原 著 論 文

音声認識装置を用いた発音検査法の確立

第1報 音声認識装置の有効性について

市場 裕康   中原 信光   薦田 淳司   堀内 政信

市川 哲雄   佐藤 修斎   羽田 勝   松本 直之

Establishment of Speech Test Using A Speech Recognition Unit

Part 1. The Utility of The Speech Recognition Unit

Hiroyasu Ichiba, Nobumitsu Nakahara, Junji Komoda, Masanobu Horiuchi,
Tetsuo Ichikawa, Shusai Sato, Masaru Hada and Naoyuki Matsumoto

Abstract : To assess speech disturbances caused by newly made prostheses, visual and acoustical methods, palatograms, sound spectrograms, etc., have been used. But the analysis and assessment requires somewhat professional knowledge and experience and much time, so such methods are not necessarily easier, and have the defect of being somewhat subjective.

The purpose of this study is to develop a new speech test that is easy to use and more objective using the speech recognition unit.

First, experimental palatal plates were made for a dentulous subject, and the utility of the speech recognition unit was estimated.

The speech recognition unit and ordinary hearing were compared in utterance of ten Japanese monosyllables, /sa/, /ta/, /ki/, /hi/, /si/, /ci/, /ni/, /mi/, /ri/, /cu/.

Results :

1. The speech recognition unit made mistakes similar to ordinary hearing.
2. The recognition point that is output by the speech recognition unit corresponded well to speech clarity except some cases.
3. The speech recognition unit could detect the changes of consonant frequencies relatively well.

Key words : speech disturbances, speech test, speech recognition unit, recognition point, speech clarity

というものである。

I. 結 言

口腔内環境の変化、特に歯や顎の喪失によって起きる発音障害は義歯を装着することによってある程度回復されるが、同時に、装着補綴物によって発音障害を誘発することになる^{1,2)}。しかし、一般に、義歯による発音障害は義歯装着者が自覚的に訴えるほど他覚的に、つまり、聴覚的に認知されにくいものであり、明らかに誤聴されるほど異常音でない限り、その障害の程度を的確に表すことはできない。しかも、特別な訓練を受けていない者にとっては非常に困難なことである。

そこで、著者らは、音声認識装置を利用した簡便な発音検査法の開発を試み、その第一段階として、音声認識装置を用いて発音検査した場合と聴覚検査法によって検査した場合の両結果を比較検討したところ、興味ある知見が得られたので報告する。

II. 研究方法

1. 音声認識装置

本研究に用いた音声認識装置はサンヨー社製音声認識ボード・SRB-64A（以下、ボードと呼ぶ：図1）で、その回路構成は図2に示すとおりである。原理的には、特定話者の音声パターン（標準パターンと呼ぶ）を前もってメモリに登録しておき（最大64音）、その後入力する被験音の音声パターンが標準パターンのどれに最も類似しているかをパターンマッチング法によって認識させる

2. 被験音声

被験音として調音点、調音方法、後続母音を考慮した単音節、サ、タ、キ、ヒ、シ、チ、ニ、ミ、リ、ツの10音を選び、これらを、発音および聴覚機能に異常の認められない正常有歯顎者成人男子1名（27歳）に発音させ、被験音声とした。

3. 音声の標準パターンの登録

音声の標準パターンの登録は、図2に示すように入力音声をAGC回路とプリエンファシス回路で前処理したのち、16チャンネルのバンドパスフィルタ（200～7,000 Hz）で周波数分析後、8 bit のA/Dコンバータでデジタル信号化する。ついで、Z80マイクロコンピュータシステムでもって周波数スペクトルの時系列を表す音声パターンを作製し、メモリに格納する。今回は、標準パターンは同一被験者のものとし、後述の実験口蓋床を装着していないときの音声をボードに登録した。

4. 実験口蓋床

装着補綴物によって発音障害が誘発されやすいことを考慮にいれ、実験口蓋床として図3に示すような5形態、7種類のものを用意した。すなわち、部分床義歯を想定した部分口蓋床として、前方床、中間床、後方床および馬蹄形床の4種類を作製し、厚みはすべて2 mmになるように調製した。同様に、全部床義歯を想定した全部口蓋床についても厚さ1, 2, 3 mmの3種類を作製し、後縁

