

# 聴覚障害者の発語に関する音響音声学的検討

吉野 公喜

日本語の言語記号システムを獲得している生来性の高度あるいは重度感音難聴者の特定発語における音響音声学的解析を試みた。平均聴力レベルが105dB以上の重度難聴者にとっては、時間的構造及び基本周波数とも、高度難聴者、健聴者と異なった超分節的特徴を示していた。母音調音図にみられる分節的特徴では、聴力レベルが95dBの重度難聴者であっても、聴覚的弁別及び聴覚的識別を十分に高めている場合には、音韻照合と聴覚的フィードバックとによって調音運動の平滑化を実現し得ることが確かめられた。

キーワード：基本周波数 音声情報処理システム ホルマント 声道 母音環境

## 1. はじめに

生来的に聴覚的フィードバックを欠くか、その確立が十分ではない聴覚障害者の多くは、分節的特徴 (segmental feature) あるいは超分節的特徴 (suprasegmental feature) において、健聴者とは著しく異なった発声 (phonation) や調音 (articulation) を余儀なくされている。聴覚障害者にみられる、このような発声や調音の不明瞭性は時として deaf speech と呼ばれることがあり、その特徴は、次のように記述されることが少なくない (Hudgins & Numbers, 1942<sup>3)</sup>, House & Fairbanks, 1953, Angelocci, et al. 1964<sup>1)</sup>, Nickerson, 1975<sup>6)</sup>)。

①摩擦音の破裂音化 ②濁音の清音化 ③母音の中性音化 (neutralization) ④各種子音の母音化 ⑤語尾音の脱落 ⑥過度の鼻音化 ⑦息つぎの不安定さ ⑧リズムパターン、ピッチパターン及びイントネーションパターンの単調さ

しかしながら、近年の聴覚補償にみられる医学的、音響工学的、教育的貢献は、聴覚障害者の音声表出の実態を大きく変えてきている。

Markides (1976<sup>5)</sup>) は、重度聴覚障害児 (profoundly deaf children) の発語の明瞭さを7段階評定尺度法で評定し、彼らの77%までが「やや困難を覚えるがどうにかわかる」レベル以上に達し得ることを報告し、聴覚-口話 (auditory-oral) による教育的改善の期待できることを示してい

る。

## 2. 目的

本研究は、幼児期の早期より相応の聴能指導及び言語指導を受け、日本語の言語記号システムを獲得している高度 (severely) あるいは重度 (profoundly) 感音難聴者の特定発語 (特定の環境音に規定された語音の発語) にみられる発話音声の音響音声学的特徴を明らかにすることにある。

## 3. 方法

(1) 被験者 被験者は、良聴耳の平均聴力レベルが70dB (HS<sub>1</sub>), 95dB (HS<sub>2</sub>), 105dB (HS<sub>3</sub>) の両側感音難聴者3名と健聴者2名 (NS<sub>1</sub>, NS<sub>2</sub>) の計5名である。被験者は、実験時すべて筑波大学に在籍していた者であり、20~25歳の成人男性である。被験者のうち、HS<sub>1</sub>, HS<sub>2</sub> は、通常聴覚を主に、読話を補助的に用いてコミュニケーションを行っており、HS<sub>3</sub> は、口話及び手指を活用してコミュニケーションを成り立たせている。HS<sub>1</sub>, HS<sub>2</sub>, HS<sub>3</sub> の聴力レベル及び語音聴取閾値、語音弁別値 (日本オーヂオロジー学会, 67語表) は、Table 1に示すとおりである。HS<sub>1</sub> と HS<sub>2</sub> は、就学前に東京教育大学 (現筑波大学) 附属聾学校で、聴能教育及び言語指導を受け、その後普通小、中学校に在籍したものである。一方、HS<sub>3</sub> は、N聾学校で幼時期に言語指導を含めた口話教育を受

け、その後小学校から普通校に籍を置いて、教育を受けてきたものである。健聴者を含め、被験者は、すべて関東、東海地区で幼児期より生活し、現在にいたっているものである。

(2) 手続き 言語音の環境が一定である  $V_1PV_2$  音節 (/ipi/, /epe/, /apa/, /opo/, /upu/) を2度発語させ、2度目の音節を音声サンプルとする。音声サンプルは、防音室内で、被験者の口前15cmにセットされたマイクロホン (SONY ECM290E) を通して、テープレコーダ (SONY TC8750-2) に、テープ走行38cm/秒で録音された。音声サンプルを得るための検査語は、カード (15

