

解 説

音声研究のすすめと声道形推定の問題*

淵 一 博**

第1部 音声研究のすすめ

— 音声研究の領域と方法 —

1. 音声研究のすすめ

音声研究の歴史は古い。多くの蓄積がある。この十年にもいろいろな展開があった。だが解決しなければならない問題はまだまだ数多くある。

それには難問が多い。先人の残した難問を今からの人にすすめるのはどうであろうか。しかし、残されているのはつねに難問であろう。とともに、時代は新しい武器を提供してくれてもいる。そして新しい展開を要請する。

道具の進歩と研究の進歩が直接に対応するわけではない。しかし間接的には強い相関がある。それをつなぐのは新しい世代の役割である。

新しい力が欲しいのは音声研究だけではない。情報処理技術の研究全体がいま大きな転換期に来ていると思われる。それは計算機技術が一つの成熟期にさしかかっていることと関係がある。計算機技術プロパーの中にもこれからの課題があるのはもちろんである。だが情報処理技術全体から考えると、広い範囲での新しい展開が必要になってきているのである。

言おうとしているのは「自然情報」の処理である。画像や音声、自然言語などの我々が「自然」ともっている情報処理機構の解明である。この分野は決して目新しいものではない。過去にいろんな経緯があったことを知っている人も少なくないだろう。しかし、新しい目と新しい手でこれに再挑戦すべき時期が来ていると考えられる。しかも、自然情報処理は、単に計算機の一応用課題なのではない。これは、将来の計算機技術そのものの新しい展開と深く関連するものなの

である。

かつての苦闘の結果、絶望とタブーが世の中に広まっていたことがあった。それはまだ十年にもならない。その余韻はいまでも残っている。しかし興味深いのは、そのような時期に次の時代への芽が育ちつつあったことである。音声研究のこの十年ほどの展開はそのころに用意されたのである。

ここで私事を述べることを許していただきたい。私が音声研究にたずさわることになったのは、6,7年前である。それまでは一介の計算機屋であった(いまでもその端くれなのである)。そのころはまだタブーの雰囲気が強かった。

音声研究に参加したのは、機会あつてのことではあるが、タブーへの挑戦の気もあつて、自らその道を選択したのである。勉強を始めてから知ったことは、そのような時期でも、実は先人たちや(その頃の)若い人たちによって新しい芽が準備されつつあったということである。それからの数年、先人たちに教えられ若い人たちに学んだ時期は、実に楽しかった。そのような雰囲気だけでなく、音声そのものの魅力に深くとりつかれてしまった。そしていまでもそれから離れがたいのである。

いま音声研究はふたたび盛んになってきている。実用化の面を見ても、音声出力システムの実用化があり、また「音声認識装置」の商品化が伝えられたりする。実用の面は大事なことであり、努力すべきことではある。しかし一方、音声の場合、その「本質」の追求がまだ(ますます)必要なのである。

「技術革新の停滞」というような絶望感がいままた世の中で流行しているが、若い人たちに道を拓かせる条件を用意するのは、古い人たちの責務でもあろう。

2. 音声研究の領域

音声は、いうまでもないが、人が音を用いて言語情報を伝達するシステムであり、その音は、発声器官に

* Invitation to the Speech Research, and the Problem of Estimation of the Vocal Tract Shape from Speech Wave by Kazuhiro FUCHI (Electrotechnical Laboratory).

** 電子技術総合研究所

よってつくり出され聴覚器官によって受けとられる。音自身は物理現象であるが、それは記号のキャリアとして使われている。そこで音声現象には多くの側面があり、いろんな方面からのアプローチがなされる。

言語現象として言語学から研究がなされ、発声・聴覚器官については生理学的に研究される。「聴え」については感覚心理学的研究が行われる。音自身としては音響物理学に属するだろうが、信号としては通信工学の対象である。情報処理機構としては、一方では神経生理学的な研究が考えられ、また、情報処理工学の分野でもあろう。

音声は非常に学際的な分野であり、また音声を中心に考えれば、そこに一つの総合科学が想定されることになる。音声をプロパーに扱うのは「音声学」であるが、伝統的な音声学は、(広義の)言語学の一部門であった。そこでもっと総合的なものとして「音声学」が提唱されている。

音声の工学的利用としては、古くから伝送があり、電話は確立された伝統ある分野である。電話の帯域設定や品質保証の裏には龐大な研究の蓄積がある。さらに「帯域圧縮伝送」(ボコーダ=音声分析再合成システム)を考えると、音声の本性にもっと深く立ち入った研究が必要になる。

音声応答のためには、音声の編集・出力合成が必要であり、入力側では「音声認識」もほしくなる。「音声タイプライタ」は夢の商品であろう。

音声言語現象の一部でもあることを考えれば、自然言語処理との結びつきが必要になってくるだろう。音声応用は言語利用と結びついて意味をますとともに音声自動認識自体が「言語認識」と深くからみあっていると考えられている。

通常環境での音声利用だけでなく、特殊環境(海底など)における音声通信のための補助手段、あるいは身障者の音声的会話をたずける手段などの大きなテーマもある。

計算機と結びついて、音声利用は電話以外にも大きく広がることが実用的にも期待され、一部は実現しつつはあるが、多くは今後に残されている。

3. 音声のとらえ方

3.1 調音の観点から

音声の音(言語音)は、発声器官でつくられる。この音は声帯などで発生し、声道などで特性づけられ、外部に放射される。このような音声器官(図-1.1)を

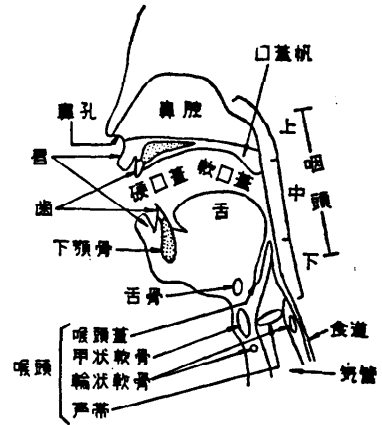


図-1.1 音声器官(正中断面図)

制御して音をととのえることを、調音といっている。そこで、この調音を記述することによって音声を記述することができる。

ところで、音声は単なる音ではなく言語音であり、記号をになうものとして、何か単一的なものから構成されていると考えられる。その単一的なものを音素と呼び、言語音として特有と考えられる性質を音韻性という。音韻が現象したものが音声であり、音韻の単位を音素と考えるわけである(この音素は、「発音記号」で表わされているものと考えておいていただければよい。音素については議論が多いが、ここでは素朴に考えることにする)。

ここでいう音声記述は、音素を調音と関連づけることである。これは古くから行われ、伝統的音声学の主要部分(調音音声学)をなしていた。

音素(発音記号)には沢山の種類があるが、それらはまたある体系をなしていると考えられる。表-1.1と表-1.2(次頁参照)はその例である。ここでは調音を基準とした整理が行われている。

ところで、音声学における調音の記述は、主として「内観」によっている。心理現象と違って、この内観的

表-1.1 母音の分類

調音点 唇の開き	前舌	中舌	後舌
狭	i - y	ɨ - u	ɯ - u
半狭	e - ø	ø	ɤ - o
		ə	
半広	ɛ - œ	ɜ	ʌ - ɔ
	æ	ɐ	
広	a		ɑ - ɒ

