原 著

日本語 5 母音の音響的同定方法と異常構音への適用について

三木 祥男¹⁾ 新川 拓也²⁾ 野原 幹司³⁾ 奥野健太郎⁴⁾

要 約: 日本語の母音を音響的に検討し, 3つの音響的キューを見出した。第1の音響的キューは, 第2フォルマント周波数と第1フォルマント周波数の比であった。第2の音響的キューは, 第3フォルマント周波数と第2フォルマント周波数の比であった。これらは単一音響管モデルの摂動理論によって理論的に説明される。

第3の音響的キューは, 第2フォルマント周波数と第3フォルマント周波数の中間領域での相対的強さと閾値から求められる。このキューは, マスキング現象に関係づけられる。

84名の音声から3つの音響的キューの規準を定めた。これらの規準により「母音の話者正規化問題」を解決できた。この方法を障害音の音声分析に適用し, 異常構音を判定するのに有効であることを確認できた。

索引用語: 日本語母音, 音響的キューの規準, 母音の話者正規化, 異常構音, マスキング現象

Acoustic Identification of Japanese Vowels and its
Application to Pathological ArticulationYoshio Miki¹⁾, Takuya Niikawa²⁾, Kanji Nohara³⁾ and Kentarou Okuno⁴⁾

Abstract: This study acoustically investigated Japanese vowels, and three acoustic cues were found. The first acoustic cue is the ratio of the second formant frequency to the first formant frequency. The second cue is the ratio of the third formant frequency to the second formant frequency. Both of these cues can be explained theoretically using perturbation theory for a singular acoustic tube model. The third acoustic cue is determined by the relative strength in the intermediate region between the second and third formant frequencies and the threshold value. This cue is related to the auditory masking phenomenon. We determined three criteria for these three acoustic cues based on the voices of 84 subjects.

1)

大阪電気通信大学医療福祉工学部医療福祉工学科²⁾:大阪大学歯学部附属病院顎口腔機能治療部³⁾:大阪大学大学院歯学研究科高次脳口腔機能学講座顎口腔機能治療学教室⁴⁾:

1)

²⁾Department of Biomedical Engineering, Osaka Electro-Communication University:³⁾Division for Oral and Facial Disorders, Osaka University Dental Hospital:⁴⁾Division of Functional Oral Neuroscience, Osaka University Graduate School of Dentistry:

2005年2月28日受稿 2005年9月8日受理

We found that the problem of speaker normalization of vowels in speech perception was solved by taking the third acoustic cue into account. These findings confirmed that this method is appropriate for evaluating pathological articulation and can be applied to acoustic analysis of pathological voices.

Key words: Japanese vowels, criterion for acoustic cue, speaker normalization of vowels, pathological articulation, masking phenomenon

はじめに

前報^{1,2)}で、100 単モーラ明瞭度検査において舌切除者の構音障害音に対する聞き手の聴取能力が音声の聴取回数を重ねると高くなることを明らかにした。このことは、明瞭度検査の都度、症例(舌切除構音障害者)と接したことのない人を被験者にしなければならないことを要求している。したがって構音障害者と接することの多い医療者が被験者になることができない。この制約は医療現場においては非常に問題である。この問題を解消するには、個々の音声の音響的特徴を抽出して、異常構音が正常な構音に比べて、どこがどのように異常であるかを評価することが必要であるとの結論にいたった。

渡辺ら³⁾と土師ら⁴⁾は第 1 ホルマント周波数 (F1) と第 2 ホルマント周波数 (F2) の相対的位置関係から 5 母音の構音異常を客観的に評価することを試みている。これらの研究においては、健常者の母音を正常例とし、それを基準にして対比している。しかしながら音声は個人差が大きいので、どの音声を正常例とするかを定めることは実際には困難である。