×20cm) に書かれ、被験者に呈示された。

各音声サンプルは、音声切り出し編集装置 (DANAC 7000)、音声情報処理システム (B&K2032, HP216, HP7470A, HP2671G) を介して、波形解析及び処理がなされた (Fig. 1)

2チャンネルFFTアナライザー (Dual Channel Signal Analyzer; B&K2032) による波形解析は、取り込み時間125ms, 時間分解能 $61.0\mu\text{s}$ , 分析周波数6.4KHz, 周波数分解能8Hzが, その設定条件とされた。A/D変換された各音声サンプルは、initial vowel ( $V_1$ ), final vowel ( $V_2$ ) 毎に、それぞれ33個 (97%アレー重複) のケプストラム\*

Table 1. Audiological outlines of hearing impaired subjects.

Subject	Age	Pure Tone Average	Hearing Level (dB: JIS-1982)						MSDS * (%)	Tested Ear **
			250	500	1000	2000	4000	8000 (Hz)		
HS 1	20	R 71.3	60	65	70	80	70	75	95	R
		L 70.0	55	60	70	80	75	75	75	
HS 2	25	R 96.3	80	90	100	100	90	85	95	R
		L 95.0	70	85	100	95	85	70	95	
HS 3	25	R 107.5	75	90	110	120	120	NR	—	R
		L 105.0	75	90	110	110	120	NR	—	

\* MSDS: Most speech discrimination score.

\*\* Tested Ear: Aided Ear.

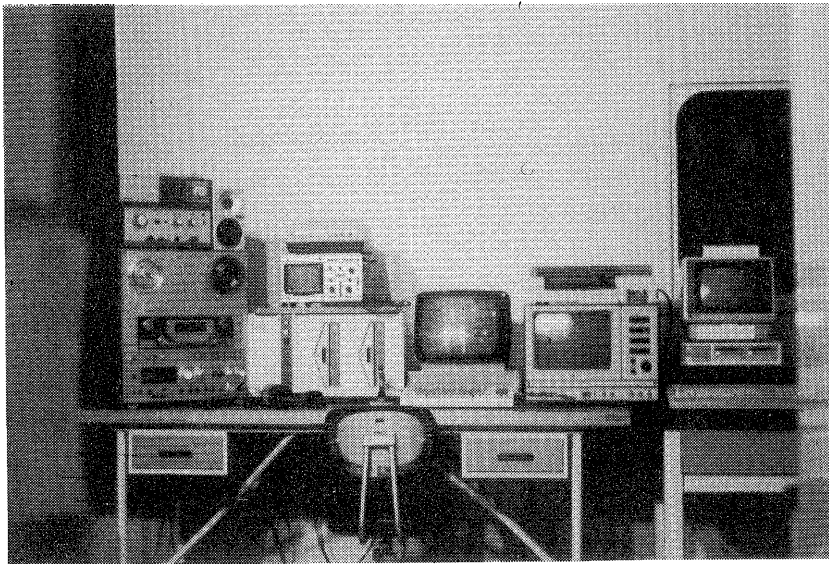


Fig. 1. Sound wave editing system and speech sound analyzing system.

(Cepstrum) 及びリフトードスペクトラム (Liftered Spectrum) として解析, 格納された。

音声サンプル /apa/ の波形の時間構造 (TD; Total Duration,  $V_1D$ ; Initial Vowel Duration, CD; Closure Duration, VOT; Voice Onset Time,  $V_2D$ ; Final Vowel Duration) と信号の開始から第10番目のリフトードスペクトラムは, Fig. 2のように示される (Fig. 2)。Fig. 3は, 健聴者 ( $NS_2$ ) の /apa/ の  $V_1$  及び  $V_2$  をリフトードスペクトラムに基づいて, 2次元重ね書きで表示したものである。音声サンプルのホルマント中心周波数は, これら33個の各ホルマントの平均をとったものである。

Fig. 3-1は, [apa] における initial vowel,  $V_1$  として Fig. 3-2は, [apa] の final vowel,  $V_2$  の時間的経過にみる周波数スペクトラムの2次元重ね書きの表示例である。横軸は周波数, 縦軸は音声開始時からの持続時間に沿った各ページにおける周波数スペクトラムの相対的強さ (log scale) を表している。

