって鼻孔から空気流（全体

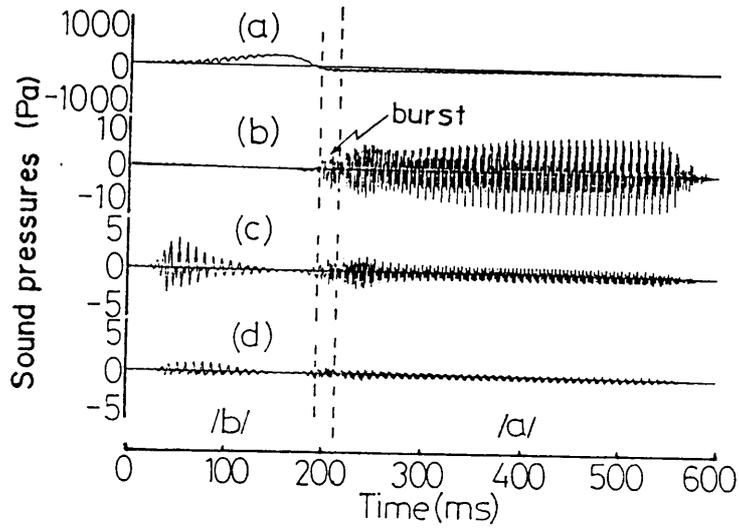


図4.5 /ba/の発声時口腔内圧と各部位の放射 (単位: Pa)  
 (a) 口腔内圧 (b) 口唇からの放射  
 (c) 鼻孔からの放射 (d) 咽頭壁からの放射

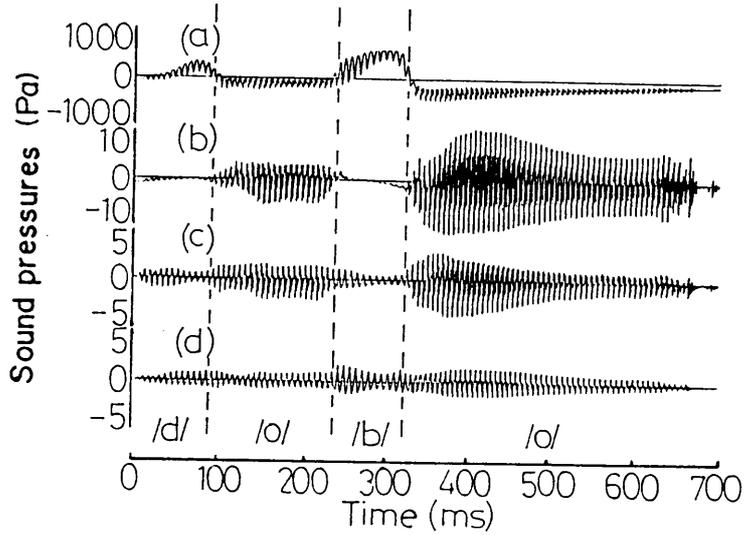


図4.6 /dobo/の発声時口腔内圧と各部位の放射 (単位: Pa)  
 (a) 口腔内圧 (b) 口唇からの放射  
 (c) 鼻孔からの放射 (d) 咽頭壁からの放射

で約 $4\text{cm}^3$ )が排出されており、口腔内圧(直流成分)が大きいほど鼻孔からの空気流が大きくなることがLubkerにより観測されている<sup>1)</sup>。それによれば、有声破裂子音のバズ区間で鼻孔放射音と鼻孔空気流の関係は次の(1)~(3)のようになる。(1)鼻孔の空気流が増大する前に放射音の振幅は減少する。(2)放射音は空気流のピークが現れる前になくなる。(3)放射音が停止した後、空気流は一定の時間に維持する。この結果の鼻孔空気流を口腔内圧の直流成分で置き換えると、文献[10]の測定結果は本研究の測定結果と一致している。このことは、空気流は口腔内圧の直流成分と同期して変化していることを示している。

Dicksonらの測定<sup>6),7)</sup>では、口腔内圧が高いほど軟口蓋は上昇することを明らかにした。鼻孔空気流は軟口蓋の上昇により引き起こされたものなので、口腔内圧の直流成分が高くなると排出空気は多くなる。一方、軟口蓋が上昇するほどそれ自身が硬くなり振動しにくくなるので、軟口蓋の振動による鼻孔放射音が少なくなる。結果的に、有声子音のバズバー区間で鼻孔の放射量が少くなり空気流の排出量が多くなるのは、双方とも口腔内圧の直流成分につれて変化するからである。文献[10]では鼻孔空気流は軟口蓋の能動的な動きによって引き起こされるものと報告されたが、鼻孔空気流と口腔内圧、鼻孔放射音圧と口腔内圧の関係より、鼻孔空気流の流れる区間での軟口蓋の動きには口腔内圧による受動的な成分もあることがわかる。

破裂子音の声道閉鎖区間で鼻孔放射音の極大値と口腔内圧(または鼻孔空気流)の極大値が時間的にずれていること、および無声破裂子音の鼻孔空気流が有声破裂子音のそれより大きいこと<sup>1)</sup>は、間接的にバズバー区間では軟口蓋に空気通路はないことを示している。第3章で考察した通り、この場合の鼻孔放射は軟口蓋の振動によって引き起こされるものである。

#### 4.4.3 無声子音における音響放射

連続音声における母音から無声子音へのわたりでは鼻孔と咽頭壁から音が放射されている。放射音の波形の例を図4.7に示す。母音/u/から無声破擦子音/ts/へのわたりで、破擦子音の発声のため舌と歯茎が一時閉鎖し、図4.7(b)にみられるように口唇から音の放射はできなくなり、口腔内圧は図4.7(a)で明らかのように高くなる。しかし、口腔内圧があまり高くないうちは声門部を空気流が流れ声帯の振動が続くので、この振動による交流成分は口腔内音圧の直流成分の上に重畳してしまう。交流成分が軟口蓋と咽頭壁を振動させたので、図4.7(c)(d)のように鼻孔と咽頭壁から音が放射される。7人のデータを調べたところ、子音/h/を除いてほとん

ど例外なく声道を閉鎖してから10~27msの区間で鼻孔と咽頭壁から放射音が現れている。

単音節の無声子音の破裂区間で鼻孔から音が放射されるものがいくつかみられた。音節／ka／の口唇放射、鼻孔放射と咽頭壁放射の波形とスペクトルを図4. 8に示す。子音／k／の破裂区間で鼻孔から音が放射されている。咽頭壁の放射音は口唇放射音と多少関連しているが、鼻孔放射音にはそういう関連性がみられない。周波数領域では鼻孔放射の200Hzのピークは口唇放射のそれより若干大きくなっている。この程度の鼻孔放射は子音の感知におそらく重要ではないが、口唇と鼻孔放射音のスペクトルの比較より、この場合の軟口蓋と鼻腔の働きは遮断周波数がほぼ200Hzにあるローパスフィルタであることがわかる。

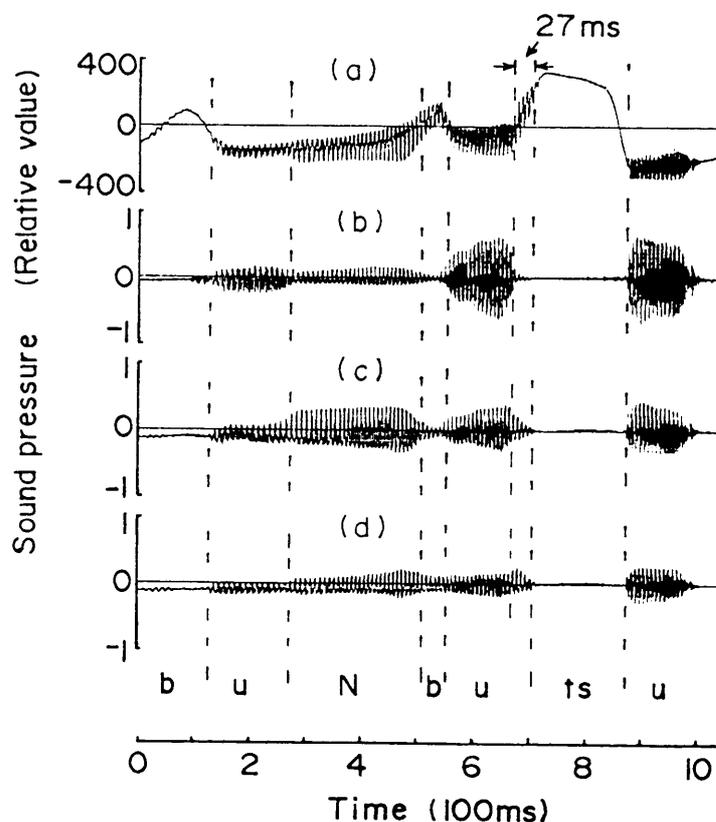


図4. 7 /buNbutsu／の発声時口腔内圧と各部位の放射（相対値）  
 (a) 口腔内圧 (b) 口唇からの放射  
 (c) 鼻孔からの放射 (d) 咽頭壁からの放射

#### 4. 4. 4 有声破裂子音の音響放射と鼻子音の音響放射

有声破裂子音のバズ区間で鼻孔から音がかなり放射されている。ここでは、バズ区間での放射音は鼻子音区間の放射音と比較し、どのような特徴を持っているかについて述べる。

有声破裂子音付きの音節 /bi/ と鼻子音付きの音節 /mi/ の波形をそれぞれ図 4. 9 と図 4. 10 に示す。有声破裂子音のバズ区間で鼻孔放射は必ず口腔内圧の直流成分の上昇に伴って行われている。鼻子音区間では口腔内圧の直流成分は上昇しない。これは、鼻子音区間では声道に空気通路があり、破裂子音区間では空気通路はないからである。

有声破裂子音の場合、声道の閉鎖点（口唇）の開放に伴う口唇放射が始まり、口腔内圧の交流成分が強くなる。この時の鼻孔放射は明らかに大きくなった。これは、破裂子音の場合軟口蓋が閉鎖しているので、鼻孔放射は軟口蓋の振動を通じて口腔内音圧の大きさに応じて変化するからである。鼻子音の場合、子音区間よりも母音区間の鼻孔放射がやや大きいがさほど変わらない。これは、鼻音化音声では、軟口蓋が開放し

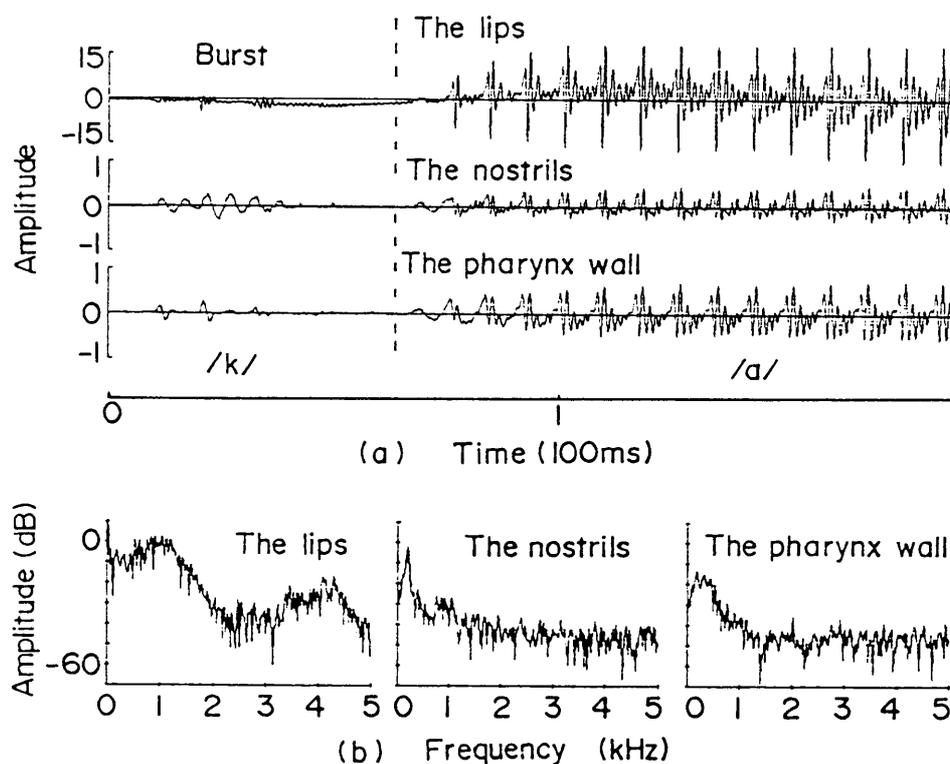


図 4. 8 無声破裂子音 /ka/ の放射音圧とスペクトル

- (a) 各部位の放射音 (Pa)
- (b) 破裂区間のスペクトル

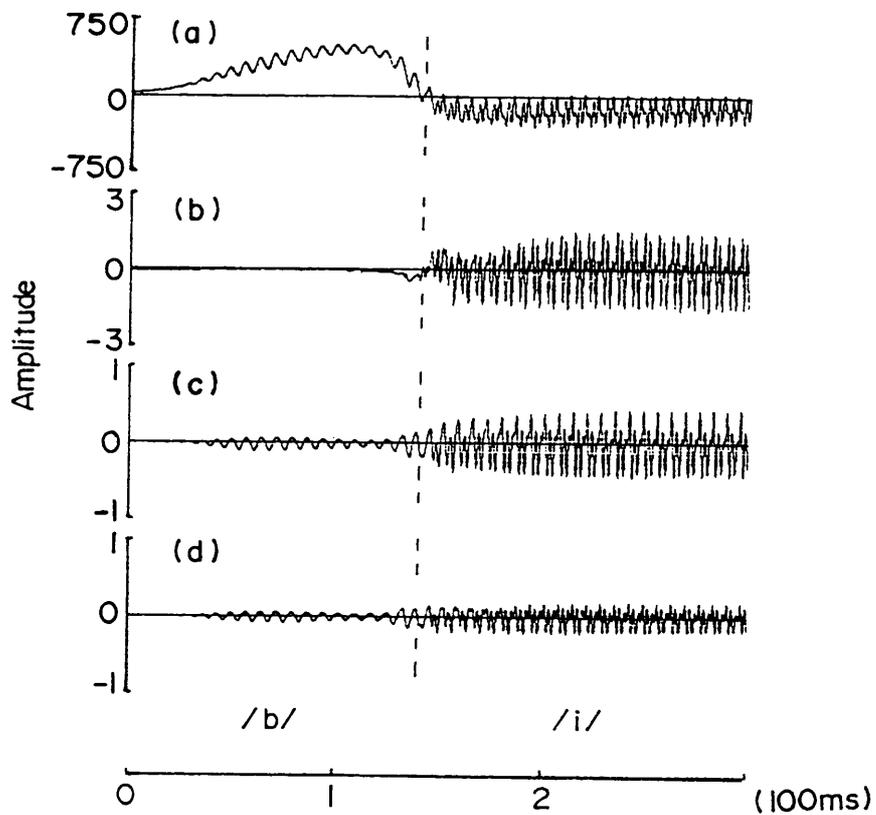


図4. 9 /bi/の発声時口腔内圧と各部位の放射 (単位: Pa)  
 (a) 口腔内圧 (b) 口唇からの放射  
 (c) 鼻孔からの放射 (d) 咽頭壁からの放射

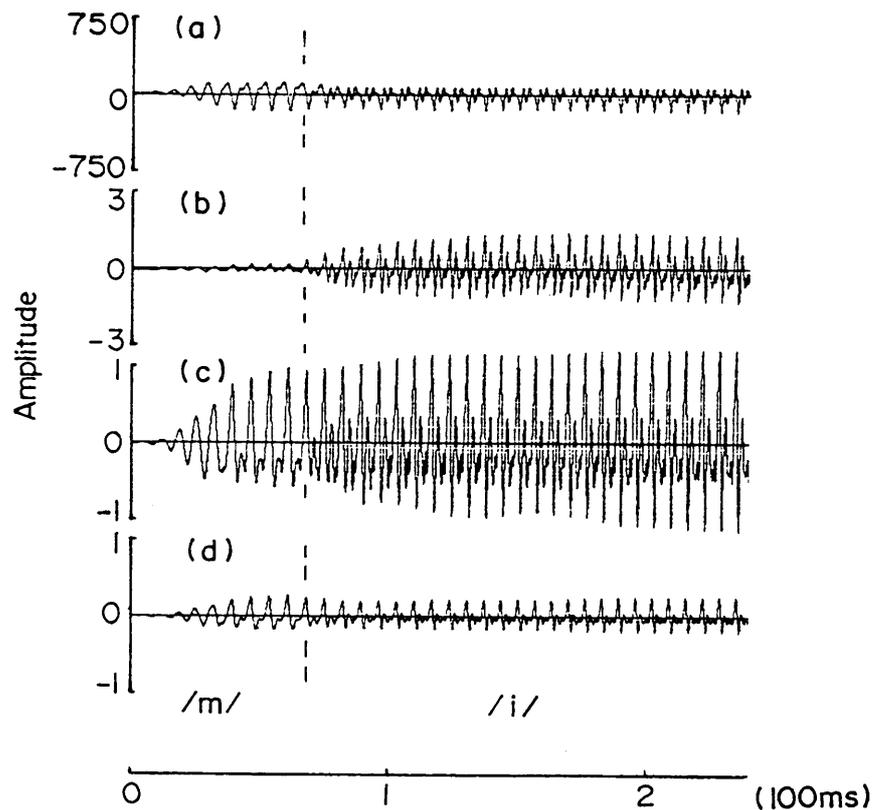


図4. 10 /mi/の発声時口腔内圧と各部位の放射 (単位: Pa)  
 (a) 口腔内圧 (b) 口唇からの放射  
 (c) 鼻孔からの放射 (d) 咽頭壁からの放射

ており、軟口蓋の開口面積の方が鼻孔放射への寄与が大きいので、口腔内圧がある程度変化しても、鼻孔放射はあまり変化しないからである。

#### 4.5 軟口蓋の動きの推定

有声破裂子音のバズ区間では軟口蓋部の口腔側から鼻孔までの伝達関数の可変要素は軟口蓋のパラメータだけなので、バズ区間の口腔内圧と鼻孔放射の関係から軟口蓋の状態を推定することができる。本節では、有声破裂子音のバズ区間における音圧の関係をj用いて軟口蓋の状態を推定する。

##### 4.5.1 軟口蓋の受動的な変化

有声子音バズバーは観測データを見る限り、いつも口腔内圧の交流成分に伴って出現している。咽頭壁放射は口腔内圧の交流成分に直接関連しているが、鼻孔放射は口腔内圧の交流成分以外の要素にも関連している。

有声破裂子音区間及び後続母音区間での口腔内圧と鼻孔放射の関係を調べたところ、被験者7名のデータにおいては鼻孔放射音圧の口腔内音圧に対する比と口腔内圧の直流成分の関係はほぼ同様である。バズ区間における鼻孔放射音圧と口腔内音圧の間のつながりを明らかにするため、口腔内圧の直流成分 $P_0$ が零である時の鼻孔放射音圧の口腔内音圧に対する音圧比を1として、その音圧比を正規化した。4人の測定データを○印で図4.11に示す。データの切出し区間はバズ区間の開始後約1/3経た点から後続母音の安定区間までの範囲にした。実測した音圧比は次のような実験式で近似できる。

$$q = a_1 + a_2 \exp(-P_0/a_3) \quad (4.4)$$

ここで、 $P_0$ は口腔内圧の直流成分、 $q$ は口腔内圧の直流成分で表示した鼻孔放射音圧の口腔内音圧に対する比、 $a_1$ 、 $a_2$ 、 $a_3$ は定数である。ただし、音圧と口腔内圧の直流成分とも単位はPaである。実測データにより求めた各定数を式(4.4)に代入して計算した値を実線で図4.11に示している。被験者4名に対して各定数、および計算した値と実測値の平均偏差を表4.8に示す。計算値と実測値の平均偏差は8%~18%となっている。有声子音バズ区間で軟口蓋が閉鎖している場合、口腔内音圧に対する鼻孔放射音圧の変化は軟口蓋の軟らかさの変化によって引き起こされたものである。よって、式(4.4)の口腔内圧の直流成分で表示した音圧比 $q$ はその直流成分による軟口蓋のパラメータの変化係数として取り扱うことができる。式(4.4)の各定数の

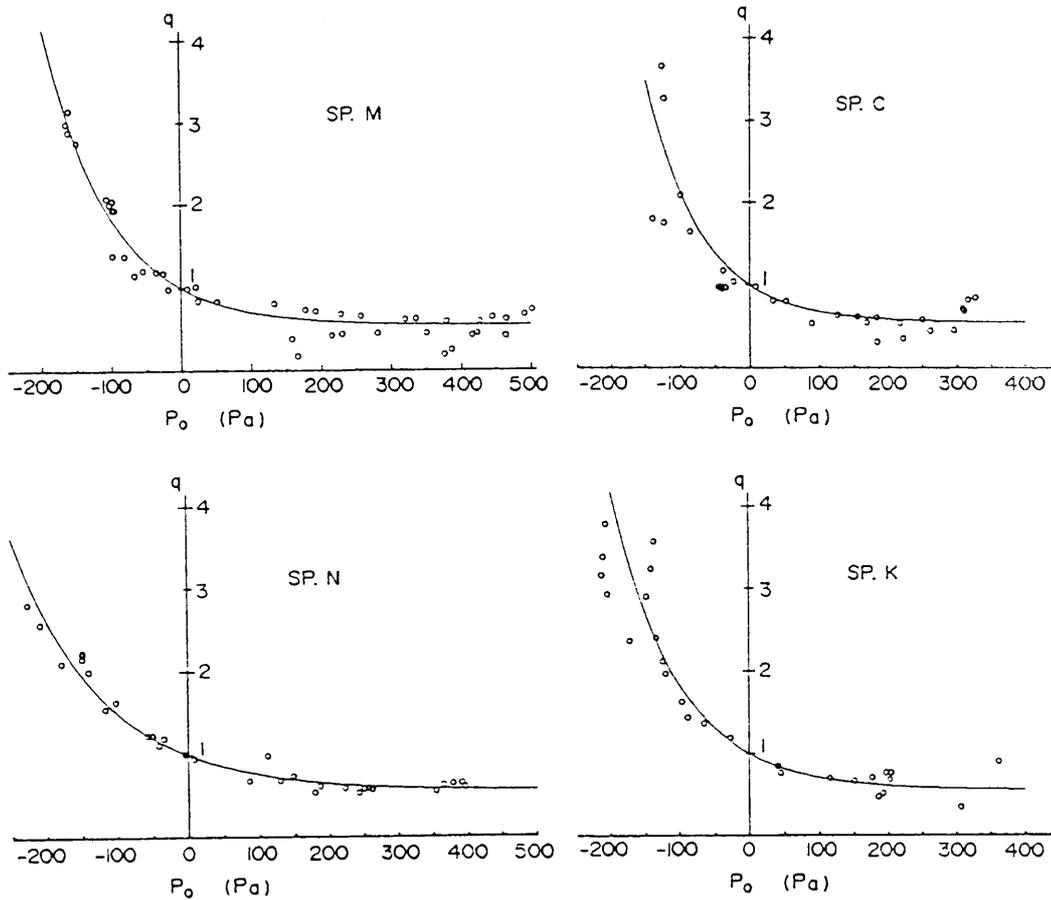


図4. 1 1 正規化した口腔内圧の直流成分( $P_o$ )による鼻孔放射音圧の口腔内音圧に対する比( $q$ )

○ : 測定値 実線 :  $q = a_1 + a_2 \exp(-P_o/a_3)$

表4. 8 式(4.4)の定数( $a_1, a_2, a_3$ )およびこれらの定数を用いて求めた値と実測値の相対誤差

	$a_1$	$a_2$	$a_3$	Relative error
Speaker M	0.55	0.45	95	17%
Speaker K	0.55	0.45	95	18%
Speaker N	0.55	0.45	130	8%
Speaker C	0.55	0.45	85	16%
Typical Val.	0.55	0.45	95	—

代表的な値としては、

$$a_1 = 0.55, a_2 = 0.45, a_3 = 95 \quad (4.5)$$

となる。式(4.5)の定数を式(4.4)に代入すると、次式となる。

$$q = 0.55 + 0.45 \exp(-P_0/95) \quad (4.6)$$

#### 4.5.2 軟口蓋の能動的な変化

有声子音の発声時における軟口蓋の能動的な変化を調べるため、話者Cの/bu/の/b/と話者Nの/zu/の/z/のデータを図4.12に示す。黒丸付きの線は口腔内圧の直流成分で、白丸付きの線は鼻孔放射音圧の口腔内音圧に対する比である。両方の関係では、口腔内圧の直流成分が高くなると、口腔内音圧に対して鼻孔放射音圧が

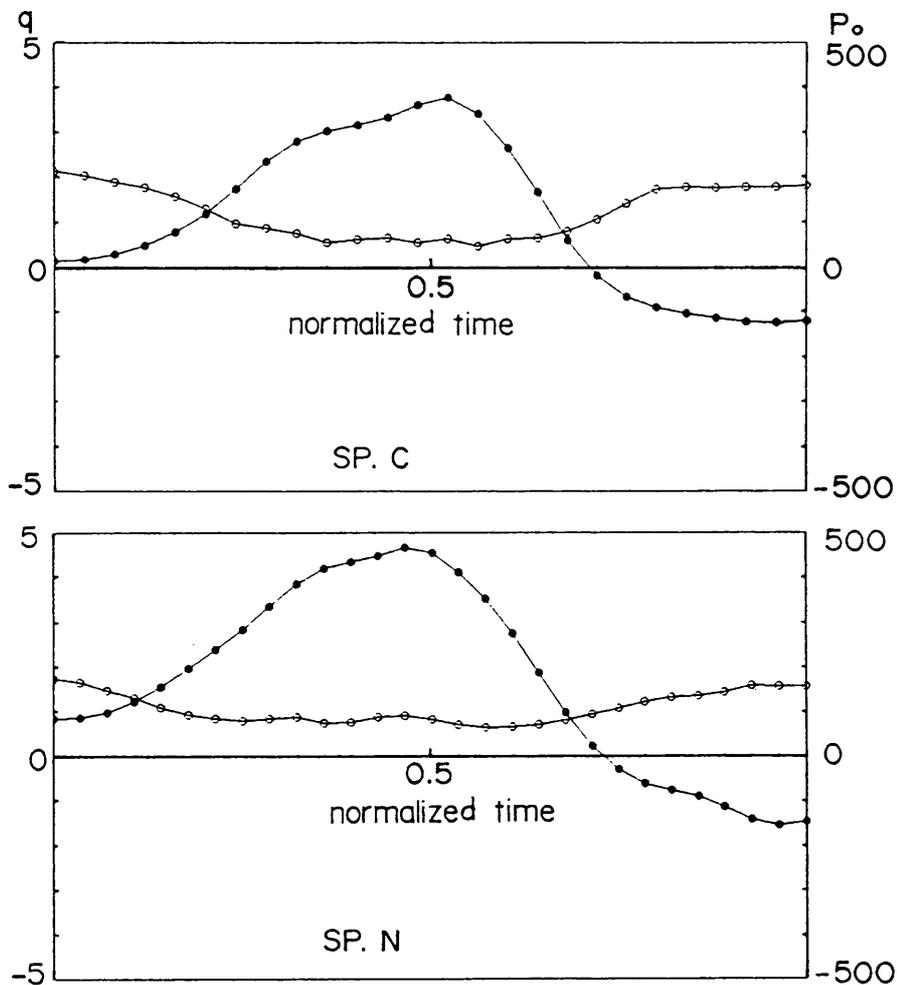


図4.12 有声子音区間での口腔内圧の直流成分と鼻孔放射音圧の口腔内音圧に対する比

- ：鼻孔放射と口腔内音圧の比(q)
- ：口腔内音圧の直流成分 (P<sub>0</sub>)

少なくなる。図の後方の約2/3の部分では、両方の関係は式(4.6)に従っている。つまり、口腔内圧の直流成分が零より大きい場合は鼻孔放射音圧と口腔内音圧の比が1より小さく、直流成分が零より小さい場合はその音圧比が1より大きい。それに対して、図の前方の約1/3の部分では鼻孔放射はかなり大きい。バズ区間の後方に比べて開始部分の鼻孔放射量は口腔内音圧に対して2倍程度になっている。このことは、バズ区間の開始部分では軟らかくて振動しやすい状態になっていることを示している。バズ区間の前半では、鼻孔放射量は他の要素によらず小さくなるのは軟口蓋の軟らかさが能動的に変化していることを示している。

鼻咽腔閉鎖強度の測定<sup>4)</sup>では、有声破裂子音の開始部分においては軟口蓋は咽頭壁に緩めに接して鼻腔を閉鎖しており、口腔内圧が高くなる同時に軟口蓋と咽頭壁との閉鎖は強くなることを示している。軟口蓋と咽頭壁が緩めに閉鎖する時は軟口蓋自身は軟らかく、強く閉鎖する（筋肉に力を入れている）時軟口蓋は硬くなる。バズバーの開始部分に鼻孔放射が大きいのは軟口蓋が咽頭壁に緩めに接してそれ自身が軟らかいからである。

#### 4. 5. 3 子音区間における軟口蓋の動きと音響放射

上述の考察により、非鼻音化子音区間における軟口蓋の動きと音響放射の関係は次のようになる。

(1) 声道閉鎖区間で口腔内圧の交流成分は発生すれば、軟口蓋と声道壁を振動させ、鼻孔と咽頭壁から音が放射される。

(2) 鼻孔放射音と咽頭壁放射音の時間的なずれは、バズバー区間で咽頭壁の筋肉の軟らかさが殆ど変化せず<sup>3)</sup>、軟口蓋の軟らかさが変化していることを示している。

(3) 非鼻音化音声区間の鼻孔放射は軟口蓋自身の状態に関連している。口腔内圧の直流成分が小さい時、つまり軟口蓋が軟らかい時、鼻孔放射は大きい。逆に直流成分が大きく軟口蓋が硬い時、鼻孔放射は小さい。軟口蓋の動きについて従来は能動説<sup>1) ~ 3)</sup>の方が多いが、今回の測定結果より、軟口蓋の状態はバズ区間の開始部分での時間的な変化は能動的な動き、口腔内圧の直流成分による変化は受動的な動きであることがわかった。