図 1

は左右の第 2 大臼歯遠心端を結ぶ線に一致させた。

5. 音声認識と録音

音声認識と録音は、被験者に実験口蓋床装着時と非装着時（以下、コントロールと呼ぶ）について、また、口蓋床装着時には簡単な会話を数分間行わせたあと被験音をそれぞれ 10～14 回ずつ発声させ、直接、音声認識ボードに認識を行わせた。同時に、データレコーダー（ソニーマグネスケール社製 KS-616、テープスピード 38 cm/s；データテープ、ソニー社製 L-750）にも録音した。なお、発声は日常会話時の大きさを指示した。

6. 分析方法

分析方法のブロックダイアグラムを図 4 に示す。

1) 認識ポイントと認識音

被験音声は、図 2 の過程を経て音声パターン化された

図 2 音声認識ボードの回路構成

のち、メモリ上の標準パターンとダイナミックプログラミング (DP) 法に基づいてパターンマッチングされ、2 つのパターン間の「距離」が認識ポイントとして算出される。そして、この距離の最も近いものが認識音として出力される。なお、DP マッチングとは、発声の仕方の違いによる時間軸方向の非線形なずれを、音声パターンの各部分を伸縮させて最も適した時間軸の整合を行い、パターン間の距離を計算する方法である³⁾。また、本ボードは、2 番目に類似度の高い標準パターンについても認識ポイントが出力可能であり、被験語を限定すれば、口蓋床装着前後の認識ポイントの差を観察することができる。

2) 認識結果の分析

実験口蓋床装着時と非装着時の音声の認識音と認識ポイントをパーソナルコンピュータ (YHP 社製、HP-87) で集計した。そして、正しく認識した語音の百分率を語音発語明瞭度（以下、語明度と表す）として表し、口蓋床装着時の認識ポイントについては、被験音ごとにコントロール時との差を求め、検定した。また、語明度、認

図 3 実験口蓋床

図 4 分析方法のブロックダイアグラム

表 1 音声認識ボードによる語明度

被験音	サ	タ	キ	ヒ	シ	チ	ニ	ミ	リ	ツ	mean	S.D.
コントロール	38.5	75	63.6	64.3	36.4	91.7	75	53.8	46.2	100	64.5	20.3
前方床	10	80	90	30	40	70	0	90	20	100	53	35.2
中間床	45.5	100	30.8	50	40	54.5	90	63.6	54.5	100	62.9	23.8
後方床	72.7	90.9	53.8	54.5	90.9	53.3	41.7	27.3	100	100	68.6	24.6
馬蹄形床	36.4	100	54.5	81.8	72.7	58.3	45.5	64.3	41.7	100	65.5	21.6
全部口蓋床(1mm)	40	60	40	40	50	20	10	80	90	100	53	27.9
全部口蓋床(2mm)	30	70	60	80	50	70	30	70	80	100	64	21.1
全部口蓋床(3mm)	20	50	20	90	70	20	10	100	50	100	53	33.5

単位：％

表 2 認識ボードの語明度の分散分析表

因子	平方和	自由度	不偏分散	不偏分散比
A(主効果)	2925	7	418	0.85
B(主効果)	25260	9	2807	5.68*
E(残差)	31134	63	494	
T(合計)	59319	79		

A：口蓋床
B：被験音

*信頼率95%で有意

識ポイントとも口蓋床と被験音を主因子とする二元配置分散分析を行った。

3) 聴覚テスト

聴覚テストは、データレコーダーに録音された音声、静かな室内で聴覚に異常の認められない11名の検者に聴取させ、①異常なし、②少しおかしい、③おかしい、④非常におかしい、の4段階で評価させ、誤聴傾向を認めた語音については聞こえたままの音を併記させた。そして、実験口蓋床装着時の各語音の語明度を、評価①(異常なし)と判定した検者数(百分率)で表し、認識ボード同様、二元配置分散分析を行った。