\* 時系列信号の対数パワースペクトルのフーリエ逆変換として定義されるパワーケプストラム

#### 4. 結果及び考察

(1) 時間的構造にみられる特徴

Table 2は,  $HS_1$ ,  $HS_2$ ,  $HS_3$  及び健聴者 ( $NS_1$  と  $NS_2$  の平均) の音声サンプルの時間構造を示したものである。 $V_1/P/V_2$  に占めるセグメント ( $V_1$ , CD, VOT,  $V_2$ ) の持続時間及び全持続時間 (TD) は, 母音それぞれによっていくぶん異なるが, 5母音の平均でみると, TD に関して  $HS_3$  は,  $HS_1$ ,  $HS_2$  よりも著しく短いことが認められる。 $HS_3$  は, TD で比較する時, むしろ健聴者に近似している。 $V_1$  と /P/ の閉鎖の解放部との間の無音区間である

Closure Duration (the silent period between the last glottal pulse of the initial vowel and the burst of the stop<sup>9)</sup>; CD) は, 聴力レベルが95dB以上の  $HS_2$ ,  $HS_3$  にあって著しく短くなっている。聴力レベルが70dBの  $HS_1$  の CD は, 140.8ms であり, 健聴者の CD (128.0ms) と類似した値を示している。健聴者及び  $HS_1$  の CD は, ともに  $V_1D$  とほぼ同じ値をとっており,  $V_1PV_2$  構造において, CD の十分に長いことが, 超分節的特徴の一つであるリズムパターンの構成に寄与していることが考えられる。無声の両唇破裂である /P/ の VOT (the period between the burst and the first

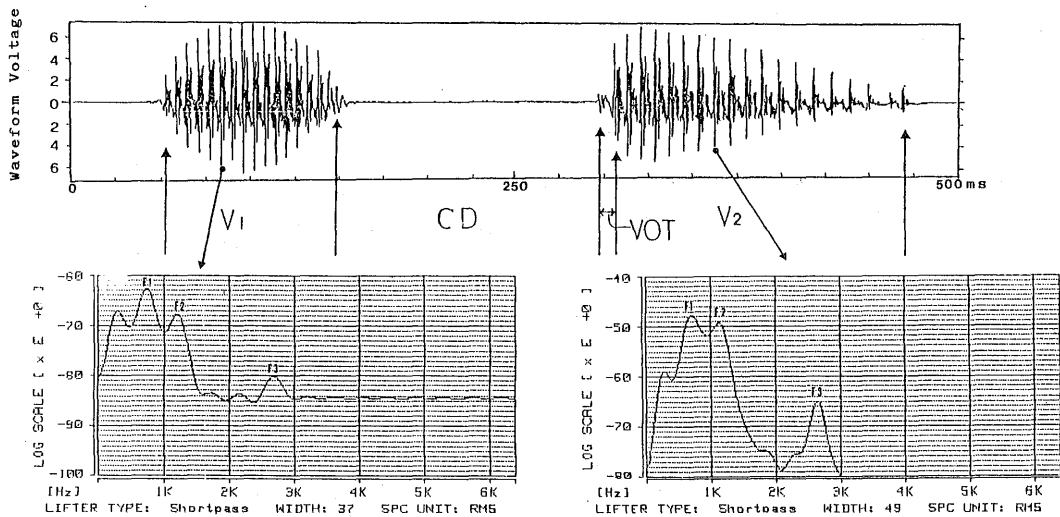


Fig. 2. Sample waveform of /apa/, with measurement boundaries indicated by arrows and operationized intervals for four segments ( $V_1$ , CD, VOT,  $V_2$ ).

Frequency spectra of the bottom column are liftered spectrum of initial vowel and final vowel respectively.

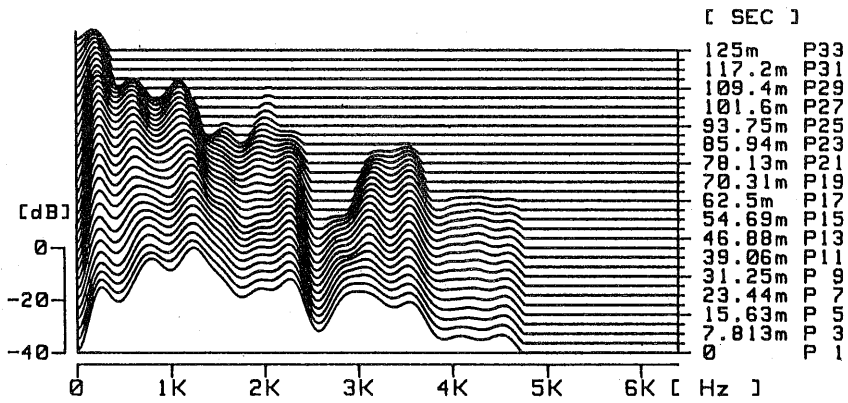


Fig. 3-1. Time-variant frequency spectrum of Initial Vowel (Initial /a/ in [apa])