ところが、母音の音響特性は個人差が非常に大きいにもかかわらず、聞き手は/a/を/a/として聴き取る。これは「母音の話者正規化問題」⁵⁾としてよく知られた問題で、現在なお未解決である。しかしながら、この問題を解決しない限り、本研究が目的とする「音響特性から構音異常を診断する」ということは不可能である。本研究は、かかる観点から計画され、本報告では日本語 5 母音を音響特性から特定する方法を提案するものである。1 症例に適用した結果、その有用性を検証できたので報告する。

実験 I 健常者の音声の検討 方 法

1. 方法 1 通常の音声

健常者 84 名を被験者として、単母音を日常会話の音程で/a/, /i/, /u/, /e/, /o/の順に約 1 秒間隔で構音し

てもらい、その音声をサンプリング周波数 16 kHz でパソコンに取り込んだ。音声は静かな通常の部屋で録音した。84 名の内訳は成人男性 59 名 (20 代 6 名, 30 代 10 名, 40 代 16 名, 50 代 23 名, 60 代 4 名) および成人女性 25 名 (20 代 11 名, 30 代 5 名, 40 代 5 名, 50 代 3 名, 60 代 1 名) である。

音声分析には (株) アニモ製の SUGI SpeechAnalyzer を用いた。分析パラメータは、フォルマント抽出条件はフレームサイズ: 32 msec, スペクトログラムの設定条件は窓のタイプ: ハミング窓, FFT サイズ: 1024 である。測定項目は音声波形, 第 1・第 2・第 3 フォルマント周波数 (F1, F2, F3), ピッチ (基本周波数 F0), 狭帯域スペクトログラムである。

2. 方法 2 音程を変えて発声したときの音声

被験者のなかの男性 1 名 (被験者 A) に「ドミソド」と音程を変えて、可能な限り低い音から高い音まで何度か発声してもらい、その音声をパソコンに取り込んだ。録音条件と音声分析法は方法 1 と同様である。

結 果

被験者 84 名の F1 と F2 を図 1 にプロットした。図 1 における F1 と F2 の間には正の相関があり、その相関係数を表 1 に示した。

被験者全員に対する F0 と F1 の関係を図 2 に図示した。両者の相関係数を表 2 の 1 段目に示す。相関係数の検定から、どの母音についても危険率 1% で相関ありという結果となった。被験者 A が方法 2 で構音したときの結果を図 3 に示した。その相関係数を表 2 の 2 段目に示した。

フォルマント周波数からピッチの影響を排除するために、次式で定義される換算フォルマント周波数 (V1, V2, V3) を導入する。

$$V1 = (F1/F0) * 100 \dots \dots \dots (1)$$

$$V2 = (F2/F0) * 100 \dots \dots \dots (2)$$

$$V3 = (F3/F0) * 100 \dots \dots \dots (3)$$

換算フォルマント周波数の意味は、ピッチとフォル

図1 F1 & F2 平面図 (対象: 男女 84 人)

図2 F0 & F1 平面図 (対象: 男女 84 人)

マント周波数の間に線形関係を仮定した場合に、F0 が 100 Hz であるときのフォルマント周波数ということになる。

84 名の音声について、V1 と V2 の関係を図 4 にプロットした。同様に V2 と V3 の関係を図 5 に示した。V2, V3 の値が大きくなったために、バラツキが大きく見えるが、相関係数を求めると表 3 のようになり、相関の強さは図 4 も図 5 も差異がない。いずれにおいても 2 つの変数間 (V1 対 V2, V2 対 V3) に強い相関が

認められる。

今、V1 と V2 の勾配を α , V2 と V3 の勾配を β と書くと、(1)～(3)式から

$$\alpha = V2/V1 = F2/F1 \dots \dots \dots (4)$$

$$\beta = V3/V2 = F3/F1 \dots \dots \dots (5)$$

(4) (5)式を使って、 α , β を計算し、図 6 にプロットした。図 6 では母音領域がお互いに重なり合わない領域を「適合領域」として、四角で囲った。これは α ,