III. 結 果

1. 認識ボードによる語明度

音声認識ボードによる語明度を表1に示す。口蓋床なしの状態、すなわち、コントロール時の語明度は平均で64.5%となり、比較的低い値を示した。口蓋床を装着させたときの語明度は口蓋床の種類と被験音によってそれぞれ大きく異なっていたが、一定傾向は認められなかった。そこで、口蓋床と被験音を主因子とする二元配置分散分析を行ったところ(表2)、口蓋床の効果は有意でなく、被験音の効果は信頼率95%で有意であった。

2. 認識ポイント

ボードによる認識ポイントの平均と標準偏差、ならびに、口蓋床装着時の検定結果を表3に示す。

被験音サ、タ、ミでは、部分口蓋床装着時の認識ポイントは、コントロール時に比べて変化は少なく、逆に、全部口蓋床装着時に大きくなった。一方、キ、チ、ツでは、口蓋床装着時の認識ポイントは、コントロール時に比べ全般的に大きな値を示した。そこで、口蓋床と被験音を主因子とする二元配置分散分析を行ったところ(表4)、口蓋床と被験音の効果はともに信頼率95%で有意であった。

3. 聴覚テストによる語明度

聴覚テストによる語明度を表5に示す。

口蓋床なしの状態、すなわち、コントロール時の語明度は平均で88.2%であった。部分口蓋床を装着したときは、前方床でサのみが45.5%と、他の被験音に比べ語明度の低下が大きく、中間床ではキ、ヒ、リの低下が、後方床ではキの低下が大きかった。馬蹄形床では、サ、キ、ヒ、シ、チ、ミ、リ、ツと、ほとんどの語音で語明度が大きく低下した。全部口蓋床では、床が厚くなればなるほど語明度は低下する傾向にあり、特にサ、キ、ヒで大きく低下した。しかし、チの場合は、床の厚みが増すに従い逆に語明度は若干良くなる傾向にあり、ミでは45.5%に低下したが、床の厚みによる変化は認められなかった。そこで、口蓋床と被験音を主因子とする二元配置分散分析を行ったところ(表6)、口蓋床と被験音の効果はともに信頼率95%で有意であった。

IV. 考 察

1. 音声認識装置について

従来より、補綴物による発音障害を客観的に把握す

表 3 認識ボードによる認識ポイント

被験音	サ	タ	キ	ヒ	シ	チ	ニ	ミ	リ	ツ
コントロール	308	366	216	255	243	209	293	271	335	234
	50	79	31	31	37	34	64	61	28	26
前 方 床	377	404	261**	275	268	245**	402**	310	339	354**
	92	91	28	26	47	20	33	36	23	52
中 間 床	319	300	333**	285	263	267**	200**	254	303**	265*
	38	62	91	43	23	36	54	39	25	33
後 方 床	334	321	515**	340**	235	280**	231*	234	303	281
	36	51	74	89	11	58	47	35	68	71
馬蹄形床	312	341	393**	278	325**	318**	272	290	423**	271**
	24	73	104	52	55	100	60	49	93	37
全部口蓋床 (1 mm)	439**	450**	449**	258	279	308**	384	335**	394	343**
	92	49	136	30	41	40	161	41	101	58
全部口蓋床 (2 mm)	404**	544**	455**	356*	298**	300**	323	319*	368**	530**
	34	71	108	115	41	33	23	29	22	56
全部口蓋床 (3 mm)	385	412	413**	333**	259	338**	345	320	413**	353**
	111	79	107	84	29	119	44	51	75	56

上段 Mean

下段 S.D.

*は信頼率95%で有意

**は信頼率99%で有意

表 4 認識ポイントの分散分析表

因子	平方和	自由度	不偏分散	不偏分散比
A (主効果)	118470	7	16924	6.48 *
B (主効果)	133877	9	14875	5.70 *
E (残差)	164545	63	2612	
T (合計)	416892	79		

A : 口蓋床

* 信頼率95%で有意

B : 被験音

るためにパラトグラム⁴⁾、磁気センサー⁵⁾、光学センサー^{6,7)}、サウンドスペクトログラム⁸⁾、線形予測分析⁹⁾、聴覚検査法¹⁰⁾など、いろいろな検査法が考案され用いられてきた。しかし、いずれの方法も専門的な知識を必要とし、また、ある程度検査者の主観が入り込むなど問題視されていた。