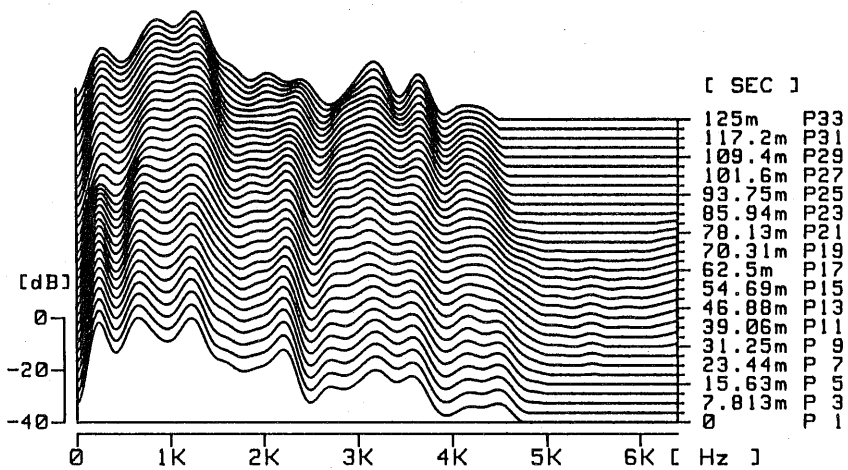


Fig. 3-2. Time-variant frequency spectrum of Final Vowel (Final /a/ in [apa])

Table 2. Mean durations and SDs for four segments (V1, CD, VOT, V2) of five V1/P/V2 syllables uttered by subjects.

	TD ( )	V1D ( )	CD ( )	VOT ( )	V2D ( )
HS 1	451.6 ms (41.8)	151.3ms (9.66)	140.8 ms (21.24)	25.0ms (8.01)	134.5 ms (15.35)
HS 2	457.0 (39.2)	198.3 (17.07)	81.9 (6.27)	15.3 (4.66)	161.5 (21.25)
HS 3	362.2 (19.18)	119.0 (3.98)	79.4 (9.58)	13.3 (5.90)	150.6 (14.95)
Normal	386.6 (15.0)	110.5 (8.33)	128.0 (19.83)	14.5 (2.70)	133.5 (15.97)

glottal pulse of the final vowel) は、HS<sub>1</sub> にやや長さがみられるものの、被験者間にほとんど差はみいだされない。V<sub>2</sub>D の TD に占める割合をみると、HS<sub>3</sub> のそれは、0.42 と大きな値を示し、HS<sub>1</sub> (0.30), HS<sub>2</sub> (0.35), 健聴者 (0.35) のいずれよりも、final vowel をひきのばして調音していることが認められる。HS<sub>3</sub> にみられるように、聴力レベルが105dB と重度であって、発話において聴覚系のフィードバックをほとんど欠如している場合には、CD が短かく、V<sub>2</sub>D が長くなることが考えられる。健聴者及びほとんど聴覚系に依拠して oral communication を高めている聴覚障害者にとっては、CD が十分に長く、V<sub>2</sub>D が相対的に短くなることのみみだされる。

### (2) 基本周波数にみられる特徴

基本周波数 (fundamental frequency; F<sub>0</sub>) は、一般に声帯音源 (voice source) における 1 秒間あたりの声帯の振動数をさす。基本周波数は、声帯を通過する気流の量、速度及び声帯筋群の緊張の度合によって影響を受ける。発声や調音にみられる音源波の基本周波数の規則的な変化は、アクセントや抑揚に関する情報の荷ない手になり、知覚上、ピッチ感覚を生じさせるものである。

Table 3 は、被験者個々における initial vowel (V<sub>1</sub>) 及び final vowel (V<sub>2</sub>) の母音毎の基本周波数を表したものである。基本周波数の抽出には、波形信号の第 1 ページ目のケプストラムが用いられた。