図3 F0 & F1 平面図 (対象：成人男性1名)

表1 F1 と F2 の相関係数					
/a/	/i/	/u/	/e/	/o/	
0.886	0.480	0.609	0.560	0.806	

表2 F0 と F1 の相関係数						
出典	対象	/a/	/i/	/u/	/e/	/o/
本報	成人男女	0.846	0.514	0.593	0.711	0.757
本報	男性1名	0.174	0.791	0.770	0.499	0.696
粕谷ら ⁷⁾ *	子供・成人	0.790	0.743	0.569	0.652	0.507

*比較のため文献7から引用

図4 V1 & V2 平面図 (対象：男女84人)

図5 V2 & V3 平面図 (対象: 男女 84 人)

図6 α & β 平面図 (対象: 成人男女 84 人)

表3 相関係数

対象図	/a/	/i/	/u/	/e/	/o/
V1 と V2 図4	0.854	0.818	0.747	0.788	0.885
V2 と V3 図5	0.850	0.914	0.685	0.941	0.825

表4 母音の適合領域

母音	α の値	β の値
/a/	1.3~2.2	1.5~2.5
/i/	5.8~10.8	1.2~1.8
/u/	2.4~4.9	1.65~3.2
/e/	3.0~5.5	1.2~1.45
/o/	1.3~2.2	2.8~4.6

β の値だけから、母音を特定できる領域である。図 6 から「適合領域」とされる α , β の値を表 4 に示した。

考 察

1. ピッチについて

音源フィルタ理論が音源とフィルタをお互いに独立したものとして取り扱ってきたことから、音声学ではピッチとフォルマント周波数の関係については、注目されることが少なかった。しかし実際の音声については、フォルマント周波数がピッチの影響を受け、そのことが母音の認識に影響することが知られている⁹⁾。

粕谷ら⁷⁾は母音について子供と成人男女の F0 と F1 を調べ、子供については両者の間に相関を認め、成人については相関を認めなかった。そこで、このような相関が生じたのは声道の長さの違いに起因し、年齢を媒体として間接的に認められる関係であるとした。

一般に、F0 と F1, F2, F3 の間に正の相関が認められるが^{6,8)}、その理由は次のように説明できる。F0 は主として声帯の膜の長さに依存し⁹⁾、短いほど高くなる。一方、F1, F2, F3 は単一音響管モデルから説明されるように声道の長さが短いほど高くなる。ところが、声帯の膜の長さも声道の長さも子供<成人女性<成人男性の関係にあるため、ピッチ(F0)とフォルマント(F1, F2, F3) は正の相関をもつことになる。

また図 3 は被験者 A が音程を変えて発声したときの F0 と F1 の関係である。/a/ については、バラツキが大きかったが、初めての発声ということで緊張したためと考えられる。このような場合でも F0 と F1 の間で正の相関が認められる。この理由は声帯や声道の長さの違いでは説明できない。しかし人体構造と単一音響管モデルの振動理論からある程度説明できる。

高い声を出すときは、輪状甲状筋を使って声帯を引き伸ばすようにする。このとき声帯は緊張し、喉頭が挙上する¹⁰⁾。これによって声門近辺に声道の狭め(狭窄)が生じると考えられる。ここで声道を単一音響管モデルで近似し、その振動理論¹¹⁾を適用すれば、声門近辺に狭めが生じるとフォルマント F1, F2, F3 が上昇する。すなわち声帯の緊張がフォルマント周波数を高める方向に働く。よって、F0 と F1, F2, F3 は正の相関をもつことになる。

2. 第 1 と第 2 の音響的キューについて

要因がどうであれ、上述のように F0 と F1, F2, F3 の間には正の相関が生じる。そこでフォルマント周波数からピッチの影響を排除するために導入した変数

が、(1)~(3)式で定義した換算フォルマント周波数(V1, V2, V3)であった。その結果、図 4 および図 5 のように、変数間に直線性の高い関係を導くことができた。