### 4. 6 有声破裂子音における音響放射のシミュレーション

#### 4. 6. 1 有声子音区間における軟口蓋の音響モデル

有声破裂子音の場合、一般的には軟口蓋が閉鎖しており口腔と鼻腔との間に空気の通路がないとされているが、実際に観測した結果では鼻孔からも音が放射されている。このように音が鼻孔から放射されるのは鼻腔と口腔が軟口蓋の振動による音響的な結合をしているからと考えて、第3章で非鼻音化母音における鼻孔放射をシミュレーションした。ここでは、3.6の軟口蓋の等価回路を用いて、バズ区間では口腔内圧の直流成分による軟口蓋の軟らかさの変化を考慮に入れる。この変化を軟口蓋のパラメータに対応させてみると、軟口蓋のステイフネスと機械抵抗は口腔内圧の直流成分の低い時には小さく、直流成分の高い時には大きいと考えられる。今回のシミュレーションでは軟口蓋のパラメータは次のようになると仮定した。

$$\begin{aligned}
 L &= 1.875 \times 10^{-2} \text{g/cm}^4 \\
 C &= q \times 9.47 \times 10^{-5} \text{cm}^4 \text{s}^2/\text{g} \\
 R &= \frac{1}{q} 6.25 \text{g/s/cm}^4 \quad (4.7) \\
 C_0 &= 1.99 \times 10^{-5} \text{cm}^4 \text{s}^2/\text{g} \\
 R_0 &= 100 \text{g/s/cm}^4
 \end{aligned}$$

となる。ここで、 $L$ 、 $R$ 、 $C$ は軟口蓋の音響質量、音響抵抗、音響コンプライアンスで、 $R_0$ 、 $C_0$ は二枚板に挟まれた粘弾性体の音響抵抗と音響コンプライアンスである（第3章参照）。 $q$ は口腔内圧の直流成分で表示した鼻孔放射音圧の口腔内音圧に対する比で、軟口蓋の軟らかさの受動的な変化を表わす。

有声破裂子音のバズ区間で声帯が振動を開始し、そしてその振動を維持することを説明するモデルがいくつか提案されている<sup>3), 7)</sup>。Flanaganら<sup>5)</sup>は声道閉鎖時に声帯振動を維持する条件は声道壁が非剛体であることを示した。Kentらの測定結果<sup>2), 3)</sup>では、有声破裂子音の場合は咽頭腔の容積が能動的に増大することを示した。また声道閉鎖期間で声帯が適当な緊張度を維持すれば咽頭腔の容積が増大する区間において声帯は振動できることを報告した<sup>3)</sup>。Rothenberg<sup>8)</sup>は咽頭腔の容積が $1 \text{cm}^3$ 増大すると声門波の持続時間は $10 \text{ms}$ 程度延びると報告している。これらの所説と上述の観測結果を考慮し、有声破裂子音のバズ区間で咽頭腔の容積が増大するモデルを作った。咽頭腔容積の増加分を次式に従って変化させる。

$$\Delta V(t) = V \cdot (1 - \cos(\pi t/T)) \times a \quad (4.8)$$

ここで、 $T$ はバズ区間の長さ、 $V$ は咽頭腔の容積で約 $55 \text{cm}^3$ 、 $t$ はバズ区間での時間変数、 $a$ は増大係数である。3.6の軟口蓋の振動板モデルと式(4.7)のパラメータを用いて、さらに、声道壁インピーダンスの単位面積あたりの機械抵抗 $B$ 、質量 $M$ 、ステイフネス

Kを次のように設定した。

$$\begin{aligned}
 B &= 3200 \text{ g/s/cm}^2 \\
 M &= 1.5 \text{ g/cm}^2 \\
 K &= 9.50 \times 10^4 \text{ dyne/cm}^3
 \end{aligned}
 \tag{4.9}$$

上述の音響モデルを用いて、有声破裂子音 /ba, bi, bu, be, bo/ の音響放射のシミュレーションを行った。

#### 4.6.2 シミュレーションの結果

/ba/ を例として、シミュレーションの結果を説明する。/ba/ を合成する時、式(4.8)の  $a=1/250$ 、 $t=0\sim 100\text{ms}$ 、 $T=100\text{ms}$  にした。この場合、咽頭腔容積の増大率は  $110 \text{ cm}^3/\text{s}$  となり、増大した容積は  $7.1 \text{ cm}^3$  程度となった。これは有声破裂子音 /b/ の実測値の範囲<sup>3)</sup>、容積の増大率  $40\sim 200 \text{ cm}^3/\text{s}$ 、増大した容積  $4\sim 9 \text{ cm}^3$  に入っている。

合成音 /ba/ の各部位の波形を図4.13に示す。バズバーの立ち上る区間では声

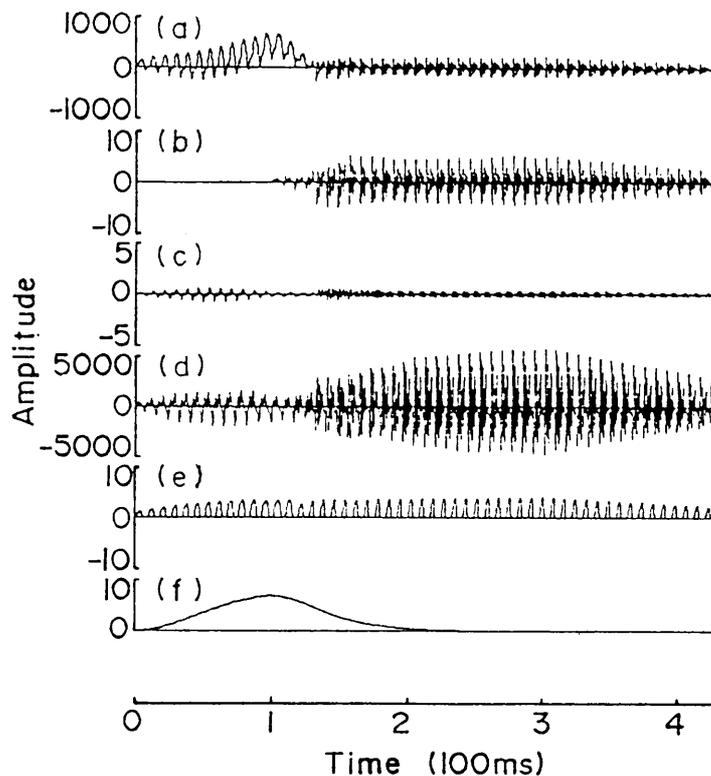


図4.13 合成音/ba/の各部の波形(音圧は相対値、  
 加速度の単位は $\text{cm}/\text{s}^2$ 、体積流の単位は $\text{cc}/\text{s}$ )  
 (a)口腔内圧 (b)(c)口唇と鼻孔からの音圧  
 (d)咽頭壁での加速度 (e)声門体積流  
 (f)咽頭腔体積( $\text{cm}^3$ )の変化

門下部と上部の間の気圧差が大きいので声帯は振動する。その後、咽頭腔の容積が式(4.8)に従って増大するので口腔内圧の直流成分はほとんど増えず声帯の振動が続く。軟口蓋が口腔内圧の交流成分につれて振動するので、鼻孔から音が放射される。式(4.8)により時間が経過すると咽頭腔容積の増大率が小さくなる。そのため、口腔内圧の直流成分が高まるので声帯振動と口腔内圧の交流成分は弱くなる。これに伴って、咽頭壁、軟口蓋の振動と鼻孔からの放射も弱くなっていく。ここで軟口蓋の軟らかさの受動的な変化を考慮に入れたのでバズ区間における鼻孔放射音と咽頭壁の加速度の極大値の時間的なずれは実音声の場合とほぼ同様である。

／ba／の実音声と合成音声のバズバーのスペクトル（ケプストラムで求めた包絡線）を図4.14に示す。鼻孔放射の特徴としては、3kHz以下に主なピークが4つある。バズバーの鼻孔放射の音響特性を矢印付きの4つのピーク（周波数の低い順でピーク1、ピーク2、ピーク3とピーク4と呼ぶ）で表せば、実音声と合成音声の各ピークの周波数と音圧を図4.15に示す。実音声の平均値と合成音声については、各ピークの周波数と音圧、およびそれらの差を表4.9に示す。合成音を実音声に比べて、各ピークの周波数の差は10%以下で、音圧差は5dB以下である。合成音と図4.14に示した代表的な実音声との相関係数が0.88となっているので、両方の特性はかなり近いことを示している。

一方、式(4.8)のパラメータを変化させてシミュレーションしたところ、容積増大率を $300\text{cm}^3/\text{s}$ 以上にすると声帯の振動が急に弱くなり、咽頭壁の振動と鼻孔の放射は

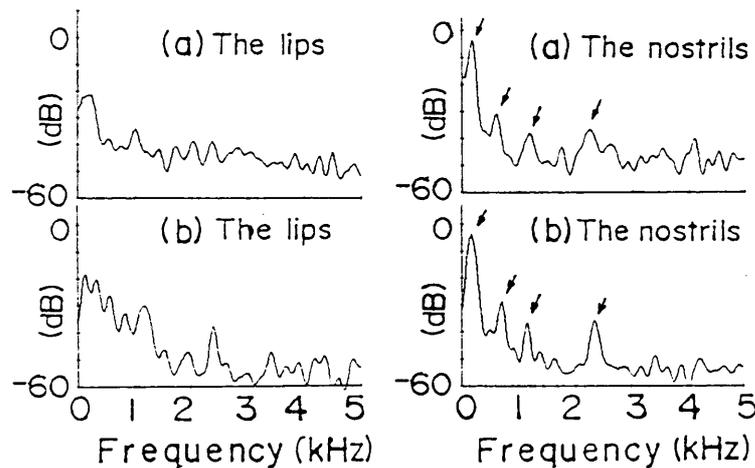


図4.14 /ba/のバズバーのスペクトル  
(a) 測定値 (b) 合成音声

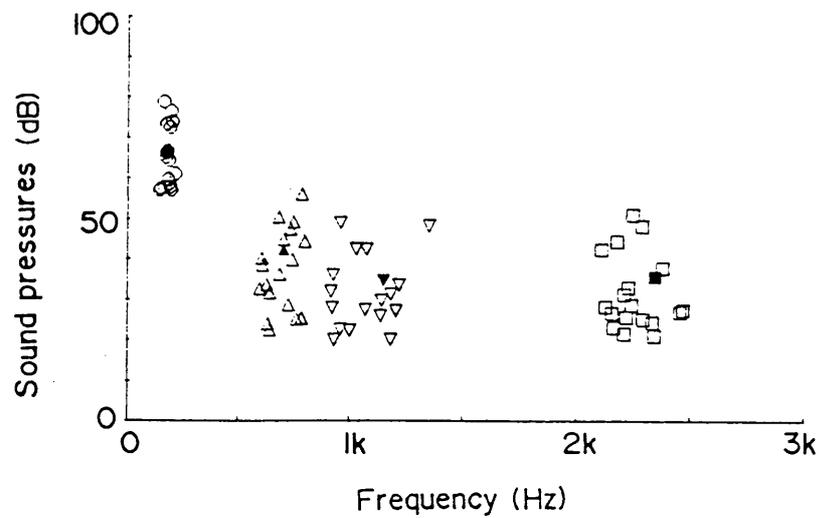


図4. 15 有声破裂子音/b/のバズバーの鼻孔放射の音響特性  
 ○△▽□：それぞれ実音声のピーク1、ピーク2、  
 ピーク3とピーク4  
 ●▲▼■：それぞれ合成音のピーク1、ピーク2、  
 ピーク3とピーク4

表4. 9 実音声と合成音の/b/の鼻孔放射のデータ  
 (括弧内の数字はピーク1との音圧差)

	Peak 1	Peak 2	Peak 3	Peak 4	Unit
Real speech	178.6	695.6	1069.1	2257.4	Hz
Synthesized speech	170.2	704.8	1152.5	2346.8	Hz
Relative error	4.9	1.3	7.2	3.8	%
Real speech	65.1	37.4 (27.7)	32.3 (32.8)	31.5 (33.6)	dB dB
Synthesized speech	66.3	42.4 (23.9)	34.6 (31.7)	35.6 (30.7)	dB dB

減少した。例えば、/ba/の場合、式(4.8)の $a=1/100$ にすると容積増大率が $350\text{cm}^3/\text{s}$ 、増大した容積が $14\text{cm}^3$ となる。合成した/ba/のバズパー部分を図4.16に示す。図4.16の30~50msの区間では、咽頭腔の容積の変化が大き過ぎるので声門波がなくなり咽頭壁の振動と鼻孔からの放射が減少している。このことより、有声破裂子音の声道閉鎖期間では声帯が適当な緊張度を維持する上で、咽頭腔容積の増大率は一定範囲の値を保つ必要があることがわかる。実音声ではバズパーの途中に波形が一瞬小さくなるかまたは消えるような現象がいくつかあったが、その区間で口腔内圧の直流成分がマイナスになる傾向は顕著ではない。今後バズ区間の発声機構を更に解明する必要がある。

#### 4.6.3 聴取実験の結果

提案した音響モデルで合成した有声破裂子音は、時間領域と周波数領域で実音声とかなり一致することが4.6.2節よりわかった。そこで、このモデルの妥当性をさらに合成音と実音声に対する聴取実験により調べる。実験では、有声破裂子音の実音声と合成音を用い、バズパーを除いた場合と除いていない場合に対して、有声子音、無声子音または判定不能の3つの選択で聴取実験を行った。被験者は正常聴力を持つ成人男性6名である。単音節のバズ区間では鼻孔放射音の方が大きいので、ここでは

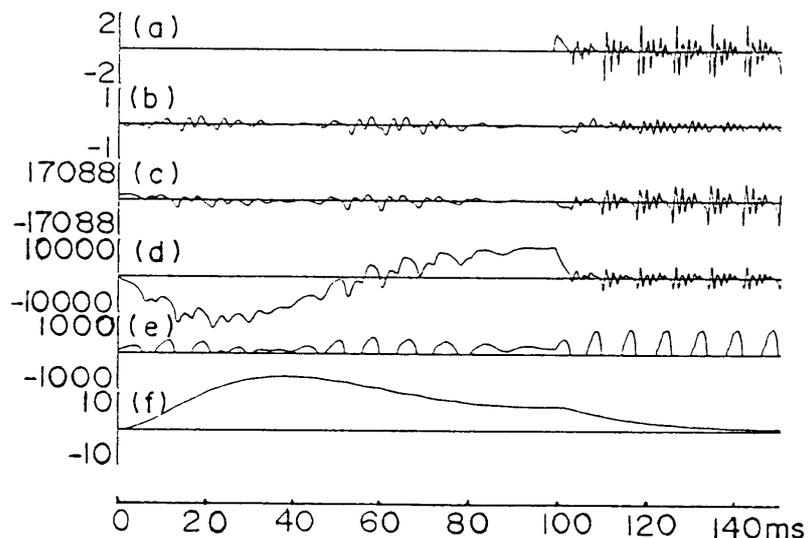


図4.16 合成音/ba/の各部の波形(音圧は相対値、  
 加速度の単位は $\text{cm}/\text{s}^2$ 、体積流の単位は $\text{cc}/\text{s}$ )  
 (a)(b)口唇と鼻孔からの音圧 (c)咽頭壁での加速度  
 (d)口腔内圧 (e)声門体積流 (f)咽頭腔体積( $\text{cm}^3$ )の変化

鼻孔放射音（鼻孔音）の有無をバズバーの有無とした。

実音声の場合、バ行の有声破裂子音の口唇音(/ba/,/bi/,/bu/,/be/,/bo/)、口唇音+鼻孔音(/ba<sup>+</sup>/,/bi<sup>+</sup>/,/bu<sup>+</sup>/,/be<sup>+</sup>/,/bo<sup>+</sup>/)、及びバ行の無声破裂子音(/pa/,/pi/,/pu/,/pe/,/po/)の音声を用いて、同一母音の音節を一つのグループにし、3種類の提示順序（たとえば/a/の場合は、/ba<sup>+</sup>/,/ba/,/pa/、/ba/,/ba<sup>+</sup>/,/pa/、/pa/,/ba<sup>+</sup>/,/ba/）で15組（3×5=15）の音声資料を用意した。音声資料と聴取結果を表4. 10に示す。被験者6名の聴取実験の結果では、バズバーのある（すなわち、口唇音+鼻孔音）場合はすべての有声子音が正確に認識されたが、バズバーを除いた（すなわち、口唇音のみ）場合は有声子音の31%は無声子音と間違えたかまたは無声子音と区別できなくなった。また、被験者の感覚によれば、口唇音のみの音声資料の多数は不自然さが感知された。

合成音の場合は、有声破裂子音の口唇音(/ba/,/bi/,/bu/,/be/,/bo/)と口唇音+鼻孔音(/ba<sup>+</sup>/,/bi<sup>+</sup>/,/bu<sup>+</sup>/,/be<sup>+</sup>/,/bo<sup>+</sup>/)を用い、同一母音の音節を一組（たとえば、母音/a/の場合は/ba<sup>+</sup>/,/ba/,/ba<sup>+</sup>/,/ba/の提示順序）にして、5組（5母音）の音声資料を用意した。音声資料と聴取結果を表4. 11に示す。同じ被験者6名の聴取実験の結果では、鼻孔音のある場合、すなわち、バズバーのある場合は有声子音の誤認

表4. 10 実測した破裂子音に関する聴取実験音声資料と実験結果

	Speech samples					Number of samples	Number of error
Group 1.	/ba <sup>+</sup>	/bi <sup>+</sup>	/bu <sup>+</sup>	/be <sup>+</sup>	/bo <sup>+</sup>	30	0
	/ba/	/bi/	/bu/	/be/	/bo/	30	9
	/pa/	/pi/	/pu/	/pe/	/po/	30	2
Group 2.	/ba/	/bi/	/bu/	/be/	/bo/	30	7
	/ba <sup>+</sup>	/bi <sup>+</sup>	/bu <sup>+</sup>	/be <sup>+</sup>	/bo <sup>+</sup>	30	0
	/pa/	/pi/	/pu/	/pe/	/po/	30	2
Group 3.	/pa/	/pi/	/pu/	/pe/	/po/	30	3
	/ba <sup>+</sup>	/bi <sup>+</sup>	/bu <sup>+</sup>	/be <sup>+</sup>	/bo <sup>+</sup>	30	0
	/ba/	/bi/	/bu/	/be/	/bo/	30	12

\*The samples with + are the sum of the radiation and the nostril radiation. The others are the lip radiation only.

識率が5%で実音声に近い。鼻孔音のない場合、すなわちバズバーを除いた場合は誤認識率が実音声の場合と同程度37%である。この結果より、本章で提案したモデルは有声破裂子音の音響的特徴をよく表わしていることがわかる。

#### 4.7 むすび

本章では、有声破裂子音などの放射特性を用いて口腔内圧と放射音圧、口腔内圧と軟口蓋の状態を考察した。放射音の結果では、有声子音だけでなく、無声子音の一部でも声道閉鎖区間に鼻孔と咽頭壁から音が放射されている。口腔内圧と放射音圧の関係より、軟口蓋の軟らかさの変化には能動的な成分と受動的な成分の双方があることを明らかにして、さらに、受動的な変化を定式化した。バズ区間での音響放射をシミュレーションするため、軟口蓋の軟らかさの受動的な変化を考慮に入れ、咽頭腔容積が能動的に増大する現象を導入して、有声破裂子音の音響モデルを作った。合成音に対する周波数特性の分析と聴取試験の結果より、このモデルは有声破裂子音の音響特徴をよく表わしていることがわかった。

表4. 11 合成した破裂子音に関する聴取実験音声資料と実験結果

Speech samples					Number of samples	Number of error
/ba/+	/bi/+	/bu/+	/be/+	/bo/+	30	2
/ba/	/bi/	/bu/	/be/	/bo/	30	12
/ba/+	/bi/+	/bu/+	/be/+	/bo/+	30	1
/ba/	/bi/	/bu/	/be/	/bo/	30	10

\*The samples with + are the sum of the radiation and the nostril radiation. The others are the lip radiation only.

## 第5章 鼻音化音声における音響放射の測定および 軟口蓋の等価開口面積率の推定

### 5.1 まえがき

発声時の軟口蓋の状態について、非鼻音化母音の場合、口腔内圧による振動を第3章で振動板モデルを作成した。有声破裂子音の場合、口腔内圧の変化による軟口蓋の軟らかさの変化を第4章で定式化した。本章では鼻音化音声における口唇放射と鼻孔放射を用いて、軟口蓋の開口面積に関するモデルを考察する。

軟口蓋の開口面積のモデルの作成には、鼻子音と鼻音化音声における音響放射特性を明らかにする必要がある。しかし、従来では鼻子音及びそれに隣接する母音区間での鼻孔放射と口唇放射の関係に関する研究が極めて少ない。本章では、まず、言語音声の文脈による放射音圧の変化を調べる。次に、分離された口唇放射と鼻孔放射の周波数特性を考察する。さらに、鼻音化音声区間における口唇放射と鼻孔放射を用いて、軟口蓋の開閉による等価インピーダンスの変化を考察し、放射音により軟口蓋の等価開口面積率の推定方法を提案する。最後に、鼻音化音声と非鼻音化音声を含む連続音声における軟口蓋の音響モデルを提案する。

### 5.2 測定装置と音声資料

#### 5.2.1 測定装置

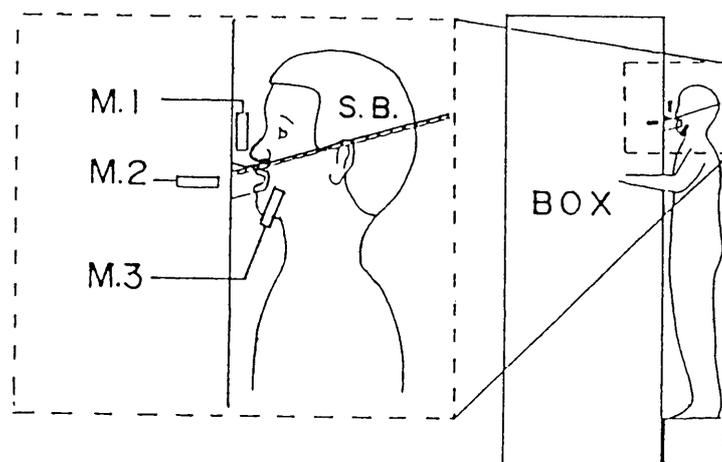
測定システムの略図を図5.1に示す。防音箱と遮蔽板を用いて、発声時の口唇放射、鼻孔放射および声道壁放射を分離して観測した。分離された放射音圧を、マイクロホン3個で同時に測定した。マイクロホンは口唇にはRion-UC26、頬にはB&K-4133であり、それぞれ精密騒音計Rion-NA60とB&K-2209に接続し、そのAC出力を取り出した。鼻孔には、もう1個のプローブマイクロホンB&K-4182で収録した信号をプリアンプで20dB増幅した。この3チャンネルの信号はデータレコータ（SONY社のDFR-3415）に収録し、遮断周波数4.8kHzのローパスフィルタを通しA-D変換して、ワークステーション（H.P.アポロDN3500）に取り込んだ。サンプリング周波数は10kHzである。

## 5. 2. 2 音声資料

鼻子音と鼻音化音声について文脈による影響を考察するため、鼻子音 $C_n$ に母音 $V$ と非鼻子音 $C$ を組合わせた音声資料を用いた。それらの組み合わせは $VC_n$ 、 $C_nV$ 、 $C_nVC_n$ 、 $C_nCV$ 、 $C_nVC$ と $CVC_n$ などである。本測定で用いた音声資料は、日本語100音節（表3.1参照）と、母音の後または前と後に鼻子音のある音節及び短文から構成されている。音声資料を表5.1を示す。これらの音声資料を、8名の成人男性に普通の速さで、特に短文では句切らないように発声させた。鼻音化音声では、頬の近くの放射音圧が鼻孔の放射音圧より10~20dB小さいので、本章では、鼻音化音声における口唇放射と鼻孔放射だけを考察する。

## 5. 3 言語音声の文脈による放射音の変化

言語音声の鼻孔放射と口唇放射を比較するため、収録した音声の始めから終わりまでを1として時間正規化した。分析フレームの幅は15msで、二分の一フレームずつ移動して放射音圧（ $P$ ）を次式で求めた。



M.:Microphone S.B.:Shield board

図5.1 測定システムの略図

M.1: 鼻孔部のマイクロホン

M.2: 口唇部のマイクロホン

M.3: 頬部のマイクロホン

$$P = 10 \log_{10} \left( \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - X)^2 \right) \quad (5.1)$$

ここで、Nは切り出し区間のサンプル数で、x(n)はn番目のサンプルで、Xは切り出し区間の平均値である。

### 5.3.1 単音節発声について

日本語の母音V、C<sub>n</sub>V型音節とVC<sub>n</sub>型音節における口唇と鼻孔からの放射音圧を図5.2に示す。

図5.2の左側の列で示した単母音Vの場合、/i/の鼻孔放射音圧が一番大きく、以下大きい順に/u/、/e/、/a/、/o/である。鼻孔放射は、口唇放射より、/i/では8dB程度小さく、他の母音では19~32dB小さくなっている。これは第3章の測定結果と一致している。

図5.2の中央の列で示したC<sub>n</sub>V型音節の場合、鼻孔放射は単母音より11~15dB強くなっている。鼻孔放射は口唇放射と比べて、/i/では2~5dB大きい、他の母音では5~15dB小さくなっている。母音区間における鼻孔放射量は、/i/では鼻音区間と同じ大きさを維持して推移するが、他の母音では口唇開口時に2~3dB下がって、その後徐々に下降している。

表5.1 実験に用いた鼻音と母音と非鼻音の組み合わせ及び音声資料

音素の組み合わせ (V: 母音, C <sub>n</sub> : 鼻音, C: 非鼻音)			
鼻音のない音節	鼻音がある前	鼻音がある後	鼻音がある両端
V	C <sub>n</sub> V	VC <sub>n</sub>	C <sub>n</sub> VC <sub>n</sub>
CV	C <sub>n</sub> VC C <sub>n</sub> CV	CVC <sub>n</sub> VCC <sub>n</sub>	C <sub>n</sub> VVVC <sub>n</sub>
音節資料			
日本語100音節, アン, イン, ウン, エン, オン, マン, ミン, ムン, メン, モン, ナン, ニン, ヌン, ネン, ノン			
単語及び短文資料			
雨, 飴, 案内, 先生, セクション, 難民, 全く, 宣伝, 全員, 全身, 水産物, 実験, ますます, 車を運転します, 窓を閉めます, 真上にあります, 飴が甘い, 雨が降っています, 名古屋に住んでいます			

図5. 2の右側の列で示したVC<sub>n</sub>型音節の場合、発声数の90%は音節の後半で鼻孔からの放射音が急激に増大する現象がある。それに関する音声波形の例を図5. 3に示す。図5. 3では、/i/から/N/へ、/e/から/N/へのわたりは1~2ピッチで完了する。全体に見るとその増大の78%は15~30msで完了する。鼻孔放射においては、増大する前は単母音より5dB程度大きく、増大した後はその前より8~15dB大きくなっている。増大する現象の61%は口唇の閉まる時期に現れ、その他は口唇の閉まる前に現れる。

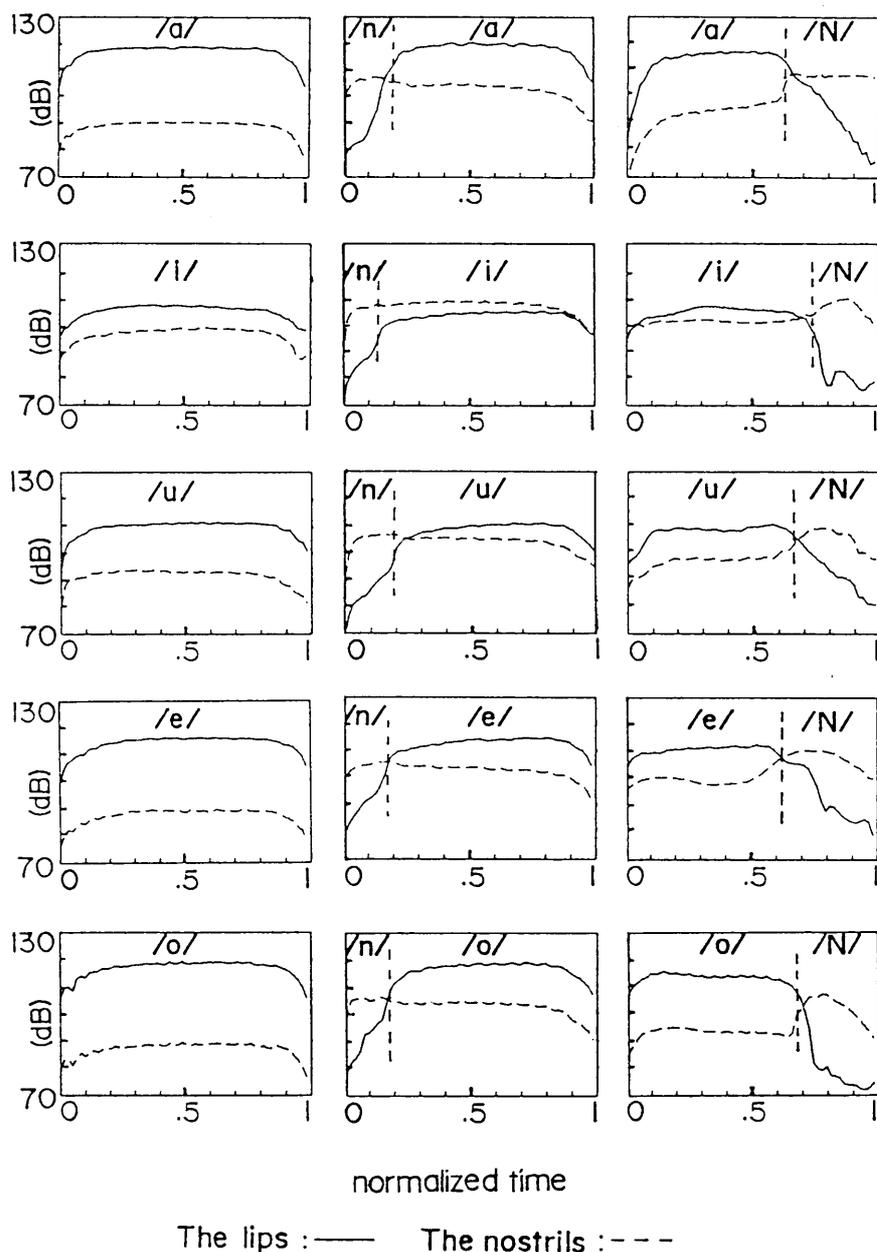


図5. 2 単音節発声時の口唇放射量と鼻孔放射量

$C_nVC_n$ の場合、母音Vの区間での鼻孔放射量は、 $C_nV$ と $VC_n$ の場合より3~10dB大きくなっている。母音Vが/i/である発声数の100%、/u/の発声数の88%は図5.4に示したように、鼻孔放射量が両端の鼻子音のそれとほぼ同じ大きさを保持しており、口唇放射量より5~10dB大きくなっている。/e/の25%、/a/の12%は鼻孔放射量は口唇のそれより大きい、/e,a,o/における鼻孔放射は両端の鼻子音のそれより2~10dB小さくなっている。