表-1.2 子音の分類

調音方式	調音位置						
	唇音	歯音	硬口蓋音	軟口蓋音	口蓋垂音	声門音	
破裂音	p, b	t, d		k, g			?
鼻音	m	n		ŋ		N	
側音		l					
ふるえ音	r						
はじき音						R	
摩擦音	φ, β	s, z					
	f, v	θ, ð	j, ʒ				h
半母音	w						

観察にはかなりの客観性が認められるが、定性的な表現にとどまらざるをえない。もっと客観的な定量的な観察はできないだろうか。

そのために、種々の装置が考案されている。その中でももっとも精密なのはX線写真によるものであろう。

X線による声道の計測の例は第2部に示した。これは Fant¹⁾ によるものであるが、このような研究の先駆が戦前の我国にあった (Chiba-Kajiyama²⁾) ことは特筆すべきことである。

X線による計測データは貴重なものであるが、問題は残る。たとえば、舌の前後、高低といった簡潔な量をうることが意外に難しい (それを何で定義すればよいか)。そして、表-1.1 のような (また図-2.10 のような) きれいな配置を定量的な処理からはなかなか導けないのである。その結果、音韻の特性づけは、調音量だけをもとにしては難しいという見解³⁾もある。

X線による計測は言うは易く、実際は大変な仕事である。被曝の問題がある。撮影条件が難しく、写真からのデータ抽出も難しい。1フレームでも大変なのである。一方欲をいえば、子音では (母音でも) 調音のダイナミクスまで知りたい。

最近、X線マイクロビームによる観察装置が開発されて (東大医)、それによる成果が期待されている (余談であるが、この装置は、音声研究以外、X線トモグラフィの新方式としても注目されている)。

そういう装置ができてX線計測はなお大変な仕事であろうが、調音の実態が本当に明らかになるのはそのような努力によるものであろう。

3.2 聴覚の観点から

人は話し出す前から音声を聞きわけ、聴くことによって学ぶ。聴覚の方が音声にとって根源的かとも思われる。しかし、この場合は記述のよりどころが明確でない。何を基準にすればよいのだろうか。

聴覚器官の生理学的研究が行われ、その上層の神経の働きまで調べられている。その結果、神経回路網の

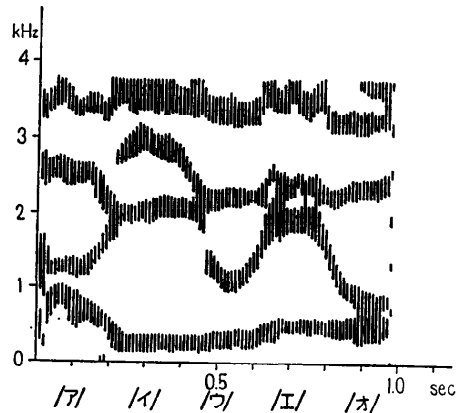


図-1.2 ホルマントのパターン

ある層のところで、周波数分析が完成されるということが明らかになった。これも大変な研究である。

一方、聴取という感覚心理学的実験によっても、聴えにとってパワースペクトルが主要であることが確立されている。

スペクトログラム (時間的に変化するパワースペクトルの様子を濃淡図形で表示したもの) を観察すると、そこに特徴的な濃いしま模様が見えてくる。これはパワの大きい周波数成分を示している。図-1.2 にそれを模式的にしたものを示す。この濃い部分は「ホルマント」と呼ばれる。各母音について特徴的な配置が見てとれるだろう。

そこでスペクトログラムのようなパターンを聴覚のパターンと考えて、音韻を整理することが考えられるだろう。この整理の試みは Jakobson-Fant-Halle⁴⁾ によってなされた。表-1.3 (次頁参照) はその結果である。

この表は Jakobson の「弁別素性」という考えにもとづいている。前に、音素を単位的なものと考えた。しかし、数多くの音素はもっと基本的な要素から構成されるのではないか。Jakobson は、2 値的な基本的な特徴があると考え、それを弁別素性と呼んだ。

表-1.3 は、そのような考えをもとにし、スペクトログラムの観察から、その印象的特徴によって弁別素性を定め、音韻を整理したものである。だが、表-1.1、表-1.2 と同じように定性的な記述である。

ホルマントの重要性は、聴取実験によって確かめられた。スペクトログラムを図-1.2 のような 2 値的な濃淡図形 (ホルマントパターン) によって近似し、それから音を再合成 (パターンプレイバック) してみる。そうしても、もとの音韻性はあまり損われない。ホル

表-1.3 英語音楽の弁別素性

弁別素性	o	a	e	u	ɪ	i	l	ɒ	ʃ	ʒ	ʔ	k	ʒ	ʔ	g	m	f	p	v	b	n	s	θ	t	z	ð	d	h	ʔ
1 母音性/非母音性	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 子音性/非子音性	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
3 集約性/拡散性	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
4 低音調性/高音調性	+	+	-	+	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-
5 変音調性/常音調性	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
6 鼻音性/口音性	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	
7 緊張性/弛緩性	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-	+	+	-	-	+	+	+	+	-	-	-	+	-
8 連続性/中断性	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	+	-	+	+	+	-	+	+	-	-	
9 粗雑性/円熟性	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	+	-	+	-	+	+	-	+	+	-	-	-	

弁別素性名の訳は竹村、藤村による。
記入のないところは冗長であることを意味する。

マントをベースにして、その他詳しい感覚心理学的研究がなされた。

ホルマントは、図でいえば下から3個ほどが重要であり、それらが音韻性をよく代表している。もとのスペクトルから見れば、(音韻性を失わず)大幅な情報圧縮になっている。しかも、縮約はされたが定量的な量である。そこで、ホルマント周波数の自動抽出が大きな研究テーマになった。

なお、ホルマントの周波数が重要であるというのは次のことを意味している。音声知覚的には、ホルマントのバンド幅はあまり重要でない。また、パワースペクトルの傾きはあまり重要でなく、ピークの位置が大事だということである。これは聴取実験によって確かめられている。

3.3 調音か聴覚か

聴覚的な手がかりとしてホルマントがえられた。しかし、ホルマントも現象的なものである。音韻との関係でもその法則性はあまりあらわではない。動的な部分での変化のパターンもそれ自身では説明のつけようがない。考えてみるとこれはやはり「音響量」なのである。真の「聴音量」についてはいまの所手がかりはない。ホルマントは発声時の調音器官の特性を反映しているものである。

Fant は表-1.3 の研究のあと、前述のX線による声道計測をベースに、声道の音響理論を立て、ホルマントを声道の共振周波数と規定した。言語学の方面では Chomsky-Halle⁹⁾ が、調音的記述による弁別素性に戻している。

ホルマントの聴取実験の結果の分析をもとにして、Liberman らは、調音器官上に戻って考えた方がより法則的であることから、聴覚においても、聴手は、発声時の調音状態を推定してから知覚しているのであろうという仮説「調音参照説」を立てた。さらにいえ

ば、調音状態というのも現象である。その元の運動指令がより根源的であり、そこに音韻性のベースがあると考えられる。運動指令に還元して知覚しているという仮説が「運動指令説」である。