図 4 を図 1 と対比してみると、各母音ごとに観測点は原点を通る直線の周りに分布していることがわかる。もしすべての観測点が直線上に乗っているなら、その直線の勾配から母音を同定できることになる。実際には、声道の個人差やその場の緊張などの要因が絡み、理想的に 1 本の直線に乗ることはなかったが、本質的なことは「V1 と V2 の勾配が母音を決定している」という点である。複数の観測点があるとき、その V1 や V2 の値が何であれ、V1 と V2 の勾配が同じなら同じ母音に属するといえる。

では、なぜ観測点が原則的に 1 本の直線上に乗り、その勾配で母音が決まるのであろうか? この点を単一音響管モデルとその振動理論から考察してみたい。単一音響管モデルによれば、管内に生じる第 n フォルマント周波数を F_n と書くと、次式で表わされる。

$$F_n = (2n-1)C/4L \dots\dots\dots (6)$$

n: 整数, C: 音速, L: 管長

さて、声道の一部に局所的狭めが生じると、フォルマント周波数が変動する。その効果を表わすために(6)式に補正係数 K_n を導入すると、

$$F_n = K_n(2n-1)C/4L \dots\dots\dots (7)$$

ここに K_n は狭めの位置だけの関数で、管内に生じる狭めの影響(振動)を表わす。

よって(4)(5)式から次式を得る。

$$\alpha = 3K_2/K_1 \dots\dots\dots (8)$$

$$\beta = 5K_3/3K_2 \dots\dots\dots (9)$$

上記の式は、勾配 α , β が声道の長さには依存せず、狭めの位置、すなわち声道の形状だけに依存していることを表わしている。このことはまさしく母音の産生そのものに依存していることを意味している。

以上のところから、図 4, 5 で観測点が母音ごとに 1 本の直線上に乗るというのは、偶々のことではなく必然性がある。しかし実際の観測点では、いろいろの要因が絡んで完全には 1 本の線上には乗らず、直線の周辺で重なりが生じる。たとえば図 4 では/e/と/u/および/a/と/o/の直線が接近しすぎて重なりが生じる。

もし F1, F2, F3 の値から α , β を計算し、それがある母音(たとえば/a/)の「適合領域」内にあれば、年

齢・性別・個人差・緊張度合いなどに左右されることがなく、その語音はその「適合領域」が表わす母音（たとえば/a/）と特定できる。これこそ「母音の話者正規化問題」への答えとなる。言語中枢は、F2とF1の比（ α ）とF3とF2の比（ β ）さえ知覚できれば、どんな年齢の人のどんな人の声であろうと、母音を正しく特定できることになる。

ここで注目されるのは、図6からわかるように α 、 β の間に相関がないことである。このことが母音の同定をやさしくしている。たとえば、図1のような図形を手掛かりに母音を同定するとすると、まずその図形の認識をしなければならず、言語中枢が母音の持続時間（0.3秒程度）のうちに認識するには複雑すぎる。実際にそのようなメカニズムで行われているとは考えにくい。上述のところによれば、 α 、 β の値がいくらということを知覚できさえすれば、母音を同定できるのであるから、同定処理はきわめて迅速に行える。また α 、 β の値は男女間で偏ることもない。このことは(8)(9)式に声道の長さのようなファクターを含まないことから予想できることである。このようなファクターに左右されないからこそ、「母音の話者正規化問題」を解決できるわけである。

もう一つ特筆すべきことは、図1の図形が原点方向に向いた楕円状になることを説明できる点である。(4)式から $F2 = \alpha F1$ と書き直してみると、 α が一定であれば原点を通る直線を表わす。本来 α は特定の母音ごとに一定の値をとるべきであるが、個人差などにより、多少ばらつく。その結果、直線の周りに観測点が集まって楕円状を形成するのである。