鼻音化された子音/g/（すなわち、/ŋ/）における鼻孔放射は口唇放射より5~10dB強くなっているが、口唇からほとんど放射しない場合もある。鼻孔放射量は鼻子音/m,n/と同じ大きさである。

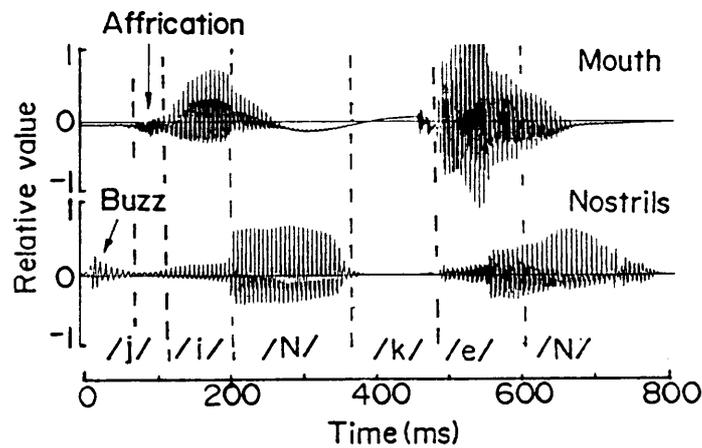


図5.3 単語（「人權」）の口唇放射と鼻孔放射の波形

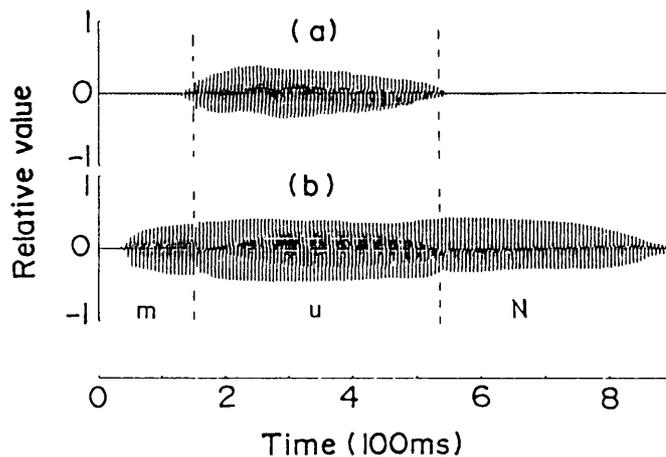


図5.4 完全鼻音化音節/muN/の口唇放射と鼻孔放射の波形  
(a) 口唇放射 (b) 鼻孔放射

### 5.3.2 連続音声について

連続音声における口唇放射と鼻孔放射の例を図5.5と図5.6に示す。

$C_nV$ の場合、図5.5の/mā/と/ni/のように鼻孔放射量は単音節に比べて、2~4dB小さく、下がる傾きが大きくなっている。

$C_nVC$ の場合、発声数の33%は両図の/mās/のように母音Vの途中で鼻子音の影響がなくなってしまう。

$C_nVV$ の場合、図5.5の/nia/のように鼻子音の影響はその隣接母音Vの終わりにま

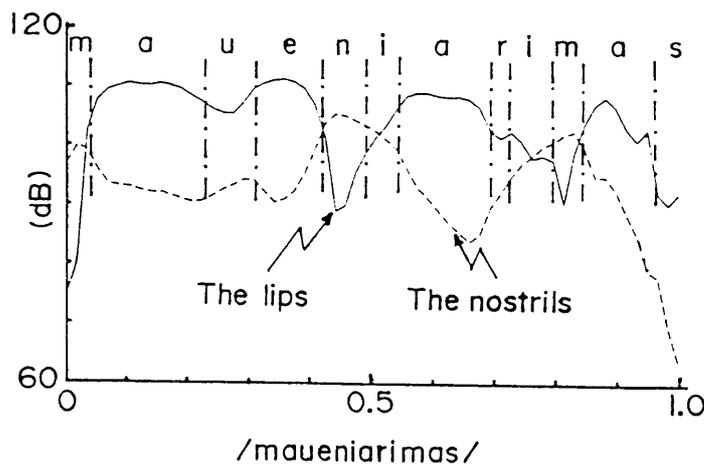


図5.5 短文（「真上にあります」）発声の口唇放射と鼻孔放射

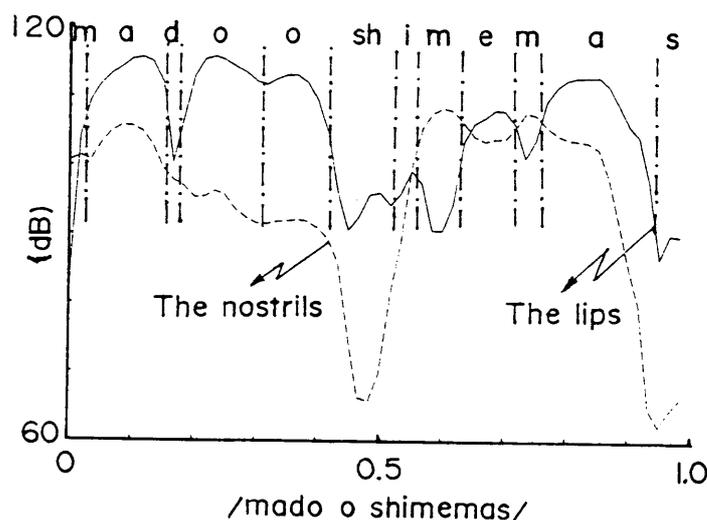


図5.6 短文（「窓をしめます」）発声の口唇放射と鼻孔放射

で維持しており、発声数の半分程度がその影響を次の母音に及ぼす。この例では、 $/a/$ の鼻孔放射が単母音より10dB程度大きくなっている。

$C_nVVVC_n$ の場合、発声数の55%は $/mauen/$ のように鼻子音の影響を中心の母音に及ぼす。この場合、中心母音 $/u/$ は単母音より鼻孔放射が5dB程度大きくなっている。

$VC_n$ の場合、発声数の81%は、図5.5の $/eni/$ と図5.6の $/ime/$ のように母音区間で鼻孔放射量の急激に増大する現象が見られる。その現象の出現頻度は単音節に比べてやや低いが、増大量と完了時間は単音節とほぼ同じである。

$CVC_n$ の音声資料について、 $C$ が有声破裂子音である場合、発声数の67%は図5.3に示したようにバズ区間での鼻孔放射量が後続母音区間の前半での鼻孔放射量より強くなっている。また、図5.6の $/mado/$ のような $C_nV_1CV_2$ 型( $C$ は有声破裂子音)では鼻子音が有声破裂子音の後続母音 $V_2$ に及ぼす影響はほとんど見られないが、鼻子音と有声破裂子音に挟まれた母音 $V_1$ の全区間に鼻孔放射は大きくなっている。これは、後続有声破裂子音の発声のため軟口蓋は軟らかくなって振動量が増えるので、軟口蓋の開口面積が減っても鼻孔放射量があまり減らない為と考えられる。

非鼻子音と鼻子音との組み合わせ、 $C_nCV$ の場合は鼻子音の影響が非鼻子音の種類によって多少異なる。すなわち、非鼻子音が無声子音のとき、鼻子音が非鼻子音を越えて母音区間での鼻孔放射量に影響を与えている例は発声数32の中に1例も見られない。非鼻子音が有声破裂子音である場合、鼻子音が非鼻子音の後続母音に対して明らかに影響しているのは、発声数28の中2例しか見られない。

#### 5.4 口唇放射と鼻孔放射の周波数特性

これまで鼻子音と鼻音化音声の周波数特性に関する研究がいくつかなされてきたが、それらの研究のほとんどは鼻孔放射と口唇放射とが混合した音に対して考察したものである。本節では、分離された口唇放射と鼻孔放射に対してそれらの周波数特性を考察する。

##### 5.4.1 鼻孔放射音と口唇放射音の比較

鼻孔放射音が口唇放射に比べて、どの程度の割合を占めているかを調べるため、同一母音に対して非鼻音化音声と鼻音化度の異なる鼻音化音声における口唇放射と鼻孔放射を考察する。

日本語母音では、 $/i/$ の鼻孔放射が一番強いので例として母音 $/i/$ と $/iN/$ 及び

／mi／の／i／の口唇放射音と鼻孔放射音及び両者の和の周波数スペクトル（ケプストラムで求めた包絡線、以下同）を図5. 7に示す。それらの鼻音化度は小さい順で／i／、／iN／と／mi／となる（便宜上、論ずる音素に下線を付ける）。図5. 7の左列は非鼻音化母音／i／である。この場合、鼻孔放射と口唇放射の和の第一ホルマンの振幅が口唇だけのそれより2dB程度大きくなっている。これは、鼻孔放射が音声知覚にある程度の影響を与える可能性を示している。このことより、鼻孔放射の強い母音（例えば狭母音の／i／と／u／）に対して、非鼻音化母音でも精度の高い音声分析と音声合成には鼻孔放射音は必ずしも無視できない。

図5. 7の中央列の／iN／の場合、口唇放射と鼻孔放射の和の第一ホルマンは、口唇だけの放射音のそれに比較すると振幅が3dB程度大きくなっている。図5. 7の右列の／mi／の場合、鼻孔放射の1kHz以下のピークは口唇放射のそれより大きくなっている。1.9kHzのピークは口唇放射と比べて、和の方が4dB大きくなっている。測定結果より、鼻孔放射の影響は主に2kHz以下に現れ、鼻音化度が大きくなると影響の及

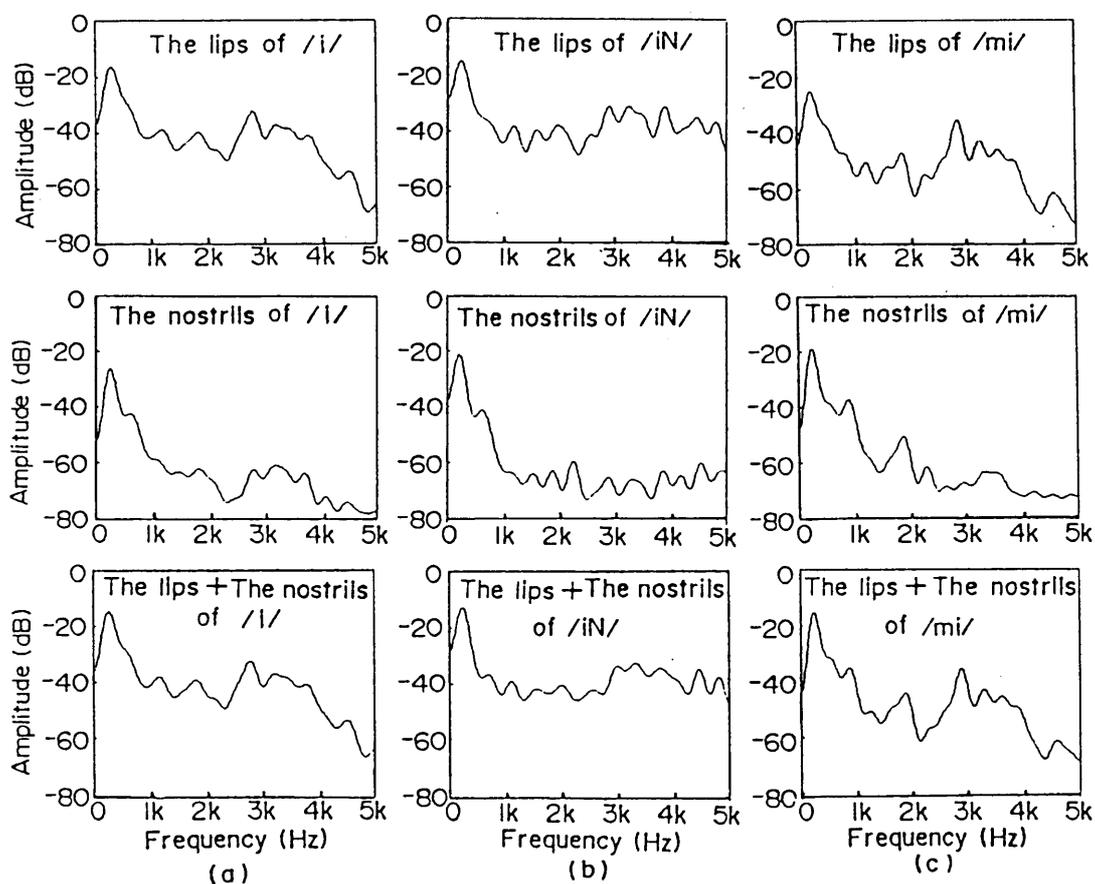


図5. 7 単母音/i/と/iN/, /mi/での/i/の口唇放射、鼻孔放射及び両放射の和のスペクトル

ばす周波数範囲は広くなることがわかる。

#### 5. 4. 2 口唇放射と鼻孔放射の極零

鼻孔放射と口唇放射における極零を調べるため、すべての音節が鼻音化されている /naNmiN/ の /miN/ に対して、防音箱を用いて分離された口唇放射と鼻孔放射、そして防音箱を用いず一個のマイクロホンで収録した音声について考察する。それらの三次元スペクトルを図5. 8に示す。図5. 8(a)は一個のマイクロホンで収録した音声で、図5. 8(b)と(c)はそれぞれ口唇放射と鼻孔放射である。母音 /i/ の区間は図5. 8の約90ms~300msである。図5. 8(a)の /i/ の区間では、第一ホルマントと第二ホルマントの間に二つの共振ピークが800Hzと1300Hzに現れている。口唇放射と鼻孔放射が分離された場合、図5. 8(b)のように、/i/ の口唇放射では1300Hz付近に共振ピークが現れたが、800Hz付近に共振ピークは現れていない。それに対して、図5. 8(c)の鼻孔放射では800Hz付近の共振ピークがはっきり現れ、そのレベルは口唇放射の対応周波数成分より4dB程度大きくなっている。/i/ の第一ホルマント付近の振幅においては、鼻孔放射が口唇放射より8dB程度大きくなっている。/i/ の第二ホルマントでは鼻孔放射が口唇放射より3dB程度小さくなっている。/i/ の第三ホルマントは鼻孔放射には現れていない。

従来の音声生成モデルでは、鼻子音以外の音が口唇だけあるいは主に口唇から放射されるとされている。その考えによれば、鼻音化母音の口唇放射と鼻孔放射では極の周波数は相等しく、零の周波数だけが異なる<sup>1)</sup>。すなわち、鼻音化母音では、鼻腔の寄与は零点だけとなる。上述の観測結果では、この考えは強く鼻音化された母音では必ずしも正しくないことを示した。実際には、強く鼻音化された母音区間では鼻孔放射が口唇放射より大きいケースが多い(5. 3. 1と図5. 4参照)。この場合、声道系には口腔と鼻腔の双方は別々の極と零に寄与する。

一方、Fant<sup>7)</sup>は軟口蓋の開口面積を変化させて口唇放射と鼻孔放射を合成音で考察した。文献[7]では、軟口蓋の開口面積を2.6cm<sup>2</sup>に設定して合成した鼻音化母音 /i/ は、観測された実音声 /miN/ の /i/ と同じく、鼻音化の度合いが大きい方に属する。この場合、合成した /i/ の口唇放射と鼻孔放射のホルマントの振幅は本研究の実測値とよく一致している。すなわち、/i/ の第一ホルマントでは鼻孔放射が口唇放射より8dB程度大きく、第二ホルマントでは鼻孔放射が口唇放射より3dB程度小さくなっている。

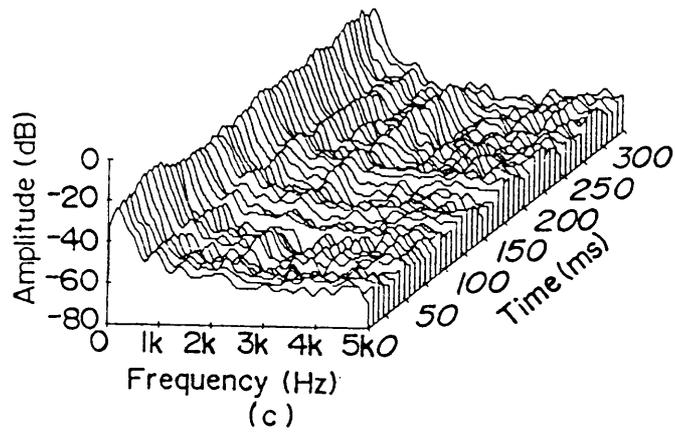
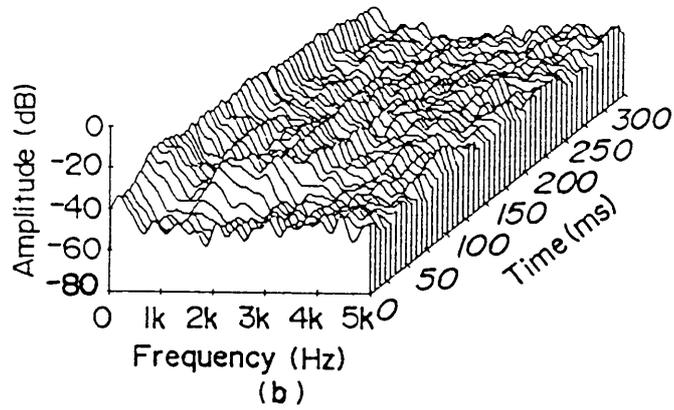
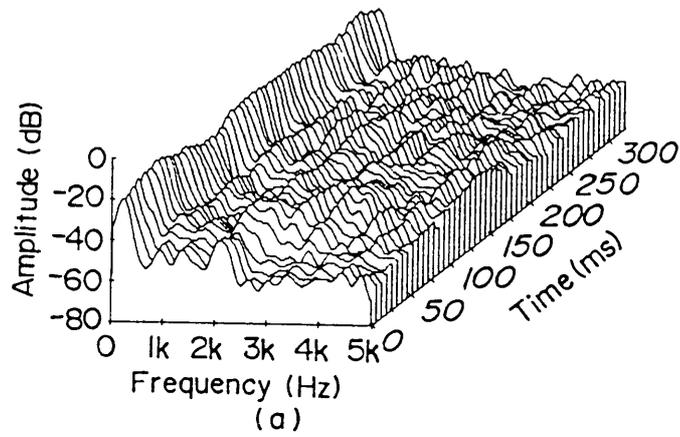


図5.8 /naNmiN/の/mi/の一部の三次元スペクトル表示  
 (a) 分離していない場合の音声  
 (b) 口唇の放射音  
 (c) 鼻孔の放射音

### 5. 4. 3 ホルマントに対する鼻孔放射の影響

従来では、自然音声（鼻孔放射と口唇放射とが混合したもの）を用いて鼻音化度が強くなると母音のホルマントが変化することを考察した<sup>2), 3)</sup>が、ここでは、口唇放射と鼻孔放射が分離された状態で鼻音化度のホルマントに対する影響を考察する。

まず、鼻音化音声における口腔と鼻腔のホルマントの連続変化を考察する。例として/eN/の口唇放射と鼻孔放射の三次元スペクトルを図5. 9に示す。図5. 9(a)と(b)はそれぞれ口唇放射と鼻孔放射である。パワーの小さいスペクトルを観測するため、図5. 9(a)の時間軸を逆にして図5. 9(c)に示す。鼻子音の影響のため、図5. 9(b)に示したように、約150ms~250msの区間で鼻孔放射は大きくなった。同じ区間では、図5. 9(c)に示したように、口唇放射の第一、第二ホルマントの振幅は同程度に減少している。口唇放射の減少に伴い、鼻孔放射での母音第一ホルマントにあたる周波数の成分は強くなり、230msのあたりで口唇放射より4dB程度大きくなっている。鼻音化の度合の小さい母音区間では、鼻孔放射には一番強いピークは鼻腔のホルマントでなく、母音の第一ホルマントの近傍にある。鼻音化の度合が大きくなると共に、鼻腔のホルマントは徐々に主要な成分になってきた。230ms以後の区間で250Hzの鼻腔のホルマントが一番強い成分となった。

次に、鼻音化度の異なる言語文脈における同一母音のホルマントにどんな変化が起こるかを観測する。例として、/aN/と/ma/の/a/のスペクトルを図5. 10に示す。/ma/の鼻孔放射は/aN/より大きくなっている。/ma/の鼻音化度が大きい。/ma/の第一と第二ホルマントの振幅差は、口唇放射では/aN/のそれと同じである。口唇放射と鼻孔放射の和のスペクトルでは、/aN/の場合第一と第二ホルマントの振幅差は口唇放射のそれに比べ変わらないが、/ma/の場合その差は口唇のそれより3dB程度大きくなっている。/ma/の口唇放射と鼻孔放射の和のスペクトルを口唇だけの場合と比較すると、第二ホルマントの大きさは変わらず、和の第一ホルマントの振幅は3dB程度増大したことがわかる。この振幅差の3dBの分は、鼻孔放射量の増大により口唇放射と鼻孔放射の和の第一ホルマントが増大したものである。

上述の考察より、母音の鼻音化度が強くなると共に、口唇放射では第一ホルマントの振幅は第二ホルマントと同程度に減少しているが、鼻孔放射における母音の第一ホルマントにあたる周波数成分が強くなるため、口唇放射と鼻孔放射の和では母音の第一ホルマントの振幅は第二ホルマントに比べて大きくなっている。弱く鼻音化される母音では、鼻孔放射には一番強いピークは母音の第一ホルマントの近傍にある。鼻音化の度合が大きくなると共に鼻腔のホルマントは徐々に主要な成分になってきた。

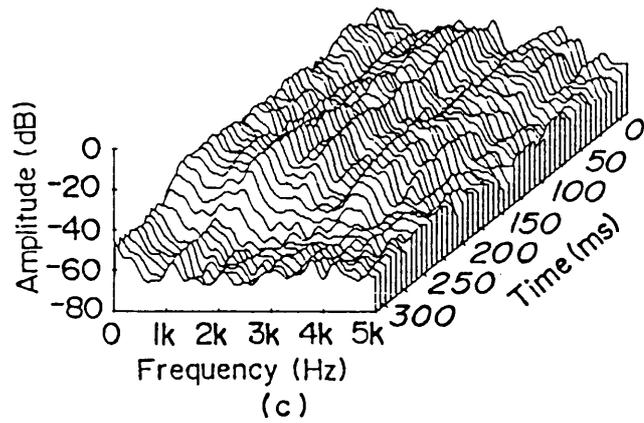
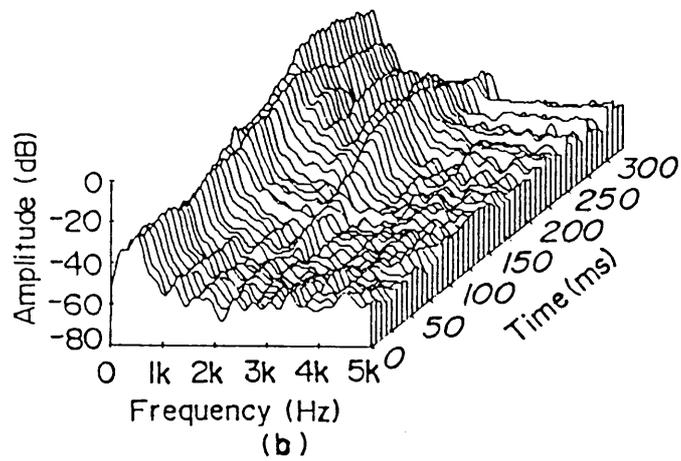
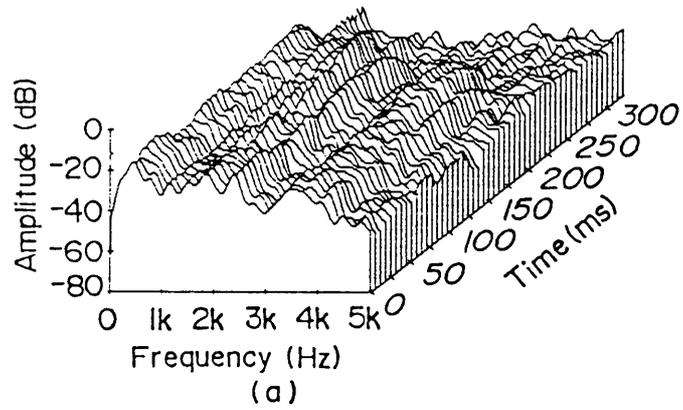


図 5. 9 /eN/の発声の三次元スペクトル表示

- (a) 口唇の放射音
- (b) 鼻孔の放射音
- (c) 口唇の放射音 (時間軸が逆になっている)

## 5.5 軟口蓋の等価インピーダンス

鼻音化音声では、鼻腔と口腔とが軟口蓋の開口面積および軟口蓋の振動によって結合している。声道の音響等価回路の略図を図5.11に示す。ここで、軟口蓋部のインピーダンスは軟口蓋振動の等価インピーダンスと軟口蓋開口面積の等価インピーダンスの並列とした。前者は、主に非鼻音化音声の場合であり、第3と第4章で検討した。鼻音化音声の場合では、軟口蓋はある程度開くので軟口蓋の振動と開口面積の等価インピーダンスは同時に働いている。本節では、軟口蓋開口面積の等価インピーダンス、及びそれと軟口蓋振動の等価インピーダンスの関係を検討する。

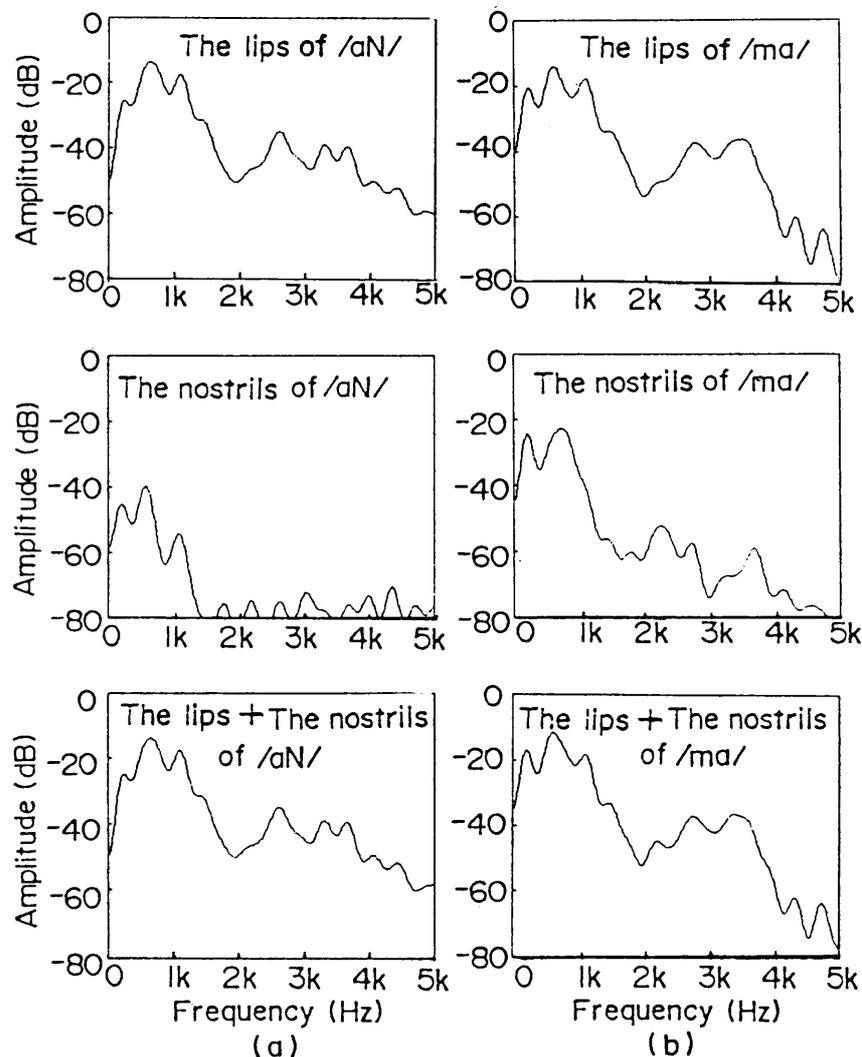


図5.10 /aN/と/ma/での/a/の口唇放射、鼻孔放射および両放射の和のスペクトル

### 5.5.1 軟口蓋の等価インピーダンスと放射音圧

声道モデルの等価回路は一般に線形回路として扱われているので、口唇放射  $p_m$  と軟口蓋部の音圧  $p_v$ 、鼻孔放射  $p_n$  と鼻腔側軟口蓋部の音圧  $p_{vn}$  の関係を次式

$$p_m = a p_v \quad (5.2)$$

$$p_n = b p_{vn} \quad (5.3)$$

で表わせる。ただし、 $a$  と  $b$  は比例係数である。軟口蓋部の口腔側の音圧  $p_v$  と鼻腔側の音圧  $p_{vn}$  の関係は

$$p_{vn} = \frac{Z_n}{Z_n + Z_v} p_v \quad (5.4)$$

となる。ここで、 $Z_v$  は軟口蓋振動の等価インピーダンス  $Z_{vv}$  と軟口蓋開口面積の等価インピーダンス  $Z_{vo}(t)$  の並列で、 $Z_n$  は軟口蓋部の鼻腔側から見た鼻腔のインピーダンスである。

第3章で導入した鼻腔の断面積関数と第4章の軟口蓋振動の等価インピーダンスを用い、軟口蓋の開口面積を  $0.5 \sim 2.5 \text{ cm}^2$  変動させて、鼻子音 /m/ を合成した。次式を用いて合成した音声データから  $Z_v$  と  $Z_n$  を求める。

$$\left. \begin{aligned} Z_n &= P_{vn} / U_n \\ Z_v &= P_v / U_n - Z_n \end{aligned} \right\} (5.5)$$

ここで、 $P_{vn}$ 、 $P_v$  と  $U_n$  はそれぞれ  $p_{vn}$ 、 $p_v$  と  $u_n$  の実効値である。軟口蓋の開口面積を