一方、音声の認識および合成技術は大規模メモリ、LSIなどの半導体技術の進歩と相まって近年著しく発達してきた。しかも、これらの音声処理技術は音声によるマン・マシーン・インターフェースを可能にしてくれるため¹¹⁾、教育(訓練)、計測など工業部門の幅広い分野において応用価値が高く、より優れたシステムの開発をめざして研究がすすめられている¹²⁻¹⁴⁾。

しかし、音声認識は、一般に単語音声認識(単語のように1つひとつ区切って発音されたものを認識するもの)と連続音声認識(会話的な音声で、技術的には前者に比べて著しく難しい)の両方から研究が進められてい

るが、現在、単語音声で実用になっているものでも特定話者認識(話者を限定した場合)であれば数百単語程度、不特定話者用であれば数十単語程度の音声認識が限度であり^{15,16)}、まだまだ改良の余地が残されている。このような音声認識装置を用いた研究報告は、歯科領域ではすでに藤井らが全部床義歯装着者の音声について行っているが¹⁷⁾、当時の認識技術が現在ほど進歩していなかったこともあり、また、分析対象が母音に限られていたために著しい成果はあげられなかったようである。最近では、村上¹⁸⁾が、実験口蓋床が子音発音に及ぼす影響をみるために単音節音声認識ボードを用い、口蓋床の影響を受けやすい語音を抽出しているが、音声認識を発音検査法として利用した報告は現在のところ見当たらない。

そこで、われわれは、聴覚ほど複雑な判断はできないが、操作が簡単で、しかも、① サンプリング間隔の切換え(5 ms と 10 ms)などによってパターンの特徴を表す重要な子音部分を通常の割合よりも多く取り込むように情報圧縮を行っている、② 線形マッチングによる予備選択を行ったのち非線形(Dynamic Programming)マッチングで最終選択を行うため、従来型の認識装置に比べて約 1/80 程度に識別処理時間が短縮されている¹⁹⁾、などの多くの改良点をもつサンヨー社製音声認識装置 SRB-64A を用いて発音検査への応用の可能性について実験を行った。

表 5 聴覚テストによる語明度

被験者	サ	タ	キ	ヒ	シ	チ	ニ	ミ	リ	ツ	mean	S.D.
コントロール	95.5	90.9	81.8	68.2	95.5	90.9	95.5	90.9	86.4	86.4	88.2	7.9
前方床	45.5	81.8	81.8	72.7	90.9	90.9	90.9	90.9	90.9	81.8	81.8	13.5
中間床	90.9	100	36.4	45.5	81.8	90.9	100	100	63.6	90.9	80.0	22.2
後方床	90.9	90.9	27.3	63.6	100	81.8	100	90.9	81.8	90.9	81.8	20.7
馬蹄形床	54.5	100	45.5	9.1	54.5	72.7	90.9	54.5	63.6	63.6	60.9	23.7
全部口蓋床(1mm)	81.8	100	54.5	72.7	100	72.7	90.9	45.5	81.8	81.8	78.2	16.9
全部口蓋床(2mm)	72.7	90.9	0	27.3	100	81.8	63.6	45.5	100	54.5	63.6	31.0
全部口蓋床(3mm)	18.2	90.9	0	36.4	81.8	90.9	81.8	45.5	54.5	63.6	56.4	29.8

単位：%

表 6 聴覚テストの語明度の分散分析表

因子	平方和	自由度	不偏分散	不偏分散比
A(主効果)	9674	7	1382	5.02*
B(主効果)	21262	9	2362	8.59*
E(残差)	17332	63	275	
T(合計)	48269	79		

A：口蓋床

*信頼率95%で有意

B：被験音

2. 語音発語明瞭度について

聴覚テストによる語明度は、今回の場合、コントロール時で平均 88.2% と、通常報告されている数値よりも 10% 程度低い値を示した（表 5）。これは、聴取実験が ORDER 法¹⁾で行われたこと、また、検査者が被験者の肉声をよく知っていることなどの点からかなり厳しい評価がなされたものと思われる。また、評価基準を 4 段階に設定したことも判断を厳しくせざるを得なかった要因かも知れない。