ところで、調音状態から音を作り出す方は理論化され、いまでは工学的応用(合成)も行われている。しかし、その逆過程はどうであろうか。音から調音状態を再現できるであろうか。調音参照説・運動指令説は調音か聴覚かという問題を止揚しようという野心的な仮説である。これを直接(神経生理学的に)証明するのは大変な難事であろう。その前に、少なくとも工学的にそのことが可能でなければならないだろう。それが可能なら、仮説の傍証にはなるだろうし、「音声認識」への新しい手がかりが得られよう。

「声道形推定の問題」はそこに位置する。これについては、第2部でその研究の一つの流れを解説する。

4. 音声波の分析

音声波の分析は音声研究の基本的立場の一つである(音響音声学ともいわれる)。とくに工学的立場のものにはもっとも重要である。

波(形)自身の観察は基本的に重要である。分析に問題があるときは波形まで戻らなければならない。

スペクトル分析は波の観察の一手段でもある。フーリエ変換については高速アルゴリズム(FFT, 最近では WFT)が利用される。音声の(パワ)スペクトルは、母音の場合、音源の周期を示す微細構造と、その上についたいくつかのピーク構造、それに全体的傾斜からなりたっている。微細構造を除いたものをスペクトル包絡という。ピーク構造はホルマントを示している。しかし、スペクトルをベースにホルマントを(自動)抽出するのはなかなか難しい。

パワースペクトルの(逆)フーリエ変換は自己相関関

数である。スペクトルの対数のフーリエ変換はケブストラムと呼ばれる。これらは、音源周期の分離にすぐれ、スペクトル包絡に関する情報が低次の項に集約される。これらについては教科書⁹⁾を見られたい。

音声分析で70年代のもっとも大きな展開は、「線形予測法」の導入である。音声の分野でこの手法を創めたのは Itakura-Saito といってよい。同じころ Atal からもこの手法を導入した。「線形予測」という名称は Atal らによる。

線形予測法は、音声生成についての一つのモデルを考え、そのパラメータの推定を行うもので、線形システム推定論の応用とすることができる。詳細は、本特集号の別解説および参考書⁹⁾にゆずるが、その基本的考えは簡単である。

いま (サンプリングされた) 波形を x_i とする。それは、それ以前の数個の値と入力との線形和

$$x_i = \sum_{l=1}^N \alpha_l x_{i-l} + \varepsilon_i$$

で表わされるというモデルを立てる。ここで、入力 ε_i は誤差であると考え、 x_i の観測値からパラメータ $\{\alpha_l\}$ を、「最小自乗法」によって推定する。その解を求めることは簡単な練習問題で、自己相関 (あるいは共分散) 係数の行列を係数とする連立一次方程式が得られる。それを解けばよい。

$\{\alpha_l\}$ をもとにその特性を計算すれば、スペクトル包絡が得られる。それはピークをむしろ強調したものであり、ピーク検出あるいは特性方程式の求根から、ホルマント抽出の道もひらかれる。

この簡単なモデルが音声分析にとって、以前のどの手法よりも有効であったことは驚くべきことである。コロンブスの卵のようにも見える。しかし、それは創始者たちの音声に対する洞察にもとづいたものというべきであろう。これは線形システム論では「全極モデル」といわれるものであり、推定法は別途示されていた。それを知っていたとしても普通は途中であきらめたであろう。音声に適切な分析項数 N は、教科書の例題よりはるかに大きかったからである。

もちろん簡単な原理だけでなく詳しい理論的展開もあった。効率的なアルゴリズムも導入されている。

Itakura-Saito が同時に導入したパラメータ (偏自己相関係数-PARCOR) は、その効率的なアルゴリズムの中に現われ、 α より根元的で、よりよい性質をもつものであった。そのパラメータは、音響管の反射係数とも解釈できることが示された。

線形予測法は分析だけでなく、(もともと)合成のモデルを与えている。PARCOR による合成回路は、安定性にすぐれ、それ以前、声道をシミュレーション合成系から導びかれていた合成回路と (形式的) に一致していたのである。

線形予測法は、音声分析に新時代をひらいた。そしてそれはたんなる新手法ではなく、もっとも基本的な手法として使われていくものと考えられる。

5. 音声理解についてのコメント

かつてのボトムアップ的な音声認識の研究の体験から、言語情報の役割が見なおされた。言語認識と深く結びつけた音声認識が構想され、「音声理解」と名付けられた⁶⁾。それはその頃(70年代初頭)、人工知能の分野で進んだ自然言語の意味処理システム「言語理解」の成果をとり入れ、それを音声の成果に結びつけようとしたものである。

言語情報の役割は、たとえば、非常に大きな雑音環境の下では、意味が了解されると同時に声も聴えてくるというような心理的感覚を思い出しただくとよいかもしれない。

ARPA の「音声理解システム」プロジェクトが発足しようというころ、私自身も音声研究と言語理解研究を (それぞれ) プロモートしようとしていた。そのことからすれば、音声理解の提案に全面的に賛成すべきであったかもしれない。音声理解の問題意識は大変もともとでもある。しかし、私にはそれは時期早尚と感じられた。私には、まだしばらくは、それぞれを大事に育てるべきだと感じられたのである。

言語情報の役割を示すために、スペクトログラムの読取実験が行われ、言語情報を使わないかぎり、人間の読取能力も非常に低いということが示された。

片や次の事実がある。意味にも構文にも関係のない「無意味単語」の聴取を考える。良好な音響条件では、人間の聴取能力は高い。それと同じ能力を達成する方法はまだ開発されていない。

読取実験の結果は、聴覚には組込まれているはずの音声の法則が、まだ自覚的に取出されていないというにすぎないのではないか。

とすれば、我々自身の不完全な「音声理解」を言語情報で補うということになる。それでいいだろうか。

一方、言語理解の研究も、70年初頭の成功には大きな意義があるが、それは問題の解決ではなく、それまで手がつかないと思われた問題への足がかりを与えた

ものである。しまわれていた難問が(再び)提示されたのである。それ自身を育てるのにまだ大いに時間がかかると思われた。

ARPAの音声理解グループには、音声そのものの解明についての絶望感と、言語理解についての楽観があったのではないか。

5年間の現実の展開はそれを裏付けていると思われる。成功であったと公式の報告書にはある。内部での率直な反省では、音声処理、言語処理のそれぞれをもっと充実させるべきだという意見が出されている。個々の成果を見ると新しい芽が見うけられる。それを育てることが、本当は大事だったのではないか。

音声、言語それぞれに適切な問題設定をしなければならぬ。結合すべき時期が来ればそれをしなければならぬ。そのためには絶望感や楽観を越えた冷静な眼が必要であろう。

6. おわりに——参考書など

音声について勉強するための参考書を紹介しておきたい。幸いに我国には優れた参考書がある。

- 1) 藤村 靖編著：音声科学，東大出版会（1972）
- 2) 電気通信学会編：出版，聴覚と音声（1966）
- 3) 中田和男：音声，コロナ社（1977）
- 4) Flanagan, J. L.: Speech Analysis Synthesis and Perception, 2nd ed., Springer-Verlag (1972)
- 5) Markel, J. D. and Gray Jr., A. H.: Linear Prediction of Speech, Springer-Verlag (1976)