以上の考察から、本報では α を「第1の音響的キュー」、 β を「第2の音響的キュー」と呼ぶことにする。表4の値は α 、 β が適合領域にあるか否かを判断する基準になるものであるから、これらを「第1の音響的キューの規準（criterion）」「第2の音響的キューの規準」と呼ぶことにする。

3. 第3の音響的キューについて

/u/と/e/および/a/と/o/を音響特性面から分離することは、粕谷ら⁷⁾や鈴木¹²⁾が指摘するように非常に難しい。

しかし音声の知覚には特別な特徴がある。その一つは遮蔽（マスキング）と呼ばれる現象である。ある周波数成分の振幅が他の周波数成分のそれよりも大幅に低いと、その周波数成分は聴こえなくなるという¹³⁾。

図7に被験者のなかのある男性（被験者B）の母音の音声波形と狭帯域スペクトログラムとスペクトルを

示した。これは成人男性の典型的パターンである。スペクトル図には縦軸に周波数、横軸に振幅の相対的強さをとってある。相対的強さは最大ピーク値を基準に、その相対値で表わしてある。また周波数の測定位置は音声波形上に縦線で示してある。スペクトル図には、線形予測分析から求めた包絡線が描かれている。本報告で採用したフォルマント周波数は包絡線から求めた値である。またスペクトログラムについては、母音は発音中の音声波形変化が少ないので周波数分解能の高い狭帯域スペクトログラムを表示した。

図7のスペクトルとスペクトログラムを見ると、母音ごとに特有の形をしている。この図形を特徴づける第1の特徴は、包絡線の第1ピーク（F1）と第2ピーク（F2）の間隔である。それは前述の α （ $=F2/F1$ ）の値によって表わされる。第2の特徴は第2ピーク（F2）と第3ピーク（F3）の間隔で、これは β （ $=F3/F2$ ）の値で表わされる。第3の特徴は、/u/と/e/および/a/と/o/を分離できる特徴であろう。スペクトログラム上で/u/と/e/のパターンを比べると、F2とF3の間の谷間の濃淡がお互いに逆であることがわかる。また/a/と/o/でも、同じことがいえる。

/u/と/o/では、F2とF3の間の谷間の相対的強さが低い。/a/と/e/では高い。そこで第2ピーク（F2）と第3ピーク（F3）の間の領域で、その相対的強さが別途定める閾値以上である領域の割合 γ を「第3の音響的キュー」と定義する。

そうすると図7から、① $\gamma > 0.5$ ならば、母音は/a, i, e/のいずれかである。つまり聴取音は/a, i, e/のいずれかに聴こえる。

② $\gamma < 0.5$ ならば、母音は/u, o/のいずれかである。

この判別式を「第3の音響的キューの規準」と呼ぶことにする。

ここでマスキング現象によって、/a, i, e/の音はF2とF3の中間域の周波数成分が聴こえ、/u, o/の音では聴き取りにくいという差が生じる。すなわち/u/と/e/あるいは/a/と/o/を明確に区分できる。

「第3の音響的キュー」というのは、聴覚印象的に「その母音らしさ」を強調する仕掛けと考えられる。ただ閾値の大きさについては、個人差が大きいかも知れない。このようなマスキング現象の存在は、聴覚の欠陥を意味するものではなくて、逆に人が言語を確実に聴き分けるために獲得した特別の能力と考えられる。

この規準の妥当性を検証するために、閾値を-35 dBにとって、スペクトログラムとスペクトル図から目視で84名の音声について調べた。その結果「第3の

図7 ある成人男性の母音＜音声波形と狭帯域スペクトログラムとスペクトル図＞

図8 α & β 平面図 (対象：成人男女 84 人)
第3の音響的キューの規準から外れている観測点の分布

図9 母音同定プロセス

音響的キューの規準」が大部分の人で成立したが、それに合わない人もいた。そこで、そのような人だけを抜き出して図8にプロットした。/a/で9点、/i/で1点、/e/で1点あった。これらの観測点は上記の規準が適用できない例外点である。しかしこれらの例外点はすべて「適合領域」にあった。したがって、これらの観測点は、「第3の音響的キューの規準」から外れるが、どの母音に属するかを特定できる。