表 7 に聴覚テスト時、誤聴傾向を認めた被験音と誤聴音について表示した。ただし、検査者がおかしいと評価しながらも誤聴音を記入しなかったものについては表から除外した。

一般に、口蓋床を装着した場合、装着部位付近に調音点をもつ子音の語明度が低下するといわれているが、その点、表中の誤聴傾向からもある程度口蓋床の種類を推測することができると考えられる。これに対して認識ボードの場合は、語明度のみで口蓋床の種類を判別することは不可能であった。また、コントロールにおいても 64.5% と低い認識率にとどまっていた。これは、今回使用したボードが、特に OA 機器の音声入力など、モーラ数が 2～5 程度の単語音声に最大の効果が発揮できるように実験的に組み込まれたボードであること¹⁹⁾、そのため、今回のように被験音が単音節（1 モーラ）の場合は音声パターンとなったときの情報量が乏しくなり、音声

認識が正確にできなくなったためと考えられる。

表 8 にボードの認識結果と判定語音を示した。誤認識音は後続母音が同一音に限られており、後続母音の認識に占めるウェイトが大きいことがわかる。また、同一後続母音を含む語音内では、キ、ヒ、チ、シ間、あるいは、ニ、ミ間での誤認識が多かった。つまり、調音点が近接していたり、調音様式が同じ音同士の誤認識が多く、聴覚とよく似た誤聴傾向を示した。

3. 認識ポイントについて

ボードによる認識ポイントと、聴覚テストによる語明度とを対比させながら考察した。

被験音サは、聴覚テストによる語明度は、前方床で 45.5%、馬蹄形床で 54.5%、3 mm の全部口蓋床で 18.2% と、比較的前方部の床と関連して低下する傾向にあった。一方、認識ポイントは、前方床装着時には有意ではないが大きくなる傾向が、また馬蹄形床装着時にはほとんど変化は認められなかった。全部口蓋床装着時にはすべての厚さにおいて認識ポイントは増大したが、その程度は、床の厚さが減るに従って逆にやや増大していた。

タは、語明度は、全体的にみて大きな変化は認められなかったが、認識ポイントは、全部口蓋床装着時に著しい増加を示した。

キは、前方床を除くすべての口蓋床装着時に語明度が低下し、特に後方床（27.3%）と、2 mm と 3 mm の全部口蓋床装着時（ともに 0%）に著しかった。認識ポイントは、すべての口蓋床装着時に有意に増大していたが、中でも、後方床装着時が最も大きく、前方床装着時は比較的小さな値を示した。この結果は発音時の舌接触部位とも関連し、聴覚同様、認識ポイントが口蓋床の影響を表しているものと考えられる。

ヒは、語明度は、馬蹄形床で 9.1%、2 mm と 3 mm の全部口蓋床で 27.3% と 36.4%、中間床で 45.5% と低

表 7 聴覚テスト時に誤聴した被験音と誤聴音について

被験音	サ	タ	キ	ヒ	シ	チ	ニ	ミ	リ	ツ
コントロール										
前 方 床	ソァ	ソ	バ		キシ	チ				
	2	1	1		1 1	1				
中 間 床				キ					ディ	
				1					2	
後 方 床									ディ	
									1	
馬蹄形床					シ			ニ		
					4			3		
全部口蓋床(1 mm)				ヒ	イ			ニ		
				3	2			2		
全部口蓋床(2 mm)	ツァ	サ		ヒ	イ		ビ	ニヒイ		トゥ
	1	1		5	4		1	2 2 1		2
全部口蓋床(3 mm)	ツァ		ヒビティ	ピ						トゥ
	4		5 5 1	1						1