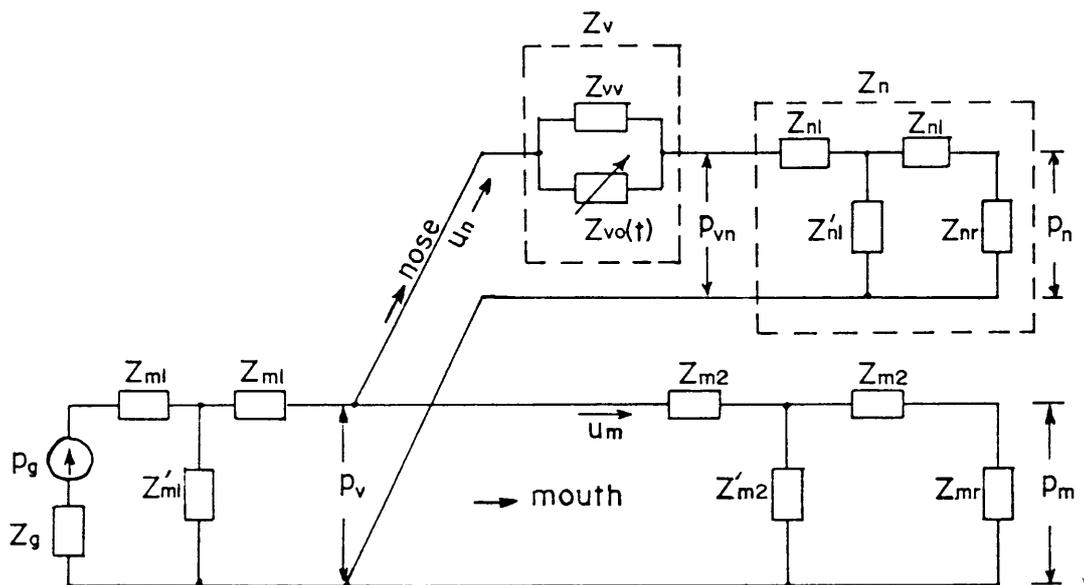


図5.11 声道の音響等価回路の略図

2.0cm<sup>2</sup>にした場合、求めたZ<sub>v</sub>とZ<sub>n</sub>は

$$\left. \begin{aligned} Z_v &= 26.3 \text{ g/s.cm}^4 \\ Z_n &= 2.6 \text{ g/s.cm}^4 \end{aligned} \right\} (5.6)$$

となる。シミュレーションの結果より、軟口蓋の開口面積を2.0cm<sup>2</sup>以下にした場合、次式

$$Z_v \gg Z_n \quad (5.7)$$

が成立する。

式(5.7)を利用し、式(5.4)を(5.3)に代入すると

$$p_n = \frac{Z_n}{Z_v} b p_v \quad (5.8)$$

となる。式(5.2)と(5.8)から

$$\frac{p_m}{p_n} = \frac{a Z_v}{b Z_n} \quad (5.9)$$

となる。

非鼻音化音声では、軟口蓋が閉鎖しており、軟口蓋開口面積の等価インピーダンスZ<sub>v0</sub>(t)は非常に大きく、Z<sub>v</sub>=Z<sub>vv</sub>で近似できるので、式(5.9)は

$$\frac{p_m(v)}{p_n(v)} = \frac{a Z_{vv}}{b Z_n} \quad (5.10)$$

となる。ここで、p<sub>m</sub>(v)とp<sub>n</sub>(v)は母音Vの関数である(表5.2参照)。鼻音化音声の場合、Z<sub>v</sub>はZ<sub>vv</sub>とZ<sub>v0</sub>(t)の並列となっているので

$$\frac{p_{mn}(t)}{p_{nn}(t)} = \frac{a}{b Z_n} \frac{Z_{vv} Z_{v0}(t)}{Z_{vv} + Z_{v0}(t)} \quad (5.11)$$

となる。但し、p<sub>mn</sub>(t)とp<sub>nn</sub>(t)は鼻音化音声の口唇放射と鼻孔放射で、時間の関数である。音声発声時には鼻腔の形が変わらないので、式(5.10)と(5.11)の係数bは同

表5.2 非鼻音化母音の口唇放射と鼻孔放射及びそれらの差(dB)

	/a/	/i/	/u/	/e/	/o/
口唇	105.1	92.7	96.7	100.4	104.2
鼻孔	81.0	85.0	82.8	79.9	80.8
差	24.1	7.7	13.9	20.5	23.4

じ定数である。鼻音化音声と非鼻音化音声では同じ母音を発声する時同様な声道の形をしていると仮定すると、式(5.10)と(5.11)での係数  $a$  は同一母音に対して同じ定数となる。ここで、式(5.10)の両辺を式(5.11)の両辺でそれぞれ割り算して整理すると

$$\frac{p_m(v)/p_n(v)}{p_{mn}(t)/p_{nn}(t)} = \frac{Z_{vv}+Z_{vo}(t)}{Z_{vo}(t)} \quad (5.12)$$

となる。式(5.12)より、非鼻音化母音の場合の口唇放射と鼻孔放射の音圧比と、鼻音化音声の場合の音圧比を用いて、軟口蓋開口面積の等価インピーダンス  $Z_{vo}(t)$  と軟口蓋振動の等価インピーダンス  $Z_{vv}$  との関係を表わすことができる。

### 5. 5. 2 軟口蓋開口面積の等価インピーダンスと軟口蓋振動の等価インピーダンス

鼻子音区間での鼻孔放射を  $p_{nN}$ 、開口面積が最大となった時の軟口蓋開口面積の等価インピーダンスを  $Z_{vo}'$  とすると、式(5.8) より

$$p_{nN} = \frac{bZ_n}{Z_v} p_v = \frac{bZ_n(Z_{vo}' + Z_{vv})}{Z_{vo}'Z_{vv}} p_v \quad (5.13)$$

である。有声破裂子音のバズ区間の鼻孔放射を  $p_{nE}$  とすれば

$$p_{nE} = \frac{bZ_n}{Z_{vv}} p_v \quad (5.14)$$

となる。鼻音化音声の場合、軟口蓋部の開口面積が増大すると共に、軟口蓋振動の等価インピーダンス  $Z_{vv}$  は若干大きくなる。しかし、振動板としての軟口蓋の変化様子がまだ明らかにされていないので、ここでは、軟口蓋の閉鎖した場合と開放する場合の軟口蓋振動の等価インピーダンス  $Z_{vv}$  を同じ値と仮定する。式(5.13)と(5.14)を整理すると

$$\frac{p_{nN}/p_v}{p_{nE}/p_v} = \frac{Z_{vv}}{Z_v} = \frac{Z_{vo}' + Z_{vv}}{Z_{vo}'} \quad (5.15)$$

実測した鼻子音区間と有声破裂子音バズ区間の口腔内音圧と鼻孔放射音圧を表5. 3に示す。平均値としては、/m/のnasal murmurでは鼻孔放射と口腔内音圧の差は-39.2dBとなり、/b/のバズ区間では鼻孔放射と口腔内音圧の差は-50.3dBとなる。線形スケールで表示すると

$$\frac{p_{nN}/p_v}{p_{nE}/p_v} = 3.6 \quad (5.16)$$

となる。式(5.16)を式(5.15)に代入すると

$$Z_{vv}=2.6Z_{v_o}' \quad (5.17)$$

である。

表5.3 有声子音バズ区間と鼻子音区間の口腔内圧と鼻孔放射音圧(dB)

後続母音		/a/	/i/	/u/	/e/	/o/	平均値
/b/	内圧	128.7	124.7	134.0	129.5	125.7	128.5
	鼻孔	84.6	73.0	84.6	76.0	72.8	78.2
/m/	内圧	128.2	131.5	130.8	132.2	130.0	130.5
	鼻孔	89.5	90.6	91.9	93.1	91.6	91.3

## 5.6 軟口蓋の等価開口面積率

鼻音化音声では軟口蓋の開閉による開口面積が連続的に変化するが、十分なデータがないので従来の音声生成モデルでは軟口蓋はONとOFFの二状態をもつスイッチとされている。軟口蓋の動きの測定がいくつかなされてきた<sup>3)~6)</sup>が、それらの測定はほとんど定性的なものである。音声発声時の軟口蓋の開口面積を正確に測定することはかなり困難なので、ここでは、直接測定の代わりに鼻孔放射と口唇放射を用いて軟口蓋の等価開口面積率の推定方法を提案する。

### 5.6.1 軟口蓋の等価開口面積率の定義

式(5.17)と式(5.12)に代入して整理すると

$$\frac{Z_{v_o}'}{Z_{v_o}(t)} = \frac{1}{2.6} \left( \frac{p_m(v)/p_n(v)}{p_{mn}(t)/p_{nn}(t)} - 1 \right) \quad (5.18)$$

である。式(5.18)からわかるように、非鼻音化音声の場合の口唇放射と鼻孔放射の音圧比と、鼻音化音声の場合の音圧比を用いて、鼻音化音声区間の軟口蓋開口面積の等価インピーダンス $Z_{v_o}(t)$ と鼻子音区間（軟口蓋の開口面積が最大となっていると仮定する）の軟口蓋の最大開口面積の等価インピーダンス $Z_{v_o}'$ との関係を求められる。

声道インピーダンスを図5.11のT型回路で表示すると

$$L = \rho / S$$

$$C = S / \rho c \quad (5.19)$$

$$R = \sqrt{2\omega \rho \mu / (rS)}$$

ここで、L、C、Rはそれぞれ単位長さあたりの音響質量、音響コンプライアンスと音響抵抗である。Sとrはそれぞれ声道の断面積と半径であり、 $\rho$ は空気の密度、 $\mu$ は空気の粘性係数である。式(5.19)により、声道インピーダンスと声道断面積はほぼ逆比例にしているため、式(5.18)の軟口蓋部のインピーダンスの関係を用いて軟口蓋の等価開口面積率の変化を表わすことができる。したがって、軟口蓋の等価開口面積率  $S_r(t)$  は次式

$$S_r(t) = \frac{1}{2.6} \left( \frac{p_m(v)/p_n(v)}{p_{mn}(t)/p_{nn}(t)} - 1 \right) \quad (5.20)$$

となる。ただし、 $p_m(v)$ と $p_n(v)$ の関係は表5. 2に示す。

### 5. 6. 2 軟口蓋の等価開口面積率の推定結果

表5. 1に示した音声資料を式(5.20)で分析して次の結果を得た。ここで、いくつかの例を示して説明する。

音声の波形と推定した軟口蓋の等価開口面積率の例（／車を運転します／の一部／(kur)umaouNteNshimas／）を図5. 12に示す。推定した軟口蓋の等価開口面積率には、図5. 12の／maouN／のように、二つの鼻子音に挟まれた3つの母音区間で、前後の鼻子音の影響によって軟口蓋が開放している。

連続音声での $C_nV$ 型音節では、図5. 12の／ma／のように、鼻子音から母音へのわたりで鼻孔放射量はスムーズに減少する。それと反対に、 $VC_n$ 型の母音から鼻子音へのわたりでは、図5. 12の／uN／と／eN／のように、鼻孔放射はある小さい値から最大値に急激に増大する。推定した等価開口面積率は二段階に変化しており、母音の開始部分では軟口蓋の開口面積がゆっくり増大してくる。二段階目は鼻子音発声のため口唇が閉じると共に軟口蓋の等価開口面積率は急激に増大する。同様な結果を図5. 13にも示す。この結果は他の測定方法で得られた結果<sup>4)</sup>と同じである。

軟口蓋の開口面積の変化は言語音声の文脈にも関係している。例えば、図5. 12の／ima／の／i／のように、鼻子音の前の母音が狭母音で後ろの母音が広母音なら、この場合の狭母音はほかの場合（例えば図5. 13の／jiN／の／i／の場合）より強く鼻音化される。すなわち、その母音は開始部分から後ろの鼻子音の影響を強く受け

て軟口蓋の開口面積率は他の文脈より先に増大するのである。類似な結果は文献[5]で得られている。

母音前後の鼻子音の影響は発声の速さによって異なる。/moN/を速く発声した場合と遅く発声した場合をそれぞれ図5.14と図5.15に示す。図5.15に示したように、ゆっくり発声する場合、前の母音の影響は後ろの母音のそれより大きい。それに対して、はやく発声する場合は図5.14に示したように、前の鼻子音と後ろの鼻子音の影響はほぼ同じである。普通の速さで発声する場合、前の鼻子音の影響は後ろの鼻子音のそれよりやや強くなっている。

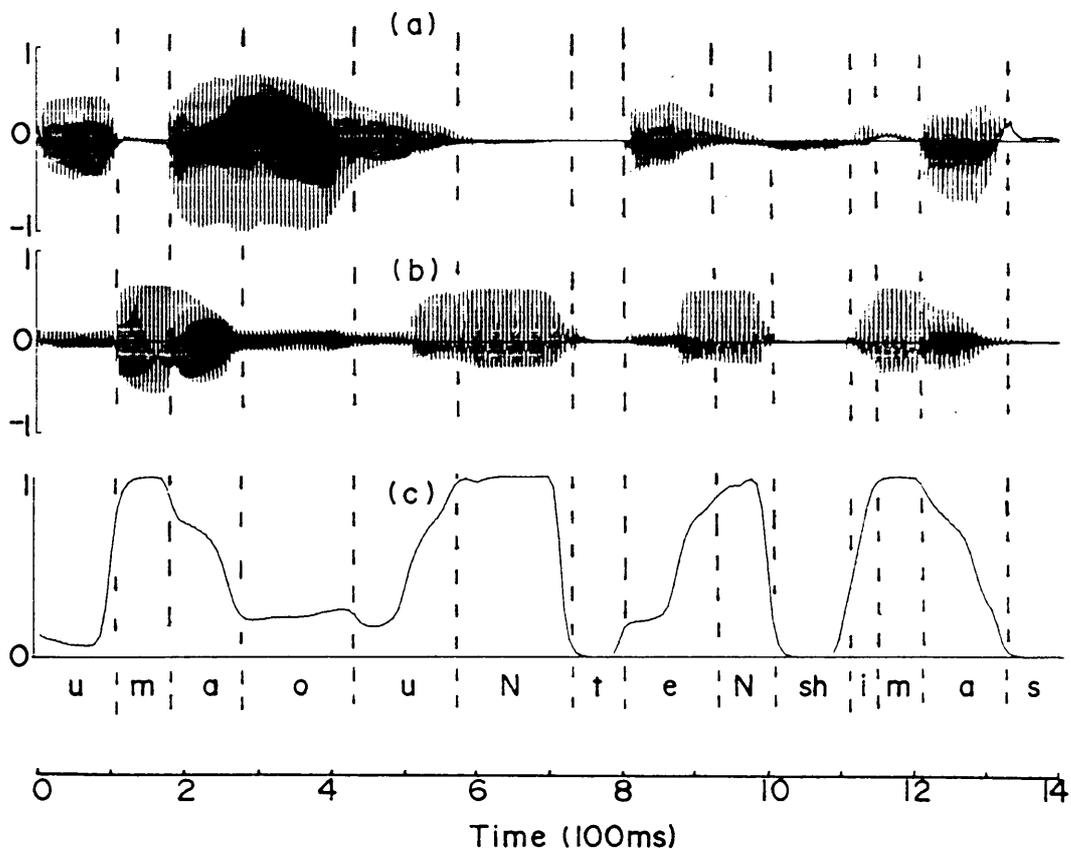


図5.12 短文(「車を運転します」)の一部の波形と  
推定した軟口蓋の等価開口面積率  
(a) 口唇放射 (b) 鼻孔放射  
(c) 推定した等価開口面積率

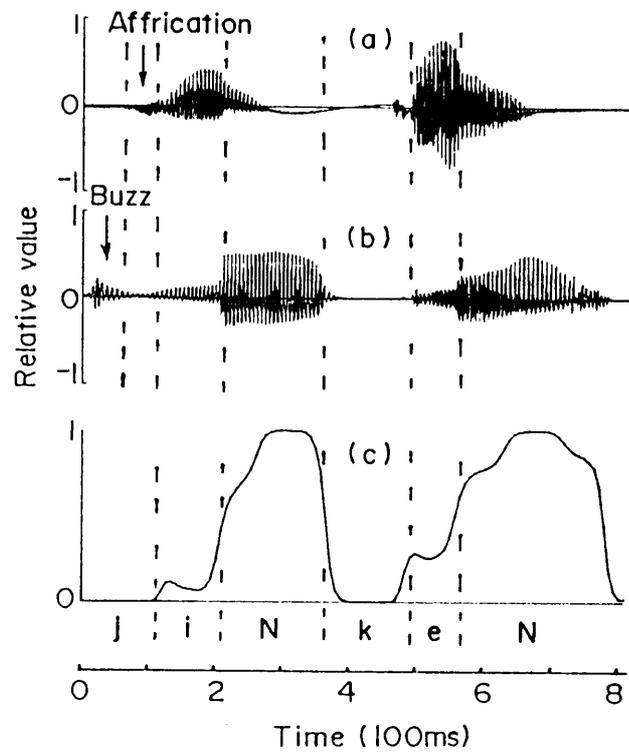


図5.13 単語(「人権」)/jikeN/の波形と  
推定した軟口蓋の等価開口面積率  
(a) 口唇放射 (b) 鼻孔放射  
(c) 推定した等価開口面積率

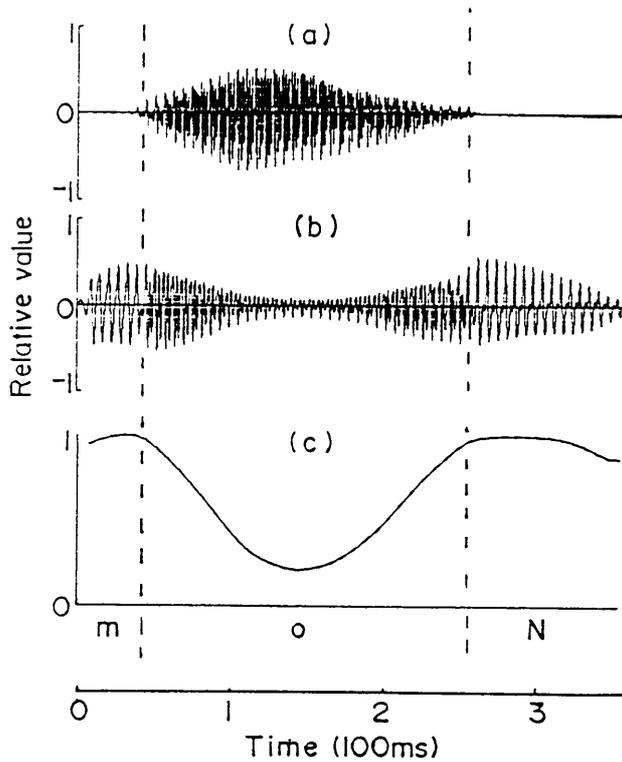


図5.14 速いスピードで発声した/moN/の波形と  
推定した軟口蓋の等価開口面積率  
(a) 唇放射 (b) 鼻孔放射  
(c) 推定した等価開口面積率

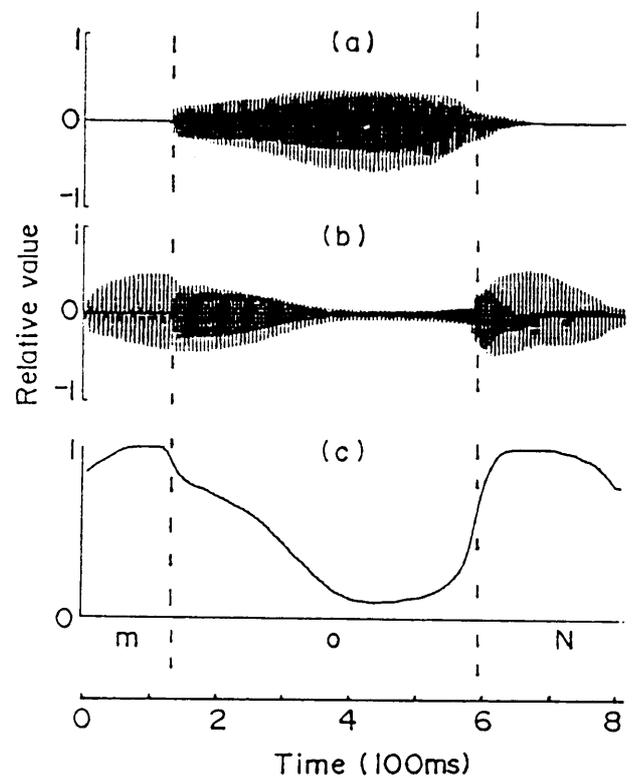


図5.15 遅いスピードで発声した/moN/の波形と  
推定した軟口蓋の等価開口面積率  
(a) 唇放射 (b) 鼻孔放射  
(c) 推定した等価開口面積率

## 5. 7 連続音声における軟口蓋の音響モデル

連続音声においては、非鼻音化音声や、鼻音化音声や、非鼻音化音声から鼻音化音声へまたは鼻音化音声から非鼻音化音声へのわたりなどのさまざまな場合があるので、軟口蓋の種々の動きを考慮する必要がある。このため、第3章では、非鼻音化母音における軟口蓋をダッシュポットとばねでつながる二枚の振動板としてモデル化した。第4章では、口腔内圧の激しい変化による軟口蓋の軟らかさの変化を定式化して軟口蓋の振動板モデルに導入した。さらに、本章で鼻音化音声における軟口蓋の等価開口面積率を推定した。上述の考察と推定に基づいて連続音声における軟口蓋の音響モデルをまとめて、図5. 16に示す。このモデルの下部のT型分岐は軟口蓋の振動板モデルの音響等価回路である。ここで、 $L$ 、 $R$ と $C$ は軟口蓋の音響質量、音響抵抗と音響コンプライアンスで、 $R_0$ と $C_0$ は二枚の板に挟まれた粘弾性体の音響抵抗と音響コンプライアンスである（第3章参照）。 $q$ は口腔内圧による軟口蓋の軟らかさの変化係数である（第4章参照）。

軟口蓋モデルの上部の分岐は軟口蓋の開口面積の等価回路である。 $S_r(t)$ は等価開口面積率で、 $L_{v_0}$ と $R_{v_0}$ は最大開放する時の軟口蓋の等価音響質量と音響抵抗である。非鼻音化音声（ $S_r(t)=0$ となる）の場合では、下部の振動板モデルだけ働く。鼻音化の度合いが小さい場合では、軟口蓋の振動板モデルの分岐と軟口蓋の開口面積モデルの分岐とは同時に作用する。鼻音化の度合いが大きい場合あるいは鼻子音区間では、軟口蓋の開口面積モデルの分岐のインピーダンスが振動板モデルのそれより小さいので、軟口蓋の開口面積モデルの分岐が作用する。

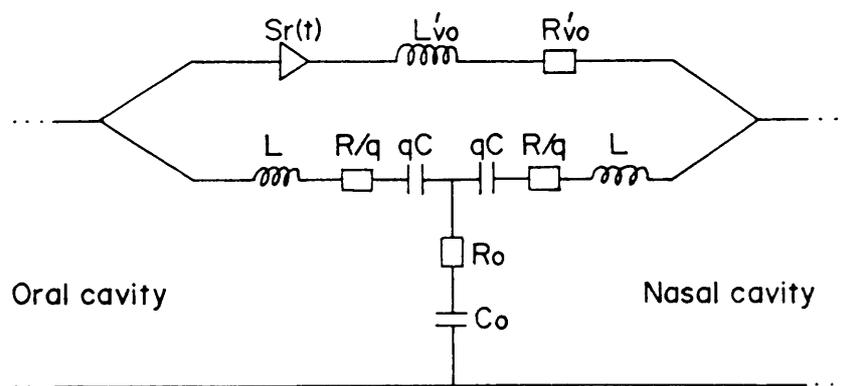


図5. 16 連続音声における軟口蓋の音響モデル

## 5. 8 むすび

本章では、音節や短文を発声した時文脈による放射音の変化を考察した。また、分離された口唇放射と鼻孔放射の周波数特性を分析して、その結果で、Fant<sup>7)</sup>の理論的な計算結果を確認した。さらに、鼻腔のホルマントと母音のホルマントの動的な変化を観測した。

上述の観測結果を用いて、軟口蓋開口面積の等価インピーダンスと軟口蓋振動の等価インピーダンスの関係を考察した上で、軟口蓋の開口面積率を放射音圧の関数と結び付けた。口唇放射と鼻孔放射による軟口蓋の等価開口面積率を推定した。最後に、非鼻音化母音、有声破裂子音のバズ区間と鼻音化音声での推定結果をまとめて、連続音声における軟口蓋の音響モデルを提案した。

## 第6章 声道内音圧及び頬の振動加速度による 頬インピーダンスの測定 — および合成音のパラメータ整合法で 声道壁インピーダンスの推定

### 6.1 まえがき

声道壁インピーダンスは音声生成モデルの構成のために必要な基本的物理量である。声道壁の効果を音声生成モデルに入れるために声道壁インピーダンスを正確に測定することが必要である。それゆえ、本章では、直接測定法と合成音のパラメータ整合法（すなわち、声道壁インピーダンスが音声のホルマント周波数と帯域幅に影響を与えることに着目して合成音声が実音声に合うようにインピーダンスのパラメータを推定する方法）の2つの方法を用いて、それぞれ声道壁インピーダンスを推定する。

声道壁インピーダンスの推定はこれまでにいくつか行われてきた。代表的な研究としては、Ishizakaら<sup>1)</sup>はインピーダンスヘッド内蔵型加振器で声道壁を振動させて、20Hz～160Hzの周波数範囲で声道壁の機械インピーダンス $z_w$ を直接測定した。鈴木<sup>2)</sup>は従来の声道壁インピーダンス $z_w$ によっては高圧下のホルマント周波数の上昇が説明できないことから出発して、高圧気体環境で発声した音声のホルマント周波数と、声道の断面積関数の等価回路から計算したホルマント周波数を比較することによって $z_w$ を求めた。当研究室<sup>3)</sup>ではIshizakaらと同様に加振器を用いて、20Hz～6000Hzの周波数範囲で声道壁の機械インピーダンス $z_w$ を直接測定すると共に、1気圧空気と31気圧ヘリウム混合気体環境での発声の場合、声道壁インピーダンスの変化が音声のホルマントの周波数と帯域幅にどんな影響を与えるかを検討した。Kamiyamaら<sup>4)</sup>は音響管を用い、頬の音圧反射特性を測定して頬のインピーダンス $z_w$ を推定した。これまでに発表された代表的な値を表6.1に示す。このうち、Flanagan(1965)の値は声道壁ではなく胃壁の測定値から推定されたもので、彼自身も声道壁インピーダンスとしては適切でないようであると述べている<sup>5)</sup>。

これまでの測定では主な問題点が二つある。一つは、実際の発声では駆動音源は声道内にあるが、測定の駆動源は声道壁の外側に置かれた。然し、声道壁の外側から内側への特性は内側から外側への特性と必ずしも同じではない。もう一つは、文献[1]と[3]の測定では、加振器と加速度ピックアップは声道壁の表面に接触しているので周波数が高くなると接触は悪くなる可能性がある。文献[3]では6kHzまで測定したが、

200Hz以上の範囲で声道壁の機械インピーダンスの実数部と虚数部とも小さくなるのは接触が悪くなった可能性が大きい。ところで、文献[1]では声道壁インピーダンスを160Hzまでしか測定しなかった。

音声発声時の声道壁の振動と放射は声道内の音圧によって引き起こされたものである。このことに基づいて、本章では従来と異なる頬インピーダンスの測定方法を提案した。この方法では、外部に置いた正弦波音源からの音を細い導音管で口腔内に入れ、口腔内の音圧と頬からの透過音圧及び頬の振動加速度を同時に測定した。それらの関係を用いて頬インピーダンスを求めた。

## 6.2 測定装置

口腔内の音圧と、頬からの透過音圧及び頬の振動加速度を測定する装置を図6.1(a)に示す。音源を口腔内に入れるために、防音箱（音源防音箱と呼ぶ）に入れたスピーカの音を導音管により取り出した。音源防音箱は大きさが W19×H16×D25cmで、厚さ1.5cmの合板製である。スピーカ(FOSTEX社のFE103)は直径10cmで、周りに吸音材をつめた。導音管は軟質プラスチックチューブで、外径10mm、内径7mm、長さ35cmである。音源防音箱の透過損失は75dB、導音管の壁の透過損失は53dBである。導音管の

表6.1 従来代表的な声道壁インピーダンスの値  
( $Z_w=B+j(\omega M-K/\omega)$ )

	B(g/s· cm <sup>2</sup> )	M(g/ cm <sup>2</sup> )	K(dyne/ cm <sup>3</sup> )	Append
Flanagan(1965)	6500	0.4	-	stomach surface
Ishizaka, French & Flanagan (1975)	800 1060 2320	2.1 1.5 2.4	84.5×10 <sup>3</sup> 213.2×10 <sup>3</sup> 491.0×10 <sup>3</sup>	relaxed cheek tensed cheek neck
J.Suzuki(1977)	1400	1.6	-	F <sub>1,1</sub> VS F <sub>1,3,9</sub>
H.Suzuki & T.Nakai (1987)	1600 1150 1170	2.12 1.96 1.87	154.9×10 <sup>3</sup> 127.2×10 <sup>3</sup> 142.8×10 <sup>3</sup>	Sub.1 Sub.2 Sub.3
Kamiyama, Miki, & Nagai (1990)	700 1900	0.5 0.3	- -	f<325Hz f>325Hz

先端開口部以外からの音の漏れを十分防ぐために、音源防音箱と導音管をさらに第2章で述べた大きな防音箱の中に入れた。大きな防音箱の透過損失が31dB以上であるから、二重に箱に入れた場合の透過損失は84dB以上となる。

頬の振動加速度の測定に用いた加速度ピックアップはENDEVCO社の2250A-10であり、円形で外径5.84mm、重さ0.4gである。図6.1(b)に示したように、加速度ピックアップを直立姿勢の被験者の片側の頬に面積約9cm<sup>2</sup>のテープで貼り付けて固定した。マイクロホンM2は加速度ピックアップを貼り付けていない側に、頬面から約3mmの位置に置いた。マイクロホンM1はB&K社である。口腔内のマイクロホンM1はB&K社のType 4182プローブマイクロホンである。プローブマイクロホンのプローブは音源防音箱から出ている導音管に沿わせて口腔内に挿入した。プローブの先は導音管の出口より5mm出ている。音源波形は正弦波で20Hz~3200Hzの範囲を1/3オクターブごとに測定し