健聴の成人男性、NS<sub>1</sub> 及び NS<sub>2</sub> の 5 母音の平均基本周波数は、V<sub>1</sub> で 157.9Hz、V<sub>2</sub> で 126.1Hz となっている。/V<sub>1</sub>PV<sub>2</sub>/ である環境音の自発的発話における V<sub>1</sub> の基本周波数は、母音の単独発話における値よりも高いことが認められる。V<sub>1</sub> の F<sub>0</sub> と V<sub>2</sub> の F<sub>0</sub> では、一貫して V<sub>1</sub> の F<sub>0</sub> がより高い数値を示している。無声破裂音 [P] に続く V<sub>2</sub> の F<sub>0</sub> は、

/a/ 及び高舌母音である /i, u/ において高く、/e, o/ で低くなっている。

各母音と F<sub>0</sub> の関係において、無声破裂子音と接続する母音の場合、/i, u/ において F<sub>0</sub> が高いことは、英語圏での報告 (House & Fairbanks, 1953<sup>2)</sup>) と一致するところであるが、/a/ に関しては、英語圏の /æ/, /a/ は両方ともに 122.0Hz、121.0Hz と低く、日本語母音と逆になっている (Peterson & Barney, 1952<sup>7)</sup>)。

聴覚障害者の F<sub>0</sub> は、HS<sub>2</sub>, HS<sub>3</sub> にあって比較的高いことがみられるが、聴力レベルあるいは聴能との一義的關係は認められない。HS<sub>1</sub>, HS<sub>2</sub> の F<sub>0</sub> は、5 母音間で変動が大きく、健聴者にみられた V<sub>1</sub> の F<sub>0</sub> が V<sub>2</sub> の F<sub>0</sub> よりも一貫して高くそして /a/, /i, u/ の F<sub>0</sub> が /e, o/ の F<sub>0</sub> よりも高いという傾向はみいだされなかった。高度あるいは重度聴覚障害者は、聴能をかなりなまでに高め、発声及び調音機能を定着させてはいるものの、母音調音と F<sub>0</sub> との関係において一貫性あるいは規則性を確立させることの困難さをうかがわせる。

### (3) 母音調音図 (F<sub>1</sub>-F<sub>2</sub> 図) にみられる特徴

母音ホルマント周波数は、声道 (vocal tract) の長さ、声道にみられるせばめ (constrictions) の位置や狭さの程度に影響される。すなわち、母音ホルマントは、口唇の開き加減、舌の高さ等の調音運動の際の声道の型 (actual vocal tract shape) による伝達関数を反映する (Pickett, 1980<sup>8)</sup>)。一般に、各母音の母音らしさ (韻質) は、声道内の共振が強く表われる第 1 ホルマント (F<sub>1</sub>) と第 2 ホルマント (F<sub>2</sub>) の周波数帯域で規定される。

Fig. 4 は、/V<sub>1</sub>PV<sub>2</sub>/ 環境で発せられた日本語 5 母音の V<sub>1</sub> 及び V<sub>2</sub> の被験者毎の母音調音図 (F<sub>1</sub>-F<sub>2</sub> 図) である。Fig. 4-1 は、V<sub>1</sub> の F<sub>1</sub>-F<sub>2</sub> 図であり、Fig. 4-2 は、V<sub>2</sub> の F<sub>1</sub>-F<sub>2</sub> 図である。図中

Table 3. Fundamental frequency (first page) of initial and final vowels of subjects.

	i /p/ i	e /p/ e	a /p/ a	o /p/ o	u /p/ u
HS1	135.4 Hz 141.2 Hz	146.3 Hz 145.0 Hz	140.0 Hz 137.7 Hz	150.3 Hz 136.5 Hz	138.9 Hz 141.2 Hz
HS2	244.6 244.6	227.6 182.0	224.5 178.1	178.1 160.6	230.8 244.6
HS3	182.0 109.2	173.4 120.5	160.6 134.3	167.2 135.4	180.1 124.1
NS1	157.6 133.2	163.9 127.0	143.7 142.5	159.1 118.7	151.7 135.4
NS2	162.2 123.2	160.6 117.0	160.6 127.0	157.6 113.0	163.9 124.1