1)が標準的な教科書であり、反復熟読すべきものである。当時の最新成果までが総合的に解説されている。2)はその一世代前までの総合である。1)にはない材料も豊富にあり、合せて参考にすべきである（なお改訂新版が近く出される。旧版の良さも保存されることを期待したい）。3)は最新刊であり、線形予測法を全面的に取入れ、問題提起もあるフレッシュな書である。4)は1)に相当する標準的な教科書。音声知覚等に関しては1)より詳しい。1)には2)にない材料もある。線形予測法に関しては5)がよい。ていねいに書かれ、プログラムも記載してあり、実用性が高い。理論的部分も明快であり、研究の歴史も書かれている。

新しい研究成果は、日本音響学会の会誌、音声研究会資料、大会予稿集、また電子通信学会の論文誌、電気音響研究会資料、大会予稿集を見ればよい。米国ではJASA（音響学会誌）があるが、分析法についてはIEEE Tr. ASSP（前身 Tr. Au）の方に論文が出る。

書物における知識も大事だが、音声そのものを自分自身で扱い、自然に学ぶことが基本でありたい。

音声実験は、最近の電子産業の発展で非常に楽になってきた。情報量からいって、画像処理などより格段に気楽なはずである。実時間処理はまだ少し苦しいが、すぐ可能になってくるだろう。

参 考 文 献

- 1) Fant, G. C. M.: Acoustic Theory of Speech Production, Mouton (1960).
- 2) Chiba, T. and Kajiyama, M.: The Vowel: Its Nature and Structure, Tokyo Kaiseikan (1941).
- 3) Ladefoged, P.: Preliminaries to Linguistic Phonetics, Univ. of Chicago Press (1971).
- 4) Jakobson, R., Fant, G. and Halle, M.: Preliminary to Speech Analysis (1952).
竹村 滋・藤村 靖訳：音声分析序説，研究社（1965）。
- 5) Chomsky, N. and Halle, M.: The Sound Pattern of English, Harper & Row (1968).
- 6) Newell, A., Barnett, J., Forgie, J. W., Green, C., Klatt, D., Licklider, J. C. R., Munson, J., Reddy, D. R. and Woods, W. A.: Speech Understanding System—Final Report of a Study Group, North Holland/American Elsevier (1973).
- 7) Petrick, R. P.: Report on 1977 ACL Panel on Speech Understanding and Computational Linguistics—A Critical Examination of the ARPA Project, AJCL Microfiche 64: 2.

第2部 音声研究の一断面

——声道形推定の問題——

1. はじめに

第1部で述べたように、音声の記述をどの領域で行うかは大きな問題である。この場合、音響パラメータから発声側のパラメータを得ることの困難さがこの問題の統一的な解決をはばんでいる。ここで扱う「声道形推定」の問題はそこに位置する問題であり、特殊な問題に見えるかもしれないが、その成否は音声研究の方向にも大きな意味をもつ。

はじめに「摂動論による解」を述べる。これは歴史的なものであり、また問題のありかを良く示してくれるものである。

ここでの主眼は「線形予測法」の出現とともに始まった研究である。その代表的なものは、Wakitaの研究とNakajimaの研究である。これに私自身の体験を追加して、諸手法の根拠や問題点を考えてみる。

2. 問題の設定——声道断面積関数

声道形といえ、ことは通りには立体構造まで含めた形状ということになる。しかし実際にはもっと近似的に扱われる。

図-2.1 を見よう。これは X 線写真から声道形を計測した例 (Fant による) である。横から見た正中断面がトレースされ、縦断面が推測されている。これから、長さ方向への断面積の分布が計算される。これを声道断面積関数 (略して面積関数) という。

あとで論じるように、音声の有効な周波数範囲では断面積だけが、しかもその相対値だけが有意味なのである。そこで、正確には、声道 (相対) 断面積関数の推定がここで問題であり、図-2.1 の下のような図形 (もっと粗く近似されることがあり、縦軸は相対値とする) を、音声波の分析 (音響量) をベースに得ることがそのゴールなのである。

声道を音響回路として扱やすくするために、さらに近似がなされる。その仮定を列挙すると、

- (1) 声道自身は線形受動系と考える。
- (2) 声道内部では損失がないと考える。
- (3) 声道の中での音波は平面波の方程式に従う。
- (4) 分岐管はないものとする。

(5) 断面積の異なる、単位長の管を連続したものの (単位長素子連続不均一音響管) を考える。

(6) 単純な境界条件で考える。

仮定 (1) と (2) は、普通の電気回路でのリアクタンス回路に相当するものである。仮定 (4) は、ここでは母音の場合だけを考えるということである。仮定 (5) は線形予測法との関連で用いられる。

これらの近似がおおむね妥当なことは、そのようなモデルによって合成した音を聴取評価することによって確かめられる。音声知覚を考慮することは、モデルを過度に簡略化しないためにも、また過度に精密化して問題をいたずらに複雑化しないためにも必要である。

そこで音声知覚について (とくに母音について) 主要な性質をまとめておこう。周波数特性でいうと、

- (1) パワ特性が重要であり、位相特性にはあまり感じない。
- (2) パワスペクトルの全体的な傾きは、音韻性にあまり関係しない。
- (3) ホルマントでいえば、その周波数が重要であり、そのバンド幅はあまり感じない。
- (4) 零成分 (アンチホルマント) にはあまり感じない。
- (5) 4 kHz 以上の領域、周波数零のあたりは、音韻性にあまり関係がない (電話帯域は、300 Hz ~ 3.4 kHz である)。

ここでもっとも大きな仮定である平面波の仮定 (3) について考えてみよう。それは波長と声道の寸法との関係による。3.5 kHz での半波長は約 5 cm である。一方声道長は典型的には 17 cm 程度である。そこで分布定数回路を考えることになる。一方、声道の太さは、半波長よりかなり小さいと考えられる。これが平面波の仮定の根拠である。しかし、この仮定は 4 ~ 5 kHz 以上ではくずれてくる (5 kHz で半波長は約 3.5 cm)。

平面波の仮定に立つかぎり、その分解能は自ずと制限される。縦断面の形状は問題にならず、その断面積 (しかもその相対値) だけがきいてくる。長さ方向についても、曲り具合などは問題にできない。断面積関数の変化についても微細な変化は問題にできず (図-2.1 の図形はすでに細かすぎる)、ある程度以上滑らかな形しか問題にできない。

問題のゴールが断面積関数であり、その推定精度も制限されるのは、そのような事情によるのである。

このような平面波の仮定の領域が、音韻知覚の特性

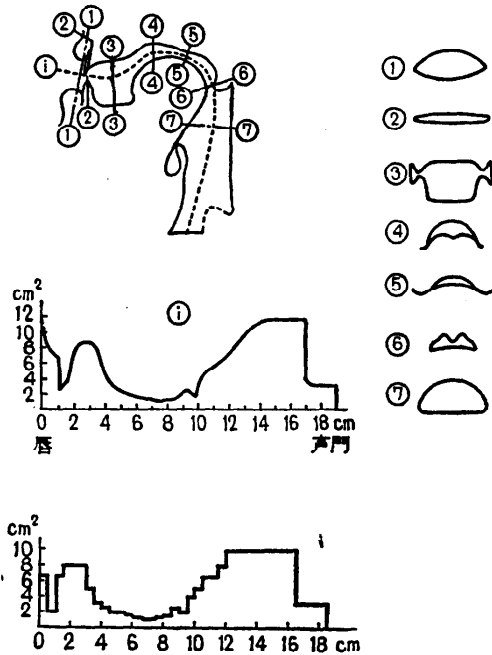


図-2.1 声道断面積関数の計測 (Fant による)