上段：誤聴音

下段：回数／11回

表 8 認識ボードの認識結果と判定語音

被験音	サ	タ	キ	ヒ	シ	チ	ニ	ミ	リ	ツ
コントロール	タサ	タサ	キヒチ	キヒ	チシ	チシ	ニミキ	ミニリ	リチヒキ	ツ
	8,5	9,3	7,3,1	9,5	7,4	11,1	9,2,1	7,3,3	6,3,2,2	11
前 方 床	タサ	タサ	キヒ	キヒ	チシ	チシキ	ミキリ	ミニ	ヒリチ	ツ
	9,1	8,2	9,1	7,3	6,4	7,2,1	5,3,2	9,1	7,2,1	10
中 間 床	タサ	タ	ヒキチ	ヒチキリ	チシ	チシ	ニリ	ミニリ	リヒチキ	ツ
	6,5	11	6,4,3	6,4,1,1	6,4	6,5	9,1	7,2,2	6,2,2,1	10
後 方 床	サタ	タサ	キヒチリ	ヒチリ	シチ	チシ	ミニリ	ニミリ	リ	ツ
	8,3	10,1	7,3,2,1	6,3,2	10,1	7,6	6,5,1	5,3,3	11	13
馬蹄形床	タサ	タ	キヒチ	ヒチ	シチツ	チシ	ニミリ	ミリニ	リヒチキ	ツ
	7,4	11	6,4,1	9,2	8,2,1	7,5	5,4,2	9,4,1	5,3,3,1	11
全部口蓋床(1 mm)	タサ	タサ	キヒチ	キヒ	シチ	シチ	ミリニ	ミニリ	リヒ	ツ
	6,4	6,4	4,3,3	6,4	5,5	8,2	7,2,1	8,1,1	9,1	10
全部口蓋床(2 mm)	タサ	タサ	キヒチ	ヒキリ	シチ	チシ	ミニ	ミニリ	リヒキ	ツ
	7,3	7,3	6,3,1	8,1,1	5,5	7,3	7,3	7,2,1	8,1,1	10
全部口蓋床(3 mm)	タサ	タサ	ヒキチ	ヒリ	シチ	シチヒ	ミニリヒ	ミ	リヒチ	ツ
	8,2	5,5	7,2,1	9,1	7,3	7,2,1	7,1,1,1	10	5,4,1	10

上段：判定語音

下段：判定回数

い値を示したが、認識ポイントも同様に増大傾向を示していた。

シは馬蹄形床装着時に語明度が 54.5% と低下する以外、全体的に良く、認識ポイントも馬蹄形床以外、比較的低い値を示した。

チは、語明度は、馬蹄形床で 72.7%、1 mm の全部口蓋床で 72.7% と比較的良い結果が得られたが、認識ポイントは逆に、すべての口蓋床で有意に増大し、特に、馬蹄形床と全部口蓋床装着時に大きな値を示した。

ニは、語明度は、2 mm の全部口蓋床装着時に低い値

を示した以外は相対的に高い値であり、中間床と後方床では100%を示した。認識ポイントは、前方床では増加に有意性が、全部口蓋床では増加傾向が認められたが、中間床と後方床では逆に有意に減少し、聴覚と良く対応していた。

ミは、馬蹄形床と全部口蓋床装着時に語明度が54.5%、45.5%と低くなったが、認識ポイントも同床形態において大きくなる傾向が認められた。

リは、語明度は、中間床(63.6%)、馬蹄形床(63.6%)、3mmの全部口蓋床(54.5%)装着時に特に低く、認識ポイントは、馬蹄形床、3mmの全部口蓋床装着時に有意に増大し、逆に、中間床装着時には有意に小さくなった。また、2mmの全部口蓋床装着時の認識ポイントが有意に増大しているのに対して語明度は100%であるなど、語明度と認識ポイントの間には判断上のズレが多く認められた。これは、リのような弾音の場合、発音の難易度が高く¹⁾、また、調音状態が微妙で個人差が大きい²⁰⁾こと、単音節のために破裂音「ディ」に誤聴されやすかった(表7)ことなどが原因と考えられる。

ツは、語明度は、馬蹄形床で63.6%、2mmと3mmの全部口蓋床で54.5%と63.6%となり、他に比べ低くなったが、全体的に高い値を示した。認識ポイントはほとんどすべての口蓋床で有意に大きくなり、特に、2mmの全部口蓋床装着時に最大値を示した。