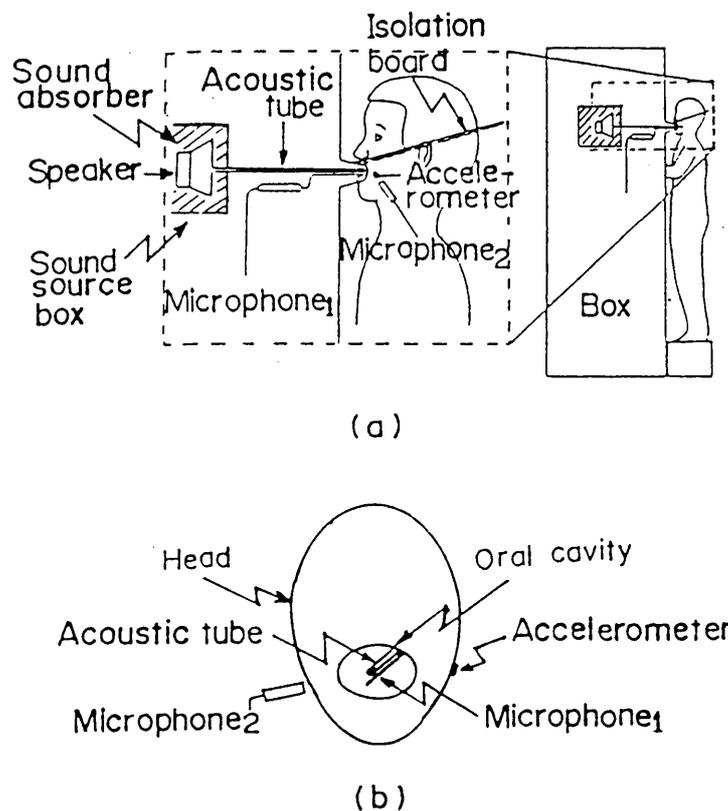


図6.1 測定システムの略図

(a) 測定システムの側面図

(b) マイクロホンと加速度計の相対位置

た。音声発声時の口腔内音圧レベルは120～130dBに達するので、口腔内の測定音圧を120dB程度にした。測定する時、被験者は唇で導音管をしっかりとくわえ一定の声道形（/u/の発声時に近い）を保つようにした。1kHz以上の周波数については、声道の共振と反共振周波数を避けるために、その周波数の0～10%範囲内で入力が一固定のもとで出力の最大値と最小値を避けて中間の値を持つ周波数にした。測定実験は無響室で行った。3チャンネルの信号はデータレコーダ(SONY社のDFR-3415、遮断周波数は5kHz)に収録して、サンプリング周波数20kHzでA-D変換し、ワークステーション(H.P.アポロDN3500)に取り込んだ。被験者は成人男性3名女性1名である。

### 6. 3 測定等価回路とインピーダンスの推定方法

#### 6. 3. 1 測定等価回路

粘弾性体の軟部組織でできている頬は第一次近似としては質量、ステイフネス、機械抵抗の直列接続と考えられる<sup>1)～5)</sup>。頬に加わる空気負荷のインピーダンスは、通常、質量と機械抵抗の直列接続で近似される。頬の放射面積をSとすると、機械インピーダンス(B:単位面積あたりの機械抵抗、M:単位面積あたりの質量、K:単位面積あたりのステイフネス)と音響インピーダンス( $R_w$ :音響抵抗、 $L_w$ :音響質量、 $C_w$ :音響コンプライアンス)の関係は

$$\begin{aligned} B &= R_w S \\ M &= L_w S \\ K &= S / C_w \end{aligned} \quad (6.1)$$

となる。本測定法の音響等価回路は図6. 2となる。ここで、 $P_1$ 、 $P_2$ とUは、それぞれ

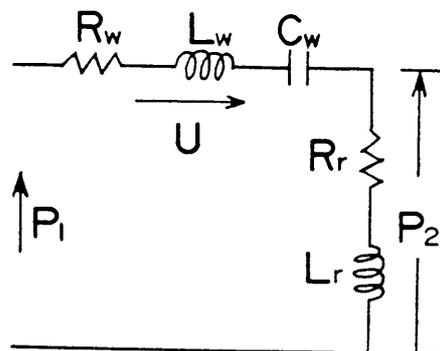


図6. 2 測定システムの音響的等価回路図

れ口腔内音圧、頬からの透過音圧及び頬の体積速度である。それらを次のように定義する。

$$\begin{aligned} P_1 &= |P_1| \exp(j(\omega t - \phi_{p1})) \\ P_2 &= |P_2| \exp(j(\omega t - \phi_{p2})) \\ U &= |U| \exp(j(\omega t - \phi_u)) \end{aligned} \quad (6.2)$$

$R_w$ 、 $L_w$ と $C_w$ は頬の音響抵抗、音響質量、音響コンプライアンスで、 $R_r$ 、 $L_r$ は頬に加わる空気負荷インピーダンスである。振動加速度は

$$A = |A| \exp(j(\omega t - \phi_a)) \quad (6.3)$$

と定義する。体積速度 $U$ と振動加速度 $A$ の関係は

$$U = \frac{AS}{j\omega} \quad (6.4)$$

となる。

図6.1のマイクロホン $M_1$ で音圧 $P_1$ 、マイクロホン $M_2$ で音圧 $P_2$ 、加速度ピックアップで加速度 $A$ を測定することによって、体積速度 $U$ は式(6.4)から求められ、図6.2の等価回路から頬に加わる空気負荷インピーダンス $Z_r$ は

$$Z_r = R_r + j\omega L_r = P_2/U \quad (6.5)$$

となり、頬インピーダンス $Z_w$ は

$$Z_w = \frac{P_1}{U} - Z_r = \frac{P_1 - P_2}{U} \quad (6.6)$$

から求められる。式(6.4)、(6.5)を式(6.6)に代入して整理すると

$$Z_w = \frac{j\omega}{S} \left( \frac{P_1}{A} - \frac{P_2}{A} \right) \quad (6.7)$$

となる。

### 6.3.2 インピーダンスの推定方法

図6.2の等価回路によると頬インピーダンスは

$$Z_w = R_w + j \left( \omega L_w - \frac{1}{\omega C_w} \right) \quad (6.8)$$

となる。式(6.7)を(6.8)に対応させ式(6.1)を利用すると、頬の機械インピーダンスの実数部 $B(\omega)$ と虚数部 $X(\omega)$ は

$$B(\omega) = \frac{\omega |P_1|}{|A|} \cos\left(\phi_{p1a} + \frac{\pi}{2}\right) - \frac{\omega |P_2|}{|A|} \cos\left(\phi_{p2a} + \frac{\pi}{2}\right) \quad (6.9)$$

$$X(\omega) = \frac{\omega |P_1|}{|A|} \sin\left(\phi_{P_1/A} + \frac{\pi}{2}\right) - \frac{\omega |P_2|}{|A|} \sin\left(\phi_{P_2/A} + \frac{\pi}{2}\right) \quad (6.10)$$

となる。但し、 $\phi_{P_1/A} = \text{Arg}(P_1/A)$ 、 $\phi_{P_2/A} = \text{Arg}(P_2/A)$  である。

$P_1/A$ と $P_2/A$ の振幅と位相を精度良く推定するため、データ系列に4096点のハミング窓を掛け、2048点ずつ移動して、FFTを5回とり、それらの平均値を求めるという方法を用いた。ここで、振動加速度Aと音圧  $P_i$  のFFTをそれぞれ $A(n)$ と $P_i(n)$  ( $i=1,2$ )とし、振動加速度 Aのパワースペクトルを $A_{aa}(n)$ 、振動加速度Aと音圧 $P_i$ のクロスパワースペクトルを  $S_{aP_i}(n)$ とすると

$$A_{aa}(n) = \frac{1}{N} A^*(n)A(n) \quad (6.11)$$

$$S_{aP_i}(n) = \frac{1}{N} A^*(n)P_i(n) \quad (6.12)$$

$$\frac{P_i(n)}{A(n)} = \frac{S_{aP_i}(n)}{A_{aa}(n)} \quad (6.13)$$

となる<sup>7)</sup>。ここで、 $n=0, 1, \dots, N-1, N=4096$ である。

式(6.9)で求められる $B(\omega)$ は類の機械インピーダンスの抵抗成分である。機械インピーダンスの虚数成分 $X(\omega)$ をMとKで表わせば

$$\begin{aligned} X(\omega) &= \omega M - K/\omega \\ &= \sqrt{MK} \left( \omega/\omega_0 - \omega_0/\omega \right) \end{aligned} \quad (6.14)$$

である。ただし、 $\omega_0$ は測定値 $X(\omega)=0$ となる角周波数で次式となる。

$$\omega_0 = \sqrt{K/M} \quad (6.15)$$

周波数 $\omega$ でのMとKを $X(\omega)$ と $\omega_0$ で表わせば

$$M = \frac{X(\omega)}{\omega_0(\omega/\omega_0 - \omega_0/\omega)} \quad (6.16)$$

$$K = \frac{\omega_0 X(\omega)}{(\omega/\omega_0 - \omega_0/\omega)} \quad (6.17)$$

となる。

#### 6.4 予備実験

原理的に、6.3で述べた測定方法は、声道（または管）の内部圧力と外部圧力、

及びそれらの差による声道壁の半径方向の動きを用いて単位面積あたりの等価質量、等価ステイフネス及び等価機械抵抗を推定する方法である。この方法の有効性を確認するため、材質の軟らかいプラスチック・ゴム製の管（以下被験管と記す）に対して予備実験を行った。

#### 6. 4. 1 機械工学的な方法での測定

実験対象としての被験管（管A）はプラスチックとゴムの混合材でできており、内径32mm、外径40mm、肉厚4mm、長さ15.5cmである。まず、被験管Aの面積と秤で測った重さから単位面積あたりの質量 $M'$ を求めた。管Aの単位面積あたりの質量 $M'$ は $0.49 \text{ g/cm}^2$ である。

次に被験管の中心軸に垂直方向で輪を切り出し、引張り試験（付録A参照）で、管壁の円周方向でのヤング率 $E$ と単位面積あたりのステイフネス $K'$ を求めた。試験ピースの形を図6. 3に示す。試験ピースのサイズは標点間距離（ $L_0$ ）19.25mm、標点間の部分の厚さ（ $t$ ）3.75mm、平均幅（ $d$ ）4.68mmである。標点間の伸びを $\Delta L$ とすると、加えた力 $F$ と伸び $\Delta L$ との関係は図6. 4となる。それらの関係は

$$S=td \quad (6.18)$$

$$E= \frac{F/S}{\Delta L/L_0} \quad (6.19)$$

$$K' = \frac{F/S}{\Delta L} \quad (6.20)$$

となる。図6. 4に示したように加えた力 $F$ と伸び $\Delta L$ との関係曲線は980kdyneの近くに切り替えるカーブがある。本実験では、内圧を120dB（ $200 \text{ dyne/cm}^2$ ）程度にしたので点線の範囲内でステイフネスの平均値を求める。式(6.18)～(6.20)に従って求めたヤング率と単位面積あたりのステイフネスはそれぞれ $7.22 \times 10^6 \text{ dyne/cm}^2$ 、 $3.75 \times 10^6 \text{ dyne/cm}^3$ であった。ここでは、求めた単位面積あたりの質量とステイフネスを静的な値と称し、それぞれ $M'$ と $K'$ で表示する。その方法を機械工学的な方法と称する。

#### 6. 4. 2 内圧と壁の振動による測定

6. 3で述べた方法で管Aの壁の機械インピーダンスを64Hz～1280Hzの周波数範囲で1/3オクターブ毎に求めた。正弦波音源をプローブマイクロホン付きの導音管で被験管の中に9cm挿入した。被験管の奥の一端は完全に蓋をし、導音管を差し込む側の一端は半開放状態になっている。導音管の先のあたり、すなわち奥から6.5cmの被験

管の側壁に加速度ピックアップを貼付し、反対側で管壁から約3mmの位置にマイクロホンM2を置いた。加速度ピックアップを面積約9cm<sup>2</sup>のテープを用いて貼り付けた。実験システムの周波数特性を調べたところ管Aに導音管を奥から6.5cmまで挿入した状態で実測した共振周波数は420Hz、525Hzと990Hz、反共振周波数1300Hzであった。空気負荷インピーダンス $Z_f$ を被験管の壁インピーダンスに比べ無視できる程度に小さいと仮定し、式(6.9)と(6.10)の第二項を零とおいて、管壁の機械インピーダンスを求めた。

本測定法では被験管の半径方向の等価質量、等価ステイフネスと等価機械抵抗を測定したが、機械工学的な方法で測定した値と比較するため本測定法で測定した等価機械インピーダンスを静的な値に変換した。静的な値を用いて比較するのは、機械イン

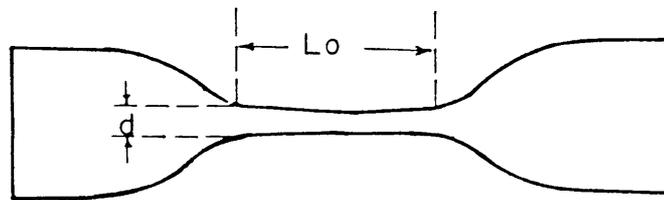


図6.3 引張り試験用の試験ピースの外観  
長さ $L_0=19.25\text{mm}$ , 幅 $d=4.68\text{mm}$ , 厚さ $t=3.75\text{mm}$

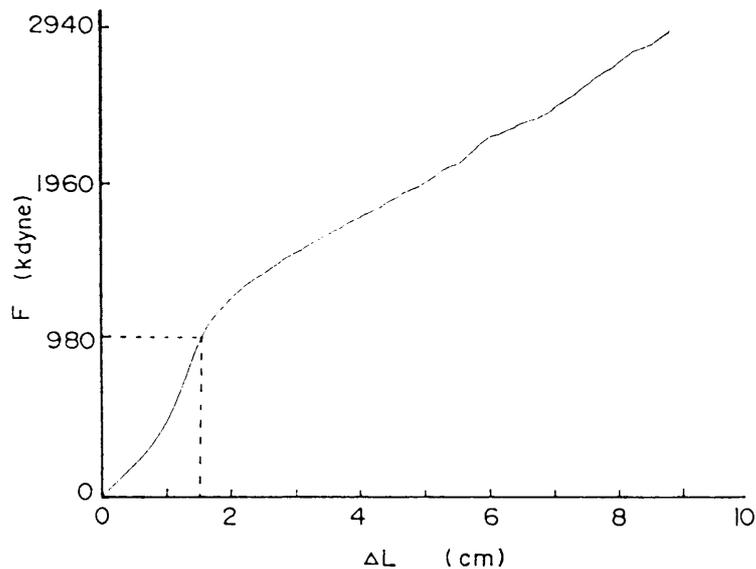


図6.4 加えた力(F)と伸び(ΔL)の関係

ピーダンスのパラメータ、特に単位面積あたりの質量と加速度計による単位面積あたりの附加質量の関係を明確に表わせるからである。

被験管Aは、管壁の厚さ(0.4cm)が管の内半径(1.6cm)に比べかなり小さく管の外側の音圧が管内の音圧より60dB程度小さくなっているため、内圧だけを受ける薄肉円筒<sup>8)</sup>として近似することができる。そうした場合、管Aの等価機械インピーダンスから静的な値への変換係数(付録B参照)は3.2となる(以下予備実験で本測定法で得られた値は静的な値に変換して表わす)。その結果を図6.5に示す。得られた値は周波数によって変化するが、その平均値と標準偏差は単位面積あたりの機械抵抗 $2.38 (\pm 0.34) \times 10^3 \text{ g/s} \cdot \text{cm}^2$ 、単位面積あたりの質量 $0.58 (\pm 0.05) \text{ g/cm}^2$ 、単位面積あたりのステイフネス $4.18 (\pm 0.39) \times 10^6 \text{ dyne/cm}^3$ であった(表6.2参照)。括弧内の数字は測定周波数範囲での標準偏差である。管Aの単位面積あたりの質量について秤で測った値と本測定法で求めた値の差は $0.09 \text{ g/cm}^2 (0.58-0.49)$ である。この差は加速度ピックアップによる単位面積あたりの附加質量と空気の負荷効果による質量を含む測定誤差からなっている。その測定誤差を確かめるため、もう一つのゴム管(管B、内径50mm、外径56mm、長さ16cm)について同様な測定を行ったところ、本測定方法では $0.40 (\pm 0.04) \text{ g/cm}^2$ であり、秤で測った単位面積あたりの質量は $0.35 \text{ g/cm}^2$ であった。その二つの管について、単位面積あたりの質量を、本測定法で測定した値と秤で測った値とを比べると、その差は $0.05 \sim 0.09 \text{ g/cm}^2$ となった。よって、管Aと管Bの場合では、加速度ピックアップによる単位面積あたりの附加質量は $0.05 \sim 0.09 \text{ g/cm}^2$ と見られる。頬の測定時、加速度ピックアップを被験管の場合と同様に貼り付けたので、頬に加わった単位面積あたりの附加質量は被験管と同程度と考えられる。

機械工学的な方法で測定した値と変換された本測定法で測定した値を表6.2に示す。機械抵抗が機械工学的な方法で測定されていないので機械抵抗に対して本測定法の測定誤差はわからないが、管Aの質量とステイフネスについて本測定法の相対誤差はそれぞれ8.2%と11.5%である。これらの結果から、我々の測定方法は頬インピーダンスの測定に対しても有効であろうと考えられる。

## 6.5 測定結果と頬インピーダンスの推定

### 6.5.1 測定した音圧と振動量

音源は正弦波で、20Hz~3200Hzの範囲を1/3オクターブごとに、口腔内音圧 $P_1$ 、頬からの透過音圧 $P_2$ および頬の振動加速度 $A$ を同時に測定した。測定したデータ系列に

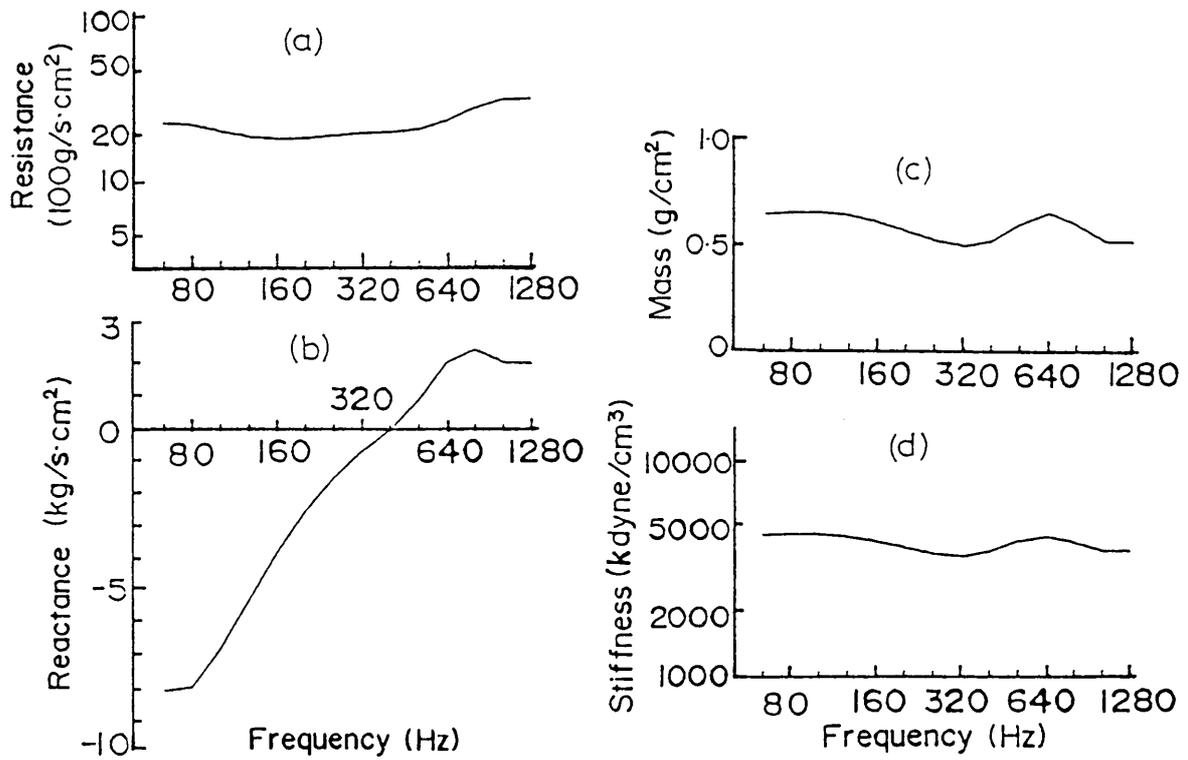


図6.5 推定したプラスチック・ゴム混合管の壁の  
静的な機械インピーダンスの周波数特性  
(a) 単位面積あたりの機械抵抗  
(b) 機械インピーダンスの虚数成分  
(c) 単位面積あたりの質量  
(d) 単位面積あたりのスティフネス

表6.2 機械工学的な方法と本測定法で測定した  
プラスチック管の壁インピーダンス

$$Z=B+j(\omega M-jK/\omega) \quad B: g/s \cdot cm^2 \quad M: g/cm^2 \quad K: dyne/cm^3$$

		mechanical method	present method	mass of Acc. loaded on tube	error %
Tube A	B	—	$2.38 \times 10^3$	—	—
	M	0.49	0.58	0.05~0.09	8.2
	K	$3.75 \times 10^6$	$4.18 \times 10^6$	—	11.5
Tube B	M	0.35	0.40	0.05~0.09	11.4

対して式(6.11)～(6.13)を利用して、 $P_2/P_1$ 、 $A/P_1$ と $P_2/A$ のスペクトルの振幅と位相差を求めた。加速度ピックアップと各マイクロホンは音源からの距離がそれぞれ異なるので、それらの位相は音源からの距離に応じて補正した。

頬からの透過音圧 $P_2$ の口腔内音圧 $P_1$ に対する振幅比 ( $|P_2/P_1|$ )と位相差 ( $\phi_{21}$ )を図6. 6に示す。頬の振動加速度 $A$ と口腔内音圧 $P_1$ の振幅比 ( $|A/P_1|$ )と位相差 ( $\phi_{ap1}$ )を図6. 7に示す。図6. 7では振幅比は40～80Hzの間で高めになっている。これは声道壁の共振周波数の付近では頬が振動しやすくなったと思われる。周波

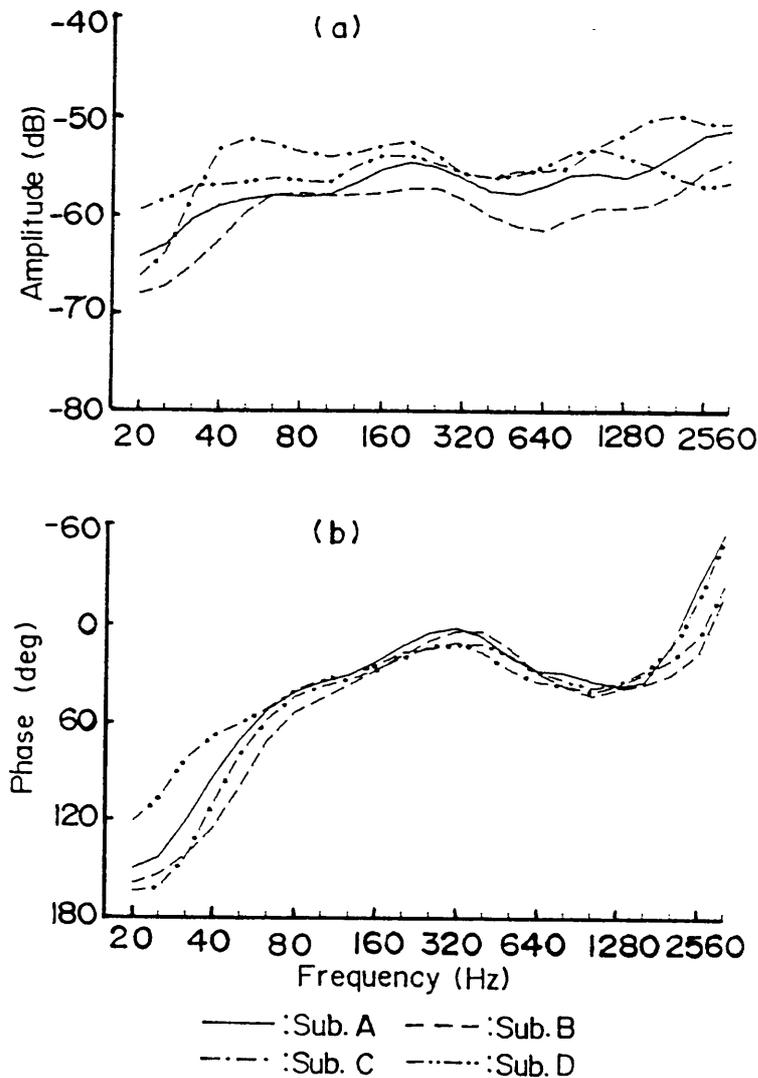


図6. 6 透過音圧( $P_2$ )と口腔内音圧( $P_1$ )の振幅比と位相差  
 (a) 振幅比 ( $|P_2/P_1|$ )  
 (b) 位相差 ( $\phi_{21} = \text{Arg } P_2 - \text{Arg } P_1$ )

数が高くなると図6.7の振幅比はやや小さくなる。これに対して図6.6に示した振幅比は下がる傾向がなくほぼ平らである。これは、頬に加わる空気負荷インピーダンスが周波数の上昇に伴って大きくなる為と考えられる。

頬からの透過音圧 $P_2$ と頬の振動加速度 $A$ の振幅比( $|P_2/A|$ )と位相差( $\phi_{p_2 a}$ )を図6.8に示す。その比は周波数が高くなるにつれて全体として上昇する傾向があり、位相差は周波数が高くなるとばらつきが大きくなる。この $P_2/A$ は式(6.6)と(6.7)からわかるように、放射インピーダンス $Z_r$ に関係する量である。その振幅比は、-40dB~

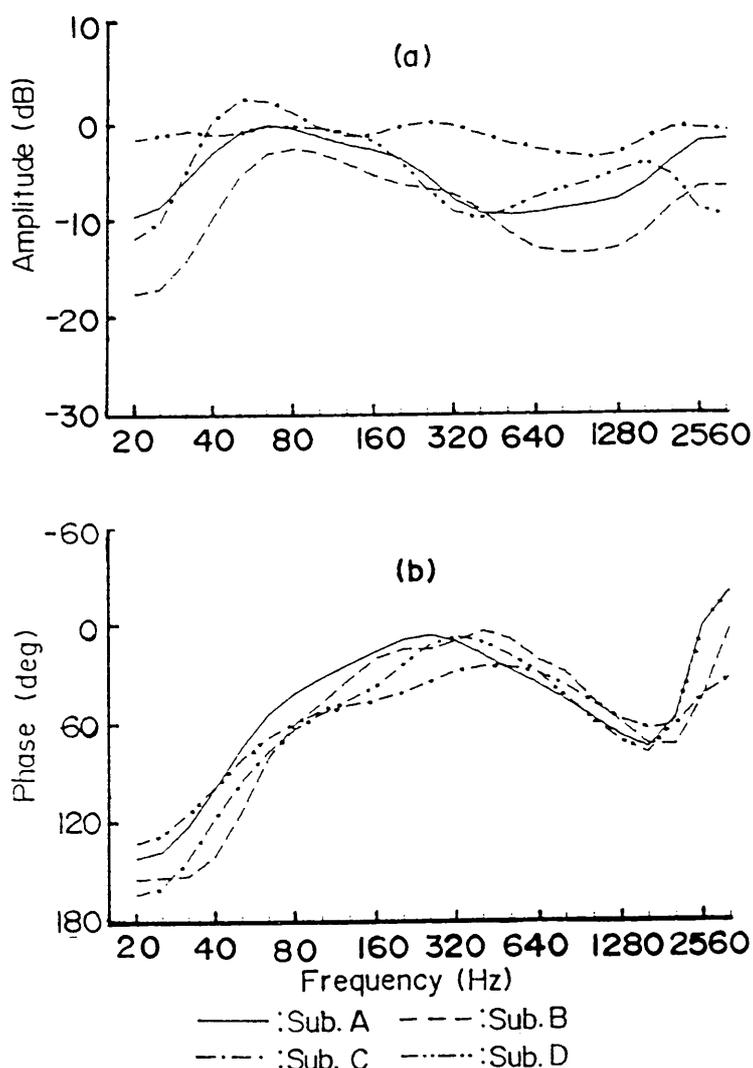


図6.7 振動加速度(A)と口腔内音圧( $P_1$ )の振幅比と位相差  
 (a) 振幅比 ( $|A/P_1|$ ) ( $0\text{dB}=1 \text{ cm}^2/\text{g}$ )  
 (b) 位相差 ( $\phi_{ap1}=\text{Arg } A - \text{Arg } P_1$ )

-60dBであって、 $|P_1/A|$ に比べてはるかに小さい。

### 6.5.2 頬インピーダンスの推定

振動加速度  $A$  と音圧  $P_2$  との比  $P_2/A$  は、式(6.7)により頬インピーダンス  $Z_w$  にも寄与するが、図6.7と図6.8から見られるように  $|P_2/A|$  は  $|P_1/A|$  より50~60dB小さい。よって、この分を無視して頬インピーダンスの実数成分と虚数成分を求める。ここで、式(6.9)と式(6.10)及び式(6.16)と(6.17)を整理すると単位面積あたりの機械インピーダンス  $B$ 、 $M$  および  $K$  は次式で与えられる。

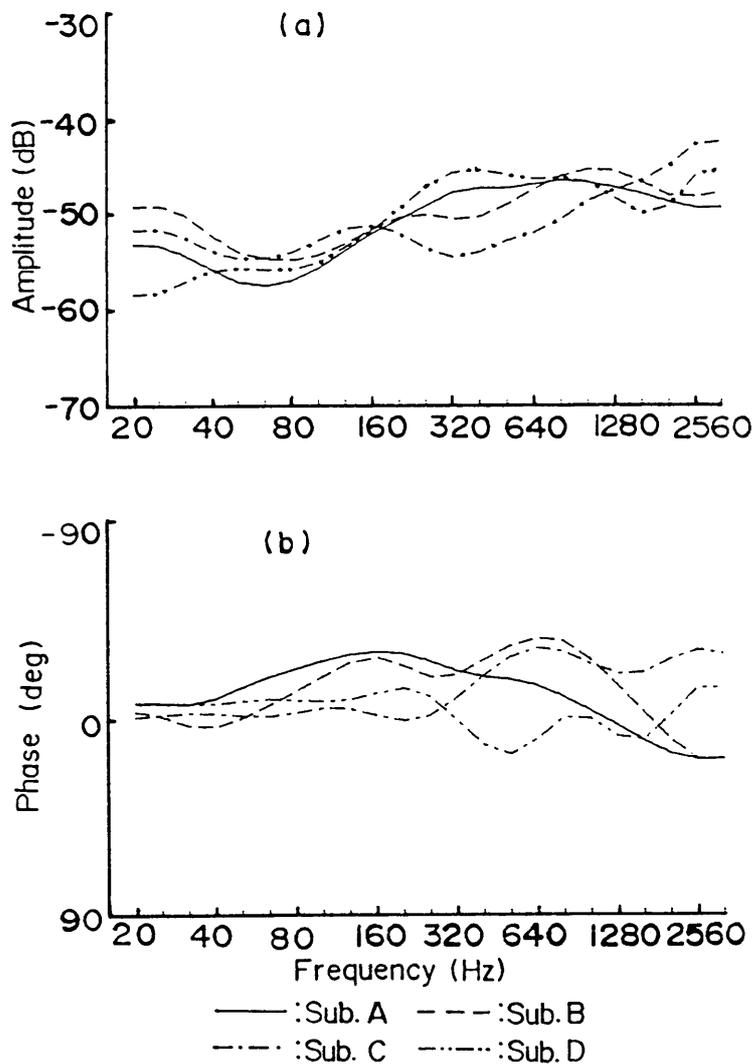


図6.8 透過音圧( $P_2$ )と振動加速度( $A$ )の振幅比と位相差  
 (a) 振幅比 ( $|P_2/A|$ ) ( $0\text{dB}=1 \text{ g/cm}^2$ )  
 (b) 位相差 ( $\phi_{P2A}=\text{Arg } P_2 - \text{Arg } A$ )

$$B(\omega) = \frac{\omega |P_1|}{|A|} \cos\left(\phi_{p1a} + \frac{\pi}{2}\right) \quad (6.21)$$

$$M(\omega) = \frac{\omega |P_1| \sin(\phi_{p1a} + \pi/2)}{|A| \omega_0 (\omega/\omega_0 - \omega_0/\omega)} \quad (6.22)$$

$$K(\omega) = \frac{\omega_0 \omega |P_1| \sin(\phi_{p1a} + \pi/2)}{|A| (\omega/\omega_0 - \omega_0/\omega)} \quad (6.23)$$