(5)の領域とほとんど重なるのは面白いことである。

声道形推定問題では、もう一点注意すべきことがある。音声の場合、音源にも強い特性があるが、観測されるのは出力だけなのである。声道形推定問題がシステム推定論の一練習問題ではありえないのは、このことによる。

ここで言訳をしておきたい。これから結果の評価に「リーズナブル」という形容詞をよく用いる。これは定性的な表現でしかないが、声道断面積関数の実測データを自由にすることは、いまでも非常に困難であり限られたデータとの定性的な比較をせざるをえないのである。それとともに、これまで定性的にせよ評価にたえる結果をうる道を探ることが問題であった（またX線計測にしろ、断面図は推測であり、精度の限界があることは指摘しておくべきだろう）。

3. 摂動論による解

ここでは、Schroeder-Mermelstein による摂動論^{1),2)}を述べる。仮定(5)は用いない。

まず、断面積が一様音響管を考える。長さが与えられたとき、一端を完全開放(唇側)、他端を完全閉鎖(声門側)とする境界条件を設定すると、この音響管の共振周波数(固有値)は容易に求まる。典型的な長さとして17 cmを仮定すると、それは500, 1500, 2500, 3500, ……kHzとなる。

いま、長さが同じ(境界条件も同じ)で、断面積分布が一様管から微小変化した不均一音響管を考える。そうすると固有値(共振周波数)も微小にずれるであろう。その微小変化を断面積の微小変化と対応づけるのが摂動論である。

結果だけを示すと、対数断面積関数 $\ln A(x)$ を、cosine 級数で展開しその係数を c_m とすると、一様管の固有値 λ_{m0} からのずれ $\delta\lambda_m$ は、

$$\frac{\delta\lambda_m}{\lambda_{m0}} = -\frac{1}{2} c_{2m-1} \quad (m=1, 2, \dots, \infty)$$

この関係を使えば、固有値(共振周波数=ホルマント周波数)が与えられたとき、それを一様管からの微小変化によると考えれば、それらから(相対)断面積関数をうることができる。

図-2.2はその例である。ここで点線はX線写真からの実測値(Fantによる)であり、実線が推定値である。ホルマント周波数は実測データから計算されたもので、3個のホルマント周波数のみを用いている(残りは一様管のそれと一致していると考えられる)。

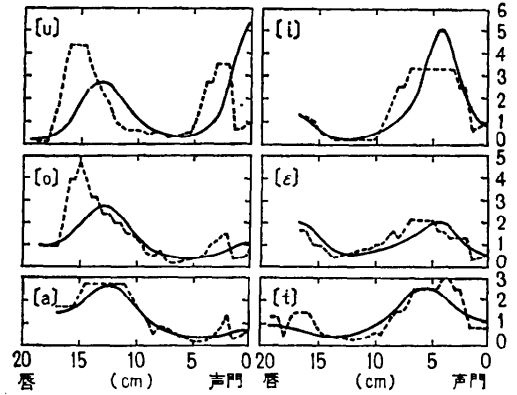


図-2.2 摂動論による推定断面積関数¹⁾
(Mermelstein-Schroederによる)

この結果は、第1次摂動論には良く合うと評価されている。(なお、点線の図形は後での結果との比較にもなるのでよく見ておいていただきたい。)

この結果にはいくつかの問題がある。第1は、各音韻の特徴はかなりよく再現されているものの、欲をいえばもっと合ってほしい。第2に、ホルマント周波数が与えられると仮定しているが、ホルマント抽出自体なかなか難しい。さらに、声道長が与えられているという仮定が問題である。声道長は音韻ごとにも変化し、また、話者によって(大きく)異なる。ここでは実測値を使っているが、音響量の中で直接それを見出すことはできない。

第1の点は摂動論にも問題がある。再現された面積関数から共振周波数を計算しなおすと、与えたものと若干異なり誤差を生じてくる。この区での面積関数は一様からの微小変化の範囲を越えているように見える。そこで、摂動論の精密化も試みられた。

だが、もっと大きな問題がある。前出の対応式をもう一度見てみよう。そこでは奇数次の展開係数しか使われていない。摂動論の結果は、展開係数の奇数次だけがホルマント周波数に影響し、偶数次の係数はホルマント周波数に影響を与えない、ということを示しているのである。

逆にいうと、偶数次項は任意に与えてもよいということであり、声道形には音響量から見ると不定の要素があるのである。これは、一意性の問題といわれる。このことは摂動論固有のことであろうか。実はそうではないのである。これについては後でまた触れるが、声道形推定問題の不可能論、否定論の論拠の一つになっている。

4. 線形予測法の出現

Itakura-Saito と Atal らによって導入された線形予測法は、70年代の音声分析に大きな影響を与えた。そして、声道形推定問題にも新しい展開をもたらした。

線形予測法はもともと、音声の生成過程を念頭に置いた分析法であったが、さらに、Itakura らが導入した PARCOR (偏自己相関) 係数は、単位長素子を連続した不均一音響管の反射係数と一致するものであった。このことは Itakura ら自身によって早くから指摘され、広く知られていた。

PARCOR 係数をベースにした(再)合成の回路が、それ以前、声道のシミュレーションから導かれていた音声合成回路と同形式であったのである。このことは、波形分析から得られるパラメータを用いて、直接的に(ホルマント抽出などの他の手段を媒介にしないで)、特性を再現する音響管を構成することができる、ということを示している。

もし、この音響管がすでに声道形(断面積関数)を再現していると認められる形のものであったら、声道形推定の問題はこの時点で終わっていたことになる。だが、その形はX線などで実測された形とは似ても似つかぬものであった(と思われた)。

その形は、消音器のように交互にはげしい凹凸のあるもので、口の閉開、舌の前後に相当する調音の定性的な特徴も見分けがたかった。うまくいかない原因は何であろうか。回路の形式的な一致ということだけでは不十分で、もっと確固とした理論が必要なのであろうか。あるいは、声道形推定はやはり不可能な問題なのであろうか。

5. 理論化と一階差分

次のステップは Wakita によって進められた³⁾。その功績の一つは、理論構成を精密化したことである。音響管における反射過程と、PARCOR の計算過程を対比し、反射係数と PARCOR 係数の対応を直接的に証明した。これは Itakura 方式を裏づけるものであった。