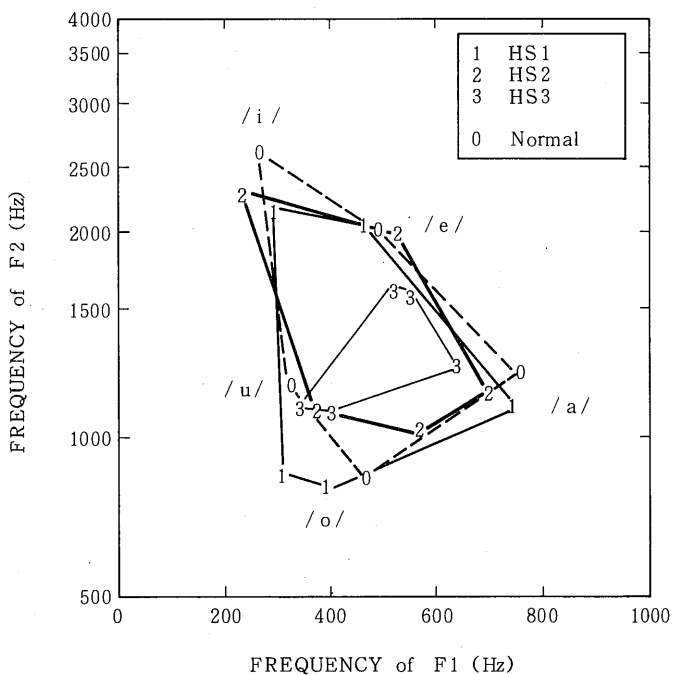


Fig. 4-1. Frequency of the second formant versus frequency of the first formant for initial vowel in V/P/V.

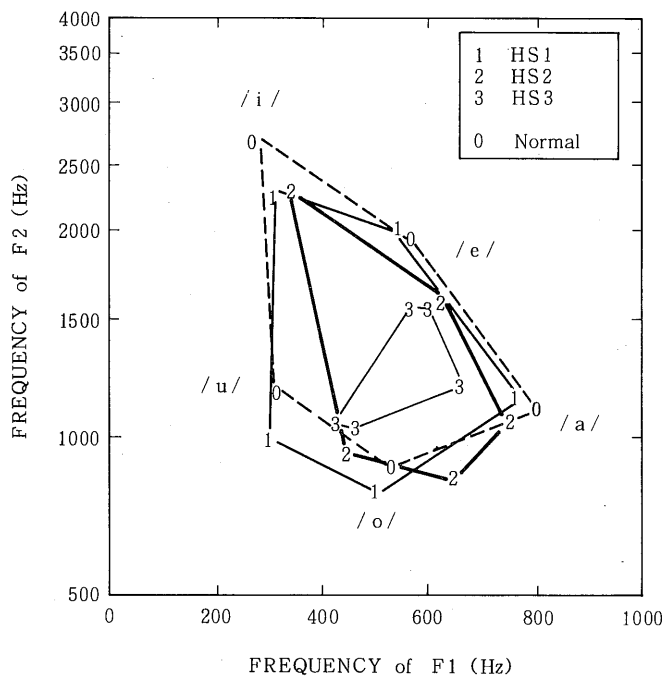


Fig. 4-2. Frequency of the second formant versus frequency of the first formant for final vowel in V/P/V.

の横軸は、第1ホルマント中心周波数を、そして縦軸は、第2ホルマント中心周波数をプロットしたものである。縦軸表示は、1000Hzまでを横軸と同じく等間隔に、1000Hz以上を対数尺度で表したものである (Kenigの軸: Kenig, et al. 1946<sup>4)</sup>)。

F<sub>1</sub>-F<sub>2</sub>図にみられる各母音の占める位置は、健聴者NS<sub>1</sub>(図中の記号O)にあって、母音間の範疇的空間距離がきわめて大きなことを示している。

HS<sub>1</sub>, HS<sub>2</sub>のF<sub>1</sub>, F<sub>2</sub>の相対的位置は、V<sub>1</sub>, V<sub>2</sub>とも健聴者であるNS<sub>1</sub>のそれと近似している。HS<sub>1</sub>, HS<sub>2</sub>とも語音弁別(67語表)は、90%以上の値をとり、周波数の弁別における相対的別閾(漸近値)を健聴者のそれに近似させているものであってみれば(図5)、語音弁別値が0%と、語音の聴覚的識別が不可能なHS<sub>3</sub>の母音調音図とは異なってくるものであろう。HS<sub>3</sub>にあっては、/i/が/e/に近く、しかも/i/も/e/も第2ホルマントの低音域への下降化の著しいことが認められる。

/i/, /e/は、日本語母音の/e/よりも、むしろ英語の/æ/に近似しており、母音の中性音化(vowel neutralization)がみいだされる。/a/は、健聴者の/a/に範疇化されるホルマント構造を示しているが、/o/はほとんど健聴成人の/u/に近く調音していることがみられる。/i/, /e/にみられる第2ホルマントの下降化と中性母音化は、V<sub>1</sub>と同様にV<sub>2</sub>にも認められる。