これらの式からわかるように、単位面積あたりの機械インピーダンスを求めるとき、頬の放射面積  $S$  は実際に必要がない。

求めた機械インピーダンスの実数部  $B(\omega)$  と虚数部  $X(\omega)$  を図 6. 9 に示す。図 6. 9(a) の実数部は単位面積あたりの機械抵抗である。それは、周波数依存性を示し、周波数が高くなるとともに大きくなる。20Hz～320Hz の周波数範囲で機械抵抗は 266～600g/s·cm<sup>2</sup> (平均値) でほぼ安定な値である。周波数 400Hz～800Hz で機械抵抗は急激に増大し、周波数がそれ以上になると約 15000g/s·cm<sup>2</sup> (平均値) の安定な値に達する。

図 6. 9(b)(c) は虚数成分  $X(\omega)$  であるが、それが零となる共振周波数は 32Hz～58Hz の範囲にある。共振周波数の近くの虚数成分の曲線の傾向はそれぞれ若干異なるが周波数の対数に対してほぼ線形になっている。20Hz～3200Hz の周波数範囲で概観すると、実数成分、虚数成分とも 4 人についてほぼ一致している。

頬インピーダンスの虚数成分に対して式(6.22)、(6.23)を利用して求めた頬の単位面積あたりの質量とステイフネスを図 6. 10 に示す。頬インピーダンスの抵抗成分と異なり、頬の単位面積あたりの質量とステイフネスは高周波領域で特に大きくなるという傾向はなく、むしろ小さくなる例も見られる。単位面積あたりの質量は、周波数 1280Hz のあたりで 4 人のばらつきが大きく、さらに周波数が高くなると共にやや小さくなる。全周波数にわたる平均値は 1.40(-0.23～+0.31)g/cm<sup>2</sup> である。単位面積あたりのステイフネスの平均値は 182(-55～+88)×10<sup>3</sup>dyne/cm<sup>3</sup> である。ここで、括弧内の数字は 4 人のデータの偏差である。測定した 4 人の値を表 6. 3 に示す。

## 6. 6 測定結果の考察

表 6. 1 に頬インピーダンスの従来 of 代表的な値を示したが、それらを本測定法で測定した表 6. 3 の値と比較し、考察する。

Ishizaka ら(1975)<sup>1)</sup> と当研究室(1987)<sup>3)</sup> による測定と今回の測定とで最も異なる点は、機械抵抗が前者では低周波で大きく高周波で小さくなるのに対し、後者では逆に

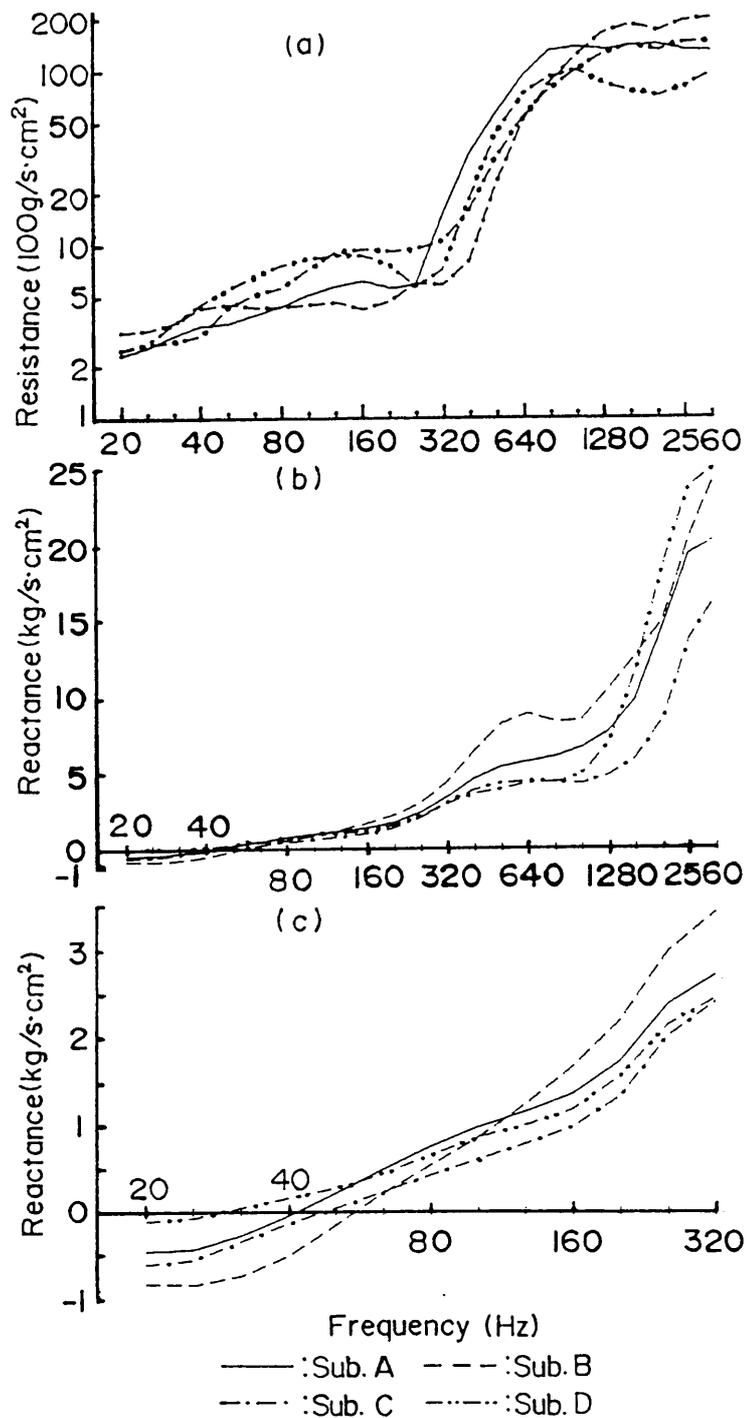


図6.9 声道内音圧と振動加速度から求めた頬インピーダンスの実数成分と虚数成分

- (a) 機械インピーダンスの実数成分 $B(\omega)$
- (b) 機械インピーダンスの虚数成分 $X(\omega)$
- (c) 20~320Hz周波数範囲での機械インピーダンスの虚数成分

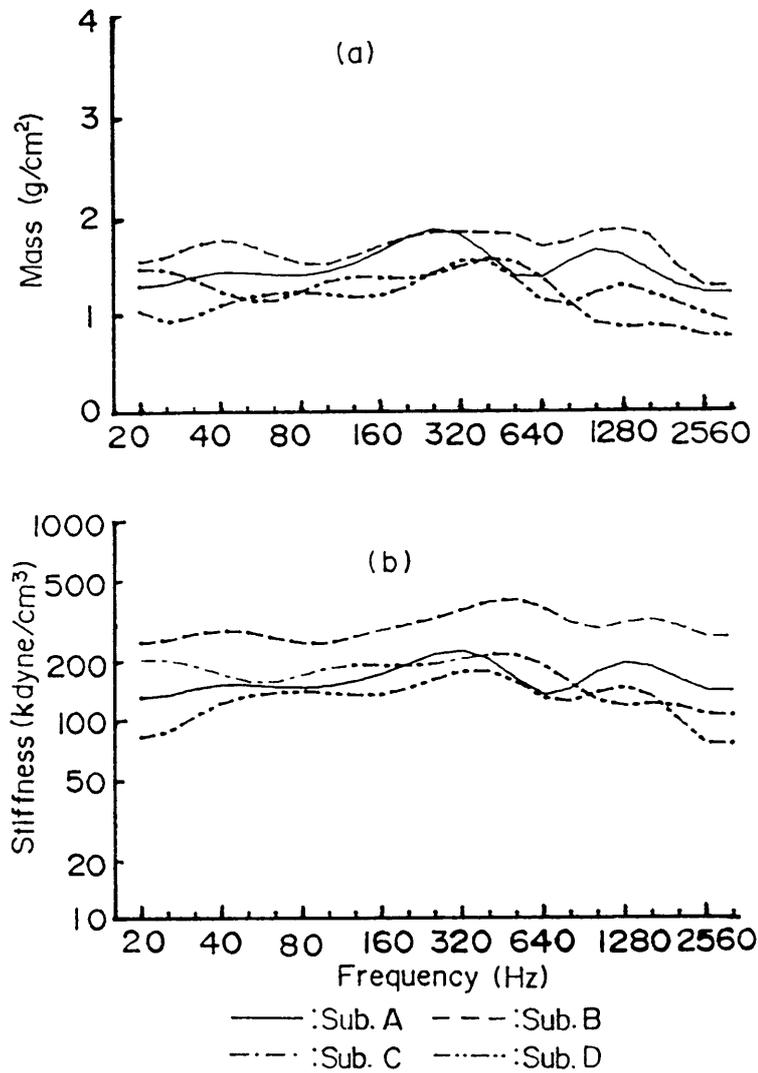


図6. 10 推定した頬の単位面積あたりの質量とステイフネス  
 (a) 単位面積あたりの質量  
 (b) 単位面積あたりのステイフネス

表6. 3 測定した受験者4人の頬インピーダンス  
 $Z_w = B + j(\omega M - K/\omega)$

	$B(\text{g/s} \cdot \text{cm}^2)$	$M(\text{g/cm}^2)$	$K(\text{dyne/cm}^3)$
Subject A	256~14400	1.49(±0.19)	155(±20.0) × 10 <sup>2</sup>
Subject B	323~20600	1.71(±0.18)	270(±26.7) × 10 <sup>3</sup>
Subject C	221~15000	1.23(±0.14)	170(±17.7) × 10 <sup>2</sup>
Subject D	267~10500	1.17(±0.12)	127(±16.9) × 10 <sup>2</sup>
Average	266~15100	1.40(±0.12)	182(±15.3) × 10 <sup>3</sup>

\* The data in ( ) are standard deviations.

低周波で小さく高周波で大きくなることである。しかし、200Hz以下の周波数範囲の平均値を比べると今回の測定値は文献[1]及び文献[3]の値とほぼ一致している。文献[3]では6000Hzまで測定したが、200Hz以上の周波数範囲での機械抵抗は今回の測定値に比べてはるかに小さい。機械抵抗の値が文献[1]、[3]と今回の測定とで逆の周波数特性を示すのは、文献[1]、[3]のような測定では一体型の駆動振動源と加速度ピックアップが頬の外側に接触しているので周波数を高くすると密着度が悪くなることと、加振器の影響質量による誤差が生ずることである。本測定方法では、もし加速度ピックアップと頬との密着度が悪くなったら求めたインピーダンスは大きくなる可能性がある。

単位面積あたりの質量では、今回の測定値は文献[1]、[3]のそれよりやや小さい程度である。単位面積あたりのステイフネスは今回の測定値は文献[1]の筋肉の緊張時の値よりやや小さく、文献[3]の測定値よりやや大きい。それは唇で導音管をしっかりとくわえているので頬には力が多少入っているからと考えられる（Ishizakaらによるtensed時の値は文献[1]では $33.3 \times 10^3 \text{ dyne/cm}^3$ となっていたが、1981年に原著者により $213.2 \times 10^3 \text{ dyne/cm}^3$ と訂正されたので表6. 1には訂正された値を示した）。

Kamiyamaら<sup>4)</sup>は5000Hzまでの範囲で頬の音響反射特性から頬インピーダンスを推定した。その測定ではステイフネス成分は測定していないが、抵抗成分は周波数が高くなると大きくなる傾向があることを指摘している点は本測定と同じである。文献[4]では頬インピーダンスの抵抗成分は二つの値で近似されており、その切り替え周波数は325Hzであった。今回測定した抵抗の値では320Hz付近に切り替えとみられるカーブが現れている。そこから800Hzまで抵抗の増大は激しく、800Hz以上の範囲でほぼ安定した値になる。320Hz以下の周波数範囲では文献[4]と本測定の抵抗値は同じオーダーであるが、高い周波数領域では本測定法で求めた値は6倍程度大きい。

単位面積あたりの質量が周波数の高くなると共に小さくなる傾向があるのは文献[4]と本測定は同じであるが、文献[4]の値は本測定の値の3分の1ぐらいである。文献[3]では1気圧空気と31気圧ヘリウム混合気体の場合で声道の伝達特性を考察した。その考察結果では、31気圧の場合、声道壁の単位面積あたりの質量を $0.4 \text{ g/cm}^2$ 以下にすると第一ホルマントに相当する山が消えてしまう。しかし、31気圧のもとで実測した音声では第一ホルマントのスペクトルの山は存在する。このような事実から、本測定の値は真値に近いものと思われる。

## 6. 7 合成音のパラメータ整合法による声道壁インピーダンスの推定

従来、合成音のパラメータ整合法による声道壁インピーダンスの推定には用いた音声資料は母音であった<sup>2)</sup>・<sup>4)</sup>。しかし、母音に対する声道壁インピーダンスの効果は必ずしも大きくないので母音の合成によるパラメータ整合法では推定精度に問題がある。一方、有声破裂子音のバズ区間では声道の空気通路が閉鎖されているので、この時の声道壁からの放射と口腔内音圧の音響特性はほとんど声道壁インピーダンスで決められる。このことから出発して、本節では、有声破裂子音のバズ区間の口腔内圧、鼻孔放射、咽頭壁放射を用いて声道壁インピーダンスを推定した。

### 6. 7. 1 実音声の音響特性

第4章では有声破裂子音のバズ区間の口腔内音圧、鼻孔放射音および咽頭壁放射音を分離して収録した。バズ区間の口腔内音圧と鼻孔放射、及び咽頭壁放射の周波数特性を調べたところ、第4章の話者7人のデータではほぼ同様な特性を示している。そのうち、話者Kと話者Sの2人のデータをそれぞれ図6. 11～6. 13に示す。スペクトルはプリアンファシスをかけ512点のケプストラムで求めたものである(以下同)。図6. 11に示したのは口腔内音圧のスペクトルである。口腔内音圧の特徴として周波数200Hz付近にピークが現れている。それは1000Hz以上の成分より23dB程度強くなっている。600Hz～700Hz付近に谷がある。900Hz～1100Hzの範囲で二番目のピークが出現する。200Hzのピークの振幅は96dB程度である。

図6. 12は鼻孔放射のスペクトルである。鼻孔放射では周波数180Hz付近に強いピークが現れている。そのピークは1000Hz以上の成分より20dB程度強くなっている。500Hz付近に谷がある。発声数の半分は600Hzの付近に2番目のピークがある。鼻孔放射のピークの最大値は50dB程度で、口腔内圧に比べて46dB程度低くなっている。

図6. 13は咽頭壁放射のスペクトルである。咽頭壁の放射では180Hz付近のピークは1000Hz以上の成分より20dB程度強くなっている。そのピークの強さは48dBで口腔内圧に比べて48dB程度低い。

### 6. 7. 2 声道壁インピーダンスの推定

この推定では、声道壁インピーダンスは一様な機械抵抗 $B$ 、質量 $M$ 、ステイフネス $K$ の直列接続したものと仮定した。推定方法としては、 $B$ 、 $M$ 、 $K$ について二つのパラメータをある値にして一つの値を広い範囲で変動させ、実音声に合うような声道壁イン

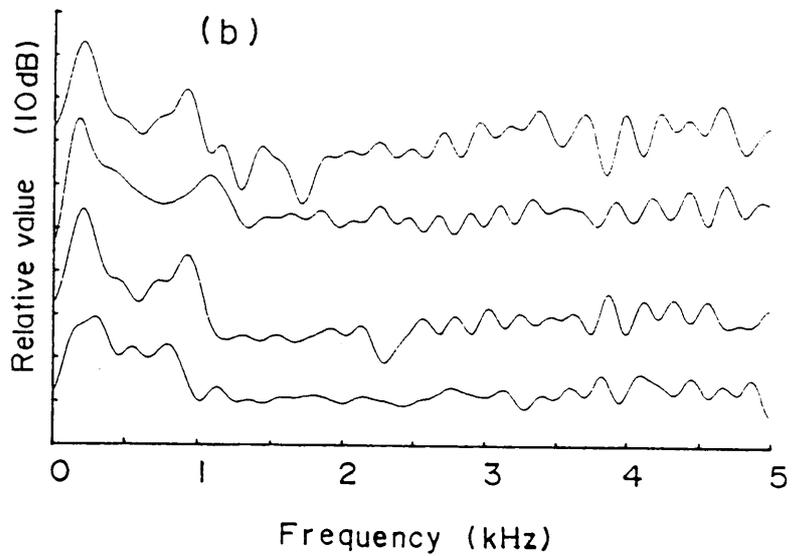
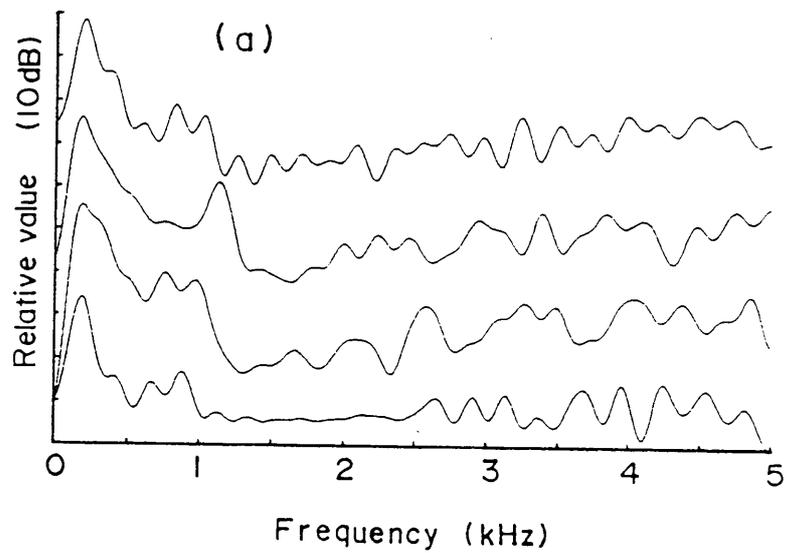


図6. 11 実測した/b/のバズ区間での口腔内音圧のスペクトル  
 (a) 話者Kの/ba/, /bi/, /bu/, /bo/ (上下順)  
 (b) 話者Sの/ba/, /bi/, /bu/, /bo/ (上下順)

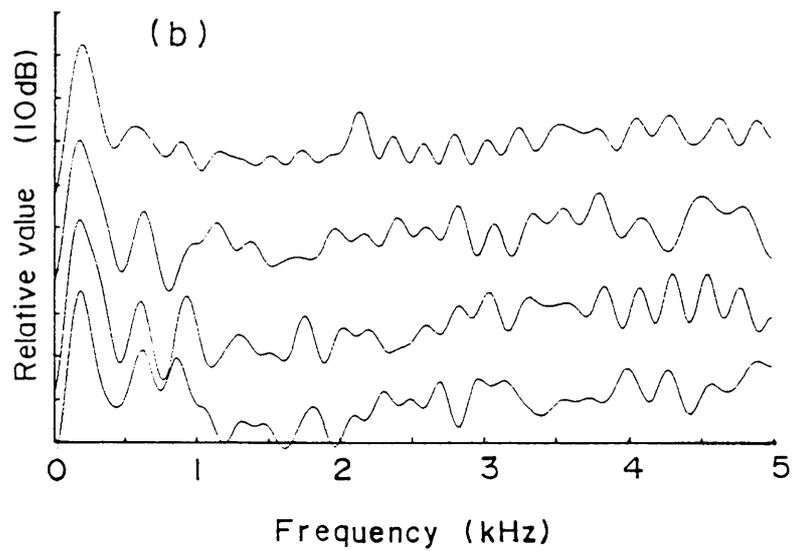
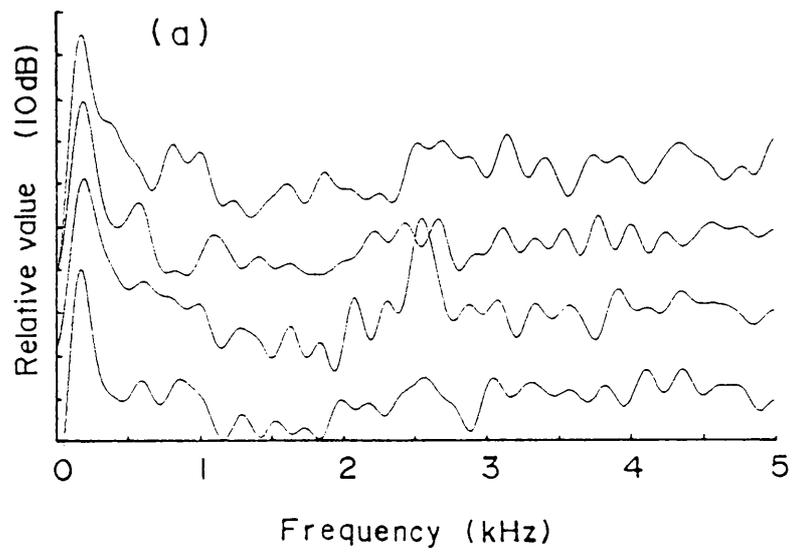


図6. 12 実測した/b/のバズ区間での鼻孔放射音圧のスペクトル  
 (a) 話者Kの/ba/, /bi/, /bu/, /bo/ (上下順)  
 (b) 話者Sの/ba/, /bi/, /bu/, /bo/ (上下順)

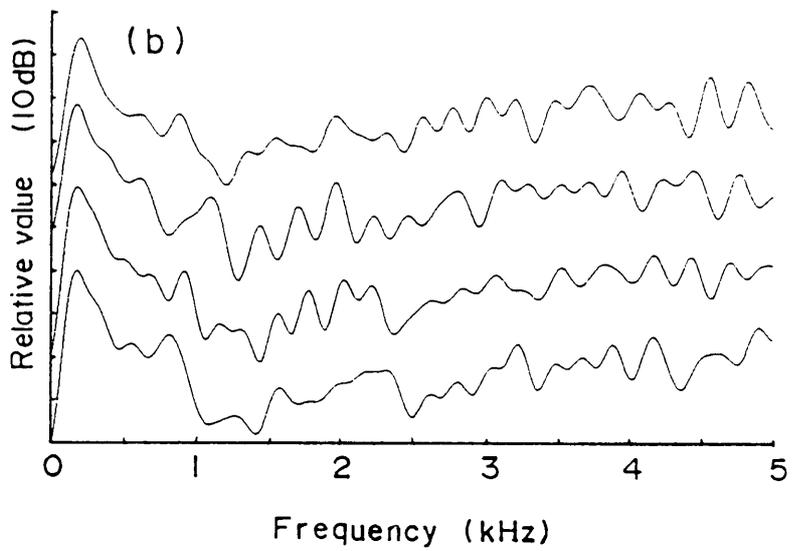
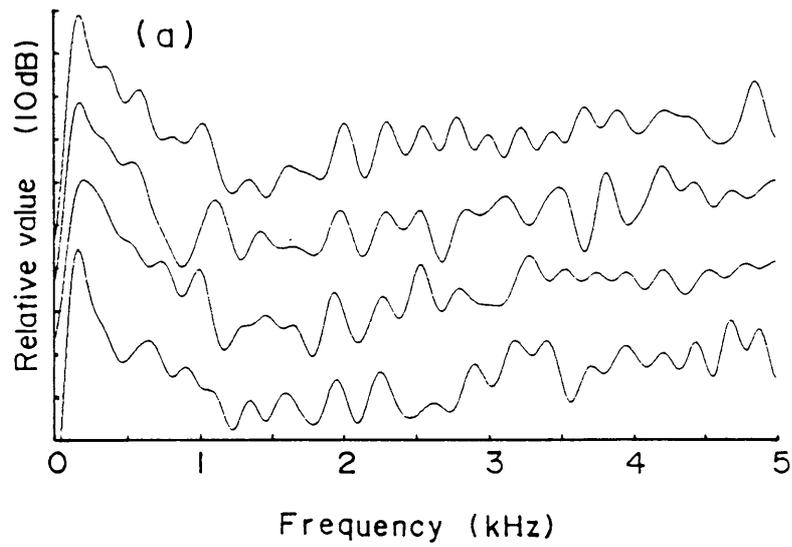


図6. 13 実測した/b/のバズ区間での咽頭壁からの放射音圧のスペクトル  
 (a) 話者Kの/ba/, /bi/, /bu/, /bo/ (上下順)  
 (b) 話者Sの/ba/, /bi/, /bu/, /bo/ (上下順)

ピーダンスを推定した。咽頭壁からの放射を正確にシミュレーションするのは難しい。第3章の咽頭壁の放射と振動加速度に関する分析によれば、咽頭壁の加速度は咽頭壁放射と1.5kHz以下では周波数スペクトルがほぼ同じで、しかも加速度の振動量は放射量に比例している。このことに基づいて、ここでは咽頭壁放射音の代わりに咽頭壁の加速度を使ってシミュレーションする。

推定方法の説明として、次の例を上げる。 $M=1.5\text{g/cm}^2$ 、 $K=120\times 10^3\text{dyne/cm}^3$ 、 $B = (4000+1000n)\text{g/sec}\cdot\text{cm}^2$ にした場合の口腔内圧、鼻孔放射、咽頭壁の加速度のスペクトルをそれぞれ図6. 14～図6. 16に示す。そこで、 $n=1,2,3,\dots,10$ で、右側の

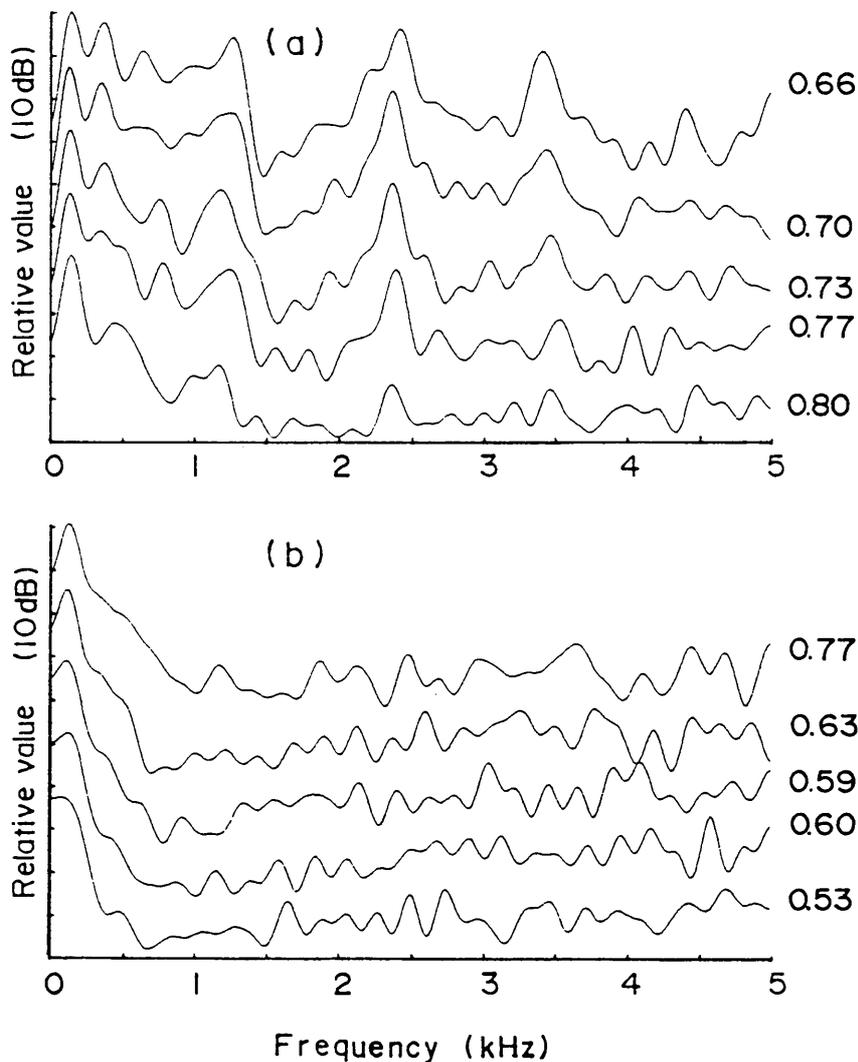


図6. 14 機械抵抗による口腔内圧のスペクトル  
 $M=1.5\text{g/cm}^2$ 、 $K=120\times 10^3\text{dyne/cm}^3$ 、 $B=(4000+1000n)\text{g/sec}\cdot\text{cm}^2$   
 (右側の数字は実音声との相関係数である)  
 (a)  $n=1,2,\dots,5$  (上下順)  
 (b)  $n=6,7,\dots,10$  (上下順)