次に、波形を一階差分(高域強調)したあと、これについて PARCOR 計算をすれば、「リーズナブル」な声道形が得られることを主張した。

図-2.3 にその結果を見よう。図の右が5つの母音についての推定結果である。左は、一階差分した波形

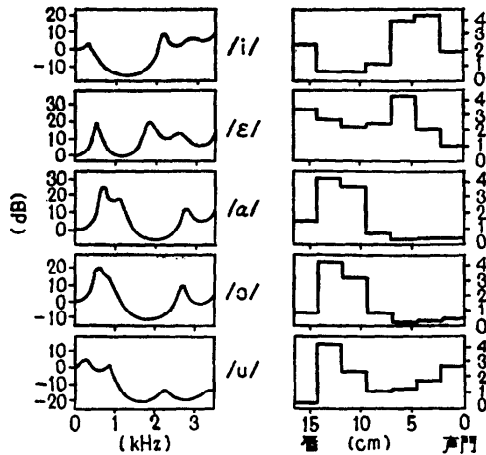


図-2.3 脇田³⁾の推定断面積関数

から線形予測法によって推定されたパワースペクトルであり、同時に左側の音響管の伝達特性でもある。図-2.2 (点線の図形)と比較して、「それらしい」ということが認められる。

これが、線形予測法をベースにした声道形推定成功の第1報であった。これは大きな前進であったが、実は部分的成功というべきであった。

いろいろなデータ、いろいろな分析条件(サンプリング周波数)で追試すると必ずしもうまくいかないことが多い。図-2.3でも前舌母音の /i/, /ε/ はよいが、後舌の /a/, /ɔ/, /u/ はいま一つの所がある。

それでは何故うまくいくことがあり、またうまくないことがあるのであろうか。Wakita の研究は理論的部分に大きな貢献があるが、アルゴリズムとしては Itakura のものと同じである。そこで、ポイントは「一階差分」にあることになる。

これは、放射と音源の特性を打消すものと説明されている。音源特性は典型的には -12 dB/oct であり、放射特性は 6 dB/oct である。一階差分によってこれを打消することができる。これが本質的なポイントである。しかし、音源特性は音韻や話者によってかなり大きく変動する。それに対して固定の一階差分でよいであろうか。Wakita はそれ以上追求していない。

摂動論のところで述べた一意性の問題はどうか。Itakura-Wakita の場合、境界条件は、唇で完全反射、声門部で無反射となっている。すべての損失を声門部に持たせているのである。このような境界条件では、PARCOR のアルゴリズムにより面積関数が一意的に定まるといのが Wakita の理論的な成果の

一つである。

ところで、線形予測係数と反射係数の対応は Atal も試みている。彼は境界条件として、声門部を完全閉鎖、唇部を無反射にしている。この設定ではどうであろうか。Wakita はそれはうまくいかないと主張している。しかし、実はそうではないのである。この点についてはあとで述べよう。

Wakita の理論と結果は Markel らの本にもていねいに紹介されている。しかし、うまくいくとだけ書かれて問題点の追求がなされていない。追試がうまくいかない（それには理由がある）ことが多いことから、かえって推定不能論を招いている。せつかくの成果が残念なことである。

6. 適応的な等化・平坦化

Wakita と同じころ、この問題を追求していた Nakajima は問題のポイントがスペクトル全体の傾きにあることを発見したり、手持ちの声道形実測データから特性を計算してみると、かなり高域強調的である。さらに言えば大体「平坦」である。

前に述べたように、放射と音源をまとめたものは音韻とか話者によって相当変動がある。これを取除くには何らかの基準が必要である。Nakajima はそれをスペクトルの平坦さに求めた。そうすると固定の高域強調ではなく、データに応じた補正が必要になる。

スペクトルの傾きを平坦化する操作（等化）をほどこしてから音響管を計算すると、(Wakita の場合より) ずっと「リーズナブル」な結果が得られる。

図-2.4 にその結果を見てみよう。図-2.2 と比べてみても、かなりもっともらしくなっていることがわかる（ただし、一方は日本語母音、他方はロシア語母音である）。図-2.3 と比べれば、後舌母音が改良されていることが分る（図-2.3 の後舌母音ではかなり傾きが残っている。これを補正すればもっとリーズナブルになろう）。図-2.4 は 12 kHz サンプルングを用いている。このため微細な凹凸があるが、これについてはあとで触れる（なお図-2.3 は 7 kHz サンプルングである）。

Nakajima の平坦化（適応等化）は聴覚的に許されるであろうか。前に述べたように、傾きの補正は音韻性にはほとんど影響を与えないのである（Nakajima は聴取実験によってこれを再確認している。）。

適応等化は、声道特性と放射・音源特性の分離と解釈され、平坦さがその基準を与えている。この点に積

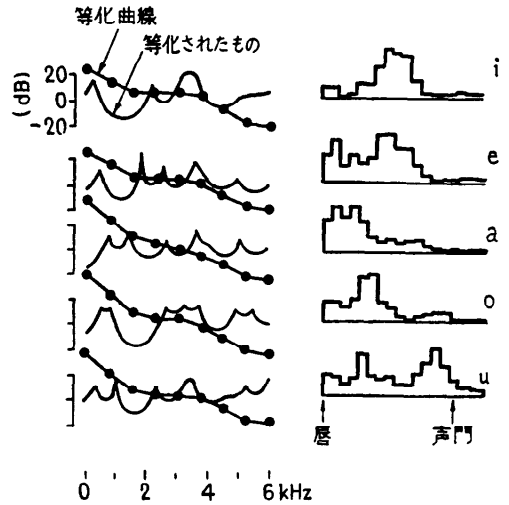


図-2.4 中島⁹⁾の推定断面積関数

極的な意義がある。

Nakajima の手法は（通常実験される成人男性の声だけでなく）、女性の声、子供の声にも適用されて良好な結果を得ている。良好な声道形（断面積関数）を安定してうる手法が Nakajima によって確立されたといっている。

ところで平坦化の操作はどうしたらよいだろうか。放射・音源特性をピーク特性のないなだらかな傾きの特性として考えてよいことは従来の研究で明らかにされている。このような特性は（臨界）制動条件のあるような特性の組合せで表現できるであろう。Nakajima はそのような制約付きの線形予測法を考えて特性を推定している。また、等化もスペクトル領域ではなく、波形領域で行っている。等化された波形からは、通常の線形予測法（共分散行列法）によって反射係数が計算され音響管が得られる。詳細は文献⁹⁾にゆずろう。

Nakajima の方法に対する根本的な問は、平坦化とは何を意味するか、そしてそれが何故よいかということであろう。

この問とともに、技術上の問題として、等化のための手法が複雑であることが、この方法の追試・普及を遅らせているように思われる。

7. 逆対称性の導入

これまで、声道形推定に関する一連の研究を見てきた。Nakajima の平坦化の基準に照らせば、Wakita の手法でうまくいく場合、そうでない場合が分る。素朴な線形予測法の適用でうまくいかなかったのは、大

きな傾きのある(低域波渡)特性を同時に実現する音響管であったためと分る。

ポイントの一つは、声道形特性、放射・音源特性をどう扱うかという点にある。声道形推定の問題の難航を理論上の不備ではないかと思える人が多かったが、実は、そのような音響回路理論の構成法より、むしろ音声本来の性質をどう利用するかということにあったといえるだろう。しかし問題は残っている。

一つは境界条件の問題である。Atal の用いた条件での推定の試みはなかったが、実はそれでも「リーズナブル」な結果が得られるのである。その場合図-2.3 や図-2.4 の図形で、その断面積を逆数にした図形になり、入出力関係は左右が入れかわる。そのようにしても、同じ程度にリーズナブルな図形になることは図からも推察されよう。ただし図-2.3 では長さ(分析項数)の足りないものがある。