平均聴力レベルが95dBのHS<sub>2</sub>は、HS<sub>1</sub>と同様に、言語音に対する弁別能及び識別能を十分に高めており、音韻照合と聴覚・音声フィードバックによって、調音運動の平滑化を実現している。これに対して、HS<sub>3</sub>は、母音調音における運動プログラムは可能であっても、調音に際してのoral tract形状及び舌運動の平滑化に困難のあることを示している。

## 5. まとめ

幼児期より、聴覚補償ならびに言語指導を受け、日本語の言語記号システムを獲得している高度あ

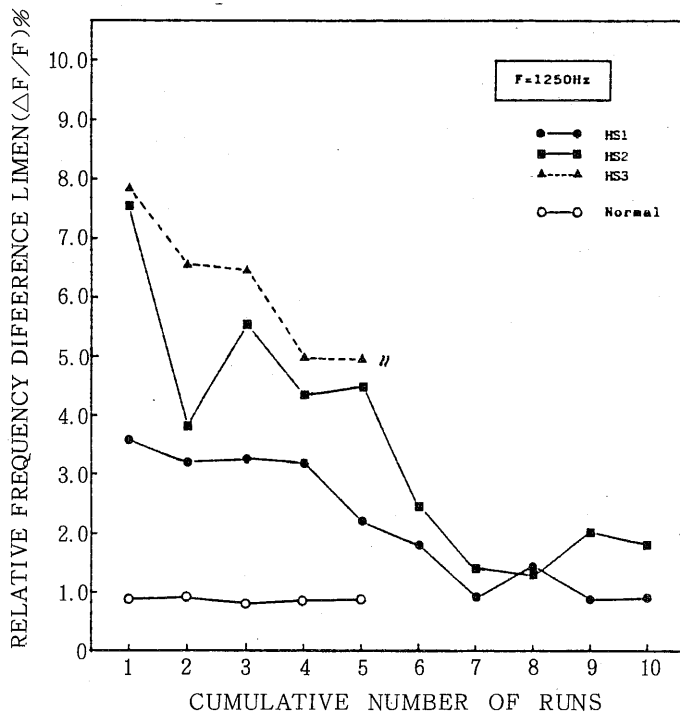


Fig. 5. Asymptotic effect of performance in frequency discrimination test at 1250 Hz. Relative frequency difference limen ( $\Delta F/F$ ) is plotted as dependent on practice sessions.

るいは重度感音難聴者3名を対象に、/VCV/音節を発語させ、音声切り出し編集装置及び音声情報処理システムによって、彼らの音響音声学的特徴の解析を行った。

(1) 時間構造にみられる特徴

/V<sub>1</sub>PV<sub>2</sub>/に占めるセグメントのうち、CDは聴力レベルが95dB以上の重度難聴者において著しく短くなっていた。健聴者と聴力レベルが70dBの高度難聴者のCDは、V<sub>1</sub>の持続時間と近似するほど十分に長かった。聴覚的フィードバックの確立が困難な重度難聴者においては、V<sub>2</sub>DのTDに占める割合がきわめて大であった。

(2) 基本周波数にみられる特徴

聴力レベルが95dB以上の重度難聴者の基本周波数(F<sub>0</sub>)は、健聴者及び高度難聴者と比較して高いことが認められるが、聴力レベルとF<sub>0</sub>の値は、必ずしも一義的ではない。高度及び重度難聴者のF<sub>0</sub>は、5母音間で大きな変動を示していた。

(3) 母音調音図にみられる特徴

聴力レベルが95dB以上の重度難聴者であっても、聴能を高め、聴覚的弁別、聴覚的識別に高い能力を示している症例においては、母音調音を健聴者と同様に範疇化し得ることが、音響音声学的に確かめられた。聴覚的フィードバックを欠く場合には、母音調音における運動プログラムは可能であっても、口腔の形状と舌運動の平滑化に困難のあることが示された。