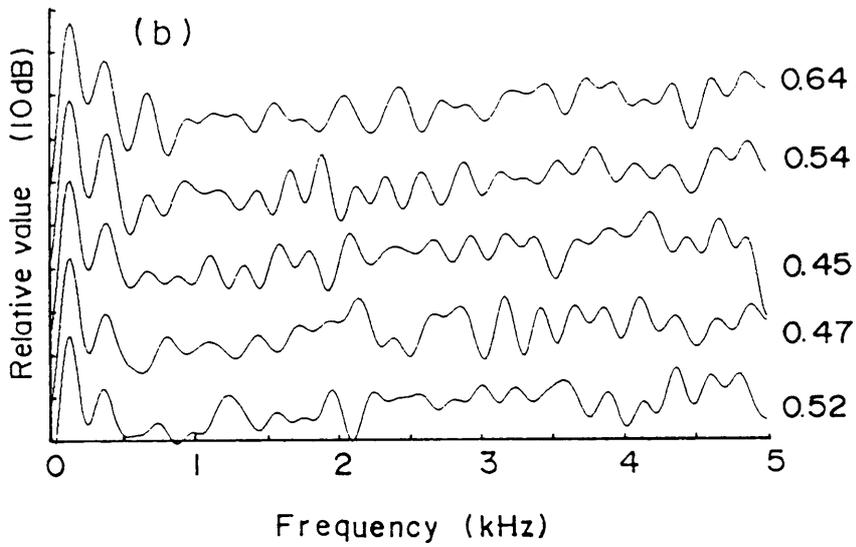
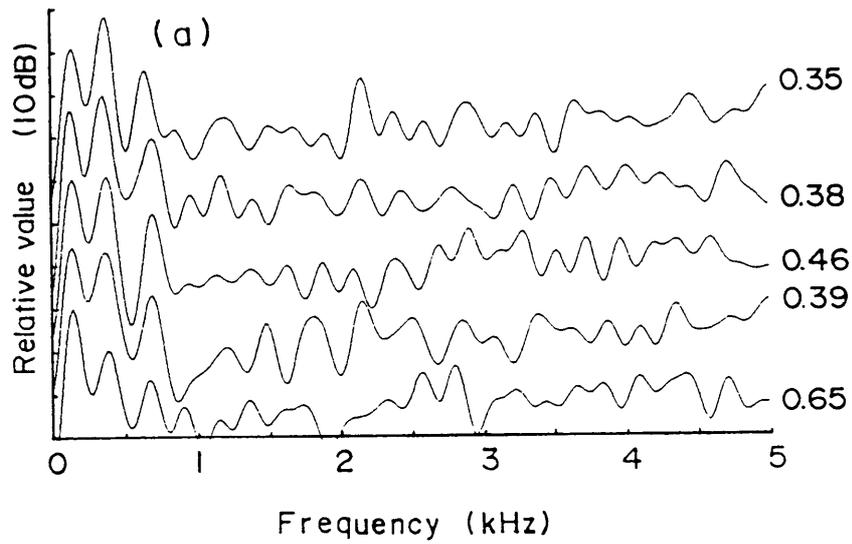


図6. 15 機械抵抗による鼻孔放射音圧のスペクトル  
 $M=1.5\text{g/cm}^2, K=120 \times 10^3 \text{dyne/cm}^2, B=(4000+1000n)\text{g/sec.cm}^2$   
 (右側の数字は実音声との相関係数である)  
 (a)  $n=1, 2, \dots, 5$  (上下順)  
 (b)  $n=6, 7, \dots, 10$  (上下順)

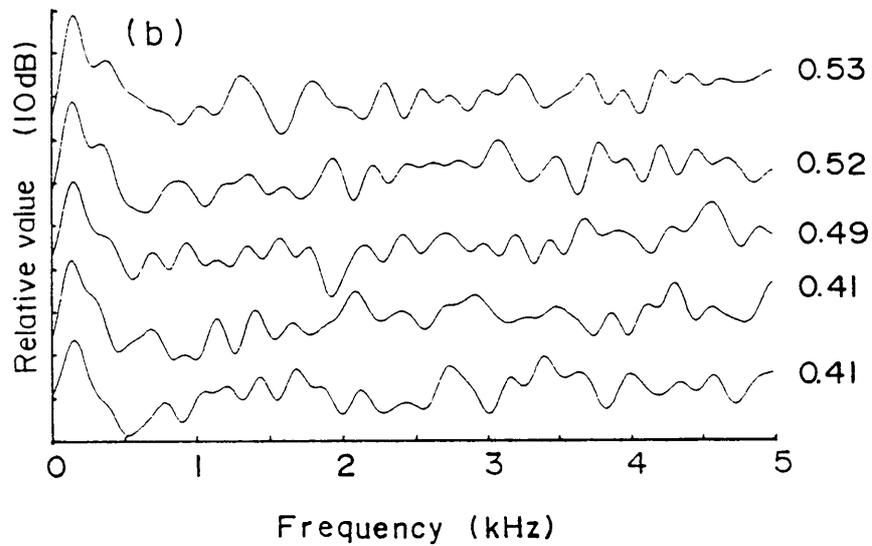
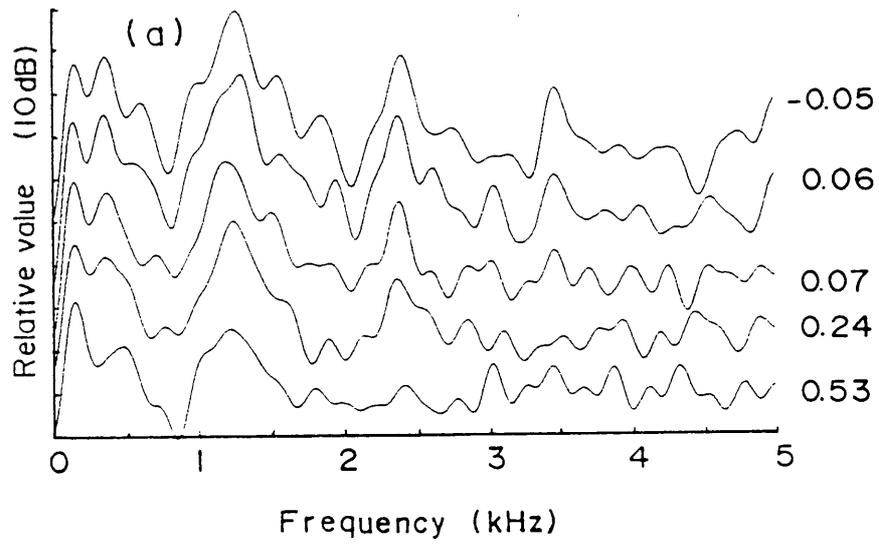


図6. 16 機械抵抗による咽頭壁の加速度のスペクトル  
 $M=1.5\text{g}/\text{cm}^2, K=120 \times 10^3 \text{dyne}/\text{cm}^3, B=(4000+1000n)\text{g}/\text{sec} \cdot \text{cm}^2$   
 (右側の数字は実音声との相関係数である)  
 (a)  $n=1, 2, \dots, 5$  (上下順)  
 (b)  $n=6, 7, \dots, 10$  (上下順)

数字は代表的な実音声（話者Sの/bu/）との相関係数である。実音声に比べて口腔内圧、鼻孔放射および咽頭壁の加速度のいずれも抵抗の小さい場合は高周波成分が高過ぎ、抵抗の大きい場合は高周波が低過ぎる。推定では、合成音と実音声との相関係数が口腔内圧、鼻孔放射、咽頭壁の加速度の場合とも高くなっているとき、つまり合成音の特性と実音声に一番近い時の機械抵抗は推定値となる。この例では、抵抗9000 g/sec·cm<sup>2</sup>である場合の相関係数は口腔内圧では0.80、鼻孔放射では0.65、咽頭壁加速度では0.53である。他の場合と比べて、この場合は相関係数とそれらの和が比較的に大きいので、推定値は9000g/sec·cm<sup>2</sup>となる。

Mを1.1~1.5g/cm<sup>2</sup>、Kを53~120×10<sup>3</sup>dyne/cm<sup>3</sup>の一つの値とし、Bを200~14000g/sec·cm<sup>2</sup>の範囲で変動させたところ、実音声に近いスペクトルを持つ合成音声の機械抵抗の代表的な値を表6.4示す。推定した機械抵抗は8500~9600g/sec·cm<sup>2</sup>となる。

同様な方法で推定した質量Mの代表的な値を表6.5に示す。そこで、Bを4000~9000g/sec·cm<sup>2</sup>、Kを53~120×10<sup>3</sup>dyne/cm<sup>3</sup>の一つの値とし、Mを0.5~2.5g/cm<sup>2</sup>の範囲で変動させたところ、推定した質量は0.6~1.4g/cm<sup>2</sup>となった。

表6.4 各設定値のもとで推定した機械抵抗B

set value	apitimum range of B	estimated value
M=1.5 K= 53.3×10 <sup>3</sup>	O: 7600-9600 N: 9600-12000 P: 9600-12000	9600
M=1.5 K= 9.47×10 <sup>3</sup>	O: 7600-9600 N: 9600-12000 P: 9600-12000	9600
M=1.5 K= 120×10 <sup>3</sup>	O: 8000-9000 N: 9500-12000 P: 9000-10000	9000
M=1.3 K= 120×10 <sup>3</sup>	O: 7500-9000 N: 9000-10000 P: 8000-11500	9000
M=1.1 K= 120×10 <sup>3</sup>	O: 7000-9500 N: 8000-10000 P: 7000-9000	8500

\*O:口腔内圧 N:鼻孔放射 P:咽頭壁の加速度

推定したスティフネスKの代表的な値を表6. 6に示す。そこで、Mを1.1~1.5g/cm<sup>2</sup>、Bを4000~9000g/sec·cm<sup>2</sup>の一つの値にし、Kを13~534×10<sup>3</sup>dyne/cm<sup>3</sup>の範囲で変動させたところ、推定したスティフネスは95~120×10<sup>3</sup>dyne/cm<sup>3</sup>となった。

表6. 4~6. 6を総合的にみると、代表的な値としては、Bは8500~9600g/sec·cm<sup>2</sup>、Mは1.0~1.4g/cm<sup>2</sup>、Kは95~120×10<sup>3</sup>dyne/cm<sup>3</sup>である。声道壁インピーダンスの値をその範囲の値にすれば適切である。ここで、Bを9000g/sec·cm<sup>2</sup>、Mを1.4g/cm<sup>2</sup>、Kを120×10<sup>3</sup>dyne/cm<sup>3</sup>にして合成した有声破裂子音/b/のバズバーのパワースペクトルを図6. 17に示す。合成音と実音声との相関係数は口腔内圧では0.83、鼻孔放射では0.71、咽頭壁の加速度では0.66である。このことより、推定した声道壁インピーダンスを用いて合成した音声の周波数スペクトルは実音声のそれとかなり近いことを示している。

この推定では、1kHz以下の範囲のピークと谷の周波数、および高周波成分の低周波成分に対する相対的な大きさにおいて合成音声を実音声に合うようにした。推定した単位面積あたりの質量とスティフネスは6. 5の実測値とほぼ同じである。6. 5で測定した頬の機械抵抗が周波数に強く依存するが、推定では用いた声道モデルは声道壁のパラメータが周波数に対して変化できないので、推定値が声道壁インピーダンスを全周波数領域で十分表わすのは難しい。推定した機械抵抗は、実測した低周波領域の値との差が大きいが、実測した高周波領域の値に近い。

表6. 5 各設定値のもとで推定した質量M

set value	aptimum range of M	estimated value
B=4000 K= 53.3×10 <sup>3</sup>	O: 0.6-0.7 N: 0.5-0.6 P: 0.5-0.6	0.6
B=6000 K= 94.7×10 <sup>3</sup>	O: 1.0-1.1 N: 0.9-1.0 P: 0.9-1.1	1.0
B=9000 K= 120×10 <sup>3</sup>	O: 1.4-1.6 N: 1.1-1.4 P: 1.4-1.5	1.4
B=9000 K= 140×10 <sup>3</sup>	O: 1.4-1.8 N: 0.8-1.2 P: 1.0-1.5	—

\*O:口腔内圧 N:鼻孔放射 P:咽頭壁の加速度

表6. 6 各設定値のもとで推定したスティフネスK

set value	apitimum range of K	estimated value
M=1.5 B=4000	O: 408- ? $\times 10^3$ N: 490- ? $\times 10^3$ P: 490- ? $\times 10^3$	
M=1.5 B=6000	O: 206-265 $\times 10^3$ N: 333-370 $\times 10^3$ P: 333-369 $\times 10^3$	
M=1.5 B=9000	O: 83-134 $\times 10^3$ N: 83-120 $\times 10^3$ P: 120-179 $\times 10^3$	120 $\times 10^3$
M=1.3 B=9000	O: 64- 95 $\times 10^3$ N: 83-134 $\times 10^3$ P: 95-179 $\times 10^3$	95 $\times 10^3$
M=1.1 B=9000	O: 53- 83 $\times 10^3$ N: 83-120 $\times 10^3$ P: 95-179 $\times 10^3$	

\*O:口腔内圧 N:鼻孔放射 P:咽頭壁の加速度

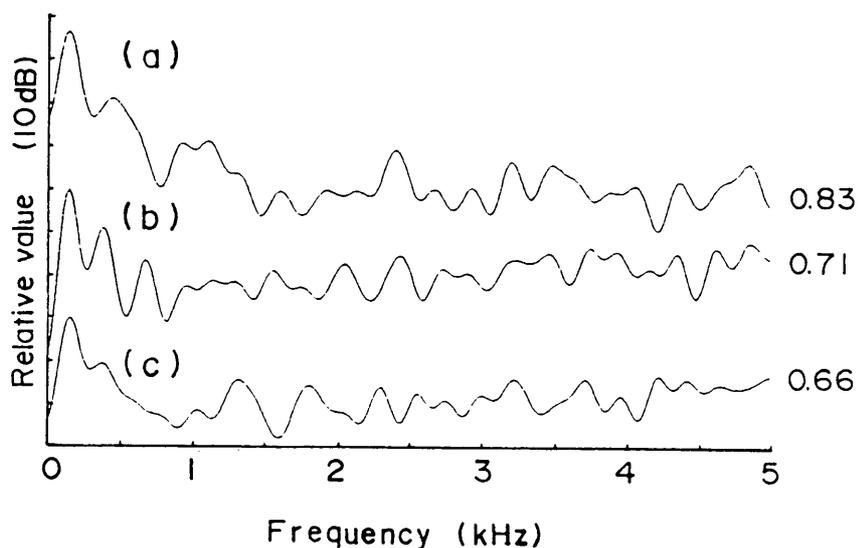


図6. 17 合成した/b/のバズバーのパワースペクトル  
 (右側の数字は実音声との相関係数である)  
 (a)口腔内音圧 (b)鼻孔放射 (c)咽頭壁の加速度

## 6. 8. むすび

音声を発声する時、声道壁の振動と声道壁の放射は声道内音圧によって引き起こされたものである。このことを基づいて新しい頬インピーダンスの測定方法を提案した。なお、予備実験として、プラスチック・ゴム製の管に対して本測定法と機械工学的な方法で測定した値を比較し、ほぼ妥当な値が得られることを確認した。実測した頬インピーダンスに対して簡単な等価回路を当てはめて推定した結果では、単位面積あたりの機械抵抗は高周波領域でかなり大きくなり、一方、単位面積あたりの質量とステイフネスは周波数と共に若干変動するが、ほぼ一定の値である。従来提案された値と比べて本測定法で得られた値は十分に信頼できる。

一方、合成音のパラメータ整合法で声道壁インピーダンスを機械抵抗、質量、ステイフネスの直列接続と仮定して推定した単位面積あたりの質量とステイフネスは、直接測定法で測定した結果とほぼ一致している。推定した機械抵抗は測定した高周波領域の値に近い。

## 第7章 結論

本論文では、防音箱と遮蔽板を用いて発声器官各部位から放射音を分離して測定した。測定結果により、発声器官のパラメータを推定して、声道壁と軟口蓋の振動を考慮した音声生成機構の音響モデルを作成した。

第1章では、音声生成と音声合成に関する研究の流れについて論じ、現状と問題点を考察した。さらに、本研究での測定と推定に関する主な発声器官の働きと、母音の分類と子音の分類について解説した。

第2章では、本研究で用いた音響測定システムについて述べた。まず、防音箱と遮蔽板の遮音特性を測定した。防音箱の透過損失は内側から外側への場合31dB以上、外側から内側への場合30dB以上である。遮音特性を上昇させるために遮蔽板を二枚の板で重ね合わせたものにして、透過損失は20dB程度となった。防音箱内側の口唇放射に対する影響と外壁の鼻孔放射及び咽頭壁放射に対する影響を予備実験で確認した。各部位の放射音を比較するため、各部位のマイクロホンの出力に対して校正を行った。上述の考察より、防音箱と遮蔽板は発声時の口唇放射、鼻孔放射と咽頭壁放射の分離に対して十分な遮音特性を持っている。マイクロホンと加速度ピックアップの周波数特性は測定周波数範囲で十分に満足できる。防音箱の導入による影響は補正できることを確認した。本研究で用いた測定システムは音声の音響測定に十分有効である。

軟口蓋と声道壁の働きを考慮した音響生成モデルを作成するため、本論文では、発声時の発声器官各部位の音圧と声道壁の振動に着目して軟口蓋の働きと声道壁インピーダンスを推定した。軟口蓋の動きを、非鼻音化母音、非鼻音化子音と鼻音化音声の3つの側面から考察した（第3章～第5章）。声道壁インピーダンスを、直接測定法と合成音のパラメータ整合法の2つの方法を用いて推定した（第6章）。

第3章では、非鼻音化母音における音響放射を考察して軟口蓋の振動板モデルを作成した。放射音の測定により、非鼻音化母音においても音が鼻孔や咽頭壁部などから放射され、口唇放射では狭母音の方が広母音より小さいが、鼻孔放射では狭母音の方が広母音より大きいことがわかった。この現象を音響モデルで説明するため、まず鼻腔の特性が実音声に合うように鼻腔モデルにはいくつかの副鼻腔を導入して、鼻腔通路を含む声道モデルを導入した。次に、閉鎖している軟口蓋は口腔内音圧によって振動することに基づいて、軟口蓋を二枚の複合振動板としてモデル化した。このモデルで合成した母音の口唇放射と鼻孔放射の音圧レベルとスペクトルを実音声のそれに比

べた結果から、非鼻音化母音における鼻孔放射は軟口蓋の振動による口腔と鼻腔の音響的結合で説明できることと、軟口蓋の振動板モデルは非鼻音化母音区間における軟口蓋の働きをよく表わせることがわかった。

第4章では、非鼻音化子音特に有声破裂子音の音響放射特性を用いて軟口蓋の動きを考察した。口腔内圧と放射音の波形の観測結果では有声子音だけでなく、無声子音の一部も、鼻孔と咽頭壁から音が放射されていることを示した。咽頭壁からの放射と口腔内圧の交流成分に直接に関連しているが、鼻孔放射音は口腔内圧の交流成分以外の要素にも関連している。そこで、軟口蓋の口腔側から鼻孔放射までの伝達関数の可変要素が軟口蓋だけであることから出発して鼻孔放射と口腔内圧の関係から軟口蓋の状態を推定した。推定結果では、有声子音のバズ区間では軟口蓋の軟らかさが変化しており、この変化には受動的な成分と能動的な成分があることを示した。受動的な変化を定式化して軟口蓋の振動板モデルに導入した。さらに、従来の測定結果に基づいて、咽頭腔容積が能動的に増大する機構を定式化して有声破裂子音の音響モデルを作った。このモデルを用いて合成した有声破裂子音の周波数特性は実音声の偏差範囲に入っている。聴取実験では、合成音は実音声とほぼ同じ結果を示している。

第5章では、鼻音化音声の音響放射を考察し、軟口蓋の開口面積率を推定した。口唇放射と鼻孔放射の考察結果では、鼻子音に母音と非鼻子音を組合わせた音素系列においては、鼻子音区間における鼻孔放射は殆ど変わらないが、母音区間での鼻孔放射量は言語文脈によってかなり変化している。分離された口唇放射と鼻孔放射の周波数特性を考察したところ、鼻音化音声には鼻腔の影響は主に2kHz以下に現れ、強く鼻音化された母音区間では、口唇放射と鼻孔放射は別々の極と零を寄与している。また、口腔内圧と放射音圧を用いて、軟口蓋開口面積の等価インピーダンスと軟口蓋振動の等価インピーダンスの関係を考察し、軟口蓋の等価開口面積率を放射音圧と結び付けた。さらに、それらの関係を用いて軟口蓋の等価開口面積率を推定した。

第6章では、直接測定法と合成音のパラメータ整合法を用いて声道壁インピーダンスを推定した。音声を発声する時の声道壁の振動と声道壁からの音響放射は声道内音圧によって引き起こされたものである。このことに基づいて、本研究では従来とは異なる方法、すなわち音源を口腔の中に入れ、口腔内音圧と頬からの放射音圧および頬の振動加速度を同時に収録して頬インピーダンスを測定する方法を提案した。この方法の有効性を確認するため、プラスチックゴム管の機械インピーダンスを機械工学的手法で直接測定した値と本方法で求めた値を比較したところ10%程度の誤差で一致した。実測した頬インピーダンスでは、単位面積あたりの機械抵抗は高周波領域でかな

り大きくなり、単位面積あたりの質量とステイフネスは周波数と共に若干変動するが、ほぼ一定の値である。従来提案された値と比較検討したところ、本測定法で得られた値は十分に信頼できることがわかった。また、有声破裂子音区間の音響特性を用いて、合成音のパラメータ整合法で声道壁インピーダンスを推定した。推定した質量とステイフネスは直接測定法で測定した結果とほぼ一致している。推定した機械抵抗は測定した高周波領域の値に近い。

放射音と口腔内圧から推定した軟口蓋の動きをまとめて、本論文では、連続音声における軟口蓋の音響モデルを提案した。このモデルは、従来のONとOFFの二状態のスイッチとされた軟口蓋モデルと異なり、非鼻音化音声では軟口蓋は有限の機械インピーダンスを持つ振動板として働き、鼻音化音声では軟口蓋の開口面積が連続的に変化し、非鼻音化音声と鼻音化音声のわたりでは軟口蓋は振動板として働きながら軟口蓋の開口面積による結合が同時に働く。このモデルで合成した有声破裂子音に対して聴取実験を行ったところ、軟口蓋の振動による鼻孔放射音を除いた場合は、実音声と合成音とも30%程度の有声破裂子音は無声子音に間違えたかまたは無声子音と区別できなかった。軟口蓋の振動による鼻孔放射音を加えた場合は合成音の有声破裂子音の誤認識率が実音声に近く5%となった。このことは、軟口蓋の振動板モデルの正確さと重要性を示している。

軟口蓋の等価開口面積率は、鼻音化音声における軟口蓋の開口面積の連続的な変化をモデル化することが可能になった。軟口蓋の等価開口面積率の推定では、いくつかの近似により鼻腔への入力としての軟口蓋の開口に相当する部分と軟口蓋の振動に相当する部分の割合に誤差をもたらす可能性があるが、口唇放射と鼻孔放射に誤差をほとんどもたらさない。こういう意味で、このモデルを用いて実音声のような自然な鼻音化音声を合成することができる。

声道壁のインピーダンスの測定と推定では、声道壁インピーダンスの影響が比較的大きな音声資料を用いて、音声発声に近い状態で測定と推定を行ったので、従来の測定より高い精度の値が得られている。これにより、音声生成モデルの構築について一つの信頼度の高いパラメータを提供したといえる。

本研究は音声生成系の研究としては、人間の発声メカニズムや会話の科学的理解に役立つばかりでなく、感情などの豊富な個人情報を含む音声の合成技術にも寄与する。

以上の考察により、本研究の目的はほぼ達成したといえる。しかし、音声生成過程における発声器官の動的な特性に対する観測と定式化には多くの問題が残されている。本研究に関する発声器官のパラメータの推定と定式化には次のような課題が残されて

いるであろう。第1に、軟口蓋の受動的な変化は定式化されたが能動的な変化については定式化されていない。音声発声時に軟口蓋の形状ばかりでなく硬度が変化している様子を直接に観測することは困難であろうが何らかの方法で観測することが今後の課題の一つである。第2に、鼻音化音声における軟口蓋の等価開口面積率を各部位の放射音圧で推定する方法を提案したが、この方法を用いて多くのデータを分析して軟口蓋の等価開口面積率の変化を定式化し、さらに音声生成モデルに導入する必要がある。第3に、声道壁インピーダンスを機械抵抗、質量とステイフネスの直列として近似したが、声道壁インピーダンスの各パラメータは周波数依存性があり、口腔、鼻腔および咽頭の各部位の壁インピーダンスも必ずしも一定の値ではない可能性があるため、今後さらに精密なモデルを用いて声道壁各部のパラメータを推定する必要がある。

## 付録 A 引張り試験

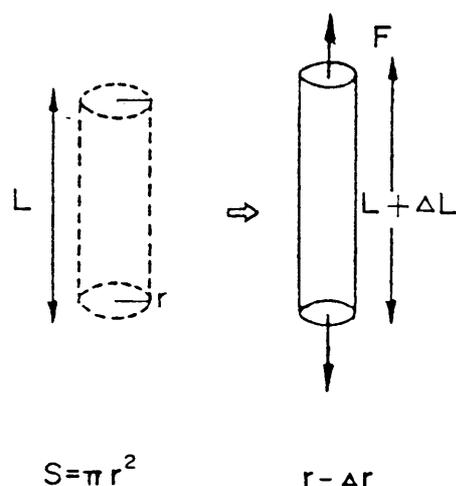
引張り試験の考えとして、物体は外力を加えられると歪みが生じる。この歪みは物体内部に応力を発生させ、この応力が外力とつりあうところで物体の変形は止まる。この時、応力と歪みとの間に一定の関係が成立し、かつ外力を除くと応力も歪みも消失してもとの形にもどってしまう性質がある時、この物体は弾性を有するという。この時微小変形に関してフックの法則が成立し、応力と歪みとの間の関係は線形関係になる。それらの関係を図A. 1に示す。

ここで、被験物のヤング率は $E$ 、長さは $L$ 、面積は $S$ 、単位面積あたりのステイフネスは $K$ とおき、加えた力は $F$ である場合、伸びは $\Delta L$ とすると、それらの関係は

$$E = \frac{F/S}{\Delta L/L} \quad (1)$$

$$K = \frac{F/S}{\Delta L} \quad (2)$$

となる。



図A. 1 引張り試験での力 ( $F$ ) と伸び ( $\Delta L$ ) の関係

## 付録 B 等価質量と等価ステイフネスから静的な値への変換

本論文で提案した測定方法は、被験管（声道）の内部圧力と外部圧力、及びそれらの差による管壁の半径方向の動きを用いて単位面積あたりの等価質量、等価ステイフネス及び等価機械抵抗を推定する方法である。この方法の有効性を確かめるため、材質の軟らかいプラスチック・ゴム製の管（以下被験管と称する）に対して予備実験を行った。予備実験では、まず被験管の面積と秤で測った重さから単位面積あたりの質量  $m'$  を求めた。次に被験管の中心軸に垂直方向で輪を切り出し引張り試験で、管壁の円周方向でのヤング率  $E$  と単位面積あたりのステイフネス  $k'$  を求めた。ここでは、上述した方法を機械工学的な方法、求めた単位面積あたりの質量  $m'$  とステイフネス  $k'$  を静的な値と称する。本測定法では被験管の半径方向の等価質量、等価ステイフネスと等価機械抵抗を測定したが、機械工学的な方法で測定した値と比較するため本測定法で測定した等価機械インピーダンスを静的な値に変換した。静的な機械抵抗が測定されていないので、ここは質量とステイフネスだけ比較を行う。

次に本論文で用いた変換方法について説明する。被験管 A の肉厚  $t_h$  は内半径  $r$  に比較してかなり小さい ( $t_h$  は 4mm,  $r$  は 16mm) ので、薄肉円筒とみられる。被験管 A を測定する時、被験管の内音圧  $P$  は管の外側の音圧より 60dB 程度大きくなっているのので被験管 A は内圧を受ける薄肉円筒<sup>1)</sup> で近似することができる。

一般に、内圧を受ける円筒には、円周応力  $\sigma_s$ 、軸応力  $\sigma_z$ 、および半径応力  $\sigma_r$  が作用する。このうち、半径応力  $\sigma_r$  は、内壁面で絶対値が  $P$ 、外壁面で 0 であるから、 $\sigma_r$  は  $\sigma_s$  と  $\sigma_z$  に対して無視できる。図 B. 1 に示すように、長さ 1 の輪を切り出し、さらにその半分を取り出して、それに働く力の釣合を考える。まず、 $y$  軸の釣合は

$$2\sigma_s t_h = 2 \int_0^{\pi/2} r P \sin \theta d\theta = 2rP \quad (3)$$

$$\therefore \sigma_s = \frac{rP}{t_h} \quad (4)$$

次に、 $z$  軸方向の釣合は

$$\pi r^2 P = 2\pi r t_h \sigma_z$$

$$\therefore \sigma_z = \frac{rP}{2t_h} \quad (5)$$

つぎに、円周ひずみ  $\varepsilon_s$  は、ヤング率を  $E$ 、ポアソン比を  $\nu$  とすると

$$\varepsilon_s = \frac{rP}{Et_h} \left( 1 - \frac{\nu}{2} \right) \quad (6)$$

内圧による半径の増加  $\Delta r$  は

$$\varepsilon_s = \frac{2\pi(r+\Delta r) - 2\pi r}{2\pi r} = \frac{\Delta r}{r} \quad (7)$$

したがって、内圧  $P$  による半径の増加  $\Delta r$  は

$$\Delta r = \frac{(2-\nu)r^2P}{2Et_h} \quad (8)$$

ただし、 $\nu$  はポアソン比で、 $E$  はヤング率です。式(8)により円周方向の伸び  $s$  は

$$s = \frac{2\pi(2-\nu)r^2P}{2Et_h} \quad (9)$$

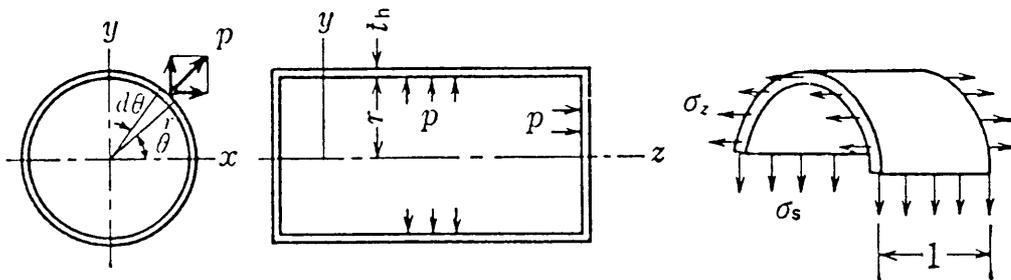
となる。長さ  $2\pi r$  の棒に同じ圧力  $P$  を加えるとその伸び  $x$  は

$$x = 2\pi rP/E \quad (10)$$

となる。式(9)と式(10)より

$$s = \frac{(2-\nu)r}{2t_h} x \quad (11)$$

運動方程式で表わすと管の場合は



図B. 1 内圧を受ける薄肉円筒 (文献 [1] による)

$$m \frac{d^2s}{dt^2} + ks = F(t) \quad (12)$$

となる。棒の場合は

$$m' \frac{d^2x}{dt^2} + k'x = F(t) \quad (13)$$

となる。式(11)を式(12)を代入して各項の係数を式(13)に対応させると

$$m' = \frac{(2-\nu)r}{2t_h} m \quad (14)$$

$$k' = \frac{(2-\nu)r}{2t_h} k \quad (15)$$

となる。ここでは、 $m'$ 、 $k'$ はそれぞれ静的な質量とステイフネスで、 $m$ 、 $k$ はそれぞれ半径方向の等価質量とステイフネスである。

式(14)と(15)によって半径方向の等価質量とステイフネスから静的な値への変換係数  $h$  は

$$h = \frac{(2-\nu)r}{2t_h} \quad (16)$$

となる。被験管 A の場合  $r$  は 16mm、 $t_h$  は 4mm で、ポアソン比  $\nu$  を 0.4 とすると  $h = 3.2$  となる。

被験管 A を厚肉円筒で近似する場合の変換係数も計算したところ、 $h = 3.56$  となった。この値は薄肉円筒の変換係数 ( $h = 3.2$ ) と約 10% の差しかないので、本実験では被験管を薄肉円筒で近似した。

---

また、ゴム管の壁の等価インピーダンスと静的なインピーダンスの比較測定は、小林(1928)<sup>2)</sup>により他の方法で行われてきた。その測定では、管内の音波伝搬特性を用いて管壁の等価インピーダンスを推定した。得られた管壁の等価インピーダンスを静的な値に変換した結果は、機械工学的な方法で得られた静的なインピーダンスとほぼ同じであった。

## 付録 C データベース「MC S D B」について

このデータベースは、実測した音声発声時の発声器官各部位の放射音圧と声道壁振動の信号に対して波形の表示、周波数の分析、音声の出力などの処理を行い、または音声信号の関する情報を提供するもので、ここで「MC S D B」(Multiple Channels Speech DataBase)を称する。MC S D Bに収録されている音声データは、防音箱と遮蔽板を用いて分離された発声器官各部位からの放射音と各部位の振動加速度である。

### 1. 音声資料とデータ構造

MC S D Bで取り扱っている音声資料は次のように4つの部分からなっている。

#### 1. 1 日本語の100音節の音声データ

この音声資料を発声する時には放射音を口唇放射、鼻孔放射と咽頭壁放射に分離して測定した。同時に声道壁の3ヶ所の咽頭壁、頬と鼻孔壁の振動加速度を抽出した。上述の6つの信号は順番にチャンネル1～チャンネル6となっている。発声者は、中井、石田、薩川、本多と千葉の5名である。データはApollo DA3500のD55/dang/data/speakerのディレクトリに保存されている(speaker:話者5名それぞれの名前である)。ディスクの都合上の問題でこのデータのbackとしてテープ1に保存されている。音声資料を表1に示す。

#### 1. 2 子音に関する音声データ

この測定では、口腔内圧、口唇放射、鼻孔放射と咽頭壁放射の音声を収録した。これらの信号は順番にチャンネル1～4となっている。話者は樋口、羽田、千葉、松村、倉本、薩川と長尾の7名である。音声データはApollo DA3500のD55/dang/kdataというディレクトリに保存されている。backデータはテープ2に保存されている。音声資料を表2に示す。

#### 1. 3 鼻音化音声に関する音声データ

この測定は、鼻音付きの音節あるいは短文発声する時の口唇放射、鼻孔放射と頬部の放射を収録してチャンネル1、2、3とした。防音箱を使っていない場合収録した音声はチャンネル4とした。発声者は長尾、松村、薩川、高橋、倉本、羽田、党と陸の8名である。音声データはApollo DA3500のD55/dang/noseに保存されている。backデータはテープ2に保存されている。音声資料を表3に示す。

Table 1 SPEECH MATERIALS

	0	1	2	3	4	5	6	7	8	9
0	a	i	u	e	o	ka	ki	ku	ke	ko
1	sa	si	su	se	so	ta	chi	tsu	te	to
2	na	ni	nu	ne	no	ha	hi	fu	he	ho
3	ma	mi	mu	me	mo	ya	yu	yo	ra	ri
4	ru	re	ro	wa	ga	gi	gu	ge	go	za
5	ji	zu	ze	zo	da	ji	zu	ze	zo	ba
6	bi	bu	be	bo	pa	pi	pu	pe	po	kya
7	kyu	kyo	sha	shu	sho	cha	chu	cho	nya	nyu
8	nyo	hya	hyu	hyo	mya	myu	myo	rya	ryu	ryo
9	gya	gyu	gyo	ja	ju	jo	bya	byu	byo	pya
10	pyu	pyo								

Table 2 SPEECH MATERIALS

(1) ga	(2) gi	(3) gu	(4) ge	(5) go
(6) za	(7) ji	(8) zu	(9) ze	(10) zo
(11) da	(12) ji	(13) zu	(14) ze	(15) zo
(16) ba	(17) bi	(18) bu	(19) be	(20) bo
(21) pa	(22) pi	(23) pu	(24) pe	(24) po
(26) ma	(27) mi	(28) mu	(29) me	(30) mo
(31) na	(32) ni	(33) nu	(34) ne	(35) no
(36) ra	(37) ri	(38) ru	(39) re	(40) ro
(41) zaba	(42) beNbetsu	(43) biNjou		
(44) dobo	(45) buNbutsu	(46) baNba		
(47) daijoubu				

Table 3 SPEECH MATERIALS

(1) aNnai (2) seNsei (3) sekusyoN  
 (4) naNmin (5) mattaku (6) seNdeN  
 (7) zeNzeN (8) suisaNbutsu (9) zeNshiN  
 (10) jiNkeN (11) masumasu  
 (12) kurumaouNtenshimasu  
 (13) madooshimemasu  
 (14) nagoyanisuNdeimasu  
 (15) aN (16) iN (17) uN (18) eN (19) oN  
 (20) maueniarimasu  
 (21) amegafurumasu (22) amegaamai  
 (23) aNa(hole) (24) ame(rain)  
 (25) aNa (26) ame(sweet)  
 (27) maN (28) miN (29) muN  
 (30) meN (31) moN (32) naN  
 (33) niN (34) nuN (35) neN (36) noN  
 (37) maN (38) miN (39) muN  
 (40) meN (41) moN (42) naN  
 (43) niN (44) nuN (45) neN (46) noN

#### 1. 4 合成音の音声データ

合成音の音声データは2つの部分からなっている。1つは、一般の音節の合成音で、もう1つは、声道壁インピーダンスの値を変えたとき合成した/b a/の音節である。前者は、Apollo DN4000のd5/dang/synth/syn\_dataに保存され、ファイル名の構造は、sy + 音節 (2文字) + パラメータのバージョン (2文字) . + 合成プログラムのバージョン (2文字) . + チャンネルの番号 (1文字) (例えば、syma05.35.1) である。合成プログラムバージョンの31と32は非鼻音化音声の場合軟口蓋に振動による口腔と鼻腔は結合するモデルを用いたが、有声破裂子音区間で軟口蓋部に漏れ面積による鼻腔は口腔と結合するモデルを用いた。バージョン33には有声子音区間での軟口蓋の振動モデルを入れた。それらに対応するパラメータバージョンは03である。合

成プログラムのバージョン36は軟口蓋の振動による口腔と鼻腔の結合により鼻孔放射のモデルを用い、無声子音のモデルも入れた。それに対応するパラメータのバージョンは06である。合成プログラムのバージョン35はバージョン36に口腔内圧の直流成分による軟口蓋の軟らかさの受動的な変化を追加したものである。これに対するパラメータのバージョンは05である。チャンネル1～6は順番に口唇放射、鼻孔放射、咽頭壁の加速度、口腔内圧、声門体積流、咽頭腔容積（あるいは口唇放射と鼻孔放射の和）である。

声道壁インピーダンスの推定に関する合成音声はApollo DN3500のd55/dang/buzz\_dataに保存されている。ファイル名の構造はbabuz1?.xx.nである。?は機械抵抗ではr、質量ではm、ステイフネスではkである。このデモンストレーション用の合成音声にはパラメータは次のように設定した。機械インピーダンスの推定では質量は1.5g/cm<sup>2</sup>、ステイフネスは120×10<sup>3</sup>dyne/cm<sup>3</sup>、機械抵抗の値は(4000+1000xx)g/s.cm<sup>2</sup>である。質量の推定では機械抵抗は9000g/s.cm<sup>2</sup>、ステイフネスは120×10<sup>3</sup>dyne/cm<sup>3</sup>、質量は(0.4+0.2xx)g/cm<sup>2</sup>である。ステイフネスの推定では機械抵抗は9000g/s.cm<sup>2</sup>、質量は1.5g/cm<sup>2</sup>、ステイフネスは120×10<sup>3</sup>dyne/cm<sup>3</sup>である。xxは1～10で、nはチャンネル番号で1～6となっている。各チャンネルの信号は上述と同じである。

## 2. MCSDBの機能

MCSDBは収録した音声信号に対して波形の表示、周波数の分析、音声の出力などの方法で処理する。

### 2.1 音声波形の表示

多チャンネルの音声波形を同時に表示できるようになっている。ユーザはチャンネル数を指定することができる。音声波形の時間軸の拡大、縮小または指定区間の表示などができる。各チャンネルの音声振幅も指定できる。

2.2 FFTの計算とスペクトルの表示。FFTの計算区間は画面の左端から512点となる。画面で表示している点数は512点未満の場合、計算区間は表示された点数となる。スペクトルは多チャンネルで同時に表示している。防音箱の影響の補正を加えるかどうかをも選択できる。

2.3 ケプストラムとLPCの次数は表示と指定ができる。画面では、計算した結果だけを表示するか、FFTのスペクトルと一緒に表示するまたは表示しないかを選択できる。

2.4 音圧レベルの表示。指定区間の平均音圧と15msずつの音圧レベルを次

式で計算して出力する。

$$P = 10 \log_{10} \left( \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - X)^2 \right) \quad (1)$$

ここで、Nは切り出し区間のサンプル数で、x(n)はn番目のサンプルで、Xは切り出し区間の平均値である。

2.5 音声の出力。音声の出力は指定チャンネルの音声の出力と、口唇放射と鼻孔放射の和の出力の2つがある。

2.6 音声パラメータの計算。音声のホルマントと帯域幅、異なるチャンネルの信号の相関係数と異なる時刻の信号の相関係数を求めることができる。有声子音バズ区間の口腔内圧の直流成分による口腔内音圧と鼻孔放射音圧の音圧比の変化、すなわち、軟口蓋の軟らかさの受動的な変化のデータを図面に表示できる。鼻音化音声における軟口蓋の等価開口面積率を推定することができる。

### 3. MCSDBの操作

APOLLO DN4000(浮動小数カードあり)のD5/DANG/ORALというディレクトリで「MCSDB」を入力してリタンすると、MCSDBは起動される。画面が次のようになる。

画面 1

```
**** SELECTION ****  
1. real speech  
2. synthesized speech  
3. random  
4. end!
```

選択1は、実音声の測定データに関するデモンストレーションである。選択2は、合成音声についてデモンストレーションである。選択1と選択2ではチャンネル数を指定した場合は必ずチャンネル1から指定チャンネルまで取り扱う。チャンネルの順序をランダムでデモンストレーションする場合、あるいはこのデータベース以外のデータを処理する場合は選択3を選ばばよい。選択4を選ぶとMCSDBは終了する。

#### 3.1 実音声波形のデモンストレーション

上の画面で1を入力すると実音声のデモンストレーションに入る。次は画面の提示にしたがって処理したいチャンネル数を指定すると、画面は次のようになる。

## 画面 2

```
**** SELECTION ****  
1. mono-syllables  
2. stop consonants  
3. nasalized speech  
4. last menu
```

ここで、選択1、2、3はそれぞれ表1、表2、表3に示した音声資料に対して種々の処理を行う。それらの音声データの説明はそれぞれ1.1、1.2、1.3節を参照する。選択4は前画面に戻る。

選択1、2、3のいずれを選択すると、それに対応する音声資料と各チャンネルの情報は画面で表示される。つぎは、話者の選択と音声(番号)の選択である。このあと、指定した各チャンネルの音声波形を同時に画面に表示する。また、画面の右側にはコマンドリストは反転で表示されている。カーソルをコマンドの領域に移動して、マウスの左ボタンを押すとこのコマンドを実行する。

コマンドリストは次のようなコマンドで構成している。

### 3.1.1 表示区間の指定

表示区間を指定する時には、コマンドリストのコマンドを選択せず、マウスの左側のボタンで表示したい区間の開始位置、マウスの右側のボタンで終わり位置を指定すると、指定区間の波形は画面に表示される。

### 3.1.2 表示区間の移動

表示区間の長さを変えずに、表示区間を移動する時、ここで、“forward”と“back”の2つのコマンドを用意している。“forward”というコマンドは表示区間の開始位置を時間増大の方向に移動する。“back”は表示区間の開始位置を時間減少の方向に移動する。移動量は双方ともデータの全区間で画面の25%ずつ移動する。

### 3.1.3 時間軸の拡大と縮小

表示された音声波形の時間軸を拡大あるいは縮小するために、“expand”と“reduce”の2つのコマンドを用意した。“expand”は表示範囲を2倍拡大する。“reduce”は表示範囲を1/2に縮小する。

### 3.1.4 音圧レベルの計算

各チャンネルの音圧を計算する時、マウスでコマンド“sound level”を選んで、3.1.1の方法で計算範囲を指定すると、指定区間における全区間の平均音圧と15msずつの音圧を計算して表示する。

### 3.1.5 周波数スペクトル

各チャンネルの音声信号の周波数スペクトル（例えば、FFT, LPC, ケプストラム）を求めるために、“FFT”というコマンドを用意した（詳細は3.2参照）。

### 3.1.6 チャンネルの指定

上述の3.1.1~3.1.5のコマンドは各チャンネルに同時に作用する。しかし、場合によって、特定のチャンネルを扱う必要がある。このとき、“CHANNEL - n-”を実行すればよい。nはチャンネル番号である。nが“A”であるとき全てのチャンネルの指定となる。マウスの左側のボタンを押すとnは増大していく。

### 3.1.7 振幅の指定

特定チャンネルの振幅の指定では、2つのコマンドを用意した。“up”は振幅を2倍ずつ増大させ、“down”は振幅を1/2ずつ減少させる。

### 3.1.8 音声の出力

指定されたチャンネルの音声をA-Dを通して音を出力する。nが“A”である場合、出力音声は口唇放射と鼻孔放射の和である。

### 3.1.9 LPF

指定したチャンネルにLPFをかける。

### 3.1.10 他の情報

“information”というコマンドを選択すると、次のような情報を得られる。

(1)有声子音バズ区間の口腔内圧の直流成分による口腔内音圧と鼻孔放射音圧の音圧比の変化、すなわち、軟口蓋の軟らかさの受動的な変化を図面に表示する。

(2)鼻音化音声における軟口蓋の等価開口面積率を推定することができる。

その操作は画面の指示に応じて行って頂く。

### 3.1.11 終了

コマンド“quit”を選択すると終了する。

## 3.2 周波数スペクトルを計算する操作

### 3.2.1 スペクトルの計算区間の移動

スペクトルの計算区間を前後移動するため、2つのコマンド“forward”と

“back”を用意している。マウスを一回押すと256点を移動する。

### 3.2.2 波形表示との変換

表示しているスペクトルに対応する波形を調べるとき、コマンド“wave”を選択すると、波形表示に変化する。

### 3.2.3 プリエンファシスの選択

プリエンファシスを選択すると、コマンド“preemph.”を実行すればよい。表示は“1”である時プリエンファシスをかけている。“0”である時かけていない。

### 3.2.4 エンベロープの選択

エンベロープの選択は3つがあり、0～2で提示している。その提示は“0”であるときエンベロープを表示しない。提示が“1”であるときエンベロープだけを表示する。提示が“2”であるときエンベロープとFFTスペクトルを同時に表示する。

### 3.2.5 LPCとケプストラムの選択

コマンド行における“LPC”の提示が“0”であるとき、表示するエンベロープはケプストラムからのもので、提示が“1”であるときエンベロープはLPCからのものとなる。

### 3.2.6 LPCとケプストラムの次数の指定

コマンド“up”と“down”を利用すると、LPCとケプストラムの次数を指定することができる。LPCの提示が“1”であるとき表示している数字はLPCの次数で、“0”である場合その数字はケプストラムの次数となる。

### 3.2.7 実行

3.2.3～3.2.6の指定が終わったら、“execute”を選択すると実行する。

### 3.2.7 他の情報

“information”を選択する場合、画面の指示に応じて操作すると次のような情報を得られる。

(1) 音声のホルマントと帯域幅の情報

(2) 異なるチャンネルと異なる時刻のスペクトルの相関係数

### 3.2.8 終了

”quit”を選択すると、終了する。

## 4. 合成音のデモンストレーション

画面1で選択2を選ぶと、MCSDBは合成音声の処理に入る。

画面 3

```
*** SELECTION ***  
1. ordinary speech  
2. buzz-bar speech
```

選択1、選択2における合成音声のファイル名の構造は1.4節に述べている。  
波形のデモンストレーションと、スペクトルの計算は3.1と3.2節を参照する。

# 参考文献

## 第1章

- 1) H. Dudley, and T. H. Tarnoczy: "The speaking machine of Wolfgang von Kempelen", J. Acoust. Soc. Am., 22, 151-166 (1952).
- 2) 齊藤 収三: "音声認識 [ I ] 音声の分析と特徴パラメータ", 信学誌, 65, 76-83 (1982).
- 3) T. Chiba, M. Kajiyama, The vowel, its nature and structure, (Phonetic Society of Japan, 1958).
- 4) G. Fant, Acoustic theory of speech production, (Mouton, 1970).
- 5) 鈴木久喜 訳, 音声の線形予測, (コロナ社, 1980).
- 6) 板倉文忠: "新しい音声合成方式'PARCOR'", 日経エレクトロニクス, 2.12, 58-75 (1973).
- 7) 板倉, 嵯峨山: "線スペクトル周波数をパラメータとした音声合成法とそのLSI化", 日経エレクトロニクス, 2.2, 128-159 (1981).
- 8) 藤崎 博也: "音声研究の現状と将来", 音響誌, 34, 117-121 (1978).
- 9) 平松啓二, 鎌田力, 角川靖夫: "X線写真による声道パターンの測定", 音響学会講論, 2-2-9 (1964.10).
- 10) 竹内章司, 粕谷英樹, 城戸健一: "鼻音のスペクトルに及ぼす鼻副鼻腔の影響", 音響誌, 33, 163-172 (1977).
- 11) M. Sawashima, H. Hirose: "New laryngoscopic technique by use of fiber optica", J. Acoust. Soc. Am., 43, 168-169 (1968).
- 12) K. Ishizaka, J. L. Flanagan: "Synthesis of voiced sounds from a two-mass model of the vocal cords", Bell System Tech. J., 51, 1233-1268 (1972).
- 13) H. Wakita: "Direct estimation of the vocal tract shape by inverse filtering acoustic speech wave forms", IEEE Trans. Audio, AU-21, 5, 147 (1973).
- 14) 白井克彦, 菅田雅彰: "音声波からの調音パラメータの推定", 信学論, 61, 409-416 (1978).
- 15) 園田頼信, 守 啓祐: "調音時の口唇運動の特徴", 音響学会講論, 103-104 (1988.10).

- 16) 鎗木時彦, 蒼田雅彰:”超音波調音観測システムを用いた舌・唇の閉鎖運動の分析”, 音響学会講論, 105-106 (1988.10).
- 17) 吉田行成, 本多清志, 垣田有紀:”韻律制御における喉頭筋の協調的活動”, 音響学会講論, 195-196 (1991.10).
- 18) 鈴木久喜, 中井孝芳:”声道壁インピーダンスとその音声パラメータへの影響”, 重点領域研究報告, PASL 62-11-1 (1987).
- 19) N. Kamiyama, N. Miki, and N. Nagai:”Measurement of acoustic reflection characteristics of the human cheek”, J. Acoust. Soc. Jpn.(E), 11, 207-214 (1990).
- 20) 大泉充郎, 藤村 靖, 音声科学, (東京大学出版社, 1978) p.12.
- 21) 日本音響学会, 音響用語辞典, (コロナ社, 1988) p.222.
- 22) 古井貞熙, デジタル音声信号処理, (東海大学出版会, 1985) p.10.

## 第2章

- 1) 竹内章司, 粕谷英樹, 城戸健一:”鼻音のスペクトルに及ぼす鼻副鼻腔の影響”, 音響誌, 33, 163-172 (1977).
- 2) 神内教博, 芳井 隆:”口腔出力と鼻孔出力の分離ピックアップ方式による摩擦音, 破擦音の識別”, 音響誌, 39, 310-320 (1983).
- 3) H. Suzuki, T. Nakai and K. Shimizu:”Measurement and analysis of speech sound radiated from vocal tract wall”, ICASSP'86, Tokyo, 1625-1628 (1986).
- 4) 清水啓司:”発声時における声道開口部と壁部からの音声放射に関する研究”, 静大・工, 修士論文 (1985).
- 5) 伊藤 毅, 音響工学, (電気書院, 1977), p.119.
- 6) 西山静男, 池谷和夫, 山口善司, 奥島基良, 音響振動工学, (コロナ社, 1979), p.11.

## 第3章

- 1) K. L. Moll, R. D. Daniloff:”Investigation of timing of velar movements during speech”, J. Acoust. Soc. Am., 50, 678-684 (1971).
- 2) L. Björk, Velopharyngeal Fuction in Connected Speech, (APPELBER BOKTRYCKERI AB, Stockholm), 44-90 (1961).

- 3) L. Björk, and B. Nylén: "The Function of the Soft Palate during Connected Speech", *Acta Chir Scand*, 126, 434-444 (1963).
- 4) L. Björk, B. Nylén: "Studies on Velopharyngeal Closure", *Acta Chir Scand*, 131, 226-229 (1966).
- 5) F. D. Minifie, J. H. Abbs, A. Atrlow and M. Kwaterski: "EMG activity within the pharynx during speech production", *J. Speech Hear Res.* 17, 497, (1974).
- 6) J. L. Flanagan, K. Ishizaka, and K. L. Shipley: "Synthesis of speech from a dynamic model of the vocal cords and vocal tract", *Bell System Techn. J.*, 54, 3, (1975).
- 7) 大泉充郎, 藤村 靖, 音声科学, (東京大学出版社, 1978), p.228.
- 8) G. Fant, *Acoustic theory of speech production*, (Mouton, 1970), p.139.
- 9) 竹内章司, 粕谷英樹, 城戸健一: "鼻音のスペクトルに及ぼす鼻副鼻腔の影響", *音響誌*, 33, 163-172 (1977).
- 10) 齊藤正男, 池田研二, 福本一郎, *医用電子工学概論* (講談社, 1981), p.28.
- 11) D. R. Dickson and W. M. Dickson: *Anatomical and physiological bases of speech*, (Little, Brown and Company, Boston 1982), p.195.
- 12) K. Ishizaka, J. C. French, and J. L. Flanagan, "Direct determination of vocal tract wall impedance", *IEEE Trans. ASSP-23*, 370-373 (1975).
- 13) 本多清志: "音声生成の生理学的背景", *音響誌*, 48, 9-14 (1992).

#### 第4章

- 1) J. F. Lubker: "Transglottal Airflow During Stop Consonant Production", *J. Acoust. Soc. Am.*, 53, 212-215 (1973).
- 2) R. D. Kent and K. L. Moll: "Vocal-tract Characteristics of the Stop Cognates", *J. Acoust. Soc. Am.*, 46, 1549-1555 (1969).
- 3) J. R. Westbury: "Enlargement of the Supraglottal Cavity and its Relation to Stop Consonant Voicing", *J. Acoust. Soc. Am.*, 73, 1322-1336 (1983).
- 4) 後藤友信: "鼻咽頭閉鎖強度とその調節に関する研究", *阪大歯学誌*, 22, 87-106 (1977).
- 5) J. L. Flanagan, K. Ishizaka and K. L. Shipley: "Synthesis of speech from a dynamic model of the vocal cords and vocal tract", *Bell System Techn.*

- J., 54, 485-506 (1975).
- 6) D. R. Dickson and W. M. Dickson, Anatomical and physiological bases of speech, (Little, Brown and Company, Boston 1982).
  - 7) F. D. Minifie, J. H. Abbs, A. Atrlow and M. Kwaterski: "EMG activity within the pharynx during speech production", J. Speech Hear Res. 17, 497-505 (1974).
  - 8) M. Rothenberg: "The Breath-Stream Dynamics of Simple Release-Plosive Production", Bibl. Phonetica 6. (1968).

### 第5章

- 1) 竹内章司, 粕谷英樹, 城戸健一: "鼻音のスペクトルに及ぼす鼻副鼻腔の影響", 音響誌, 33, 163-172 (1977).
- 2) S. Smith: "Vocalization and added nasal resonance", Folia Phoniatica, 3, 165-199 (1951).
- 3) L. Björk, Velopharyngeal Fuction in Connected Speech, (APPELBER BOKTRYCKERI AB, Stockholm), 44-90 (1961).
- 4) A. Bladon, A. Al-Bamerni: "One-stage and two-stage temporal patterns of velar coarticulation", J. Acoust.Soc. Am., Suppl. 1, 72, S104 (1982).
- 5) A. Parush, D. J. Ostry: "Superior laterral pharyngeal wall movements in speech", J. Acoust. Soc. Am., 80, 749-756 (1986).
- 6) K. L. Moll, R. D. Daniloff: "Investigation of timing of velar movements during speech", J. Acoust. Sco. Am, 50, 678-684 (1971).
- 7) G. Fant, Acoustic theory of speech production, (Mouton,1970), p.139.

### 第6章

- 1) K.Ishizaka, J.C.French, and J.L.Flanagan: "Direct determination of vocal tract wall impedance", IEEE Trans. ASSP-23, 370-373 (1975).
- 2) 鈴木誠史: "声道壁のインピーダンスの検討", 音響誌, 34, 149-156 (1978).
- 3) 鈴木久喜, 中井孝芳: "声道壁インピーダンスとその音声パラメータへの影響", 重点領域研究報告, No.PASL 62-11-1 (1987).
- 4) N.Kamiyama, N.Miki and N.Nagai: "Measurement of acoustic reflection cha-

racteristics of the human cheek”, J. Acoust. Soc. Jpn.(E) 11, 207-214 (1990).

- 5) J. L. Flanagan, Speech analysis synthesis and perception, (Springer-Verlag. Berlin. Heidelberg., New York, 1972), p.65.
- 6) H. Suzuki, T. Nakai and K. Shimizu:”Measurement and analysis of speech sound radiated from vocal tract wall”, ICASSP86 Tokyo, 1625-1628 (1986).
- 7) 谷萩隆嗣, デジタル信号処理の理論 3. 推定・適応信号(コロナ社, 1986), p.119.
- 8) 竹園茂男, 材料力学 (朝倉書店,1989), p.143.

### 付録

- 1) 竹園茂男, 材料力学 (朝倉書店,1989), p.143.
- 2) 小林勝一郎:”音響管の特性の研究”, 電子学会雑誌, 446-470 (1928,4)

## 論文目録

### 論文

- (1) 党、中井、鈴木：“連続音声発声時の鼻孔からの放射音圧の文脈による変化の測定”，信学論 J74-A,7,pp.1121-1124 (1991).
- (2) 鈴木、党、中井：“音声発声時の口唇，鼻孔，咽頭部の音と振動の測定および非鼻音化音声における鼻腔口腔間の音響漏えいのシミュレーション”，信学論 J74-A,12, pp.1705-1714 (1991).
- (3) 党、中井、鈴木：“声道内音圧及び頬の振動加速度による頬インピーダンスの測定”，音響誌，条件付き採録.
- (4) 党、中井、鈴木：“有声破裂子音における口腔内音圧と鼻孔からの放射の測定とシミュレーション”，静岡大学電子科学研究科研究報告，印刷中.
- (5) J. Dang, T. Nakai and H. Suzuki:” Investigation of sound radiations from the lips and the nostrils in nasalized speech”, J. Acoust. Jpn. (E), in submitting
- (6) 党、中井、鈴木：“破裂子音区間における口腔内圧および放射音の測定とシミュレーション”，音響誌，投稿中.

### 口頭発表

- (1) 党、中井、鈴木：“発声器官の各部位の振動と放射の測定”，日本音響学会講演論文集 pp.229-230 (1989,10).
- (2) 党、中井、鈴木：“鼻腔と口腔の相互作用及び副鼻腔の影響”，日本音響学会講演論文集 pp.259-260 (1990,3).
- (3) 中井、党、鈴木：“閉鎖軟口蓋の振動による鼻腔口腔結合を考慮した音声生成モデル”，信学技報 Vol.89 No.454 SP89-145 pp.23-30 (1990,3).
- (4) 党、中井、鈴木：“声道内音圧による声道壁の振動及び声道壁からの透過音圧の測定”，日本音響学会講演論文集，pp.273-274 (1990,10).
- (5) 党、中井、鈴木：“発声時の口腔内圧と発声器官各部位からの放射音の測定”，日本音響学会講演論文集 pp.295-296 (1991,3).
- (6) 党、中井、鈴木：“頬の振動加速度及び口腔内外音圧による頬インピーダンスの測定”，信学技報 SP90-103 pp.1-8 (1991).

- (7) 党、中井、鈴木：“連続音声発声時の鼻孔からの放射音圧の測定”，信学技報 SP90-103 pp.9-16 (1991).
- (8) H. Suzuki, T. Nakai, J. Dang and C.Lu:”Speech production model involving subglottal structure and oral-nasal coupling through closed velum”, ICSLP90 Vol.1 pp.437-440 (1990,11).
- (9) 党、中井、鈴木：“有声破裂子音バズ区間の音響特性による声道壁インピーダンスの推定”，日本音響学会講演論文集 pp.215-216 (1991,10).
- (10) 党、中井、鈴木：“連続音声における軟口蓋の音響モデル”，信学技報 SP91-126 pp.31-38 (1992)

## 謝 辞

本研究を進めるにあたり、全面的に御指導と御支援を頂いた鈴木久喜教授に心から深謝致します。また、本論文をまとめるにあたり多くの御指摘をなさってくださった後藤敏幸教授、福田明教授、深林太計志教授、多々良陽一教授に深く感謝致します。

日頃貴重な御討論と御助言を頂いた中井孝芳助手に深く感謝致します。御支援を頂いた石田敬三技官に深く感謝します。

本研究を行うに際し、御援助を頂いた大学院学生陸成湘氏、高橋信之氏、倉本克治氏、学部卒業生本多直樹氏、ならびに音声資料の収録に御協力を頂いた大学院学生河西康仁氏をはじめ当研究室の学生諸氏に感謝致します。日頃からお世話になった本学大学院電子科学研究科ならびに工学部の諸先生や同僚方に感謝の意を表わします。