とすれば再び一意性の問題が生じると見える。しかし(とくに Nakajima の平坦化をした場合)、二つの図形は左右を入れかえればほとんど重なり合うのである。これはどういうことであろうか。平坦化と何か関係があるだろうか。

断面積を逆数化し左右入れかえたものが完全に重なるとき(いずれも相対値であるから定数倍を除いて)逆対称(antimetical)な声道形といわれる。図-2.2 を見なおして見よう。ここで(奇数次の展開項によって)得られた声道形は実は逆対称であった。

ということは、図-2.4 などの図形はほとんど逆対称的だということである。

ここで回路網理論について少し復習しよう。我々の仮定での音響管はリアクタンス回路に相当している。

図-2.5 は電源側にも損失のある場合で動作状態といわれる。図のような回路の解析には進行波の考えを使った S-行列の理論が有名である。負荷側、電源側での波について、

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} S(s) & \varphi(s) \\ \varphi(s) & \tilde{S}(s) \end{bmatrix} \begin{bmatrix} b_2 \\ a_2 \end{bmatrix}$$

が成立つ。ここで S(s) は動作伝送係数、φ(s) は特性関数と呼ばれる。

図-2.6 のような音響管でも同じような定式化ができる。ただ単位長素子の遅れを基本に考えて、z-変換を用いる。そうすると変換領域が s から z に変わるだけでほとんど同じ関係式が成立つ。Atal や Wakita などの展開した理論はそのようなものである(ただし両者とも動作状態は考えていない。動作状態の音響管

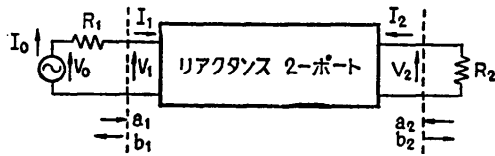


図-2.5 動作状態の回路

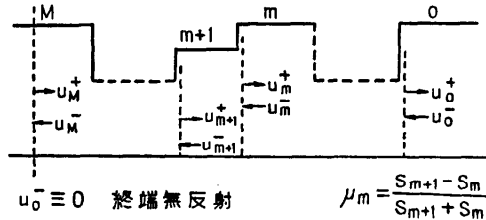


図-2.6 単位長素子継続不均一音響管

は、両端を一様管、すなわち無反射で終端することに相当する)。そこで以下は図-2.5 (こちらがなじみが多いと思うので)を使って議論をすすめる。

ところで回路網合成理論によれば、負荷 R₂ を開放した開放伝送係数、あるいは V₁ から見た電圧伝送係数が与えられると、それからはしご型回路を合成することができる。前者が Itakura-Wakita の定式化に相当し、後者が Atal の定式化に相当する(それぞれの境界条件を思い出してほしい)。アルゴリズム的に行くと、Itakura らは、伝送係数そのものからではなく、それに対応する自己相関係数から回路を導く新手法を提案していることになる。この方が波形分析との関連からすると直接的で有利である。

S(s) と φ(s) が与えられれば、それらからも回路を導くことができる。S(s) と φ(s) の間では、

$$S(s)\tilde{S}(s) = 1 + \varphi(s)\varphi(s)$$

という基本関係が成立つ。φ(s) が与えられれば、S(s) は(安定条件を満たさなければならないから)一意に決まる。その逆は成立たない。

いま、ある周波数で φ(s)=0 になるとしよう。この場合減衰がもっとも少なく、動作伝達特性(S(s)S̃(s)の逆数)上でピークをなすことになる。ここでホルマント周波数が特性関数の零点を与えるものと考えてみよう。そうすると、ピーク値はつねに1であるから、全く傾きのない(平坦な)特性が得られることになる。

一方、実周波数の零点だけで定まる φ(s) について

$$\varphi(s) = \varphi(-s)$$

が成立つ。実はこれは回路が「逆対称」であることを意味する。音響管でいえばホルマントのピーク値が一定な場合は、逆対称音響管に対応するのである。

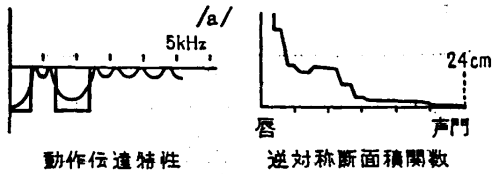
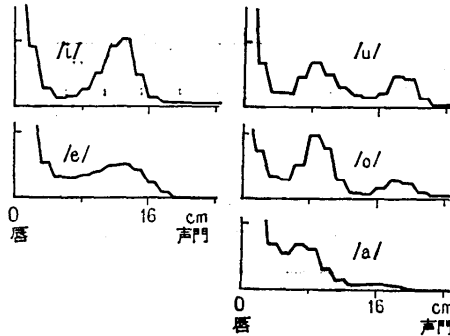


図-2.7 逆対称音響管の特性

図-2.8 推定された逆対称断面積関数⁶⁾

この手法によって、摂動論によらず、ホルマント周波数を正確に再現する（逆対称）声道形をうることができる⁶⁾。図-2.7 はそれを示している。

図-2.8 は日本語5母音について得た結果である。これを図-2.2～図-2.4 と比べられたい。一番リーゾナブルと見るのは自己満足であろうか。

この結果は、Nakajima の平坦化という基準について、その解釈を与えていると考えられる。適応等化は実は、逆対称に近づけるための操作であった。

ここで問題はまた元に戻ったように見える。

8. 問題の再考

一意性の問題を考えよう。動作特性をもとに考えると、一意的なのは逆対称のときだけである。非対称的なとき、また特性に傾きのあるときには、同じ特性を持ついくつかの音響管の形状があることになる。

Wakita らは一意になる境界条件を与えたが、一端が完全反射というのはあまりよい近似ではない。そして境界条件の選び方自身に任意性があるわけである。また、理論はともかく、よりリーゾナブルな解を得たのは、高域強調→平坦化（≒逆対称化）の方向であった。そこで逆対称解というのが、声道形推定問題の解のあり方と考えられる。

しかし、逆対称解は摂動論ではあまり評価が高くなかったのではないか。ところが図-2.2 には大きな問題がある。それは声道長の問題である。そこでは実測

値を使って合せてある。しかし推定結果をよく見ると、横にずれているものが多いことに気付く。考えてみると、放射特性などの影響で「音響的な長さ」というのは実体的な長さより長いはずなのである（開口端補正などはそれである）。音響的な長さを補正するとホルマントだけ（逆対称）でも、もっとよく合うと見られる。このことは実測の声道形自身、図-2.2 の見かけよりもっと逆対称的だということである。

実際の声道形が逆対称的であるという先験的な理由はないだろう。しかし、次のことは考えられる。逆対称回路というのは、ホルマント部でインピーダンス整合のとれた、すなわち、その成分をもっともよく伝達する一種の最適回路である。声道自身発声の能率化のため、そのようなポーズに近づくことはないだろうか。

摂動論では声道長を与えることが必要であるが、前述の逆対称声道形の推定では長さを明示的には与えていない。そして唇側にはその外の部分と見えるものがついているが、それまで合めて逆対称になっているのが興味深い。