以上のように聴取実験による語明度と、ボードによる認識ポイントとの間には多少とも判定上での意見の相違が認められる。これは、聴取実験の場合、①主観によって音韻性を判断することの難しさ、特に、検者が特別な訓練を受けていない場合に困難であること、②被験音を単音節にした場合、短いために聴き間違える可能性があることなど、聴取方法に問題点があるようである。一方、ボードの場合は、①単音節であるための情報量の不足からくる認識エラーの混入、②発声ごとに音声パターンの変動が現れるなど、本研究の結果に影響を及ぼす因子は比較的多い。しかし、認識ポイントと調音点ならびに口蓋床との関連性などを検討した場合、比較的、聴覚との対応が良く現れており、また逆に、聴覚以上に良く判断していると思われる結果なども認められ、検査用語彙の改良によって発音検査に本ボードの使用が有効であることが示唆された。

4. 周波数分析からみたボードの特性

検査用語彙の改良にあたっては、ボードの特性を知り、それに適した検査用語彙を設定する必要がある。そ

図5 キ発音時の時間—スペクトルパターン

こで、コントロール時に比べ認識ポイントが大きく悪化した後方床装着時のキについて周波数分析を行い、線形予測法²¹⁾で求めた時間—スペクトルパターン(図5)からボードの特性を検討した。なお、分析にあたっては、音声を遮断周波数7kHz、48dB/octの低域通過フィルターに通した後、サンプリング周波数14kHz、12bitで量子化し、コンピュータ(TEAC社製、PS9000モデル300)のCRT上で視察により子音部、母音部のセグメンテーションを行い、子音部では分析窓長8ms、シフト間隔5ms、分析次数7次、母音部では分析窓長25ms、シフト間隔10ms、分析次数18次の自己相関法に基づく線形予測法でスペクトルを求めた。なお、分析窓はハミングの窓を用い、母音部においてのみ一次差分による高域強調を施した。標準化周波数と時間分解能は本ボードと一致させた。図4に示すようにコントロール時のキでは、子音部/k/の第1ピークは約3kHz付近にみられ、後続母音/i/の第3フォルマントにやや周波数の下降をみせながら遷移しているのがよくわかる。これに対して、後方床装着時のキでは、子音部/k/の第1ピークは音圧が上昇し、周波数がかなり低下している。そして、このピークはやや上昇しながら後続母音/i/の第3フォルマントに遷移している。堀内は、キの子音部では第1ピーク周波数が最も重要で、歪音の場合、この第1ピーク周波数が低下することを報告している⁹⁾が、本研究においても周波数パターンに同様のことが観察された。また、母音部においても、第1、第2フォルマント周波数は両者の間に大きな差は認められなかったが、子音部から第3フォルマントへの遷移においては、その様相が両者間で大きく異なっていた。また、第4フォルマントが歪音のキでは明瞭になっていた。

以上の結果ならびに、今回の実験結果から後続母音の誤認識がきわめて少なかったことを考え合わせると、このような母音部の変化は、ボードの分析では、子音部の変化と比較した場合、あまり意味はもたないように思われる。つまり、本ボードは、子音部の変化を比較的良好

とらえることのできる特性を有しており、それ用の検査用語彙を選択する必要があると考えられる。

5. 検査用語彙の検討

この種の認識装置を発音検査に応用する場合、まず問題となるのはパターンの安定性についてであろう。音声は、発声のたびごとにいつも同じパターンで発声されるとは限らないし、また、経時的にも変動する。特に、パターンの変動は、時間方向、周波数方向、振幅方向に現れる。時間方向の変動は母音部が伸縮しやすく、子音部があまり伸縮しないという非線形な性質をもっており、本ボードの場合はダイナミック・プログラミングによってこの変動を吸収する³⁾ ようになっている。しかし、それでも可及的にこの変動を小さく押えられるような被験音を選択することが望ましい。また、周波数方向の変動は、同一話者であればほとんど無視できるが、パターンのもつ情報量を増やすためにも被験音は、単音でなく、平易な連続音や、良く慣れ親しんだ単語が良いと思われる。残る振幅方向の変動に対しては、線形な伸縮でもってある程度補うことは可能である³⁾ が、被験者がリラックスできるように検査方法にも工夫が必要である。

以上のことから、今後、適当な検査用語彙と検査方法を設定すれば、音声認識装置を用いて口腔内環境の変化に伴う発音の障害度を、客観的な物理量として表現することが可能なことが示唆された。