逆に適合領域から外れている観測点もある。これらの適合領域の外にある観測点は、すべてこの規準によって/u/と/e/および/a/と/o/のいずれであるかを判別できた。たとえば/a/と/o/の適合領域の中間域では、/a/と/o/の観測点が入り混じっている。しかし/a/の観測点では、いずれも γ が0.5以上であった。逆に、/o/の観測点では、いずれも γ が0.5未満であった。すなわち適合領域の中間域では、この規準による判定で完全に識別できた。このことは/u/と/e/の間でも成立した。

このように α 、 β の値ではどの母音に属するかを特定できない場合でも、さらに第3の音響的キューの規準によって、すべてを完全に特定できる。上記の母音同定プロセスを図9にまとめた。

4. 母音の話者正規化問題について

先行研究によってフォルマントF1とF2が母音を特定する主要なキューとなることが知られている⁵⁾。日本語の母音については、いくつか報告がある^{7,14)}。

図1の特徴は、①母音ごとに特定の領域に楕円状に分布していること、②その楕円は長軸の方向がおおよそ原点方向に向いていること、③男女の観測点が別々の塊となって分布していること、④/a/と/o/および/u/と/e/の領域に重なりが生じていることである。

この特徴は図1だけの特殊なものではなく、日本語の母音^{7,14)}ばかりでなく、英語の母音¹⁵⁾でも認められている。しかし、なぜこのような特徴ある分布を示すのかということを説明した報告は、今までない。上記の考察「第1と第2の音響的キューについて」で述べたことは一つの試みである。

「母音の話者正規化問題」には多くの研究がなされてきた。本報のようにフォルマント周波数の比に着目して、その対数値を論じるものは“formant ratio theory” (Miller⁸⁾など)と呼ばれ多大な影響を残したが、100%の認識率が得られず、結局成功していない。

従来手法では、図1の/a/と/o/および/u/と/e/の領域の重なりを分離する術がなかったからである。そのことが「母音の話者正規化問題」の解決を困難にしていた。本報では「第3の音響的キューの規準」を導入することで、その解決に成功した。

実験II 構音障害者の音声方法

1名の症例を被験者として、実験Iの方法1に従って、その音声を録音し音声分析を行った。症例は前報^{1,2)}と同一の男性で舌垂全摘手術を受けた構音障害者である。

症例は2種類の舌接触補助床を装着している。1つは硬口蓋と接するものでPAPと略称する。もう1つは口腔底に接するものでLAPと略称する。通常は両者を同時に装着しているが、本実験では非装着と装着の場合に分けて測定した。

結果と考察

図10にPAP非装着時の音響データを示した。健常

図10 構音障害者の母音（PAP 非装着）＜音声波形と狭帯域スペクトログラムとスペクトル図＞

図11 α & β 平面図（対象：構音障害者）

PAP および LAP 装着の有無による違い

者のデータ（図7）と対比すると、かなりの違いがある。図11に症例の α , β の関係図を示した。①非装着, ②PAP単独, ③LAP単独, ④両者併用の4ケースについてプロットした。適合領域から外れたところにあるのは, /i/のすべてと併用系の/o/であった。後者は「第3の音響的キューの規準」を満たしているので、正しく聴取される。「第3の音響的キューの規準」に適合しない観測点は, /i/のPAP単独とLAP単独であった。しかし/i/の非装着と併用系でも、閾値を-30 dBにとると、「規準」から外れた。冒頭で構音障害者との接触が、聞き手の異常構音に対する聴取能力を高めることを述べたが、/i/の聴取能力が向上するということは閾値が低くなることを意味するのかもしれない。図11から判断すると、いずれの/i/も/e/とも/u/とも/i/ともいえない音になっていることがわかる。前報^{1,2)}の明瞭度検査から予測されるように、症例では他の母音に比べると/i/が他の音に異聴される比率が高く、また/i/を後続母音とする「イ列音」の明瞭度が特に低い。これは図11の結果とよく符合する。一番/i/の適合領域に近いのはPAPとLAPの併用系であった。舌接触補助床の効果は、それなりに認められる。この結果は症例の日常の実情とよく一致する。