## 文 献

1. Angellocci, A. A., Kopp, G. A., and Holbrook, A. (1964): The vowel formant of deaf and normal hearing eleven-to fourteen-years old boys. *J. S. H. D.*, 29; 156-170
2. House, A. S. and Fairbanks, G. (1953): The influence of consonant environment upon the secondary acoustical characteristics of vowels, in Lehiste, I. Ed. *Readings in Acoustics*, 1967, The MIT Press, 128-136
3. Hudging, C. V. and Numbers, F. C. (1942): An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 25, 289-392
4. Koenig, W., Dunn, H. K., and Lacy, L. Y. (1946): The Sound spectograph. in Lehistie, I. Ed. *Readings in Acoustic Phonetics*. Mass.: MIT Press, 1969
5. Markides, A. (1976): Comparative linguistic proficiencies of deaf children taught by two different methods of instruction manual versus oral., *The teacher of the Deaf*, 74: 307-316
6. Nickerson, R. S. (1975): Characteristics of the speech of Deaf Persons, *The Volta Review*, 77; 342-362
7. Peterson, G. E., and Barney, H. L. (1952): Control methods used in a study of the identification of vowels. *J. Acoust. soc. Am.* 24; 175-184
8. Picket, J. (1980): *The Sounds of Speech Communication*. Baltimore: Park Press.
9. Weismer, G., and Cariski, D. (1984): On speaker's abilities to control speech mechanism output: theoretical and clinical implications. In Lass, N. T. (Ed.) *Speech and Language; Advances in Basic Research and Practice*.



## Summary

### Acoustic-Phonetic Characteristics of Vowels Utteranced under VCV Environment of Deaf Persons.

Tomoyoshi Yoshino

Congenital profound deafness prevents the normal acquisition of speech. General effect of deafness on segmental and suprasegmental aspects of conventional spoken language is not always dependedt upon the severity of the hearing level that is estimated by pure tone audiometry.

The amount and quality of speech training, the use of suitable amplification device and the intensive auditory training are of very important for the intelligible speech production.

The relationship between auditory identification and speech production may be probably inter-reciprocal.

The purpose of this article is to present some of the findings on characteristics of the speech of hearing impaired speakers who are different, in the auditory discrimination and identification ability for speech sound.

Subjects are three severely and profoundly prelingual hearing impaired adult males (HS<sub>1</sub>, HS<sub>2</sub>, HS<sub>3</sub>) and nomal hearing persons (NS<sub>1</sub>, NS<sub>2</sub>). They were undergraduate and graduate students in University of Tsukuba in Japan. Table 1 and Fig. 1 show the hearing level (JIS-1982) and the most speech discrimination score of subjects.

The speech samples utteranced by subjects were /ipi/, /epe/, /apa/, /opo/, /upu/. These samples were analyzed by sound wave editing system (DANAC 7000), speech sound analyzing system (B&K2032, HP216, HP7470A, HP2671G) and the software (TSK-FM).

#### I. Duration and temporal characteristics.

It has been considered that the speech of deaf persons differ from that of hearing speakers with respect to temporal aspects.

TD of HS<sub>1</sub> was similar to that of normal speakers, but TDs of HS<sub>2</sub> and HS<sub>3</sub> were longer than that of NS<sub>1</sub> and NS<sub>2</sub>. In such an ideomatic expression, TD may depend on the degree of hearing impairment.

Temporal aspects that constitute four segments (V<sub>1</sub>, CD, VOT, V<sub>2</sub>) and TDs were measured. TD, V<sub>1</sub>D, and VOT of HS<sub>3</sub> was shorter than that of other subjects. Closure duration and VOT of HS<sub>2</sub> and HS<sub>3</sub> were shorter than that of HS<sub>1</sub> and normal speakers.

#### II. Fundamental frequency and intonation

It is well known that fundamental frequency of voiced speech sounds varies considerably in

the speech of a given speaker. Fundamental frequencies ( $F_0$ ) of Japanese five vowels in VCV speech environment (V1/p/V2) were measured. Average fundamental frequencies of initial and final vowels of HS<sub>1</sub>, HS<sub>2</sub>, HS<sub>3</sub> were higher than that of hearing speakers. The fundamental frequency of initial vowel was higher than that of final vowel, on the average in all subjects. Deaf persons were apt to show a relatively high average fundamental frequency and a monotonous intonation.

### III. Articulation of vowel.

HS<sub>3</sub> with hearing level 105dB failed to produce appropriate vowel. Fig. 4. 1. and 4. 2 showed coordinate plots of F2 against F1 in initial each vowel and final each vowel respectively. In the case of HS<sub>3</sub> that frequency difference limen is significantly higher than that of other hearing-impaired subjects and a normal hearing speaker, the degree of overlap or proximity among the areas representing different vowels was much greater.

Severely hearing impaired persons who improved fairly well in auditory motor function and profoundly speech perception enhanced articulatory competence as same as a normal hearing person.

**Key word:** fundamental frequency, speech sound analyzing system, formant, vocal tract, vowel environment