V. 結 論

音声認識装置を利用した発音検査法の開発を目的として、成人有歯顎者 1 名に実験口蓋床を装着させ、音声認識装置の有効性を聴覚検査法と比較しながら考察したところ、以下の結論を得た。

1. 音声認識装置は聴覚と類似した誤聴傾向を示した。
2. 認識ポイントは、聴覚による語明度と比較的よく対応していたが、被験音によっては一致しない場合もあった。
3. 音声認識装置は、子音部の変化を比較的良好にとらえることのできる特性を有していた。

以上のことから、今後、音声認識装置に適した検査用語彙を選択し、検査方法を一定化すれば、口腔内環境の変化に伴う発音の障害度を客観的な物理量として表現することが示唆された。

本論文の要旨の一部は、第 4 回四国歯学会 (1984 年 12 月、徳島大学) において発表した。また、本研究の一部は、文部省科学研究費補助金 (課題番号 01771650) により行われた。

文 献

- 1) 松木教夫：全部床義歯患者の発音の研究 (その 1) 自覚的障害について、(その 2) 他覚的障害について、口病誌, 38: 252~265, 333~362, 1971.
- 2) 山縣健佑：歯科治療と発音, 日歯医師会誌, 34: 705~709, 1981.
- 3) 大西宏樹, 飯田正幸, 宮武正典ほか：音声認識ボードの開発, 三洋電機技報, 14(2): 20~26, 1982.
- 4) 市場裕康：全部床義歯装着者の S 発音時の舌の動態に関する研究, 補綴誌, 32: 787~797, 1988.
- 5) 堂脇一朗, 佐藤裕二, 松田一雄ほか：実験的口蓋床装着者 (有歯顎者) の発音時における舌の運動特性について, 広大歯誌, 19: 285~289, 1987.
- 6) 園田頼信, 許斐俊治, 守 啓祐：光学センサによる舌の調音観測, 信学技報, SP 87-58, 1987.
- 7) 市川哲雄, 薦田淳司, 堀内政信ほか：光センサによる軟組織変位計測システムの開発とその応用, 補綴誌, 33: 582~588, 1989.
- 8) 山縣健佑：発音試験用標準日本語彙に関する研究, 補綴誌, 8: 173~217, 1964.
- 9) 堀内政信：人工口蓋床の形態が発音に及ぼす影響に関する研究, 補綴誌, 33: 863~877, 1989.
- 10) Hisatoshi, T.: Influence of the prosthesis on speech articulation, 九州歯会誌, 37: 304~322, 1983.
- 11) 鈴木誠史：日本の音声研究—来し方, 行く末—, 日本音響学会誌, 45: 886~889, 1989.
- 12) 藤本潤一郎：ファジィ音声認識, 数理科学, 284: 63~69, 1987.
- 13) Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE, 77: 257~286, 1989.
- 14) Waibel, A., Hanazawa, T., Hinton, G. et al.: Phoneme recognition using time-delay neural networks, IEEE Trans. ASSP, 37: 328~339, 1989.
- 15) 斉藤収三：新世代の音声研究への期待, 日本音響学会誌, 45: 880~883, 1989.
- 16) 中津良平：音声認識・合成技術の製品化及び市場動向, 信学技法, SP 89-103, 1990.
- 17) 藤井輝久, 鈴木松美, 北本三郎ほか：或る総義歯患者の発音について, 岐歯学誌, 4: 385~387, 1976.
- 18) 村上 賢：子音の発音に及ぼす口蓋床の影響, 日矯歯誌, 44: 435~448, 1985.

- 19) 宮武正典, 大西宏樹, 田島 洵: DP マッチングの時間短縮を行った単語音声認識装置, 音声研究会資料, S 83—55, 1983.
- 20) 関根 弘, 阿部 勤, 江崎梅太郎ほか: 義歯の音声学的研究 第1報 口蓋床による調音障害について, 補綴誌, 5: 175~182, 1961.
- 21) Markel, J.D. and Gray, Jr.A.H. (鈴木久喜訳): 音声の線形予測, コロナ社, 東京, 1981.