次に高域補正の問題に移ろう。聴覚の特性からも平面波の仮定からも、ある周波数以上は意味を失う。その部分をどう処置しておけばよいか。図-2.3 のようなサンプリング周波数では問題でないように見える。しかし、子供の声などではそれでは不十分である。というのはサンプリング定理によって、折りかえしの特性が仮定されてしまうからである。一方、図-2.4 のように 12 kHz あるいはそれ以上を用いると対策が必要である。

図-2.4 の等化曲線で 4 kHz 付近にかなり鋭い折れ曲りがあるのは、その現われである（Nakajima はそのために臨界制動以下の条件を導入している）。

実はこの部分は、むしろ全くの平坦である方が望ましいのである。しかし、自然の減衰が強いため図-2.4 でもまだ特性が残っている。面積関数の細かい凹凸はそのせいである。この部分の特性のなさ（平坦さ）は、得られた図形の滑かさに対応する。図-2.7、図-2.8 はそのような特性を用いている。

「平坦化」には実は二つ問題があるのである。一つは逆対称化であり、もう一つは、高域補正である。後者のことは忘れられがちであるが、結果に対する影響は大きい。

ところで、Wakita, Nakajima の方式のよさはホルマント抽出を媒介にしないでよいという点にある。図-

2.8の結果はホルマントの考えで解説したが、実際上の手法としては、ホルマント抽出を直接用いないで済むこと⁶⁾を追記しておく。

9. 断面積関数のあと

断面積関数の推定がここでのゴールであったが、それがうまくいけば、実の所は、さらに少数のパラメータに縮約したいわけである。それが弁別素性に対応するような(調音)パラメータであれば理想的になる。

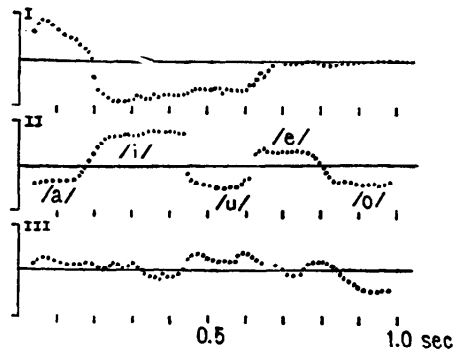
一つの方法は、さらに調音器官のモデルを立て、断面積関数をそれにあてはめて、パラメータを抽出しようとするものである。Nakajimaらの結果をベースにして、その方向の研究がいくつか進められている。

もう一つの方法は、断面積関数(またはその対数)が基本関数の和で表わされるという線形モデルを立ててみることである。このためには多変量解析の手法が使える。そのような実験によると、男声のグループだと3個程度、女声や子供の声を入れても4個程度のパラメータで元の断面積関数がほぼ復元できるという結果が得られている。これから音を再合成して聴いてみるとかなり良い音声を得られる。

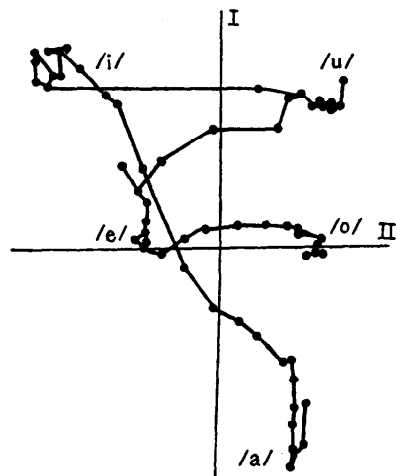
図-2.9は、(概)逆対称声道形を主成分分析した例である。男声の連続母音 /アイウエオ/ の主要平面での軌跡とパラメータの時間変動が示されている。これを図-2.10の古典音声学での「母音図表」と比べられる。この結果は今後の研究方向を示唆していると思われる。

これまで、声道形推定は母音を中心に進められてきた。それは、母音の音響的性質がかなりよく分っていたことと、母音でも推定が難しかったことによる。とはいえ、母音自体の問題が解決済なのではない。「話者正規化」と「調音結合」という二つの大きな難問が残っている。話者による違いは、音源特性の違いと声道長の違いに主として起因する。前者は適応等化によりある程度解決したと見られる。後者はほとんど解決されていない。

声道長抽出が一つの問題であるが、それとともに何が不変量であるかという問題が大きい。同じ音韻(と考えられるもの)を発声するとき、大人の声道形と子供の声道形はどのような関係にあるだろうか。相似であるようでもあり、相似からずれているようでもある。相似から少しずれているらしいというのが定説である。耳で聴けば十分判別できるものを、どのように分離すればよいか、従来のホルマント的分析でも十分



(a) 抽出された I-IIIパラメータ



(b) I-IIパラメータの軌跡

図-2.9 /アイウエオ/(男声)のパラメータ表示

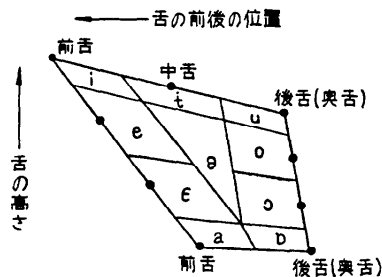


図-2.10 母音図表、調音器官のかまもと音韻との関係 (Principles of the International Phonetic Association 1949 から)

には明らかになっていない。声道形推定がそれにどう役立つか今後の課題である。

調音結合は「分節化」の問題でもある。物理的には

連続な量が離散的な記号をになっている。それとともにその連続量は前後の音の発声によって影響を受けている。そのためには調音の動態の観察が必要である。

X線による動態観察が進んできているが、声道形推定の試みはそれにどう寄与するであろうか。これも今後の課題である。

母音でも問題は本質的には解決されていない。現行の音声認識システムは、「特定話者」「特定単語」が主である。その限界を越えようとする、母音についても、すぐこのような問題につき当るのである。

10. おわりに

声道形推定の試みの一つの流れについて書いてみた。この問題については別のアプローチもある。調音器官のモデルを立て、そのパラメータを直接推定しようという試みがあるが、まだ見るべき成果を得ていないと思われる。また、壁面の損失の問題等、推定にかかわる問題の検討もある。線形システム推定論からの研究も続けられている。その一々には触れられなかった。

少し長々と書いてみた。また私見を強く書いてみた

が、どうであつたらうか。これもまた「音声研究のすすめ」になれば幸いである。

参考文献

- 1) Mermelstein, P.: Determination of the Vocal Tract Shape from Measured Formant Frequencies, *JASA* **41**, 1283~1294 (1967).
- 2) Schroeder, M.R.: Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements, *JASA* **41**, 1002~1010 (1967).
- 3) Wakita, H.: Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms, *IEEE Tr. Au-21*, 417~427 (1973).
- 4) Nakajima, T. et al.: Estimation of Vocal Tract Area Functions by Adaptive Inverse Filtering, *Bull. ETL* **37**, 462~481 (1973).
- 5) 中島隆之 ほか: デコンボリューションによる声道形の推定と適応型音声分析システム, *日本音響学会誌* **34**, 157~166 (1978).
- 6) 淵 一博: 逆対称声道形の推定と多帯域波近似, *日本音響学会音声研究会資料* S 77-05 (1977).
(昭和53年6月2日受付)