結 語

母音は、語音の基本になる一番重要な部分であるにもかかわらず、多くの先行研究によっても音響特性から母音を同定することができなかった。本研究では、単一音響管モデルの摂動理論と聴覚的マスキング現象への考察から、母音を特定するための3つの音響的キューを見出した。そして健常者84名のデータから、3つの音響的キューの規準を明らかにした。特に第3の音響的キューを見出したことで、「母音の話者正規化問題」の解決が容易になった。得られた規準を舌切除者の異常構音に適用し、実情と合致することを確認した。これによって、構音障害者の母音について、音響特性から診断する方法に目処が立った。

謝辞 文献のご提供とご所見を頂いた東北文化学園大学科学技術学部応用情報工学科 鈴木久喜先生に感謝いたします。

文 献

- 1) 三木祥男, 野原幹司, 奥野健太郎: 舌切除患者と聞き手の間の会話頻度が発語明瞭度に及ぼす影響について. 音声言語医学 47: 25-33, 2006.
- 2) 三木祥男, 荒木千佳, 野原幹司, 他: 舌切除患者の単モーラ発語明瞭度検査における被験者の聴取能力に関する検討. 音声言語医学 47: 143-154, 2006.
- 3) 渡辺 宏, 進武 幹, 松尾浩一, 他: 舌可動部半側切除後の経時的構音変化. 耳鼻, 31: 375-381, 1985.
- 4) 土師知行, 楠 敬蔵, 岸本誠司: 舌・口腔底腫瘍切除後の母音のホルマントについて. 音声言語医学, 35: 261-265, 1994.
- 5) レイ・D・ケント, チャールズ・リード: 音声の音響分析 (荒井隆行, 菅原 勉監訳), 海文堂出版, 東京, 108-109 頁, 2002.
- 6) Fujisaki H and Kawashima T: The roles of pitch and higher formants in the perception of vowels. IEEE Trans Audio Electroacoust, AV-16: 73-77, 1968.
- 7) 粕谷英樹, 鈴木久喜, 城戸健一: 年齢, 性別による日本語5母音のピッチ周波数とホルマント周波数の変化. 日本音響学会誌, 24: 355-364, 1968.
- 8) Miller JD: Auditory perceptual interpretation of the vowel. J Acoust Soc Am, 85(5): 2114-2134, 1989.
- 9) Titze I: Physiologic and acoustic differences between male and female voices. J Acoust Soc Am, 85: 1699-1707, 1989.
- 10) GJ Borden, KS Harris: 言葉の科学入門 (廣瀬 肇訳), MRC メディカルリサーチセンター, 東京, 85-86 頁, 1994.
- 11) レイ・D・ケント, チャールズ・リード: 音声の音響分析 (荒井隆行, 菅原 勉監訳), 海文堂出版, 東京, 31 頁, 2002.
- 12) 鈴木久喜: 線形判別関数を用いた母音識別. 電気通信学会雑誌, 47: 316-324, 1964.
- 13) P・ラディフォギッド: 音響音声学入門 (佐久間章訳), 大修館書店, 東京, 95 頁, 1997.
- 14) 藤崎博也: 音声認識の諸問題. 日本音響学会誌, 28: 33-41, 1972.
- 15) Peterson GE and Barney HE: Control methods used in a study of vowels. J Acoust Soc Am, 24: 175-184, 1952.

別刷請求先: