EXPERIMENTAL ADVANCEMENTS IN SPEECH SIGNAL REDUCTION TECHNIQUES

By

JAMES EDGAR BAKER

Bachelor of Science
United States Naval Academy
Annapolis, Maryland
1954

Bachelor of Science
Oklahoma State University
Stillwater, Oklahoma
1962

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1962

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1971

EXPERIMENTAL ADVANCEMENTS IN SPEECH SIGNAL
REDUCTION TECHNIQUES

Thesis Approved:

Victor W. Bolie
Thesis Adviser

S. Keith Adams

Paul A. McCollum

Jack A. Olmin

D. Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

To my thesis adviser, Professor V. W. Bolie, I offer my sincere thanks for his many hours of assistance and his patient guidance during this research. To the other members of my committee, Professor H. J. Allison, Professor P. A. McCollum and Professor S. K. Adams, my appreciation is also expressed. A special thanks is offered to Professor H. T. Fristoe.

The financial support, electronics parts and film provided by the United States Air Force is gratefully acknowledged.

This thesis would not have been possible without the encouragement and understanding of my wife, Thelma.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

LIST OF FIGURES

Figure | Page

Figure Page

Figure | Page

CHAPTER I

INTRODUCTION

The real time recognition of speech by machine is a goal which has been pursued by many of the leading research centers of the world over the past thirty years. It is a capability which will become more essential as the man-machine interface becomes more critical with increasing speeds of mechanical devices. More tasks are being assigned to machines and an increasing percentage of time is required for the machine input-output function. The time lost in getting information in and out of machines is becoming the limiting consideration in the application of computers to new areas. A voice programmable computer would make these machines available to virtually any person or job where the cost could be justified. Some of the first applications which can be anticipated are voice controlled typewriters, accounting machines and calculators. Voice controlled security devices will follow when person-to-person differences are better understood. A further major area of benefit will be the reduction in bandwidth necessary for voice communication systems by the transmission of only those characteristics of speech which are necessary to reconstruct intelligible speech which has enough speaker dependent features to be aesthetically pleasing to the listener.

In spite of the potential benefits of automatic speech recognition and the magnitude of effort expended, only very limited progress has been made. Reasonable recognition scores have been achieved only for very small carefully selected vocabularies spoken by a small number of equally carefully selected speakers.

1.1 Previous Research in Automatic Speech Recognition. The desire for an efficient voice transmission system has led to extensive research in the identification of the minimum number of parameters which must be provided for a human listener to piece together the linguistic content of speech. These bandwidth compression systems, called vocoders, which require human perception, have used various techniques such as amplitude-limiting and formant coding (Lindgren, 1). No machine, envisionable in the reasonable future can be expected to have the linguistic or semantic flexibility of the human brain and therefore these techniques are inapplicable to automatic speech recognition.

One of the first documented attempts to build an automatic speech recognizer was reported by Dreyfus-Graf (Lindgren, 2) of Switzerland in 1950. His machine used a six channel parallel filter bank, the outputs of which were used to drive six deflection coils, radially spaced around a movable recording pen. Each word was recorded as a two dimensional figure whose shape was a crude function of speech frequency content and time.

A much more practical device which attempted to recognize the digits zero through nine as spoken by a single speaker was first announced by Bell Telephone Labs in 1952 with later improvements reported in 1958 (Lindgren, 1). This technique used the frequency of spectral energy peaks (called formants) as parameters for comparison with reference words which were stored in a memory element. Such devices are referred to as formant trackers. Accuracies in the 90% range were achieved for single speakers but when no adjustments were made for speaker differences they fell to the 50% vicinity.

Another ten years of speech research with parallel advancements in miniaturization led to the development by Philco-Ford engineers of a zero through nine digit recognizer with a volume of 0.8 ft^3 which achieved 90% accuracy for 10 selected speakers. The recognition scheme abandoned the formant tracking techniques which had been employed by the Bell Labs' recognizer in favor of three new parameters of comparison. These were (1) the Single Equivalent Formant Frequency (SEF); (2) the Single Equivalent Formant Amplitude (SEFA); and (3) the presence or absence of voicing. The first two of these are defined as SEF = $(2T)^{-1}$ in which T is the time from glottal pulse to first zero crossing, and SEFA is the maximum amplitude during this interval. This recognizer was a hard-wired device with removable logic sections which were different for each word. Logic sections were built to recognize any pair of words up to three syllables in length or one word of up to six syllables (Teacher, *et al.*, 3).

Continued attempts at digit recognition by formant measurement are being made by Ewing and Taylor (4) who use the average rate of zero crossings of a speech wave as a measure of the first formant (F1) and the average rate of zero crossings of the differentiated wave as a measure of the second formant (F2). The first method employed by these authors applied an analog measure of F1 to the vertical plates and an analog measure of signal amplitude to the horizontal plates of a storage oscilloscope. The patterns were neither unique for individual sounds nor consistent for the same sound spoken by different persons. These authors then plotted the difference frequency F1 minus F2 versus time and found that a given sound spoken by various speakers showed good consistency but there was not absolute uniqueness for different sounds.

Virtually all word recognition research has been based upon the concept that all English speech is made up of about 40 basic sounds which are called phonemes. Recognition schemes have therefore emphasized the isolation of characteristics for distinguishing between these phonemes when spoken one at a time. It has been assumed that connected phoneme or word recognition would evolve from these techniques by sequential identification of phonemes. Many years of only limited progress in phoneme boundary definition of connected speech have led some recent researchers to attack the theory of phonetic speech recognition. It is impossible in most cases to indicate boundaries between one phoneme and another in the acoustical signal. In different environments a given phoneme will differ greatly because of variations in speaking rates and the time required to change the vocal tract shape from one phoneme to another. In some cases the vocal tract will begin a transition from one phonemic state to another but well before the change is complete it will begin a transition to a third phonemic state. As a result it is likely that none of the phonemes will be produced as they would if they were spoken in isolation. It thus seems reasonable that in recognition systems a trade off will be necessary between vocabulary size, word selection, rate of pronunciation and machine capabilities. The fact that acoustical phonetic boundaries are subtle and time variant led Sitton (5) to investigate a mechanical segmentation procedure. A zero axis crossing count was found to show promise for phoneme boundary definition with reasonable computer processing time requirements. A more pessimistic view is expressed by Pierce (6) who concludes that

> These are strong reasons for believing that spoken English is, in general, simply not recognizable phoneme by phoneme or word by word, and that people recognize utterances, not because they hear the phonetic features of the words distinctly, but because they have a general sense of what a conversation is about and are able to guess what has been said.

A direction for future research is indicated by the theory of Denes and Fry who have insisted upon the necessity for linguistic information in any recognition scheme. Lindgren (1) quotes Denes as concluding "... Automatic speech recognition is probably possible only by a process that makes use of information about the structure and statistics of the language being recognized as well as the characteristics of the speech sound wave." A further area of search for perceptual clues which has been pursued by several groups is the study of articulatory movements which has been referred to as the motor theory. This theory of perception is based upon the belief that a human perceives speech sounds by reference to the articulatory movements he knows are necessary to produce those sounds (Lindgren, 2). Experiments have attempted to show that the process of learning to recognize speech is greatly aided by the voicing of sounds and the association of articulatory movements with the sounds produced. Results have so far been inconclusive.

Many years of research with the cochlea of human cadavers have permitted Georg von Bekesy to determine most of today's knowledge of the physiological characteristics of the cochlea. In his experiments he cut away some of the bone surrounding the cochlea and placed a tiny window in the side wall from which measurements were made of basilar membrane displacement using a stroboscopic light. He showed that the cochlea performs a spectral separation of sound but has relatively poor selectivity with a constant $Q = 2$ for the entire length. For this work he was awarded the Nobel prize in Medicine in 1961 (Bekesy, 7; Flanagan, 8,9; Fletcher, 10). The spectral function of the cochlea has suggested that an important property of speech which contains intelligence is the frequency content. Therefore, much research has been directed toward a spectral analysis of phonemes.

Early research in spectral analysis of speech employed the method of recording the signal on a continuous magnetic tape or drum followed by multiple repetitions of the signal being applied to the input of a voltage tunable bandpass filter which permitted the recording of the signal energy as a function of the center frequency of the bandpass filter. Most recent work has employed the use of multiple bandpass filters which provide near real time spectral data. Spectral processors which seek to use formant locations in the identification of phonemes are of two general types. The first, and most common, consists of many parallel narrow band filters each of which is followed by a detector and low-pass filter which provide a voltage level proportional to the energy content of that frequency band. Formant locations are indicated by the channels which have the largest output level. Recent significant efforts using this technique include the work of Campanella and Phyfe (11), Hughes and Hemdal (12), Bobrow and Klatt (13) and Lecours and Sparkes (14). The second type of spectral processor which is used for formant locations is a parallel arrangement of wide band filters each of which is assigned to one of the formants. They are broad enough that the formant for which they are designed will be included for all phonemes. The filter outputs are fed to zero crossing counters which provide voltage output levels which are proportional to the dominant frequency component in that band. This method provided an accurate frequency measure for simple waveforms but when two or more components have comparable amplitudes serious errors can arise (Hyde, 15). This technique was employed by Gilmour (16) who used a 700 Hz low-pass filter to isolate F1, a 700 to 2100 Hz bandpass filter for F2 and a 2100 Hz high-pass filter for F3. Results were typical of other approaches with 90% recognition rate for a single speaker and a vocabulary of seven words and the digits. Stops and fricatives could not be separated. A variation of this technique, which was mentioned earlier, was that employed by Ewing and Taylor (4) which did not use any filters but rather used the zero crossings of the unprocessed speech as a measure of F1 and the zero crossings of the differentiated wave as a measure of F2. A further variation by Reddy (17) sampled a microphone output every 50 μsec with an analog to digital converter and recorded intensity level and number of zero crossings in each 10 msec interval. Intensity levels were used to group similar segments and zero crossings were used to resolve ambiguities. Limited success was claimed but a later paper (Reddy, 18) revealed that the system could not distinguish the stops and several of the vowels.

A very comprehensive narrow band filter analysis of speech, based upon the phoneme recognition concept, was performed by Hughes and Hemdal (12) at Purdue Research Foundation. The spectral analyzer consisted of 35 bandpass filters with center frequencies varying between 286 and 9500 Hz. A bandwidth of 46 Hz at the low frequency end of the spectrum was expanded to 963 Hz for the 9500 Hz filter with all filters having a selectivity (Q) of 10 or less. The single speaker approach was justified on the basis of a lack of knowledge of speaker differences. An algorithm was developed to use a general purpose digital computer to separate phonemes based upon a distinctive feature approach but no capability for adaptation was considered. Nine distinctive features were used to identify 34 phonemes. Three features were found to reasonably separate the vowels with some overlap which could be removed with the additional dimension of time. The conclusions of this research were that the low frequency spectrum has more importance than previously recognized. A set of filters with good precision below 500 Hz was recommended because of the necessity for accurate F1 tracking and attention to low frequency stress cues. This view is opposed by Stewart (19, 20) who considers the low frequency spectrum to be unimportant because of some experiments which he performed. In one experiment he used 300 Hz low-pass filtered noise as a masking signal and found that the quality of speech was not significantly affected. In another experiment Stewart passed speech through an 800 Hz high-pass filter and found that it was quite intelligible but "tinny" sounding.

Another speech recognition system for a limited vocabulary was developed by Bolt, Beranek and Neuman, Inc. for the NASA Electronics Research Center and reported by Bobrow and Klatt (13). This system attacked the problem of recognition of a limited vocabulary of spacecraft related words as spoken by several speakers. The spectrum analyzer was a 19-channel-filter-bank designed for optimum formant determination for male voices which had a glottal pulse rate from 80 to 150 Hz. In order to insure that there would be between two and four harmonics in every filter, bandwidths were used as follows: 360 Hz for center frequencies of 260 to 2780 Hz; 600 Hz at 3260 Hz; 840 Hz at 3980 Hz; and 1080 Hz bandwidth at 4940 and 6020 Hz. Each word was recorded as two seconds of data sampled at 0.01 sec intervals which provided a 19 x 200 matrix, each element of which was a number from zero to 63, in logarithmic units, covering a 45 db range of intensity. The recognition scheme separated the words into syllables based primarily on the total signal energy. Pattern recognition methods based on the word matrix were considered inapplicable because "... it is by now well known that speech spectral patterns ... do not cluster according to the spoken word". Bobrow and Klatt consider phoneme recognition to be a more difficult task than their method of word recognition which was based on spectral energy although they acknowledge that their technique is limited to small vocabularies. The conclusions of this research were that a finer frequency resolution between 300 and 1200 Hz is desirable and that more filters are required to detect sudden movements in energy concentrations.

Word recognition has also been investigated by Martin (21) who used a filter bank of 19 channels, each with a low selectivity (Q = 8). Distinctive features of phonemes were selected and tests involving 155 speakers were conducted. Recognition scores of 88 to 94% were achieved for 34,000 digits spoken singly and in pairs. A new feature which Martin found useful was the energy change with frequency.

Glaesser, Caldwell and Stewart (22) have developed and Stewart has marketed an analog model of the human cochlea which uses a passive cascade of tuned elements with between 19 and 36 pickoff points. The model simulates the cochlea well in selectivity (Q = 2) but the arbitrary choice of pickoff points is not representative of the data available in human perception. Speech recognition with the analog ear was conducted by the developers (Caldwell, 23) and is continuing by the USAF Aerospace Medical Research Laboratories where 19 pickoffs are being used in a neural model to simulate the human auditory system as closely as possible (Mundie and Moore, 24; Cannon, 25). Recent improvements to the analog ear were developed by Bolie (26) by the use of a digital computer to analyze the various network parameters, and a simulated passive network (Q = 10) of 100 sections was used in a digital computer simulation by Lake (27) to analyze the formant locations of 19 phonemes spoken by 12 people. Speech spectrograms were generated with a digital computer by Oppenheim (28).

In spite of the varied approaches to speech recognition and the many years of concentrated efforts there is still no reliable method for removing speaker dependent variations or glottal pulse fluctuations.

1.2 Summary of Thesis Content. This thesis is devoted to high resolution studies of speech sounds, including the development of the necessary hardware. The included audio-filter bank has the highest spectral resolution of any known device at the low frequencies, which have been identified by other research efforts as necessary for accurate pitch tracking. It also has the wide, high frequency, bandwidth which is required for tracking the rapid energy shifts which occur in connected speech. In addition, as a part of this thesis, a three-dimensional oscilloscope display was designed and built for the study of connected speech.

A data set of high spectral resolution was compiled for 11 phonemes as spoken by 3 male and 2 female speakers over an octave range of pitch variation for each phoneme. The importance of pitch in speech recognition is discussed, and some results of experiments in computer recognition of sustained phonemes are presented.

CHAPTER II

TUNED AMPLITUDE TRACKER SYSTEM DESCRIPTION

2.1 Introduction. A review of the literature revealed that a finer spectral resolution of speech than was possible with previous audio filter banks might provide a useful source of information in speech research. The conclusions of two recent significant efforts (Bobrow and Klatt, 13; Hughes and Hemdal, 12) were that better frequency resolution in the spectrum below 1200 Hz would probably provide valuable information for improving the understanding of speaker dependent variations and glottal pulse fluctuations. These features of speech have historically been so little understood that they have been avoided in prior experimental designs.

Advancements in solid state electronics miniaturization in the past five years have made possible the design of the high resolution Tuned Amplitude Tracker (TAT) System which is outlined in Appendix B and described in greater detail in this chapter. Figure 2.1 is a photograph of the equipment rack containing the complete system. In this chapter the functional description of each subsystem is followed by a discussion of the mathematical model of the TAT units and some detailed performance measurements of the completed TAT System. A table summarizing all of the abbreviations used in this thesis is given in Appendix A. In each figure resistances are in ohms and capacitances are in microfarads unless labeled otherwise.

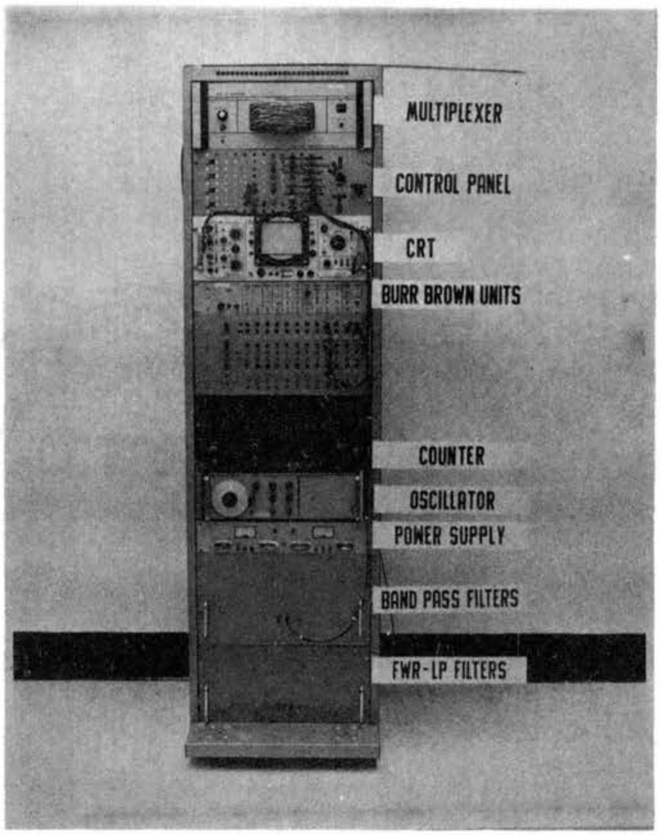


Figure 2.1. Speech Analyzer Equipment Rack

2.2 Functional Description. A block diagram of the elements of the TAT System is shown in Figure 2.2. The speech input may be applied either through a microphone or from a magnetic tape recorder to a spectral shaping pre-amplifier and a speech amplifier which drives the parallel array of TAT channels. The multiplexer samples the output of each of the 97 TAT channels at a rate determined by the clock, and provides a DC analog output voltage which is equal to the voltage on the input line being sampled. The multiplexer also provides a sync pulse for the cathode ray tube oscilloscope (CRT) which displays an image of speech energy content as a function of frequency.

2.2.1 Microphone. The Shure model 545S microphone was used as the air pressure to voltage transducer for the sounds which were spoken as a part of this thesis and was also used in the recording of the tape library which was compiled by Lake (27). Experimental voltage measurements made with this microphone revealed the typical low frequency signal amplitude to be ± 2 millivolts for a normal voice volume and for the microphone placed 8 $\pm$ 2 inches directly in front of the mouth. At this distance the signal amplitude was found to be relatively insensitive to small changes in mouth-to-microphone distance.

The detailed microphone specifications are provided in Appendix C. It can be seen from the response characteristic curve of Figure 2.3 that output amplitude is flat within -3 db and +8 db (where 0 db = 1 volt/microbar) over the 100 to 6400 Hz frequency range covered by the tuned amplitude trackers. It has a "cardioid" polar characteristic with side cancellation of 6 db and 15 to 20 db cancellation at the rear. The microphone has a dual output impedance of 150 ohms and 40K ohms. The rated sensitivity is 150 $\pm$ 4 db at 1000 Hz.

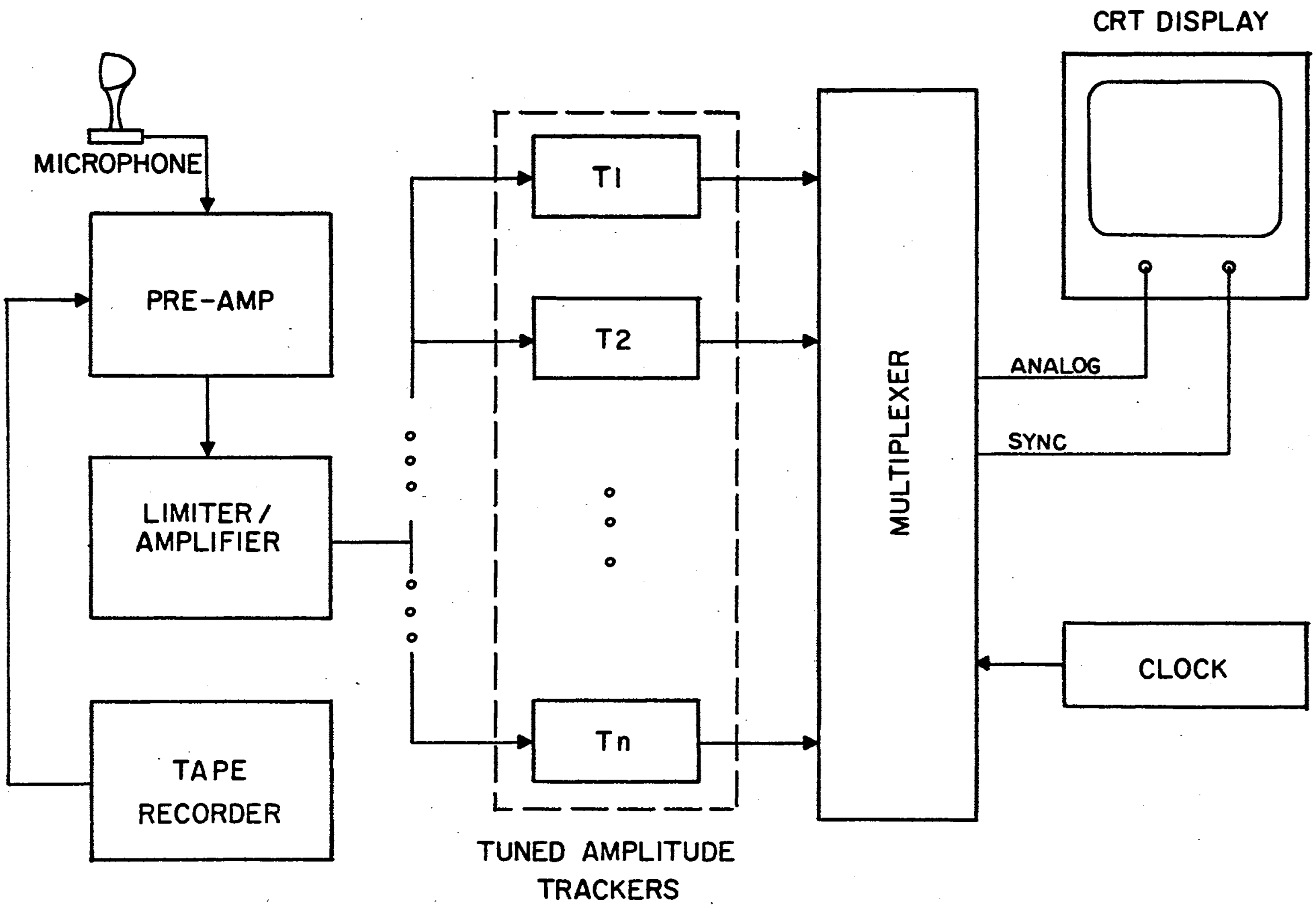


Figure 2.2. Tuned Amplitude Tracker System Block Diagram

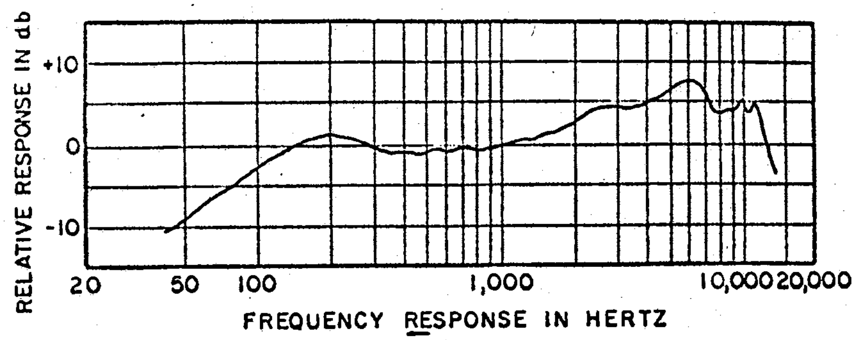


Figure 2.3. Shure Microphone Model 545S Frequency Response Characteristic

2.2.2 Magnetic Tape Recorder. The Wollensak Model 1980 Magnetic Tape Recorder was used for all sound recordings and playback of phonetic tapes. A tape speed of 7.5 inches per second and Channel A were used. The tape recorder gain setting was adjusted for each recording to obtain the maximum signal with no amplitude clipping or other distortion as viewed on an oscilloscope. The audio signal was taken from the "External Speaker" jack.

2.2.3 Speech Pre-Amplifier. The circuit which is shown in Figure 2.4 accepts the 150 ohm impedance differential output of the microphone or the attenuated output from the tape recorder. Spectral gain shaping is accomplished by three stages of modified high-pass filtering and four stages of gain. In the first stage a uniform gain of 40 db is provided by a 709N Operational Amplifier (OA) circuit. The second and third stages each incorporate a modified high-pass filter of 1100 Hz corner frequency in which the 0.29 μfd capacitor is bypassed with a 735 ohm resistor in order to limit the low frequency loss. The third stage incorporates a modified high-pass filter of 1460 Hz corner

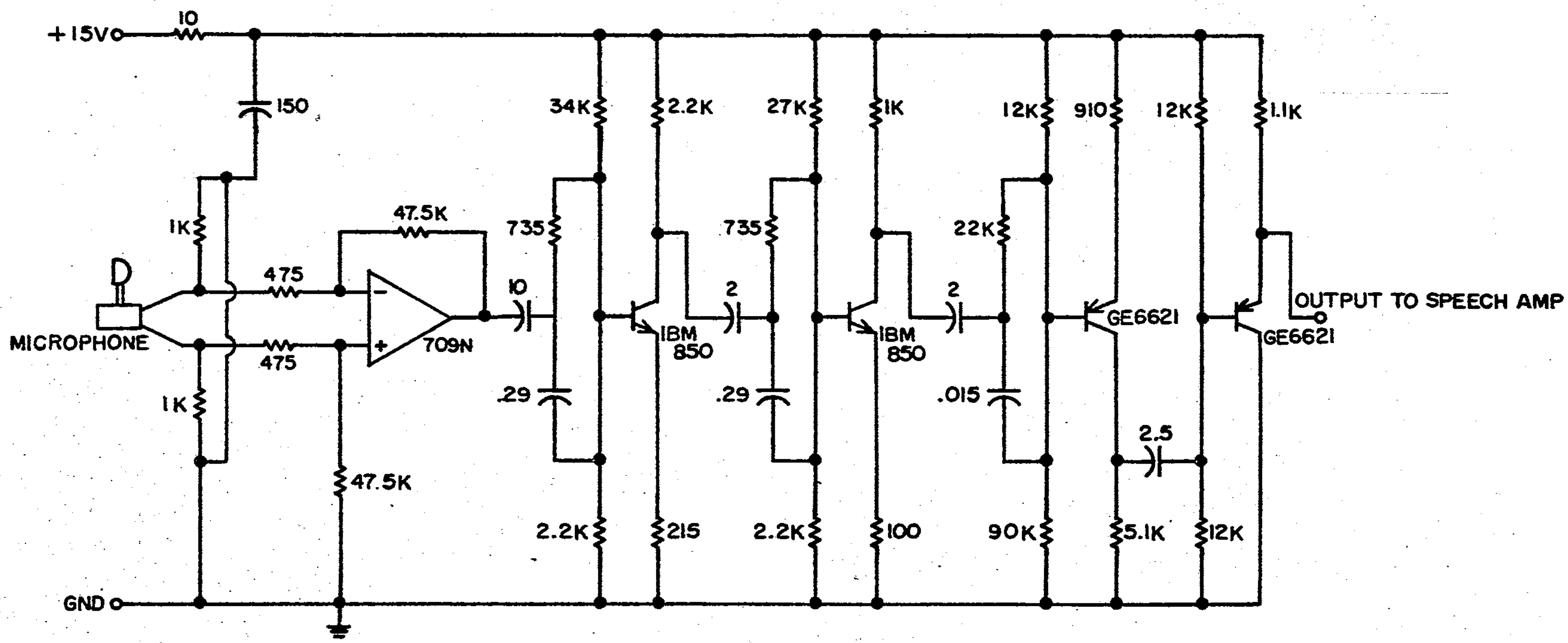


Figure 2.4. Speech Pre-Amplifier Circuit Diagram

frequency in which the 0.015 μfd capacitor is bypassed by a 22K resistor. The fourth stage is an emitter follower which prevents loading effects.

Each stage of gain shaping was designed to prevent speech clipping and to provide about the same maximum amplitude in each spectral segment at the Pre-Amplifier output. This method of shaping was selected to permit use of the full dynamic amplitude range of the spectral display. It is considerably different from the 6 db per octave gain increase which was commonly employed by previous researchers and upon which the Formant Theory of speech perception is based. This departure from former approaches was necessary to take advantage of the resolution capability of the new system. The entire Speech Pre-Amplifier is packaged in a Burr-Brown casing and is provided with a power supply filter for both the +15 volt input (shown in Figure 2.4) and the -15 volt input which is used only in the OA. Oscillations were observed before the filters were installed. The maximum signal output without clipping or distortion is ±4 volts, which is achieved only with very loud sounds voiced close to the microphone.

2.2.4 Soft Limiter. The speech signal from the Pre-Amplifier is processed by the circuit which is shown in Figure 2.5, which limits the signal amplitude and provides a low impedance drive to the Speech Amplifier. The soft limiting action begins at an input threshold of ±3.2 volts, as determined by the input voltage divider network. The 1 volt forward voltage characteristic of the 1N816 diodes rounds off the voltage peaks which exceed this threshold. The push-pull output circuit supplies an output impedance of about 200 ohms; negative feedback is used to eliminate crossover distortion. For signals less than limiting threshold the circuit gain is about 1.1 and as a result the output voltage is about ± 3.5 volts.

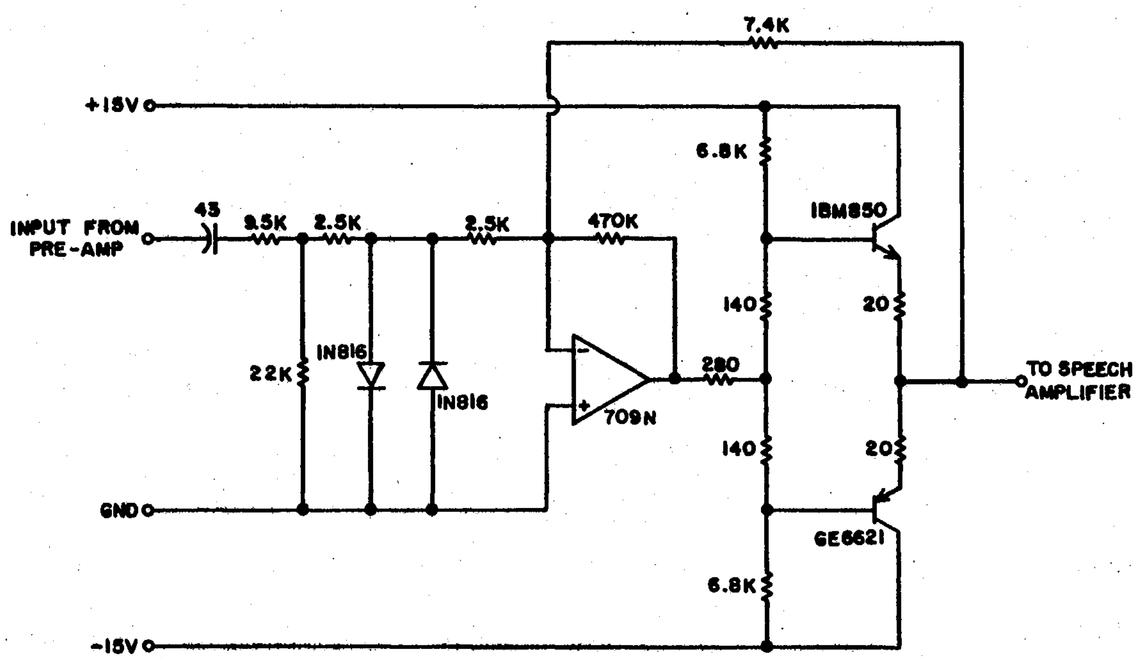


Figure 2.5. Soft Limiter Circuit Diagram

2.2.5 Speech Amplifier. The initial system made use of the Soft Limiter Circuit to drive the 97 TAT channels. During calibration it was found that each channel responded with a maximum DC output of 10 volts when its input was a ± 10 volt pure-tone signal from an audio oscillator. With speech input, however, the channels responded with only about half their maximum output capability because of the amplitude variations inherent in the speech signal. In order to provide a drive signal which would use the full dynamic range of the filter bank, the Speech Amplifier circuit which is shown in Figure 2.6 was used to drive the TAT channels. This circuit boosts the ± 3.5 volt input signal from the Soft Limiter circuit to a ± 20 volt signal at the output. The 2N328A transistors are biased into "Class A" operation and provide a current-drive for the 40322 power transistors, each of which has about 17 square inches of aluminum heat sink. The 10 ohm-56 μfd filter for the supply voltage eliminated the periodic oscillations which were observed before its installation.

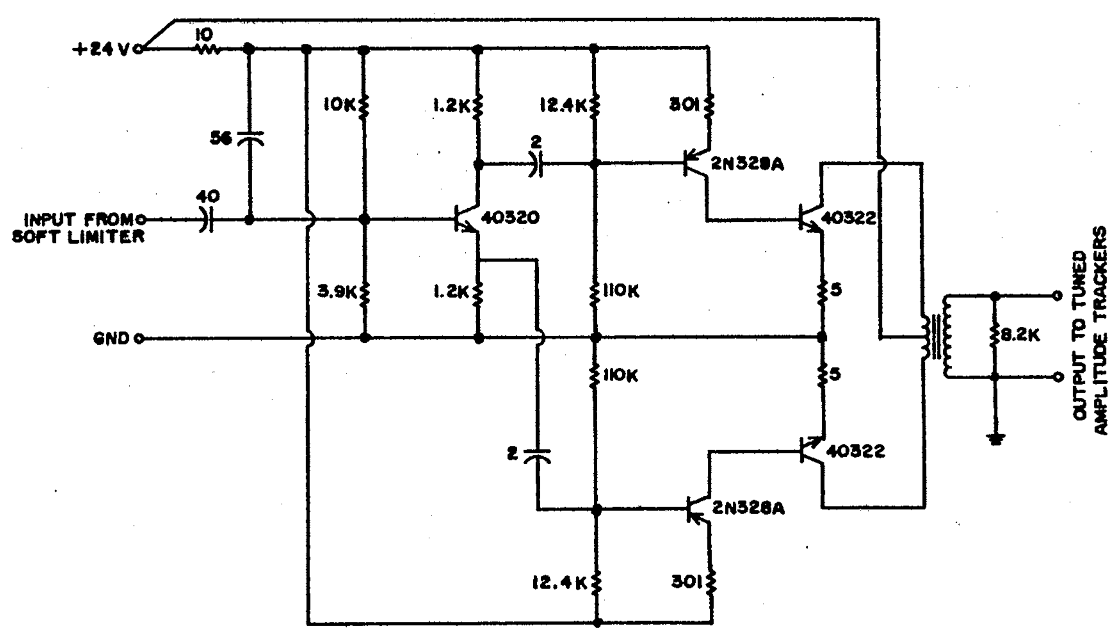


Figure 2.6. Speech Amplifier Circuit Diagram

2.2.6 Tuned Amplitude Trackers. Each of the 97 TAT channels contains the circuit elements shown in Figure 2.7. A detailed discussion of the design and operation of this circuit is included in Section 2.3. The purpose of each TAT is to convert an AC input voltage, E_K, into a DC voltage, V_K. The amplitude of V_K is proportional to the amplitude of those frequency components of E_K which fall within the selected pass-band. The pass-band of each channel is determined by the OA, U1, and its associated components. The selectivity, Q, was held constant at 25 for all channels. This provided -3 db crossover for adjacent filters whose resonant frequencies were chosen for the K^{th} channel as discussed in Appendix B by the relation:

$$F(K) = (100) \cdot 2^{[(K-1)/16]} \quad .$$

Table I contains a listing by channel number, K, of (1) resonant frequency, F(K); (2) the type of United Transformer Company inductor, L'_K; (3) the size of resonating capacitor, C'_K; (4) the bandwidth, BW; (5) the frequency of the lower -3 db half power point, LHPP; and (6) the frequency of the upper -3 db half power point, UHPP. The Q was held constant to provide the desired detail in low frequency resolution. The value of Q = 25 was selected because it permitted about three times the resolution of the Hughes and Hemdal system (12) and was compatible with the 100 input channel capability of the available multiplexer. The inductance range of each coil type is given in Appendix B.

The OA, U2, and related components perform fullwave rectification of the output signal of the bandpass filter. The OA, U3, and associated components make up a low-pass filter which converts the signal to a DC voltage, amplifies the level and provides a low-pass characteristic,

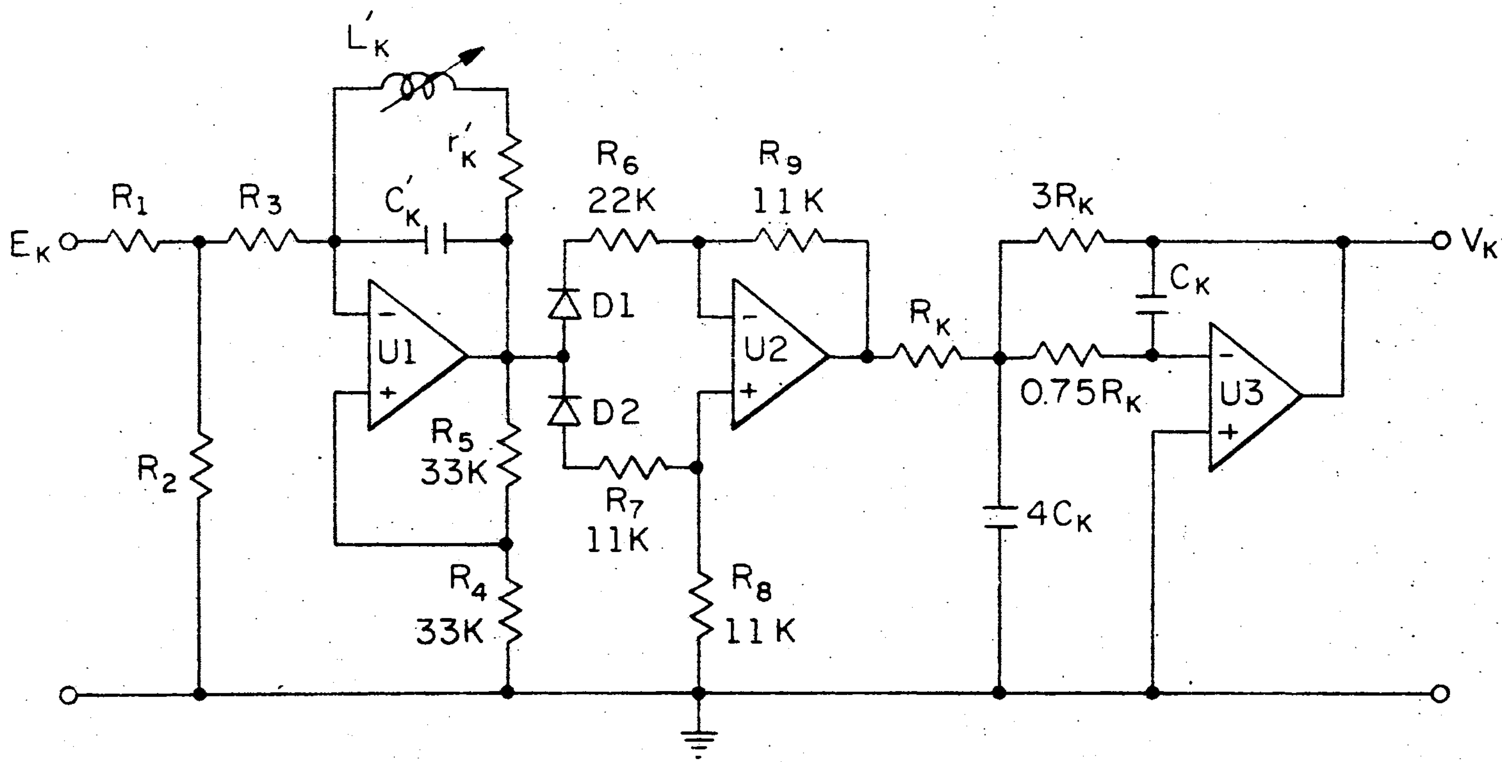


Figure 2.7. Tuned Amplitude Tracker Circuit

TABLE I

TUNED AMPLITUDE TRACKER DATA

K	F(K) (Hz)	Inductance Model No. HVC-	Resonator C'_K (μfd)	BW (Hz)	LHPP (Hz)	UHPP (Hz)
1	100	7	.5	4	98	102
2	104	7	.5	4	102	106
3	109	7	.5	4	107	111
4	113	7	.5	5	110	115
5	119	7	.5	5	116	121
6	124	7	.5	5	122	127
7	130	7	.5	5	127	132
8	135	7	.5	5	132	137
9	141	7	.5	6	138	144
10	147	7	.5	5	145	150
11	154	7	.5	6	151	157
12	160	7	.5	7	156	163
13	168	7	.5	7	164	171
14	175	7	.5	6	172	178
15	183	7	.5	8	180	188
16	191	7	.5	8	187	195
17	200	7	.47	8	196	204
18	208	7	.5	9	203	212
19	218	6	.5	9	213	222
20	227	6	.5	9	223	232
21	238	6	.5	10	232	242
22	247	6	.47	10	243	253
23	259	6	.5	10	254	264
24	270	6	.5	12	265	277
25	283	6	.5	12	276	288
26	294	6	.5	12	288	300
27	308	6	.5	12	302	314
28	321	6	.5	13	314	327
29	336	6	.5	14	329	343
30	350	6	.5	14	343	357

TABLE I (Continued)

K	F(K) (Hz)	Inductance Model No. HVC-	Resonator C'_K (μfd)	BW (Hz)	LHPP (Hz)	UHPP (Hz)
31	367	6	.5	15	359	374
32	381	6	.5	16	373	389
33	400	6	.5	16	392	408
34	416	6	.5	17	407	424
35	436	6	.5	17	428	445
36	454	6	.5	19	445	463
37	476	6	.5	19	467	486
38	495	5	.2	20	485	505
39	519	5	.2	20	508	528
40	540	5	.2	22	529	551
41	566	5	.2	22	554	576
42	588	5	.2	24	576	600
43	617	5	.2	26	604	630
44	642	5	.2	26	629	655
45	673	5	.2	26	659	685
46	700	5	.2	27	687	715
47	734	5	.1	29	719	748
48	763	5	.2	30	748	778
49	800	5	.22	32	784	816
50	832	5	.33	33	815	848
51	872	5	.22	33	857	890
52	907	5	.27	36	889	925
53	951	5	.22	37	932	969
54	990	5	.27	40	970	1010
55	1038	5	.1	41	1019	1060
56	1079	5	.27	43	1057	1100
57	1131	5	.1	46	1108	1154
58	1177	4	.22	47	1153	1200
59	1234	4	.1	48	1211	1259
60	1283	4	.1	51	1257	1308
61	1345	4	.068	50	1322	1372

TABLE I (Continued)

K	F(K) (Hz)	Inductance Model No. HVC-	Resonator C'_K (μfd)	BW (Hz)	LHPP (Hz)	UHPP (Hz)
62	1400	4	.1	56	1372	1428
63	1467	4	.1	60	1437	1497
64	1526	4	.1	61	1495	1556
65	1600	4	.082	62	1572	1634
66	1664	4	.1	67	1631	1698
67	1745	4	.082	68	1712	1780
68	1815	4	.1	72	1779	1815
69	1903	4	.082	75	1866	1941
70	1979	4	.043	79	1939	2018
71	2075	4	.068	85	2033	2118
72	2158	4	.047	86	2115	2201
73	2263	4	.068	91	2217	2308
74	2353	4	.1	94	2306	2400
75	2468	4	.05	98	2420	2518
76	2566	4	.05	103	2514	2617
77	2691	4	.056	105	2637	2742
78	2799	3	.1	112	2743	2855
79	2934	3	.056	110	2880	2990
80	3052	3	.039	124	2990	3114
81	3200	3	.05	123	3149	3272
82	3328	3	.033	133	3261	3394
83	3490	3	.05	145	3420	3565
84	3630	3	.05	141	3560	3701
85	3806	3	.039	152	3738	3890
86	3958	3	.036	158	3879	4037
87	4150	3	.033	169	4073	4242
88	4316	3	.022	172	4230	4402
89	4526	3	.033	175	4545	4620
90	4707	3	.022	188	4613	4801
91	4935	3	.033	197	4847	5044
92	5132	3	.036	210	5027	5237

TABLE I (Continued)

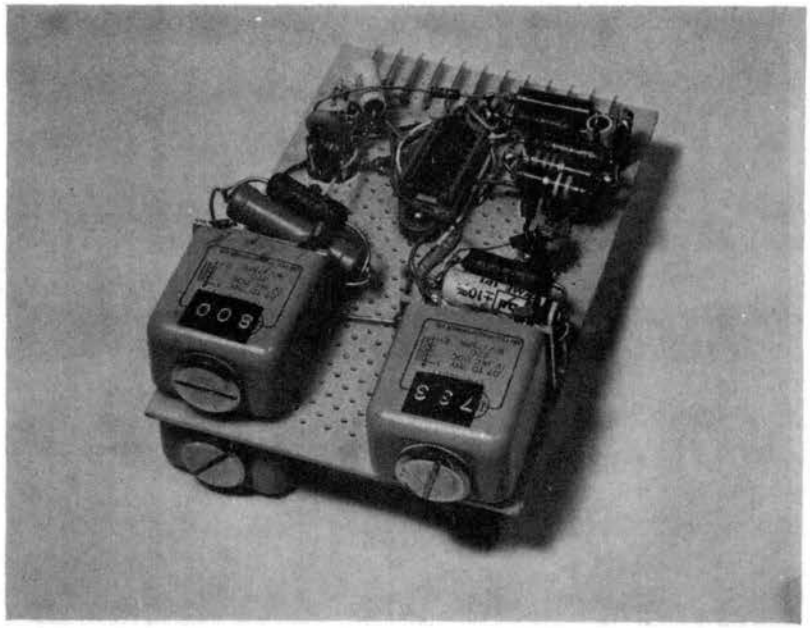
K	F(K) (Hz)	Inductance Model No. HVC-	Resonator C'_K (μfd)	BW (Hz)	LHPP (Hz)	UHPP (Hz)
93	5382	3	.033	216	5280	5496
94	5597	3	.027	219	5490	5709
95	5869	3	.02	222	5763	5985
96	6104	3	.022	252	5986	6238
97	6400	3	.015	256	6280	6536

with the frequency of the -6 db point set equal to the bandwidth of the bandpass filter. Figure 2.8 shows one of the bandpass filter Vector boards each of which contains three filters, and one of the Vector boards on which are mounted three fullwave rectifiers and low-pass filters. The complete system of 97 bandpass filters is mounted on a common chassis, with the output from each filter available to the connector on the rear of the chassis as shown in Figure 2.9. This connector and cable provide an input to the Fullwave Rectifier and Low-Pass Filter boards by means of a similar connector which is located at the right rear of the chassis shown in Figure 2.10. The connector at the center of the chassis provides the 97 channel output to the multiplexer. A final gain calibration is provided at the output of each low-pass filter by means of a resistance divider network which limits the output voltage to +5 volts, which is the maximum allowable multiplexer input.

2.2.7 EECO 765 Multiplexer. The detailed multiplexer performance data is included in Appendix C. Each of the 97 input voltage levels is sampled and applied to the analog output for a length of time determined by the clock pulses which cause the multiplexer to step from one channel to the next. The order in which the input lines are sampled is programmable by means of a patch board on the front panel. The sync pulse as shown in Figure 2.2 is a zero-going pulse superimposed on a -10 volt level. This pulse is identified by the manufacturer as the "Scan Address End Ring". The pulse remains on during the time the last programmed input channel is being sampled.

The multiplexer input impedance specifications state that the channel being sampled has an impedance greater than 100 megohms in parallel with a capacitance determined from the equation,

(A)



(B)

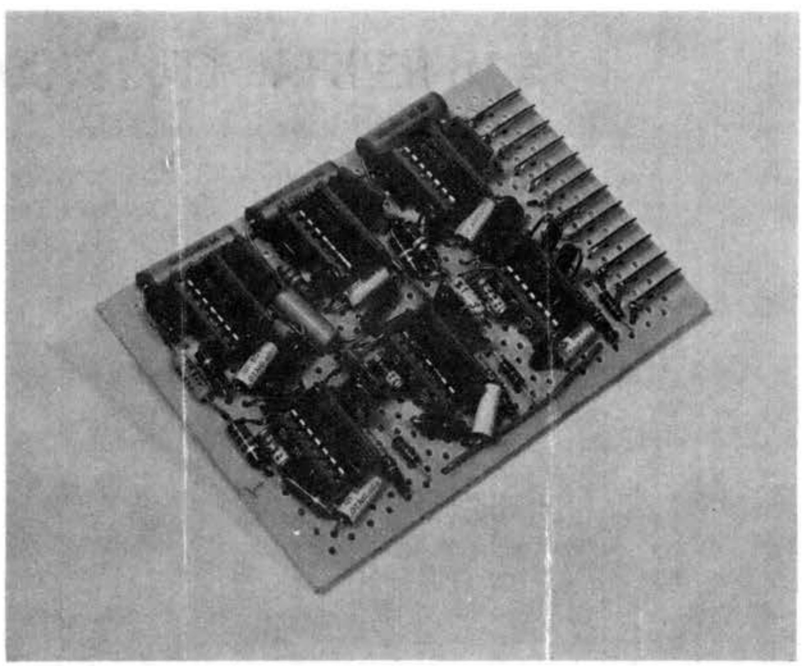


Figure 2.8. Bandpass Filter (A) and Fullwave Rectifier and Low-Pass Filter (B) Circuit Boards

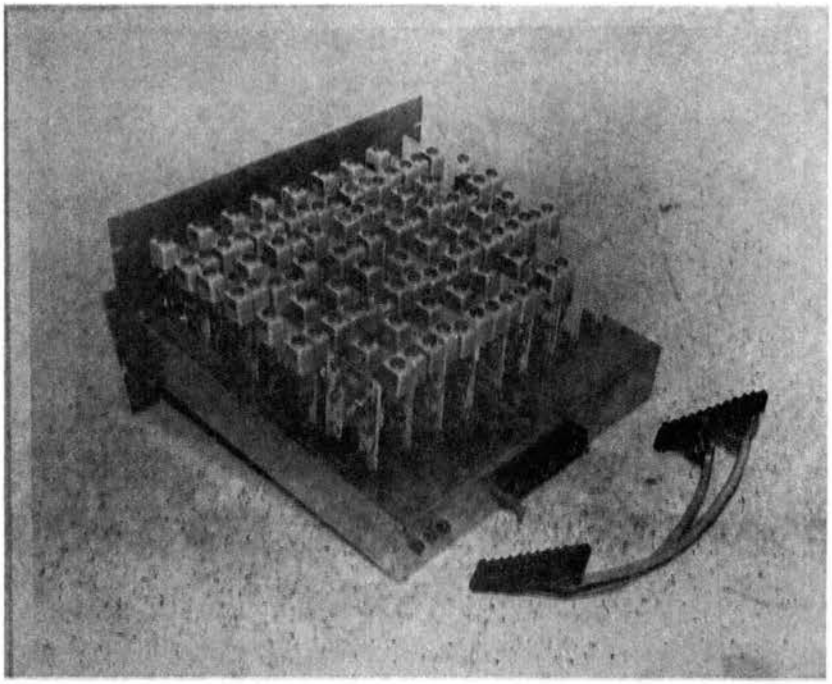


Figure 2.9. Bandpass Filter Chassis

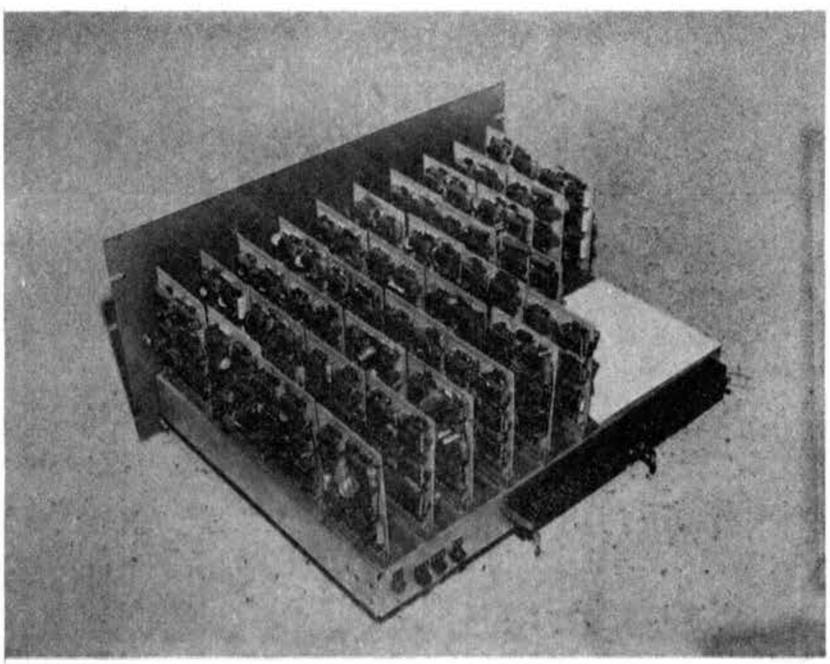


Figure 2.10. Fullwave Rectifier and
Low-Pass Filter
Chassis

$$C = [1.5(N-1) + 10] \text{ pico-farads (pfd)} = 154 \text{ pfd}$$

where, N = 97 input channels.

During the early construction stages the output of each low-pass filter was applied directly to the multiplexer input. The multiplexer output switched sharply to each new level. However, when the low-pass filter output was applied to a resistance divider which changed the output impedance as seen by the multiplexer to about 5K ohms, a time constant of about 5 μsec was observed. This effect is shown in Figure 2.11. This was corrected with the addition of a capacitor $C_{K(1)}$ in each voltage divider network, as shown in Figure 2.12. The reduction in switching time which resulted can be observed in Figure 2.13. The value of this output capacitance which should have been about 150 pfd according to the multiplexer specifications, was found to be about 0.1 μfd. Therefore 0.1 μfd was used for the compensation capacitor, $C_{K(1)}$.

2.2.8 Clock. A free running multivibrator was used as a digital clock to control the rate at which the multiplexer stepped from one input channel to the next. The circuit shown in Figure 2.14 was built into a Burr-Brown casing and had a panel mounted switch to provide a choice of an 11.1 KHz or a 50.3 KHz pulse rate. The 50.3 KHz rate was used throughout this research except when the multiplexer was controlled by the IBM 1620 Computer as discussed in Chapter V.

2.2.9 RM 564 Tektronics Storage Oscilloscope. The storage oscilloscope is a versatile visual display device which may be used to provide a variety of speech spectral plots. It can be used in a normal oscilloscope mode to reveal transient information or may provide a visual integration over several minutes. The vertical display has a

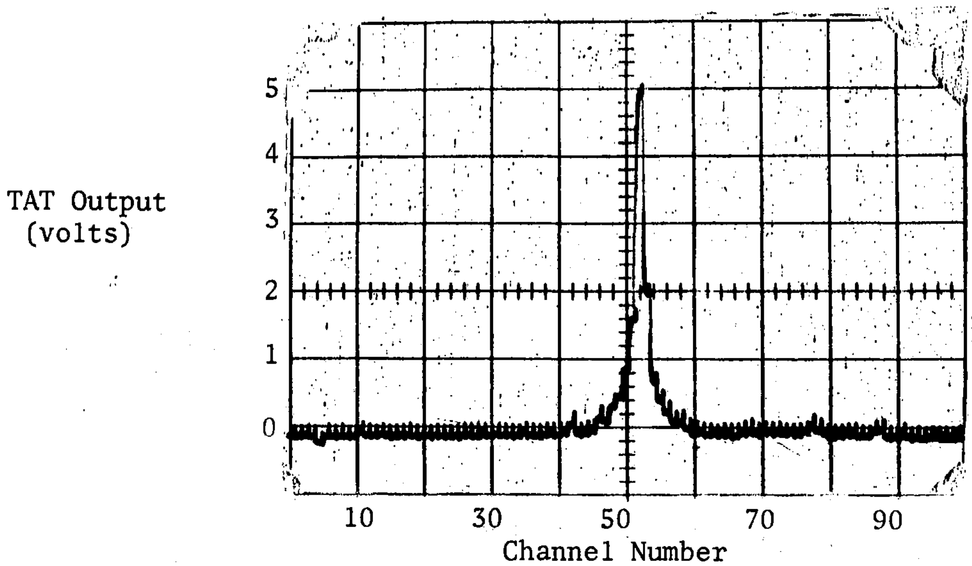


Figure 2.11. Multiplexer Output Without Input Capacitance Compensation (800 Hz Sine Wave Input)

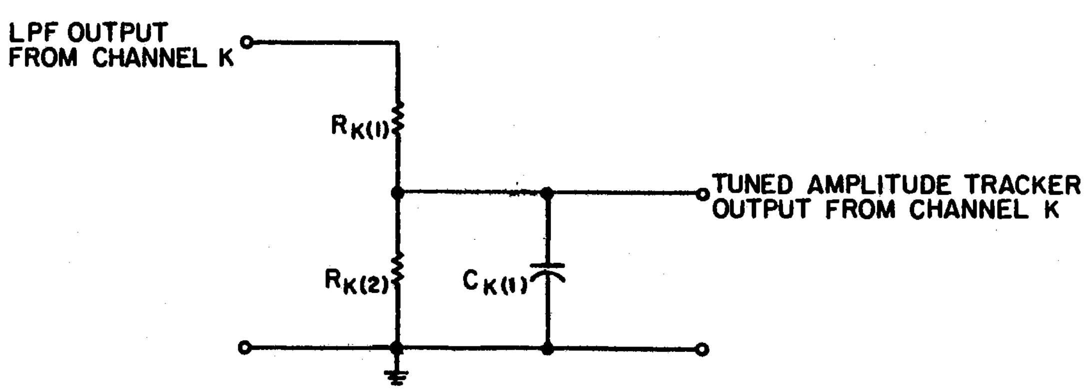


Figure 2.12. LPF Voltage Divider and Compensation Capacitor

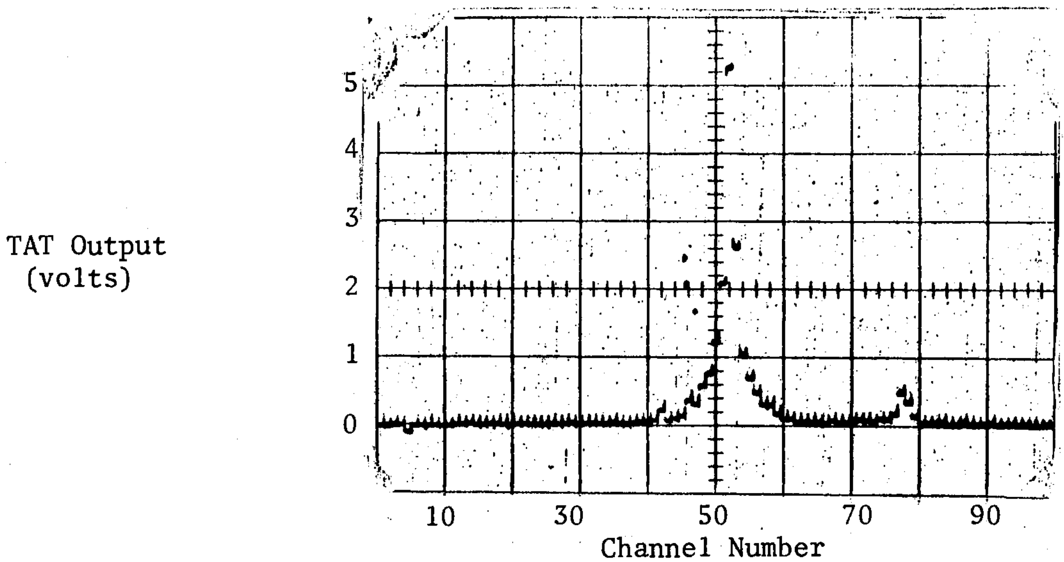


Figure 2.13. Multiplexer Output With Input Capacitance Compensation (800 Hz Sine Wave Input)

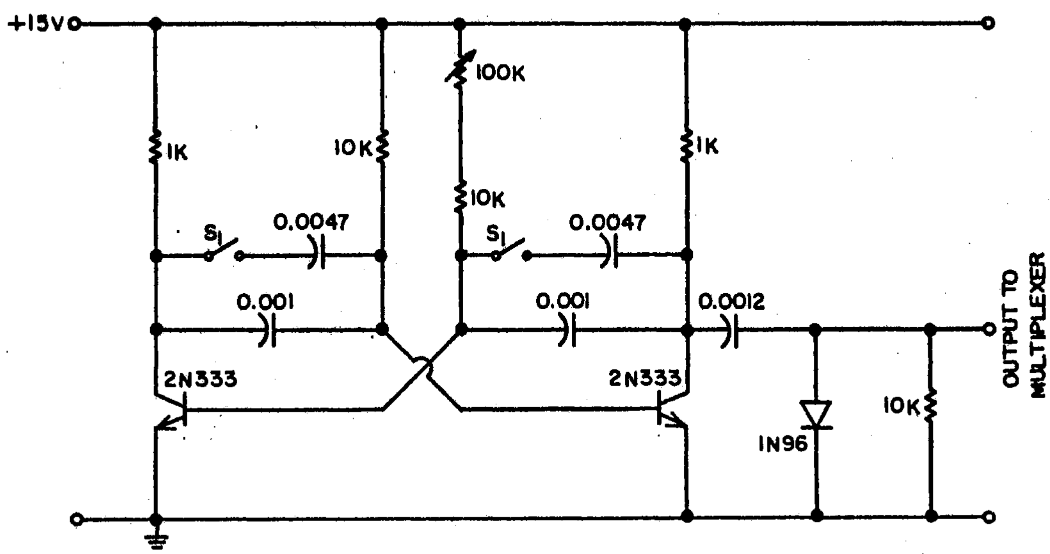


Figure 2.14. Clock Circuit

dynamic range of 100 db per centimeter with a continuous vernier control. The horizontal display can use (1) an externally generated deflection voltage as was employed in the display described in Chapter III of this thesis; (2) an internal sweep generator which is triggered by an external sync voltage as is used in the display in this chapter; or (3) an internal sweep generator which is triggered by an adjustable level from either of two vertical deflection amplifiers. A number of other modes of flexibility are also available, the details of which may be found in Appendix C.

2.3 The Mathematical Model. This section describes the theory underlying the design for the TAT channels which are made up of (1) the tuned bandpass filters; (2) the fullwave rectifiers; and (3) the lowpass filters. In the construction of the filter bank system, United Transformer Company HVC type inductors and SN72 709N OA's manufactured by Texas Instruments, Inc., were used. Specifications of both are found in Appendix C. A discussion of the OA frequency compensation, which was required in the spectral band of speech, is found in Appendix B.

2.3.1 Tuned Bandpass Filter Design. The desire for a constant Q and amplitude across the speech spectrum made it necessary to select a circuit in which these parameters could be controlled, preferably independently. Huelsman (29) has reported that an operational amplifier with control elements as shown in Figure 2.15 has a transfer characteristic given by

$$\frac{V}{E} = \frac{-Y_1Y_3}{(Y_1 + Y_2 + Y_3 + Y_4)Y_5 + Y_3Y_4} \tag{2.3.1}$$

where E = the input voltage, V = the output voltage and Y_K = the admittance of the K^{th} branch. The choice of branch components as indicated in Figure 2.16 results in a bandpass filter which has a transfer function, after simplification by Bolie (30), of

$$\frac{V}{E} = \frac{-AT(j\omega)}{Q + T\,(j\omega) + T^2Q\,(j\omega)^2}$$

where T = ARC/Q and F_0 = center frequency = $1/2\pi T$. Although this circuit was reported to be capable of obtaining a Q of about 10, an attempt was made to extend Q to 25 in order to provide better glottal pulse tracking and better formant frequency definition than has previously been possible. A mathematical analysis of the transfer function indicated that this should have been possible but in practice the best Q which could be obtained was about 12. This was undoubtedly due to an inability to adequately control component accuracy. It was therefore decided that in spite of the disadvantages of iron core inductors it would be necessary to use them if a Q of the order of 25 were desired with a small number of circuit elements, although for other applications Q values as high as 5,000 have been realized using 4 operational amplifiers in an RC network designed by Tarmy and Ghausi (31).

The desire to maintain independent control of both gain and bandwidth over the speech spectrum of 100 Hz to 6400 Hz with a minimum number of components led to the use of both positive and negative feedback in the bandpass filter which is reported in this thesis. In the network shown in Figure 2.17 and in the following equation Z is the impedance of the parallel RLC tank circuit and $k = R_4/(R_4 + R_5)$ is the ratio of positive feedback. The current I is that flowing through the

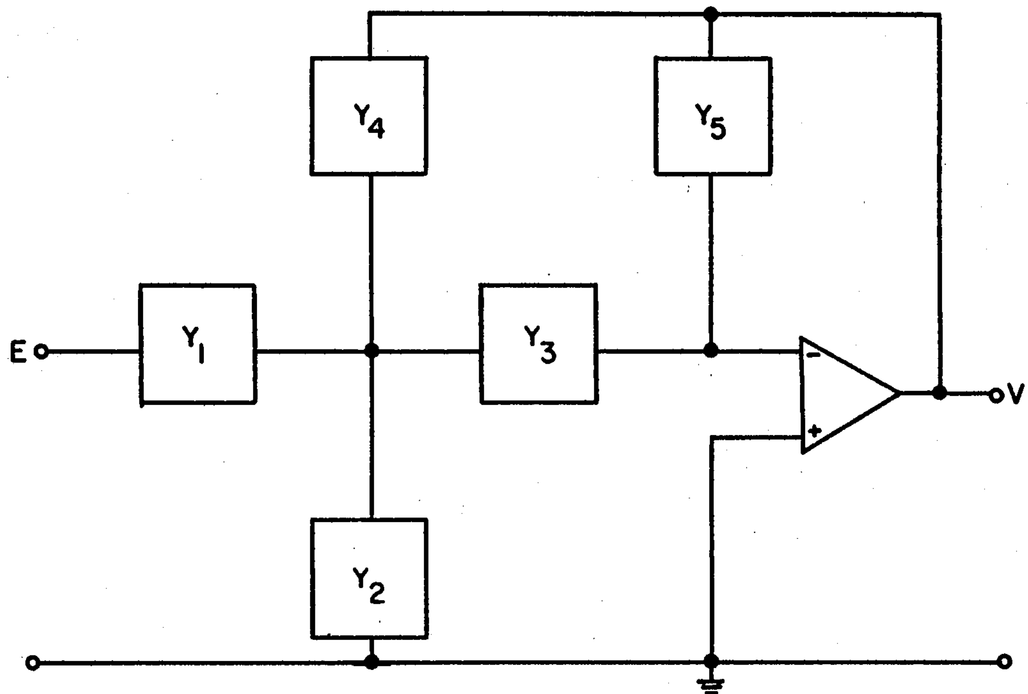


Figure 2.15. General Filter Network

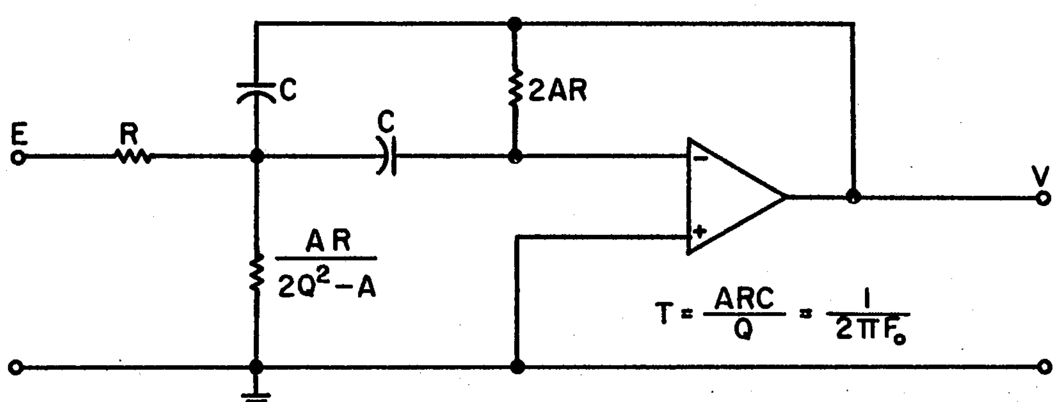


Figure 2.16. RC Active Bandpass Filter

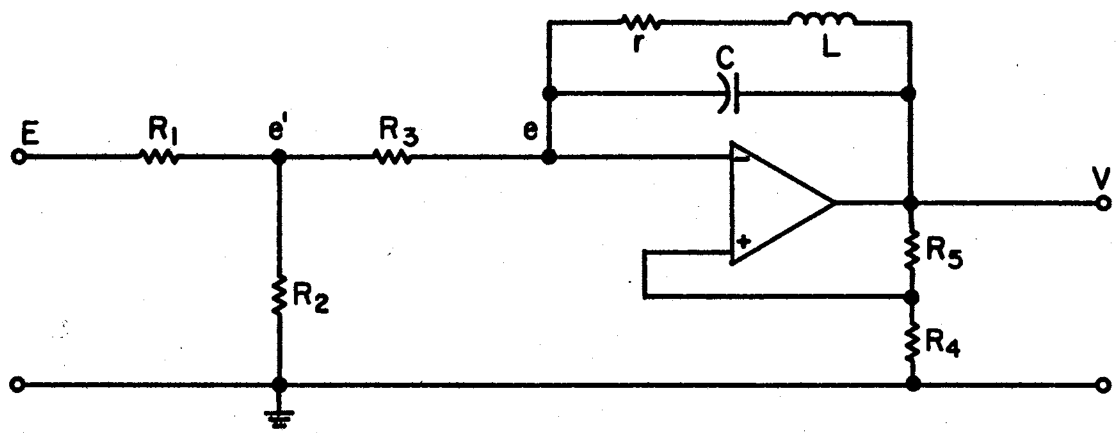


Figure 2.17. Active Tuned Bandpass Filter

parallel RLC impedance toward the OA output terminal. The current and voltage relations for the input resistance divider network can be written,

$$\frac{E - e'}{R_1} = \frac{e'}{R_2} + I \tag{2.3.2}$$

$$e' = e + IR_3 \quad . \tag{2.3.3}$$

Substituting (2.3.3) into (2.3.2) and solving for the voltage at the inverting input of the OA yields:

$$\frac{E}{R_1} = \left[\frac{1}{R_1} + \frac{1}{R_2}\right] \cdot \left[e + IR_3\right] + I$$

$$= \left[\frac{1}{R_1} + \frac{1}{R_2}\right] e + \left[1 + \left[\frac{1}{R_1} + \frac{1}{R_2}\right] R_3\right] I$$

$$e = \left[\frac{R_1 \cdot R_2}{R_1 + R_2}\right] \cdot \left[\frac{E}{R_1} - \left[1 + \frac{(R_1 + R_2)\ R_3}{R_1\ R_2}\right] I\right] \quad . \qquad (2.3.4)$$

Since the operational amplifier causes the output voltage, V, to vary such that the inverting input voltage, e, becomes approximately equal to the non-inverting input voltage, kV, then these two input voltages can be assumed to be equal. Using the relationship, $I = (k-1)V/Z$, the positive feedback voltage becomes

$$kV = \left[\frac{R_1\ R_2}{R_1 + R_2}\right] \cdot \left[\frac{E}{R_1} - 1 + \left[\frac{(R_1 + R_2)\ R_3}{R_1\ R_2}\right] \cdot \left[\frac{(k-1)V}{Z}\right]\right]$$

from which the voltage transfer characteristic V/E can be determined to be

$$\frac{V}{E} = \frac{-\left[\frac{R_1\ R_2}{R_1 + R_2}\right] \cdot \left[\frac{1}{R_1}\right]}{\left[\frac{1-k}{Z}\right] \cdot \left[R_3 + \frac{R_1\ R_2}{R_1 + R_2}\right] - k} \quad . \qquad (2.3.5)$$

It is convenient for notation to use the following substitutions in (2.3.5), noting that R_p is the equivalent parallel resistance of R_1 and R_2, R_s is a series resistance, and that $Q_1{}^2 r$ is the resonant impedance of the parallel LC combination

$$\left.\begin{aligned} R_p &= \frac{R_1 R_2}{R_1 + R_2} \\ R_s &= R_p + R_3 \\ Z &= \frac{Q_1{}^2 r}{1 + j\,Q_1\left[\frac{F}{F_0} - \frac{F_0}{F}\right]} \end{aligned}\right\} \qquad (2.3.6)$$

In these equations, Q_1 is the selectivity ($Q_1 = F_0/BW$) of the parallel LC circuit, r is the series resistance of the inductor in ohms, F is the frequency of concern in Hertz, and F_0 is the resonant frequency of the LC circuit. Insertion of these identities into (2.3.5) yields,

$$\frac{V}{E} = \frac{-\dfrac{R_p}{R_1}}{\left[\dfrac{(1-k) - R_s}{Q_1{}^2 r}\right] \cdot \left[1 + jQ_1\left[\dfrac{F}{F_0} - \dfrac{F_0}{F}\right]\right] - k} \quad . \qquad (2.3.7)$$

A convenient form of this voltage transfer characteristic is

$$\frac{V}{E} = \frac{-A}{1 + jQ\left[\dfrac{F}{F_0} - \dfrac{F_0}{F}\right]} \qquad (2.3.8)$$

where A is the gain at resonance and the other parameters are as defined above. Writing (2.3.7) in the form of (2.3.8) yields,

$$A = \frac{1}{\dfrac{(1-k)R_s R_1}{R_p Q_1^{\,2} r} - \dfrac{kR_1}{R_p}} \qquad (2.3.9)$$

and

$$Q = \frac{\left[\dfrac{(1-k)R_s R_1 Q_1}{R_p Q_1^{\,2} r}\right]}{\left[\dfrac{(1-k)R_s R_1}{R_p Q_1^{\,2} r}\right] - \left[\dfrac{kR_1}{R_p}\right]} \quad . \qquad (2.3.10)$$

Equation 2.3.10 may be solved for the ratio of the overall circuit Q divided by the RLC tank circuit Q_1, to give

$$\frac{Q}{Q_1} = \frac{1}{1 - \left[\dfrac{1}{1-k}\right] \cdot \left[\dfrac{Q_1^{\,2} r}{R_3 + \dfrac{R_1 R_2}{R_1 + R_2}}\right]} \quad . \qquad (2.3.11)$$

<u>2.3.1.1 Tuned Bandpass Filter Calibration.</u> The first step in the alignment procedure of each tuned filter is to ground the non-inverting input to the OA, apply a signal at the intended resonant frequency to the left end of R_3 with R_2 open and vary R_3 until unity gain (with inversion) is observed at the OA output. This insures that $R_3 = Q_1^{\,2} r$. The second step is to choose a value of positive feedback, k, by selection of R_4 and R_5 (greater positive feedback is required for low values of Q_1 in the parallel RLC tank circuit).

The desired circuit gain of unity and the selected circuit parameters of $R_3 = Q_1^2 r$ and $k = 1/2$ can be used in Equation 2.3.9 to find that the relation $R_1 = 2R_3$ must be satisfied. The remaining parameter, R_2, can be found by connecting the circuit as shown in Figure 2.17 and varying R_2 until the desired Q value of $Q = 25$ is observed. The approximate value of R_2 can be found from Equation 2.3.11, but the difficulty in accurately measuring and controlling the various components makes the preferable technique that of measuring Q and A. The range of control of Q, however, can be found from Equation 2.3.11, to be governed by the formula,

$$\frac{Q}{Q_1} = \frac{R_3}{R_2} + \frac{3}{2} \quad .$$

For positive component values, the maximum allowable value of Q_1 can be seen to be governed by

$$\frac{Q}{Q_1} > \frac{3}{2} \quad .$$

The desired active tuned-filter Q value of $Q = 25$ implies that $Q_1 < 16.8$, but in order to provide some measure of control with R_2, a Q_1 of the order of 10 was found to be about optimum. The Q_1 of the parallel LC branch can be reduced by adding a parallel resistance of the order of 10K. Alternately, the parallel reactance X_p may be varied in accord with the relation developed by Fristoe (32),

$$Q_1 = \frac{X_s}{r_s} = \frac{r_p}{X_p} \quad ,$$

in which r_s is the equivalent series resistance, r_p is the equivalent parallel resistance, X_s is the series reactance and X_p is the parallel reactance. The parallel reactance can be increased, and hence Q_1 can be reduced, by decreasing the capacitance and increasing the inductance. The change of reactance is the preferable method if additional inductance is available in the coil, since with near maximum inductance the tuning slug is well within the core gap and thus less susceptible to mechanical vibrations.

2.3.2 Fullwave Rectifier Design. The fullwave rectifier function is to invert the positive half cycle of the input waveform and to pass the negative half cycle, so that both half cycles have the same (negative) amplitude after processing by the rectifier circuit. For the positive half cycle of the input waveform, the positive input terminal of U2 is grounded, so U2 behaves as a simple sign reversing amplifier of 0.5 gain. For the negative half cycle of the input waveform, U2 serves as a simple follower amplifier having its' output connected to its' negative input terminal, and having its' positive input terminal excited by the R_7-R_8 voltage divider.

2.3.3 Low-Pass Filter Design. The low-pass filter converts the fullwave rectifier output into a DC voltage having an amplitude proportional to the AC input voltage to the fullwave rectifier. The circuit which was selected for this function was based on the work of Huelsman (29) who found that an OA with control components as shown in Figure 2.15 has a transfer characteristic given in general form by Equation 2.3.1. The choice of branch components as indicated in Figure 2.18 then results in a transfer function which simplifies to

$$\frac{V}{E} = \frac{-A}{(1 + j\omega T)^2} \quad ,$$

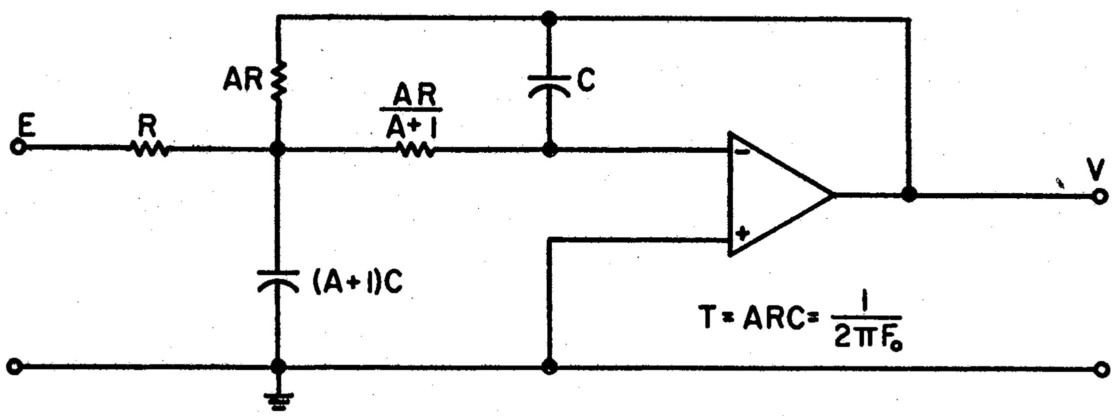


Figure 2.18. Low-Pass Filter

where A is the DC gain, $\omega = 2\pi B(K)$, $T = AR_K C_K$ (in seconds) $= [2\pi B(K)]^{-1}$, and B(K) is the frequency (in Hz) at which the output voltage is half the input voltage. The voltage gain as a function of frequency is shown in Figure 6 of Appendix B for the 800 Hz channel. The cutoff frequency (the -6 db gain point) of each low-pass filter was made equal to the bandwidth of the corresponding tuned bandpass filter in order to assure rapid response to the expected faster transients at the higher frequencies. Table II shows the component values selected for each channel as derived from the above equation.

The gain of each low-pass filter was selected to make the combined gain of the fullwave rectifier and low-pass filter equal to unity. A gain of approximately 3 is required to compensate for the fullwave

TABLE II

LOW-PASS FILTER DESIGN PARAMETERS

K (Channel No.)	F(K) (Hz)	B(K) (Hz)	$R_K C_K$ (msec)	R_K (kilohms)	C_K (μfd)
1	100	4.0	13.30	53.2	.25
2	104	4.2	12.80	51.2	.25
3	109	4.4	12.20	48.8	.25
4	113	4.5	11.70	46.8	.25
5	119	4.8	11.20	44.8	.25
6	124	5.0	10.70	42.8	.25
7	130	5.2	10.20	40.7	.25
8	135	5.4	9.87	39.4	.25
9	141	5.6	9.38	37.4	.25
10	147	5.9	9.06	36.2	.25
11	154	6.2	8.60	34.4	.25
12	160	6.4	8.30	33.2	.25
13	168	6.7	7.89	31.5	.25
14	175	7.0	7.60	30.4	.25
15	183	7.3	7.23	28.9	.25
16	191	7.6	7.00	28.0	.25
17	200	8.0	6.63	26.5	.25
18	208	8.3	6.40	25.6	.25
19	218	8.7	6.08	24.3	.25
20	227	9.1	5.88	23.5	.25
21	238	9.5	5.58	22.3	.25
22	247	9.9	5.37	21.5	.25
23	259	10.3	5.11	20.4	.25
24	270	10.8	4.93	19.7	.25
25	283	11.3	4.69	18.7	.25
26	294	11.8	4.51	18.0	.25
27	308	12.3	4.30	17.2	.25
28	321	12.8	4.15	16.6	.25
29	336	13.4	3.94	15.7	.25

TABLE II (Continued)

K (Channel No.)	F(K) (Hz)	B(K) (Hz)	$R_K C_K$ (msec)	R_K (kilohms)	C_K (μfd)
30	350	14.0	3.80	15.4	.25
31	367	14.7	3.62	14.5	.25
32	381	15.2	3.43	13.7	.25
33	400	16.0	3.32	13.3	.25
34	416	16.6	3.20	12.8	.25
35	436	17.4	3.04	60.7	.05
36	454	18.1	2.94	58.8	.05
37	476	19.0	2.79	55.7	.05
38	495	19.8	2.69	53.9	.05
39	519	20.7	2.56	51.2	.05
40	539	21.5	2.50	50.0	.05
41	566	22.6	2.34	46.7	.05
42	588	23.5	2.26	45.3	.05
43	617	24.6	2.15	43.0	.05
44	642	25.7	2.08	41.6	.05
45	673	26.9	1.97	39.4	.05
46	700	28.0	1.90	38.0	.05
47	734	29.3	1.81	36.2	.05
48	763	30.5	1.74	34.8	.05
49	800	32.0	1.66	33.2	.05
50	832	33.3	1.60	32.0	.05
51	872	34.8	1.52	30.4	.05
52	907	36.3	1.46	29.2	.05
53	951	38.0	1.40	28.0	.05
54	989	39.5	1.35	27.0	.05
55	1038	41.5	1.28	25.6	.05
56	1079	43.2	1.24	24.8	.05
57	1131	45.4	1.17	23.4	.05
58	1177	47.1	1.13	22.6	.05
59	1234	49.4	1.07	21.4	.05
60	1283	51.4	1.03	20.6	.05

TABLE II (Continued)

K (Channel No.)	F(K) (Hz)	B(K) (Hz)	R_KC_K (msec)	R_K (kilohms)	C_K (μfd)
61	1345	53.8	.98	19.7	.05
62	1399	56.0	.95	19.0	.05
63	1467	58.7	.90	18.1	.05
64	1525	61.0	.87	17.5	.05
65	1600	64.0	.83	16.6	.05
66	1664	65.8	.80	16.0	.05
67	1745	69.8	.76	76.0	.01
68	1815	72.7	.74	73.6	.01
69	1903	76.7	.70	69.7	.01
70	1979	79.1	.67	67.4	.01
71	2075	83.0	.64	63.9	.01
72	2158	86.3	.62	61.6	.01
73	2263	90.8	.59	58.6	.01
74	2353	94.1	.57	56.6	.01
75	2468	98.9	.54	53.7	.01
76	2566	102.7	.52	52.0	.01
77	2691	107.5	.49	49.3	.01
78	2799	112.0	.48	47.5	.01
79	2934	117.3	.45	45.2	.01
80	3052	124.0	.43	43.0	.01
81	3200	128.0	.41	41.4	.01
82	3328	133.0	.40	40.0	.01
83	3490	139.0	.38	38.0	.01
84	3629	145.0	.37	37.0	.01
85	3806	152.2	.35	34.9	.01
86	3958	158.2	.34	33.6	.01
87	4150	166.0	.32	32.0	.01
88	4316	172.5	.31	31.0	.01
89	4526	181.0	.29	29.3	.01
90	4707	188.2	.28	28.3	.01
91	4935	197.0	.27	26.9	.01

TABLE II (Continued)

K (Channel No.)	F(K) (Hz)	B(K) (Hz)	$R_K C_K$ (msec)	R_K (kilohms)	C_K (μfd)
92	5132	205.2	.26	26.0	.01
93	5382	215.0	.25	24.6	.01
94	5597	223.8	.24	23.7	.01
95	5869	234.8	.23	22.9	.01
96	6104	244.1	.22	21.8	.01
97	6400	256.0	.21	20.7	.01

rectifier gain of 0.5, and for the factor of 0.636 involved in converting AC to DC. A ±10 volt sinusoidal input at the bandpass filter resonant frequency causes the low-pass filter output to be about +10 volts DC. The resistance divider on the output of each low-pass filter (Figure 2.12) was used to set the multiplexer input voltage at 5 ± 0.1 volts with a ±10 volt resonant sine wave input to the bandpass filter.

2.4 Performance Verification. Prior to actual use in tests with speech signals, the complete system was tested for calibration, response speed, and linearity by use of a sine wave oscillator equipped with a frequency counter, and an accurate CRT equipped with a scope camera.

2.4.1 The Microphone, Amplifiers and Recorder. The Shure 545S microphone provides a voltage output as shown in Figure 2.19 for a normal room volume speaking voice and the microphone held about 8 inches directly in front of the mouth. The sound which is recorded is the sustained phoneme /ae/ as in at.

The Pre-Amplifier performance at 142 Hz and 6100 Hz is shown by the Lissajous curves in Figures 2.20 and 2.21. The smoothness at both frequencies shows that no significant distortion is present at either end of the speech spectrum. The oscillator was tuned through the intermediate frequencies with similar results. The phase shift of about 40 degrees at 142 Hz can be seen in Figure 2.20, although this is of only casual interest because it has no effect on the spectral filter response. At the high end of the spectrum the phase shift is observed to be less than 10 degrees.

The Soft Limiter action is demonstrated by Figures 2.22 and 2.23, for which the input was intentionally increased to ±5 volts to demonstrate the soft limiting action which begins at an input of ±3.2 volts and is completed by about ±4 volts. The softness of the limiting action is best shown by the Lissajous patterns in Figures 2.22B and 2.23B, which illustrate the smooth curvature of the corners.

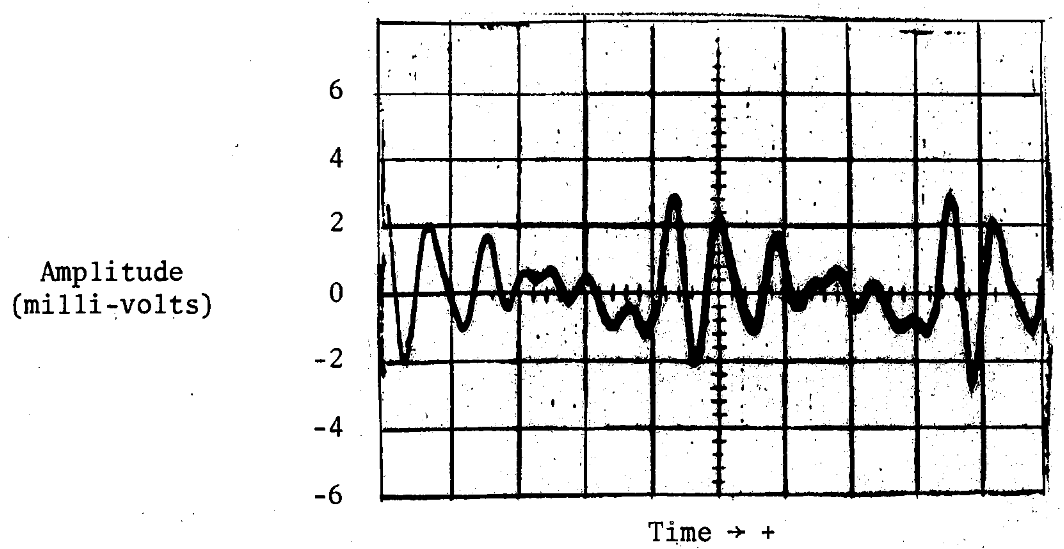


Figure 2.19. Shure 545S Microphone Output for the Phoneme /ae/ as in At (GPR=140 Hz)

The Speech Amplifier input-output characteristic is shown for 142 Hz and 6100 Hz in Figures 2.24 and 2.25. The push-pull power amplifier and transformer can be seen to cause no apparent distortion, and negligible phase shift, over the frequency spectrum of interest in speech research.

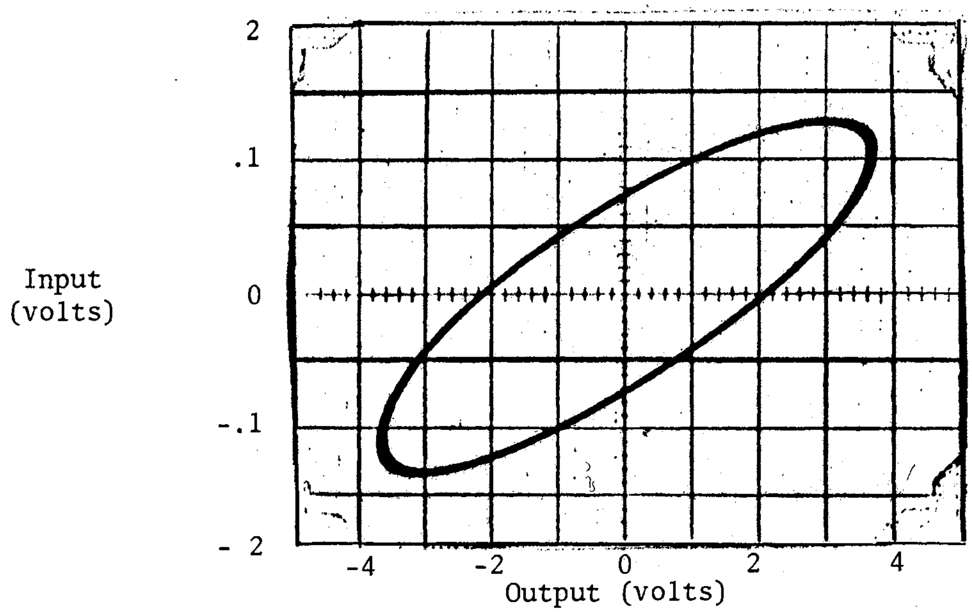


Figure 2.20. Pre-Amplifier Lissajous Pattern With a 142 Hz Sine Wave Input

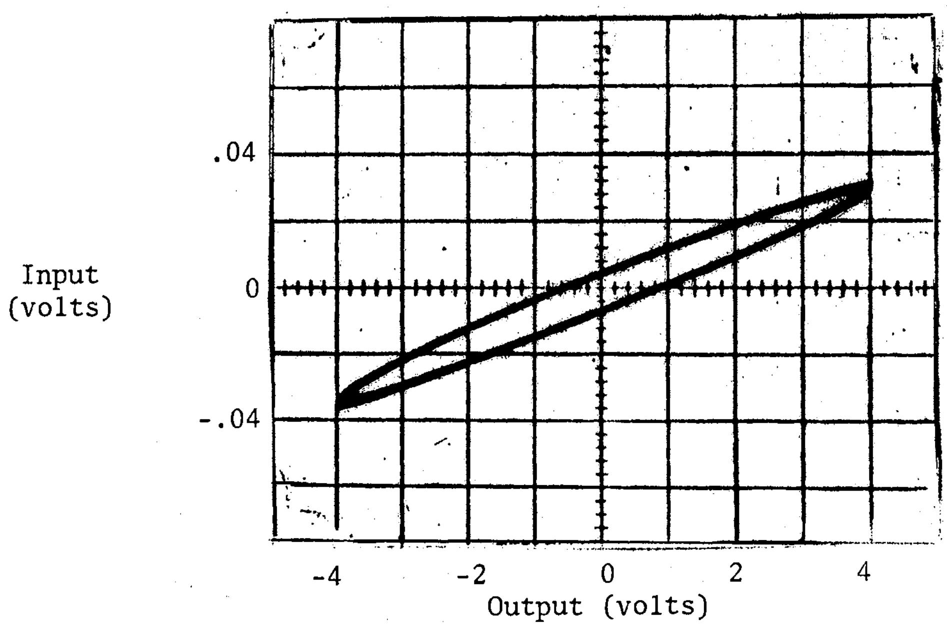


Figure 2.21. Pre-Amplifier Lissajous Pattern With a 6100 Hz Sine Wave Input

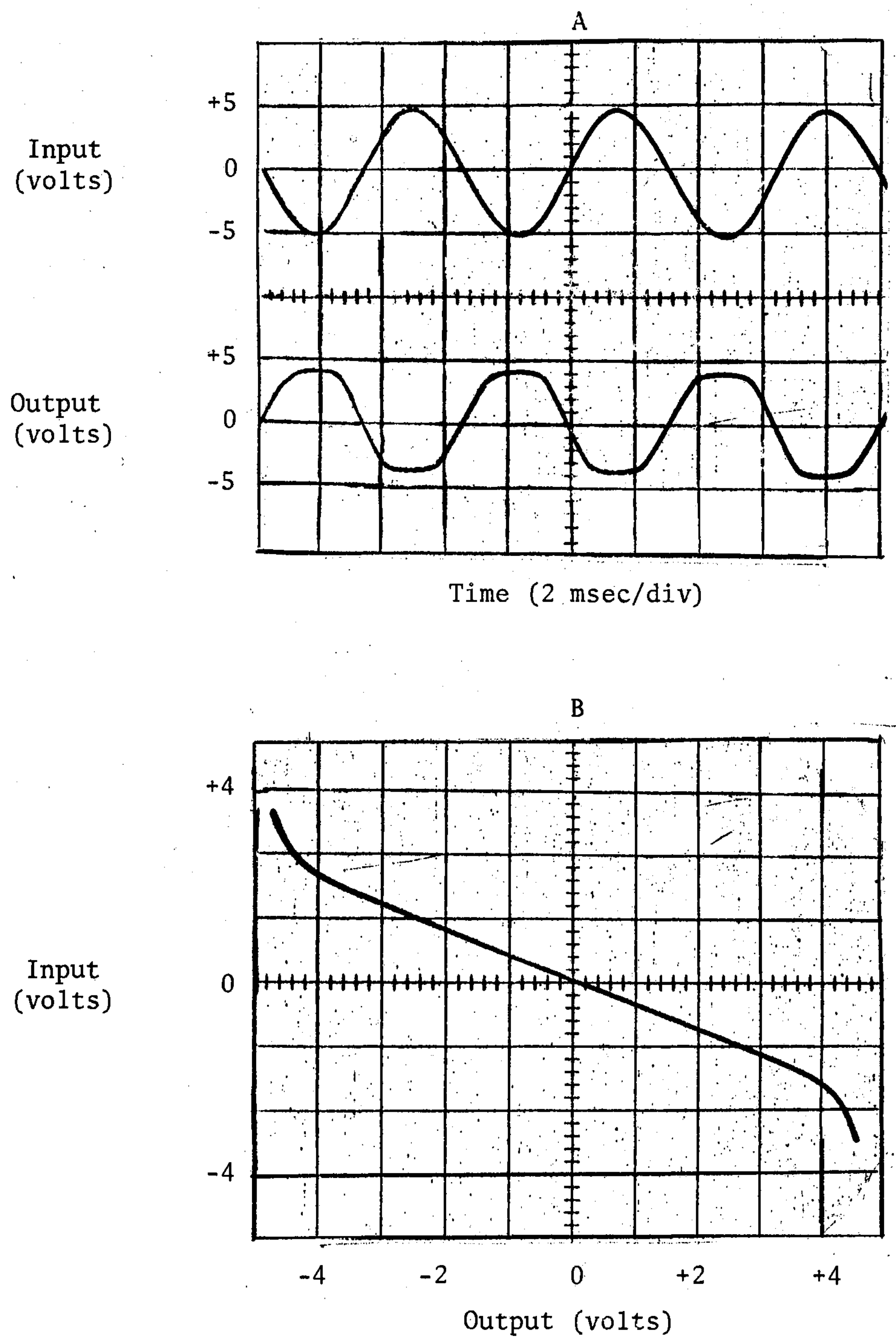


Figure 2.22. Soft Limiter Input and Output Waveforms (A) and Lissajous Pattern (B) With a 142 Hz Sine Wave Input

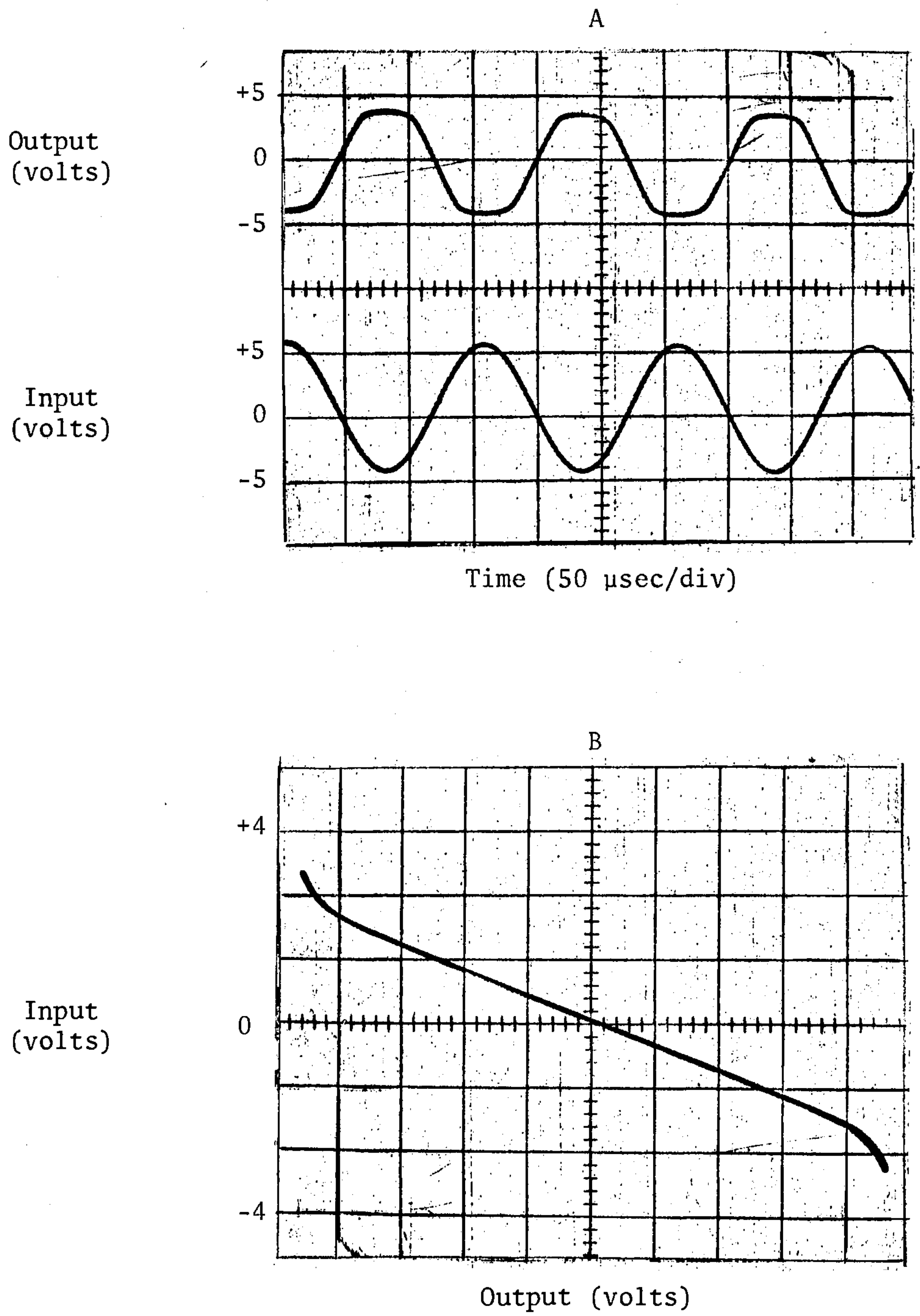


Figure 2.23. Soft Limiter Input and Output Waveforms (A) and Lissajous Pattern (B) With a 6100 Hz Sine Wave Input

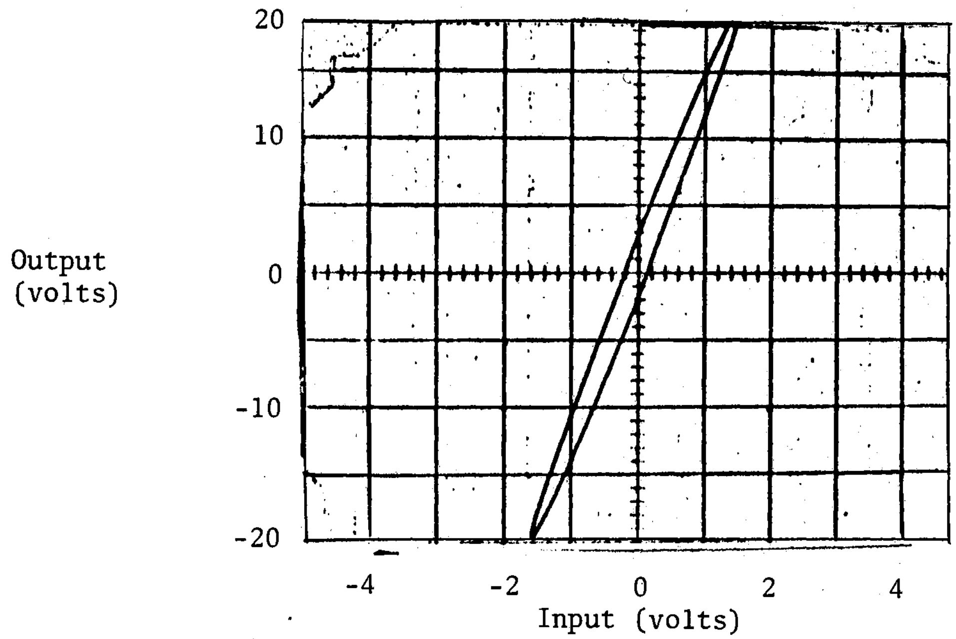


Figure 2.24. Speech Amplifier Lissajous Pattern
With a 142 Hz Sine Wave Input

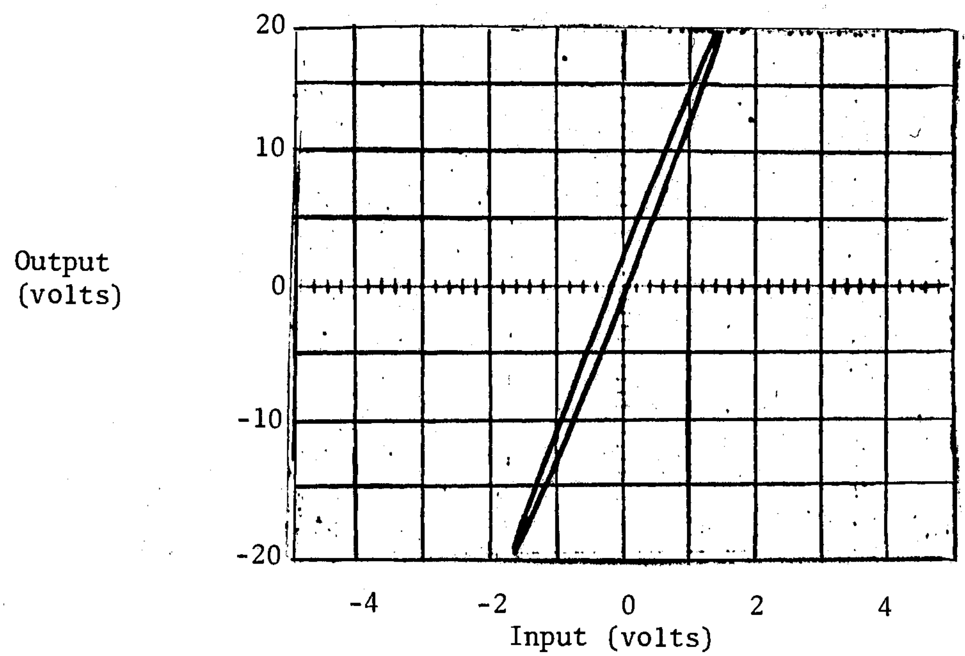


Figure 2.25. Speech Amplifier Lissajous Pattern
With a 6100 Hz Sine Wave Input

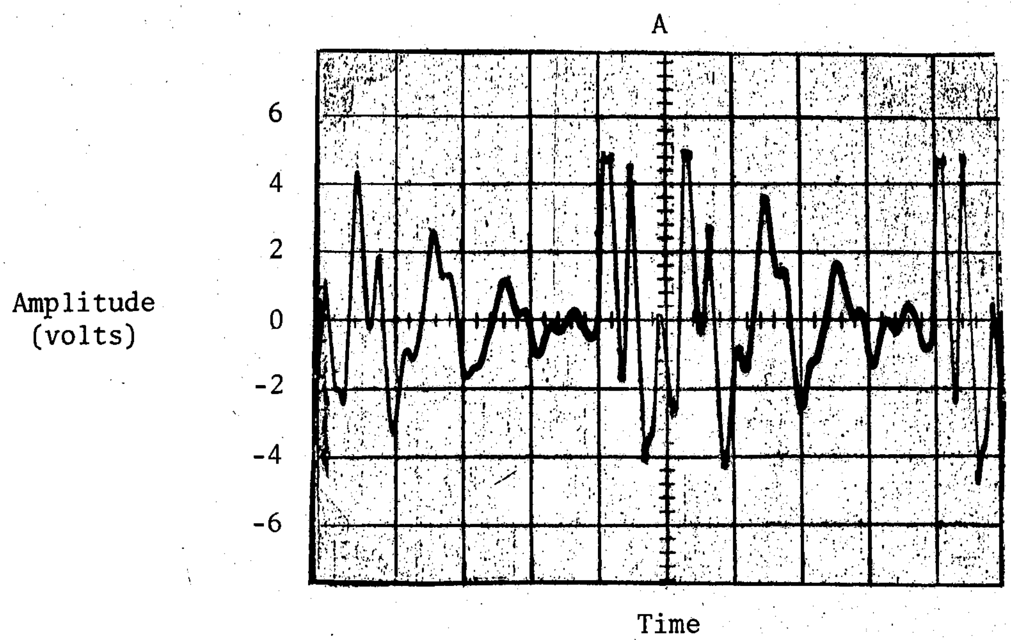
The Wollensak Model 1980 Magnetic Tape Recorder output voltage level is set by a volume control on the recorder. By observing the output wave on a CRT it was found that clipping began at different gain settings for different phoneme tapes. This clipping effect is shown in Figure 2.26A for the phoneme /ae/. The same sound is shown in Figure 2.26B with a gain setting which removed the clipping.

2.4.2 Tuned Bandpass Filters. The performance of a typical tuned bandpass filter is shown by a plot of the output voltage amplitude as a function of frequency for the 800 Hz filter and is included in Appendix B as Figure 6. The verification of this characteristic for each tuned bandpass filter was conducted as a routine part of the construction phase. The -3 db points for each filter are recorded on Table I.

During the construction of each of the tuned filters the voltage amplitude measurements were made with the Tektronix Type 564 Storage Oscilloscope which provides 8 cm of vertical viewing distance and 10 cm in the horizontal dimension. Repeated measurements revealed that the error in voltage amplitude could be held to about 2%. The measurement of frequency provided by the General Radio Model 1151-A Frequency Counter was accurate to within ± 1 Hz for one second of integration time.

The expected error in bandwidth measurement due to statistical measurement variations can be developed from the magnitude of the voltage transfer function $|V/E|$ which is

$$\left|\frac{V}{E}\right| = \left| \frac{-1}{1 + j\,Q\left[\frac{F}{F_0} + \frac{F_0}{F}\right]} \right| = \left[1 + Q^2\left[\frac{F}{F_0} - \frac{F_0}{F}\right]^2 \right]^{-\frac{1}{2}} .$$



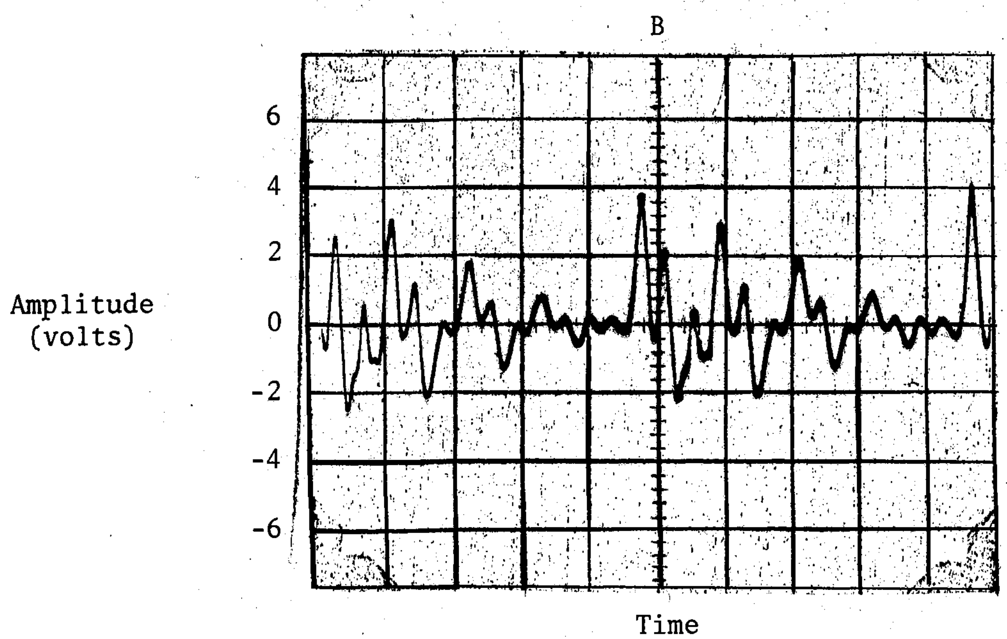


Figure 2.26. Tape Recorder Output When Overdriven (A) and When Not Overdriven (B) [Phoneme is /ae/ as in <u>A</u>t]

The slope of $|V/E|$ at the half power points is given by

$$\frac{\partial \left|\frac{V}{E}\right|}{\partial F} = \pm \frac{1}{2} \left|\frac{V}{E}\right|^{+3} \cdot Q^2 \cdot (2) \cdot \left[\frac{F}{F_0} - \frac{F_0}{F}\right] \cdot \left[\frac{1}{F_0} + \frac{F_0}{F^2}\right]$$

$$= \pm \left|\frac{V}{E}\right|^{+3} \cdot \left[\frac{F}{F_0} - \frac{F_0}{F}\right] \cdot \frac{Q^2}{F} \cdot \left[\frac{F}{F_0} - \frac{F_0}{F}\right] .$$

Evaluating the slope at the upper half power point $F_2 = F_0\left[1 + \frac{1}{2Q}\right]$ yields

$$\left.\frac{\partial \frac{V}{E}}{\partial F}\right|_{F = F_2} = -(.707)^3 \cdot \left[\frac{Q}{F_0\left[1 + \frac{1}{2Q}\right]}\right] \cdot \left[1 + \frac{1}{2} Q + \frac{1}{1 + \frac{1}{2Q}}\right]$$

$$\cong - \frac{(.707)}{2B} \cdot \left[1 + \left[1 - \frac{1}{Q}\right]\right] = - \frac{.707}{B} .$$

Thus for the 6400 Hz filter

$$\left.\frac{\partial \frac{V}{E}}{\partial F}\right|_{F = F_2} = - \frac{.707}{B} = - \frac{.707}{256} = - .00276 .$$

The standard deviation of error in frequency measurement of F_2 is then

$$\Delta F_2 = \frac{\Delta \left|\frac{V}{E}\right|}{.00276} = \frac{.02}{.00276} = 7.25 \text{ Hz} .$$

The bandwidth measurement accuracy of the 6400 Hz tuned filter is thus

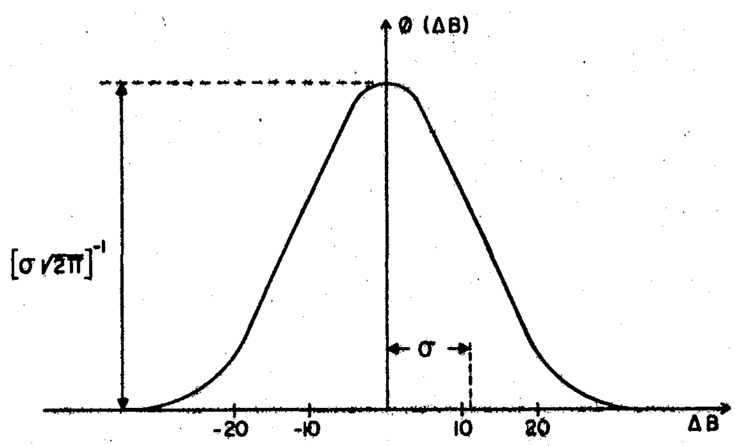
$$B = F_2 - F_1 = 256 \pm 7.25 \sqrt{2} \text{ Hz} \quad .$$

A gaussian distribution of the bandwidth measurement errors for the 6400 Hz filter is plotted in Figure 2.27. It can be shown from this figure that 31.8% of the bandwidth measurements for this filter can be expected to be in error by more than 10.24 Hz, and similarly, that 4.6% of the measurements can be expected to be in error by 20.48 Hz or more.

The standard error of measurement (ΔB) decreases as the resonant frequency is reduced, in accord with formula,

$$\Delta B = (\Delta F_2)\sqrt{2} = \left[\frac{B}{.707}\right] \cdot \Delta \left|\frac{V}{E}\right| \cdot \sqrt{2} = .0016\ F_0 \quad .$$

Table III shows the range of bandwidth measurement error and several representative values.



$$\phi(\Delta B) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \text{EXP}\left[\frac{-(\Delta B)^2}{2\sigma^2}\right]$$

Figure 2.27. Assumed Distribution of Bandwidth Measurement Error (at 6400 Hz)

TABLE III

TYPICAL BANDWIDTH MEASUREMENT ERROR

Resonant Frequency (Hz)	Bandwidth (Hz)	Standard Bandwidth Measurement Error (Hz)
100	4	0.16
200	8	0.32
400	16	0.64
800	32	1.28
1600	64	2.56
2468	98	3.94
3200	123	5.12
4935	197	7.90
6400	256	10.24

This analysis has revealed the fact that measurement error alone would restrict the control of tuned filter bandwidths to about 5%. This percentage was set as a performance standard during the tuned filter construction phase, and was used as a standard during periodic performance checks made prior to the taking of experimental data.

2.4.3 Fullwave Rectifiers. Figures 2.28A, B, and C show the input and output waveforms of the fullwave rectifiers with a pure sine-wave input at the resonant frequency of the 100, 800 and 6400 Hz channels. It is seen that the desired objective of the fullwave rectification is attained.

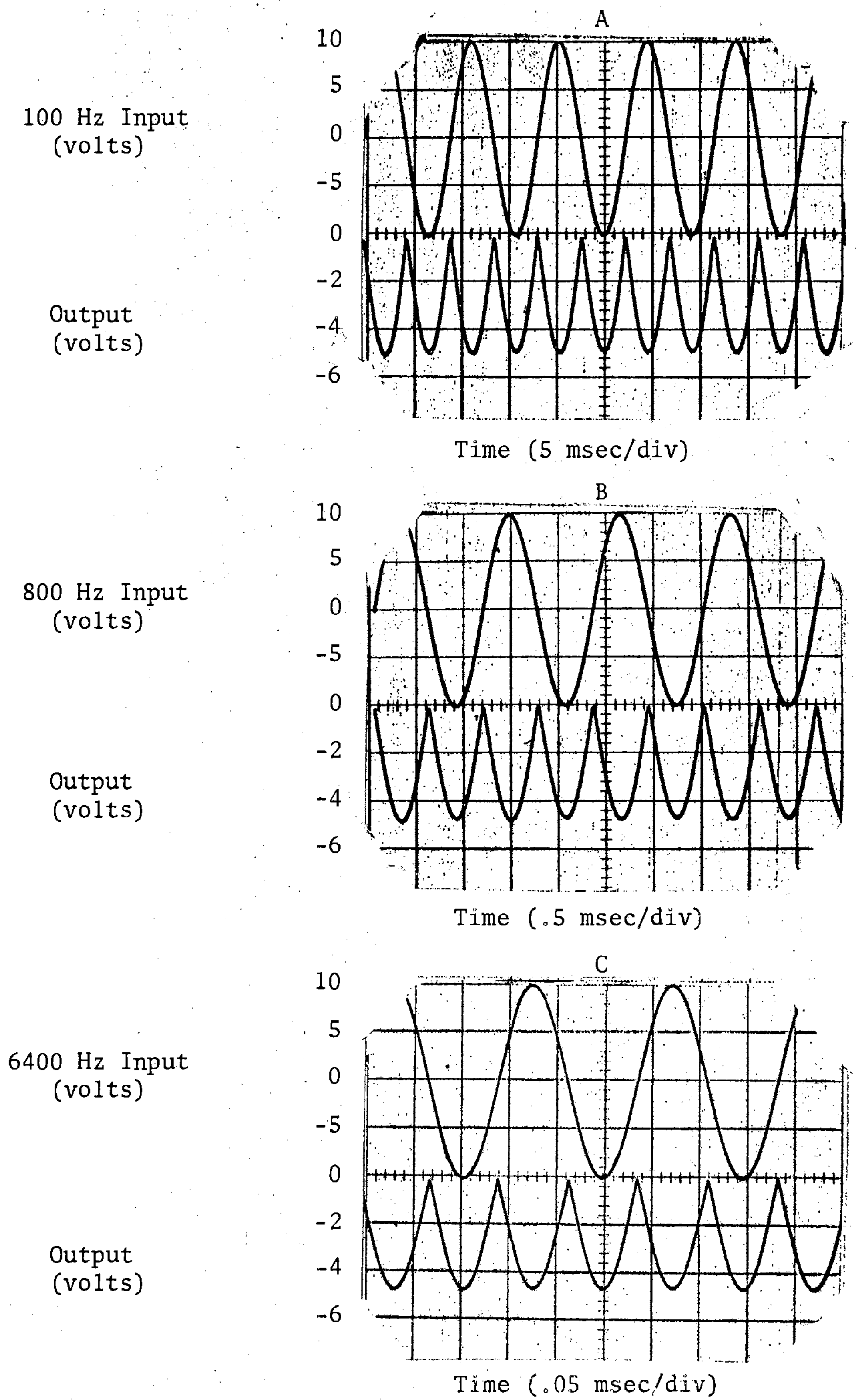
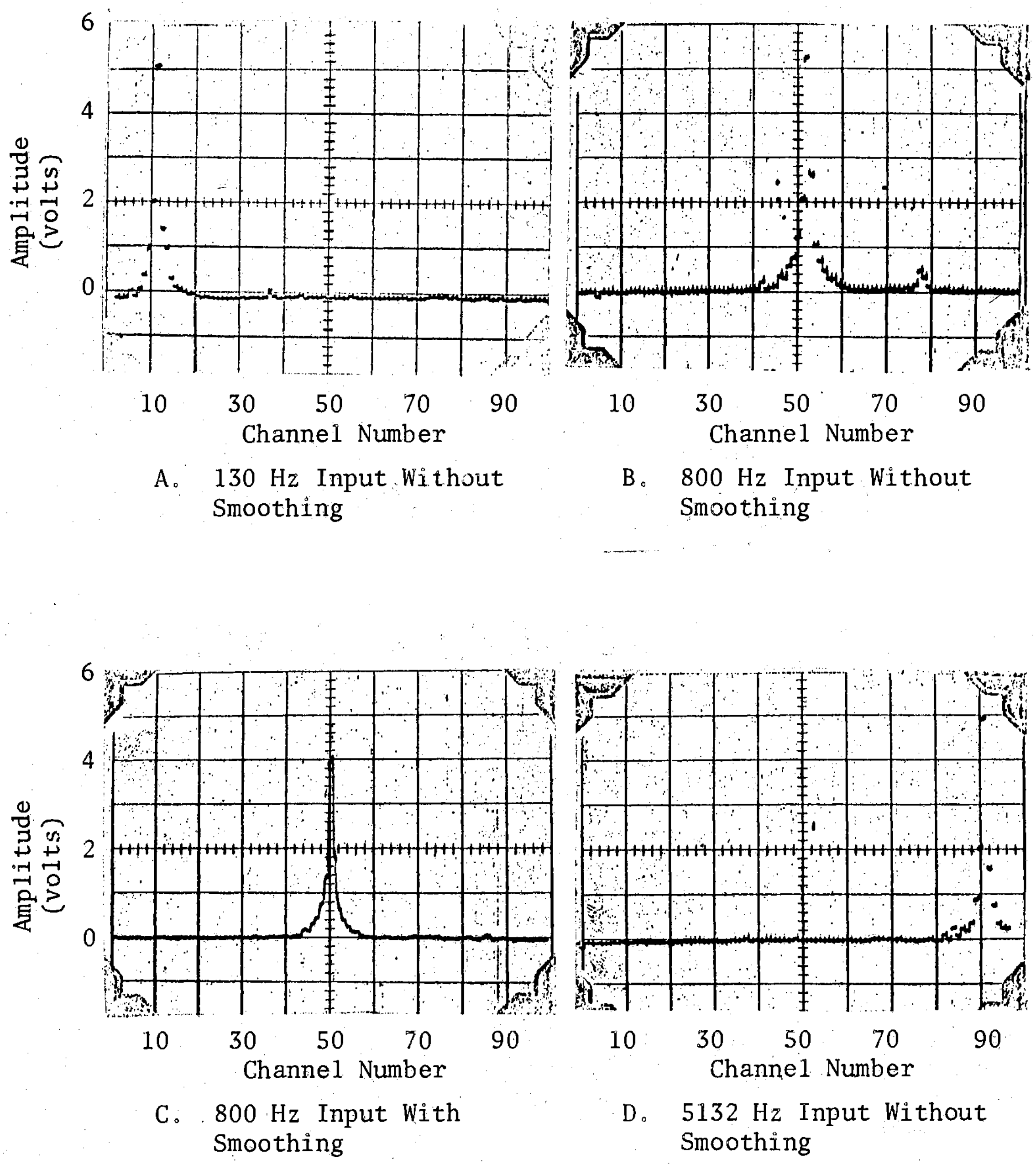


Figure 2.28. Fullwave Rectifier Input (Top) and Output (Bottom) When Driven With 100 Hz (A), 800 Hz (B) and 6400 Hz (C) Sine Waves

2.4.4 Low-Pass Filters. The low-pass filter performance is graphically illustrated by Figure 7 fo Appendix B, which shows the DC output voltage level as a function of frequency. Again, the 800 Hz channel was used, for which the -6 db point is at a frequency of 32 Hz, corresponding to the 32 Hz bandwidth of the 800 Hz tuned bandpass filter.

2.4.5 Spectral Resolution. The TAT System as shown in Figure 2.2 provides a CRT display of speech energy content as a function of frequency. The CRT horizontal axis is calibrated linearly with filter number, and thus exponentially with filter frequency, beginning with the 100 Hz filter on the left and ending with the 6400 Hz filter on the right. The vertical axis provides a linear measure of each tuned amplitude tracker channel output voltage. This level depends upon both the amplitude of that portion of the speech spectrum which is within the bandpass of the pertinent channel, and upon the bandwidth of the low-pass filter, which determines the maximum rate at which the filters will respond. The frequency resolution of the system is illustrated, for pure sine wave inputs generated by a Hewlett Packard Model 3300A Function Generator, in Figures 2.29A, B, C, and D. Figure 2.29 A is the system response to a 130 Hz signal. Figure 2.29 is the same display for an 800 Hz input. Figure 2.29C shows the display smoothing which can be achieved by placing a simple 10K-360 pfd integrator in series with the multiplexer output. Figure 2.29D shows a display with a 5132 Hz input, generated without the smoothing integrator.

2.5 Construction. During the audio filter construction phase, technical assistance was provided by several Electrical Engineering



A. 130 Hz Input Without Smoothing

B. 800 Hz Input Without Smoothing

C. 800 Hz Input With Smoothing

D. 5132 Hz Input Without Smoothing

Figure 2.29. TAT System Display With Sine Wave Inputs

students. The work of J. A. Avallone, D. J. Hansen, M. J. Johnston, W. A. Martin, J. M. Meyer, L. J. Paleck, M. D. Pelly and C. D. Shepherd is greatly appreciated.

CHAPTER III

SPEECH SAMPLER DISPLAY

3.1 Introduction. The pattern of energy content as a function of frequency, displayed as either an instantaneous sample or as an integrated sum over some time interval, may be a desirable recognition display for those phonemes which are time invariant except for small glottal pulse fluctuations. Many phonemes, however, exhibit transient spectral-amplitude characteristics which require the additional dimension of time for identification. More importantly, connected speech involves the combination of phoneme groupings in which the characteristics of each phoneme depend heavily upon both the vocal tract initial condition, which is a function of previous utterances, and upon the anticipatory muscular condition, which is a result of expected future utterances. Time is, therefore, an essential parameter of measurement in speech research. This chapter includes a description of a scheme for the display on a CRT, of spectral-amplitude speech characteristics over a time sample of 64 msec. The time sample can be started automatically upon word initiation or can be started at the discretion of the operator by means of a panel mounted switch. The 64 msec time sample was purely an arbitrary choice which was hoped to be sufficient to evaluate the technique of display. It was a convenient choice because it used the maximum 50,000 Hz sampling rate of the EECO 765 Multiplexer, it used all 97 channels of the filter bank, it could be implemented with a 6-bit binary counter and the 32 traces appeared to be an upper limit in resolution on the largest CRT display for which photographic instrumentation was available in the University Laboratory. Further experience with the display will probably suggest a more optimum trade-off.

The implementation of this scheme, the Speech Sampler, is shown in block diagram form in Figure 3.1. It differs from the Tuned Amplitude Tracker System of Chapter II only in the manner of CRT control. In Chapter II the sweep generation was provided by the CRT internal circuitry with only a sync trigger and the analog vertical deflection voltage provided by the EECO 765 Multiplexer. The Speech Sampler provides analog voltages for complete electron beam control.

The display is made up of 32 spectral-amplitude plots, each of which is identical to that which was described in Chapter II, except for size and position. A three dimensional effect is obtained by shortening the spectral and amplitude dimensions of each trace and positioning each 2 msec spectral-amplitude sample down and to the left of the previous sample. Figures 3.2A, B, and C illustrate the scheme and show a 64 msec time sample of the Speech Sampler response to a 109 Hz, an 800 Hz and a 5132 Hz pure sine-wave as generated by a Hewlett Packard Model 3300A Function Generator. The signal is fed to the Speech Amplifier for this illustration. The value of the Speech Sampler in displaying transient effects is illustrated in Figures 3.3A, B, and C where the voiced stop-consonant-vowel transitions which are used are /b/ + /ae/ as in <u>bad</u>, /g/ + /ae/ as in <u>gab</u>, and /d/ + /ae/ as in <u>dad</u>. A male voice and the microphone were used as the sound source.

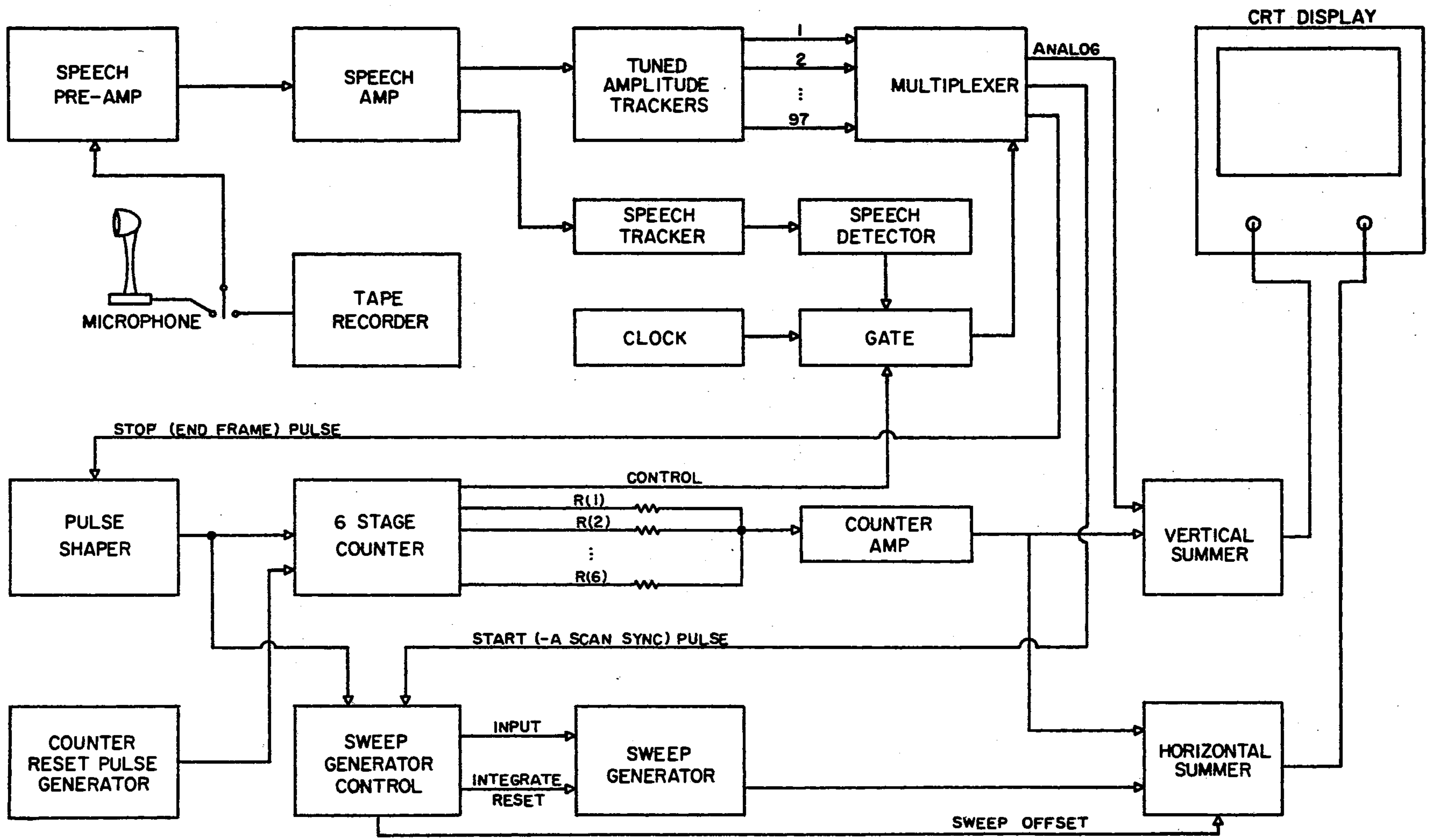


Figure 3.1. Speech Sampler Block Diagram

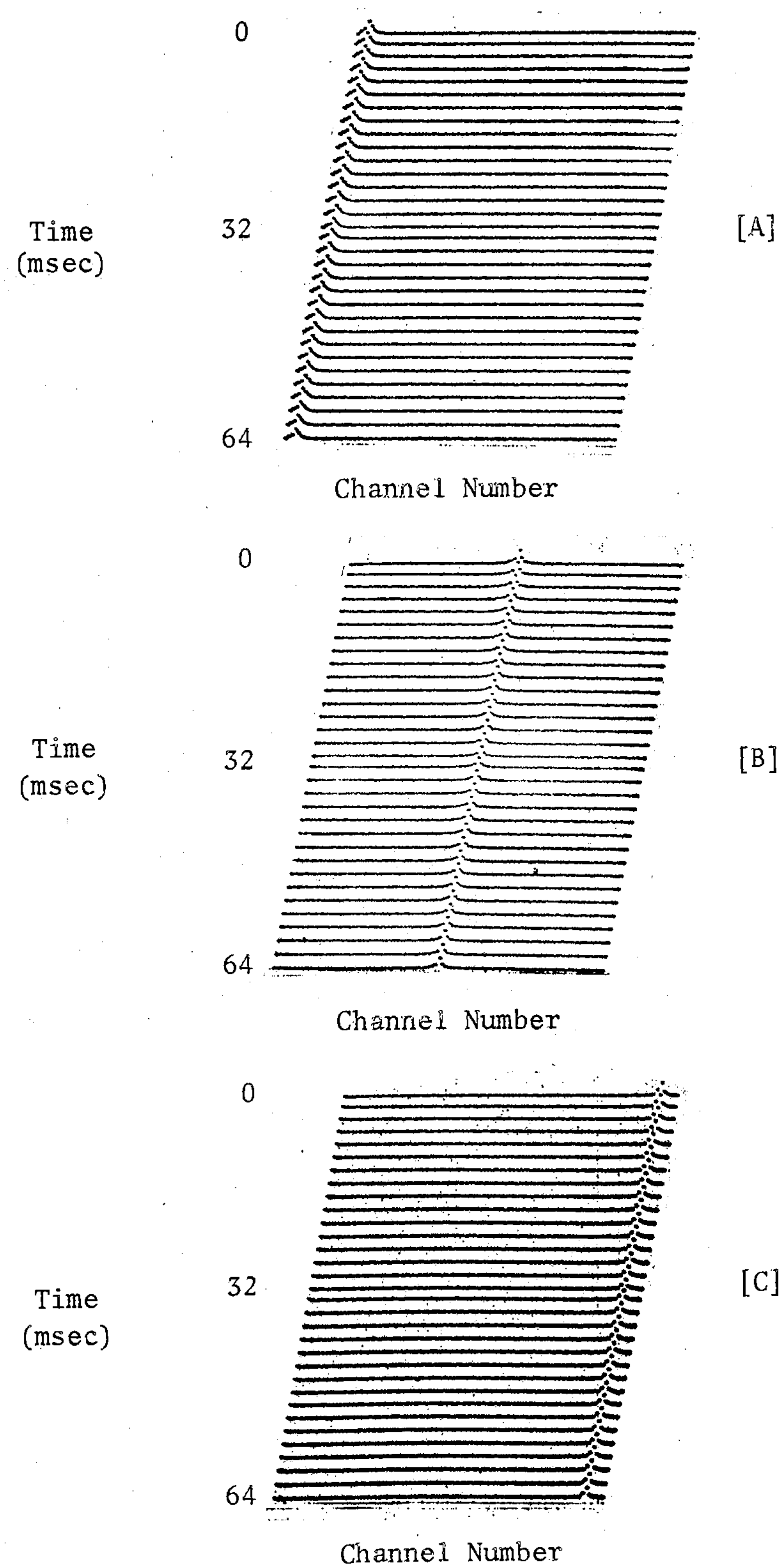


Figure 3.2. Speech Sampler Display for Sine Wave Inputs of 109 Hz (Top), 800 Hz (Middle) and 5132 Hz (Bottom)

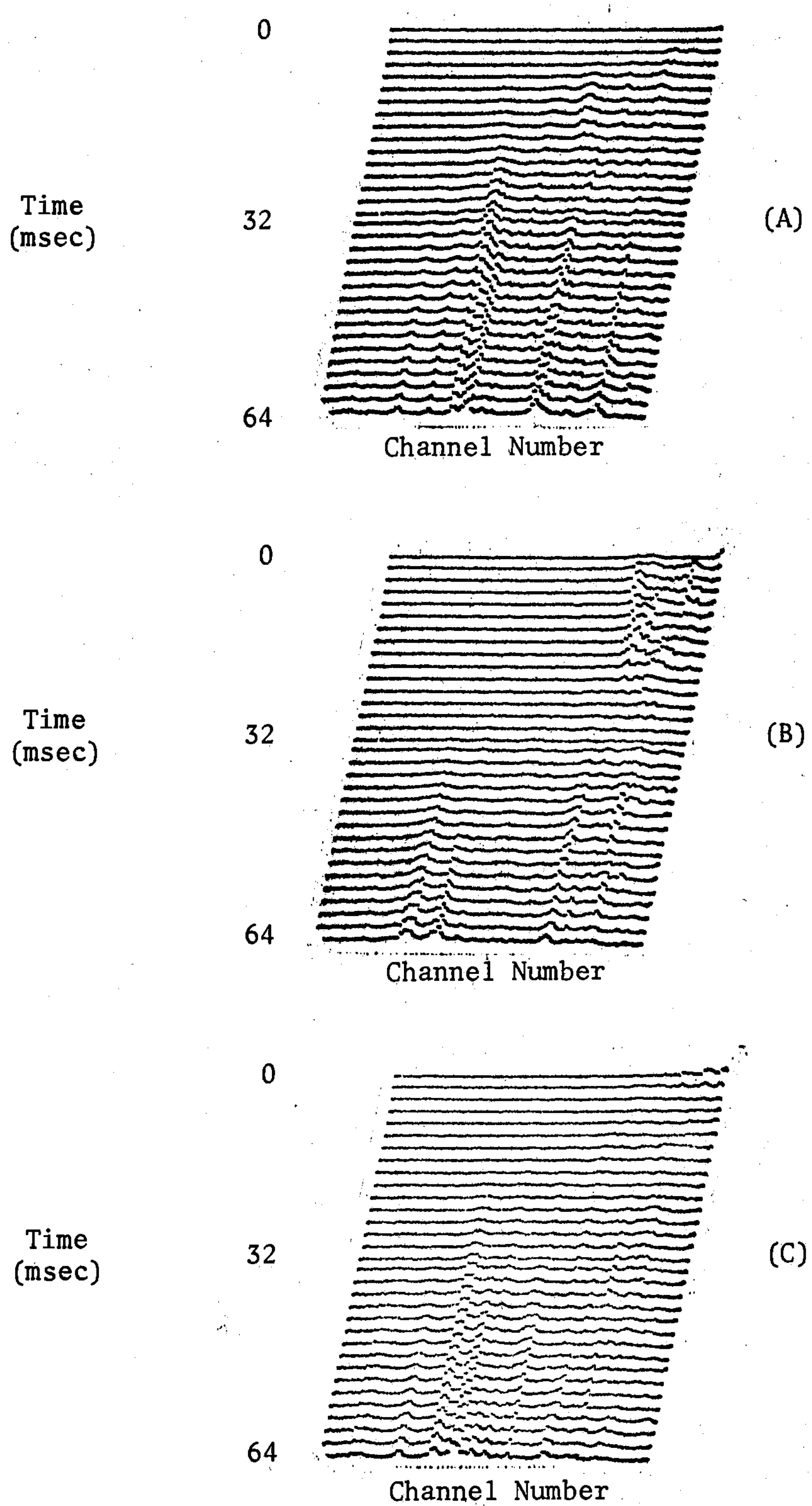
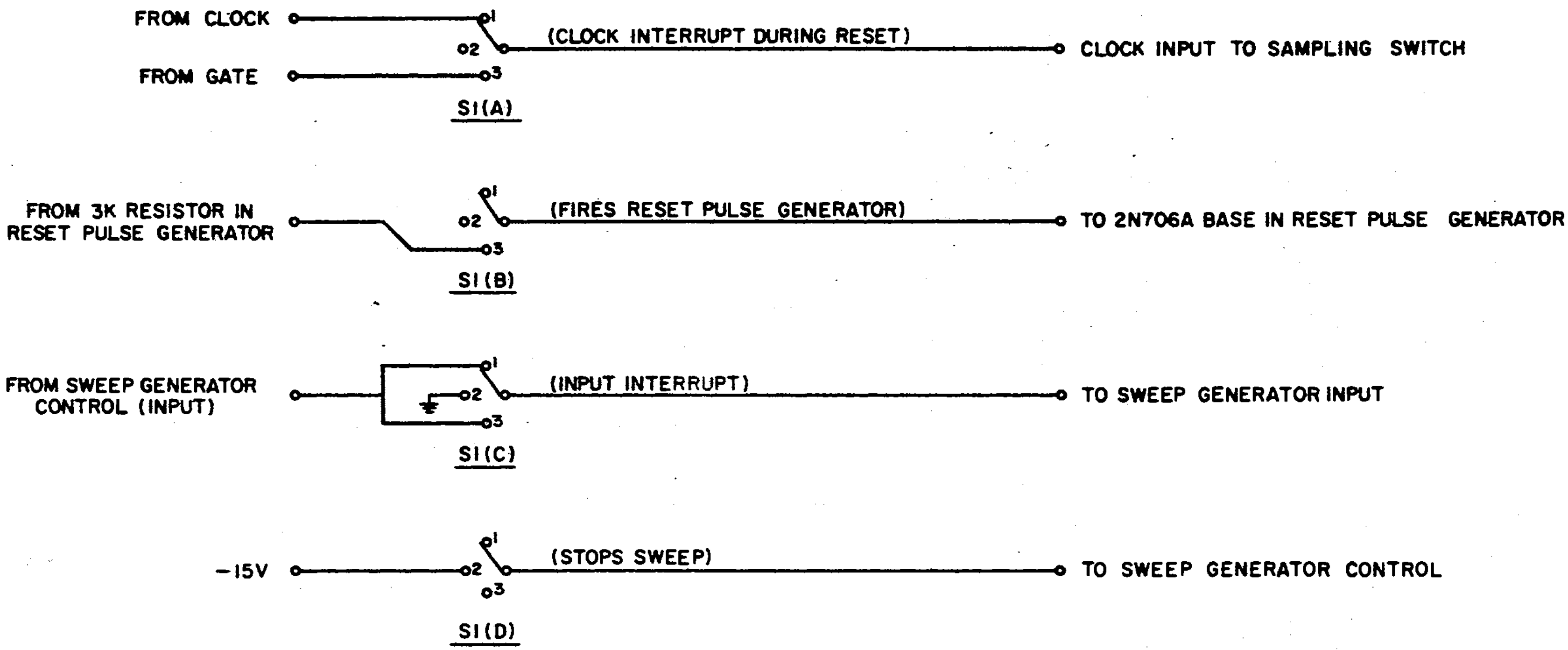


Figure 3.3. Speech Sampler Display of Phoneme Transition /b/+/ae/ as in Bat (A), /g/+/ae/ as in Gab (B), and /d/+/ae/ as in Dad (C)

3.2 Hardware Implementation. The remainder of this chapter will be devoted to a functional description of the circuitry which is used to achieve the three dimensional Speech Sampler Display. The heart of this display system is the multiplexer which, as depicted in Figure 3.1, time-multiplexes all 97 DC voltage levels from the tuned amplitude trackers onto a single output line at a rate determined by the clock pulses which are permitted to pass through the Gate. The multiplexer also provides start and stop pulses which are used to control the sweep.

3.2.1 Mode Selector Switch. Two modes of operation are possible with this display. The first is a continuously running sample as was used in Figure 3.2. In this case the Gate in Figure 3.1 is by-passed and the sweep position is determined entirely by the state of the six stage counter. The multiplexer analog output is displayed for the first 32 sweeps which represent 64 msec of speech. This is followed by a 64 msec period during which the sixth counter stage provides a blanking pulse which removes the electron beam from the CRT display area. This sequence is repeated as long as the Mode Selector Switch, S1, is in the "CONTINUOUS" position. The logic diagram for this switch is shown in Figure 3.4. In the "CONTINUOUS" position the first level, S1(A), of the four level three position switch provides a gate by-pass, for the clock pulse, to the multiplexer. The third level of the switch provides an input voltage level to the Sweep Generator.

The second mode of operation which can be selected by means of S1 is a single 64 msec sample of speech after which the CRT display remains blank. This display is obtained by first placing S1 in the "RESET" position and then placing it in the "SAMPLE" position. In the "RESET" position, S1(A) stops the multiplexer stepping by interrupting the clock



S1 POSITION	MODE OF OPERATION
1	CONTINUOUS (SWEEP)
2	RESET (PRIOR TO 32 SWEEP SAMPLE)
3	SAMPLE (32 SWEEPS)

Figure 3.4. S1, Mode Selector Switch Diagram

trigger line, S1(C) grounds the input to the Sweep Generator, and S1(D) provides a sweep stop voltage to the Sweep Generator control. When the switch is moved to the third position ("SAMPLE") S1(A) permits those clock pulses which are passed by the Gate, to go to the multiplexer. Switch level S1(B) triggers the Reset Pulse Generator, S1(C) applies the input voltage to the Sweep Generator, and S1(D) removes the sweep stop voltage from the Sweep Generator Control. The purpose of each of these signals is discussed with the circuits in which they are used.

3.2.2 Speech Tracker. The circuit which is used to provide a voltage level during the presence of speech, in order to synchronize the start of the 64 msec sample, is the Speech Tracker which is shown in Figure 3.5. It is very similar to the fullwave rectifier and low-pass filter which is used in the 6400 Hz TAT channel. It has the same 256 Hz corner frequency in order to be compatible with the response time of the fastest channel. A blocking capacitor is used in series with the input to provide DC isolation. The input to the Speech Tracker is the amplified speech, from the amplifier, which is also applied to the input of each of the 97 TAT channels. The output DC voltage level is used in the Speech Detector.

3.2.3 Speech Detector. The circuit shown in Figure 3.6 accepts the output voltage from the Speech Tracker and, when it exceeds +6 volts, generates a -10 volt pulse of 80 msec duration. This pulse opens the Gate at the start of speech if the mode selector switch is in the "SAMPLE" position. The means by which this pulse is created may be examined by reference to Figure 3.6. The left two transistors T1 and T2, make up a Schmitt trigger in which T2 is normally conducting. When

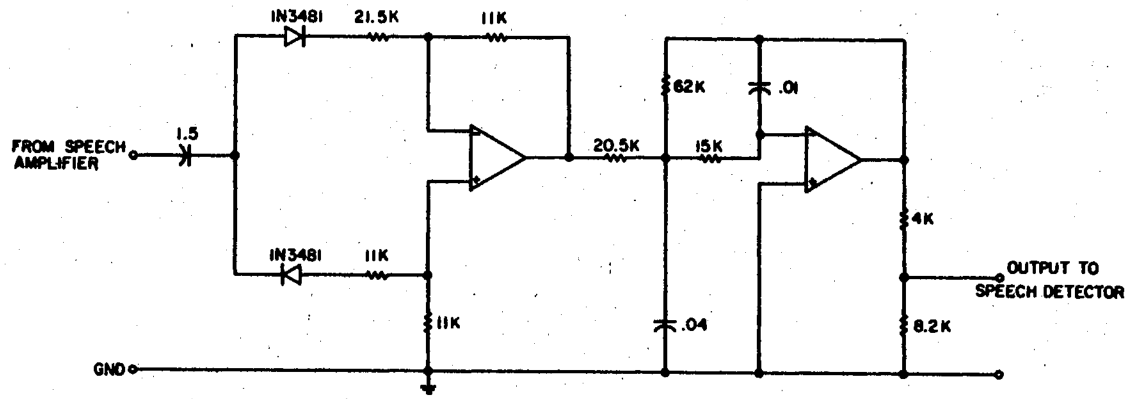


Figure 3.5. Speech Tracker

the input voltage reaches +6 volts the trigger fires (T1 turns "ON" and T2 turns "OFF") creating a positive going voltage on the collector of T2. This becomes a positive spike when passed by means of the 50 pfd capacitor and 22K resistor to T3 and T4 which make up a monostable multivibrator. The spike switches the multivibrator which generates the 80 msec pulse, the width of which is determined by the 1 μfd capacitor and the 270K resistor. The emitter follower, T5, provides isolation. When switch S2 is in the "IN" position the Detector output is passed to the Gate. When S2 is in the "OUT" position the control line to the Gate is open.

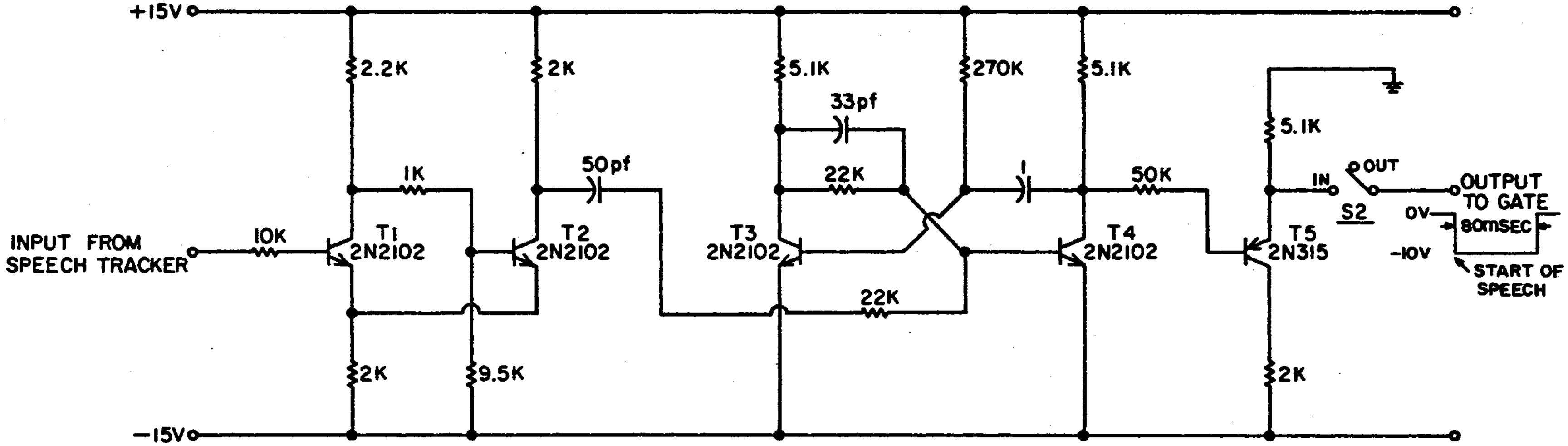


Figure 3.6. Speech Detector

3.2.4 Gate. The multiplexer steps from channel to channel any time it receives a clock pulse. Interruption of the pulse is therefore used to start and stop the display. This is accomplished by the circuit shown in Figure 3.7. The clock pluse is a negative going 10 volts from a zero level, of about 10 μsec duration. When T1 is "OFF" the clock pulse passes through T4 on the right and to the multiplexer. When T2 is "ON" the two 10K resistors form a voltage divider clamping the base of T4 at near ground potential and preventing passage of the clock pulse. The IN695 diode, which has a 0.6 volt forward characteristic, holds the base-emitter junction of T4 in a reverse bias condition, it provides a current path to keep T1 "ON", and it prevents the application of a positive voltage on the input line from the Clock.

Control voltages for the Gate are provided by the Speech Detector and the sixth stage of the Counter. A zero voltage (ground) on either of these lines will act as an inhibit voltage to prevent passage of the clock pulse to the multiplexer. The Speech Detector control line has zero volts, -10 volts, or an open circuit depending on the Speech Detector output level and the position of S2. Only the zero volt level will inhibit passage of the clock pulse. The control line from the Counter sixth stage has a zero volt level applied when the Counter is in the "1" state and a -10 volt level when the Counter is in the "0" state. Thus the clock pulse passes through the Gate only when the sixth counter is in the "0" state and when either S2 is open ("OUT" position) or the Speech Detector output is -10 volts.

3.2.5 Pulse Shaper. During the sampling of the last input channel, the multiplexer places a zero volt pulse, from a -10 volt level, on a special output line which is called the "End Frame". This pulse

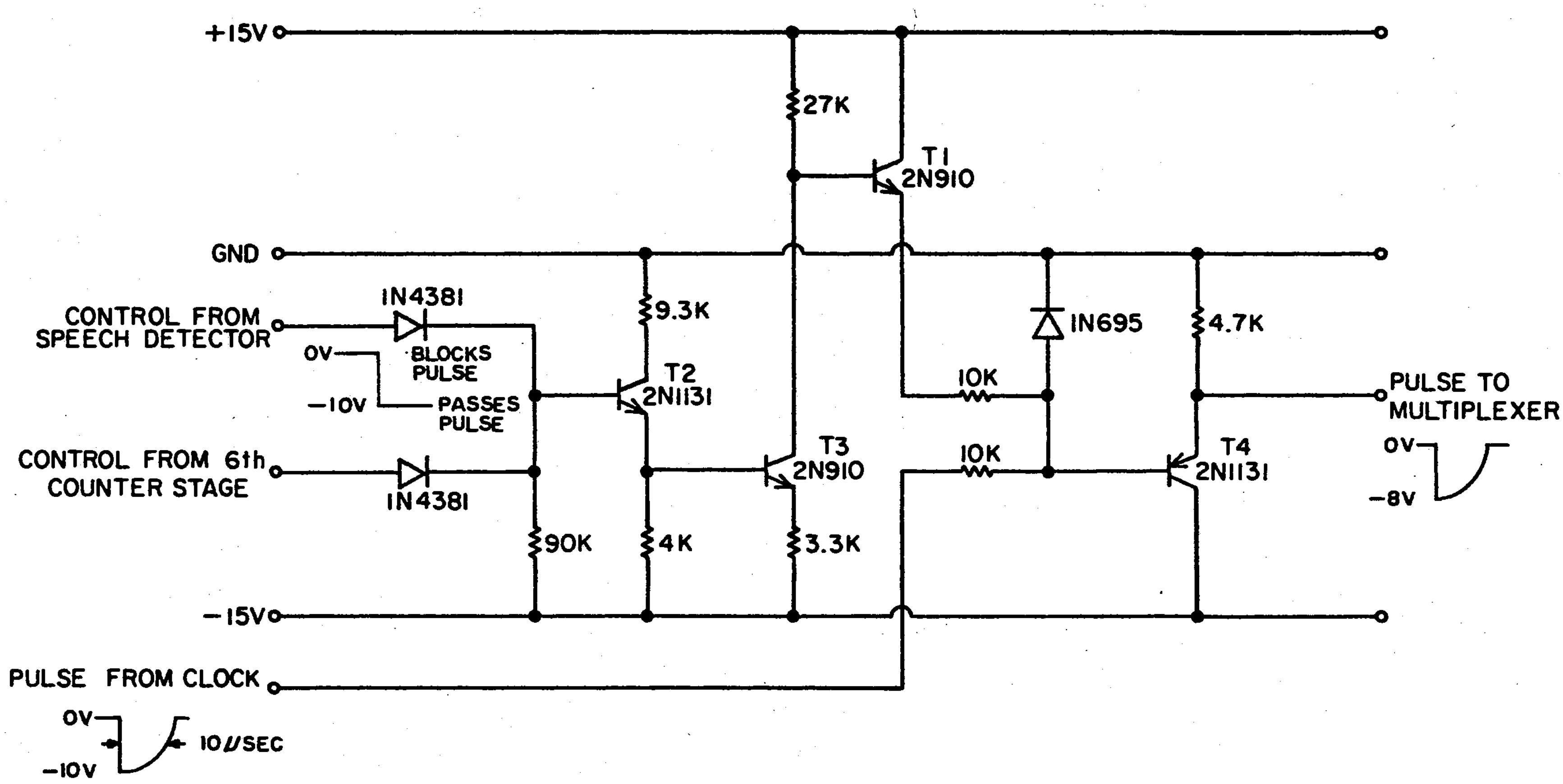


Figure 3.7. Gate

is used to advance the counters and to reset the Sweep Generator in preparation for the next sweep. When stepping at a 50 KHz rate, this pulse is about 15 μsec in width. The Pulse Shaper inverts the pulse and provides a low impedance drive for stepping the "1" bit counter stage and turning off the Sweep Generator Control multivibrator. The circuit is shown in Figure 3.8. The Germanium diodes provide a turn-on bias for the Germanium output transistors.

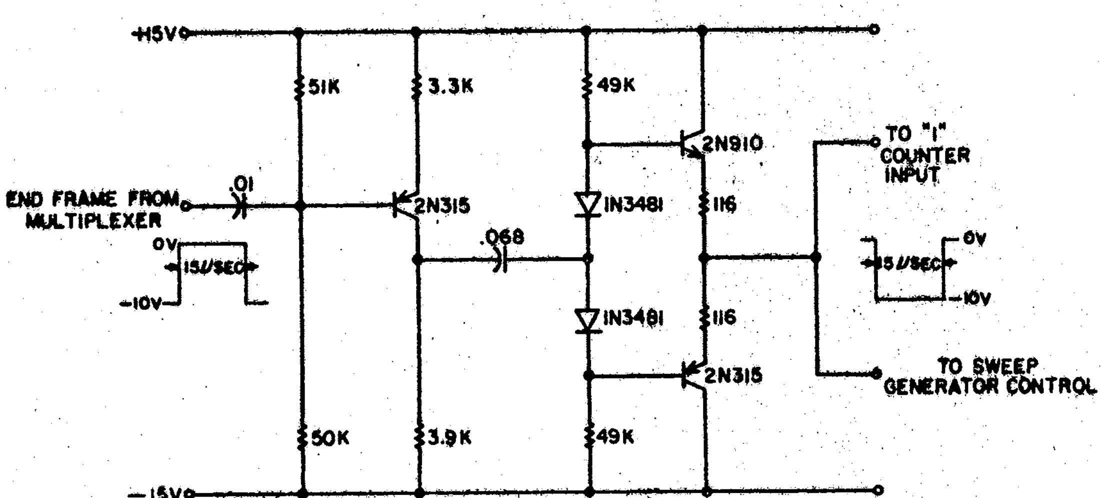


Figure 3.8. Pulse Shaper

3.2.6 Counter. The Counter consists of six bistable multivibrators with the necessary triggering and readout circuitry. Each counter is identical to that shown in Figure 3.9 except that the control signal to the gate is taken from only the sixth counter. Each stage is triggered by a negative going level change which is differentiated to a negative going spike by the 470 pfd input capacitor and the 13K resistor to ground. The negative spike passes through the input diodes and changes the state of the counter by applying a negative pulse to the base of the transistor which is "OFF".

During the "0" state T1 is "ON" and thus has a near ground potential on its' collector. This voltage is applied by means of an isolation emitter follower, T3, to a summing resistor which is shown in Figure 3.10. When the counter stage is in the "1" state the collector of T1 is at -10 volts which is also passed to the summing resistor.

The negative going level change which is used to trigger the next stage is taken from the collector of T2. This event occurs each time the counter stage switches from the "1" to the "0" state. This collector voltage level in the sixth counter is also applied through an isolation emitter follower, T4, to the OR circuit in the Gate. A "1" state in the sixth counter stage applies a ground potential to the Gate. This stops the sweep by blocking passage of the clock pulse.

3.2.7 Counter Amplifier and Vertical Summer. Each counter which is in the "1" state applies a -10 volts on the left end of its' summing resistor which is shown in Figure 3.10. The resistors are chosen to be inversely proportional to the weighting value of the counter stage to which they are connected. The result is that the 2N910 transistor acts as a current summing device and a current to voltage converter. The

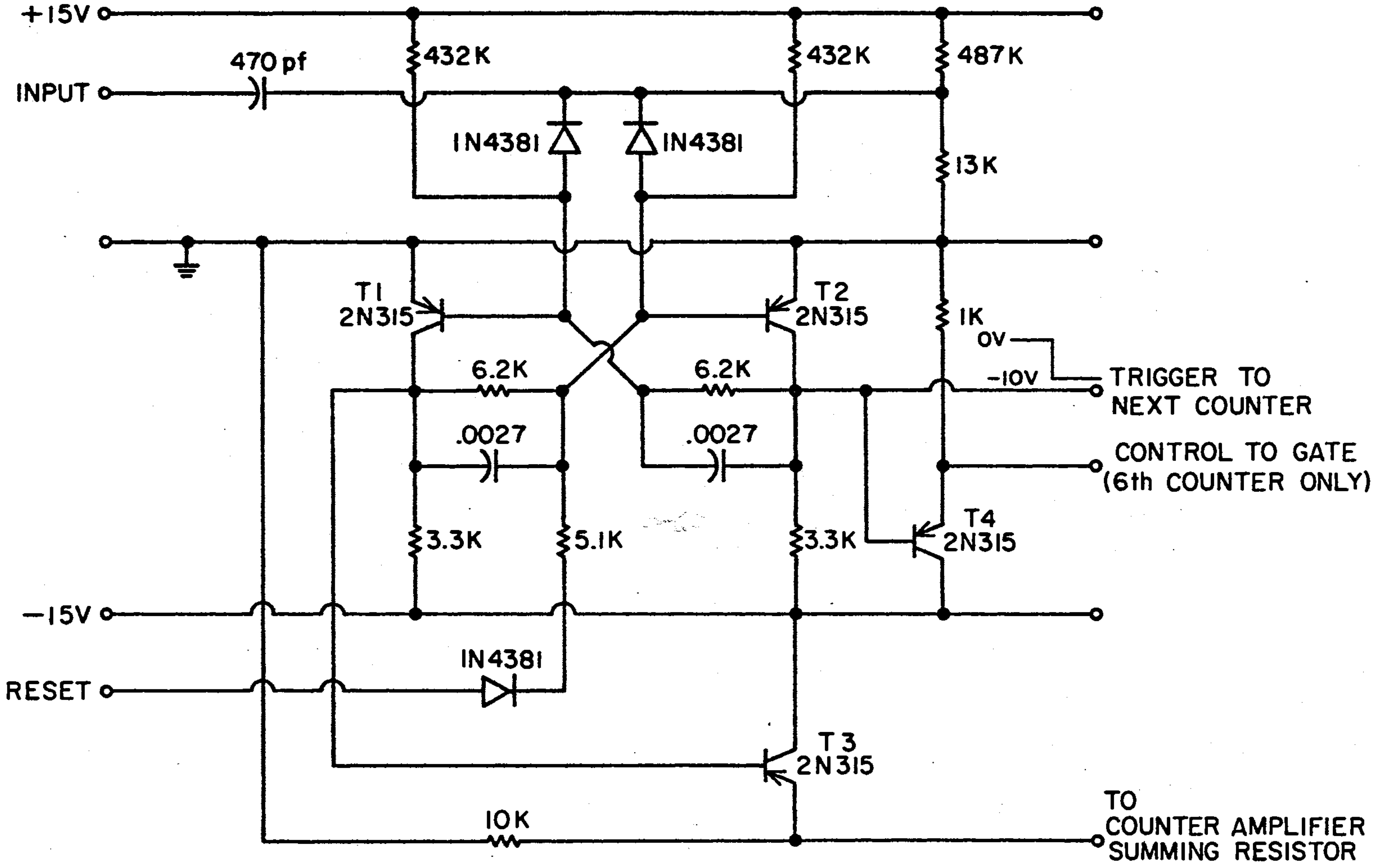


Figure 3.9. Counter Stage

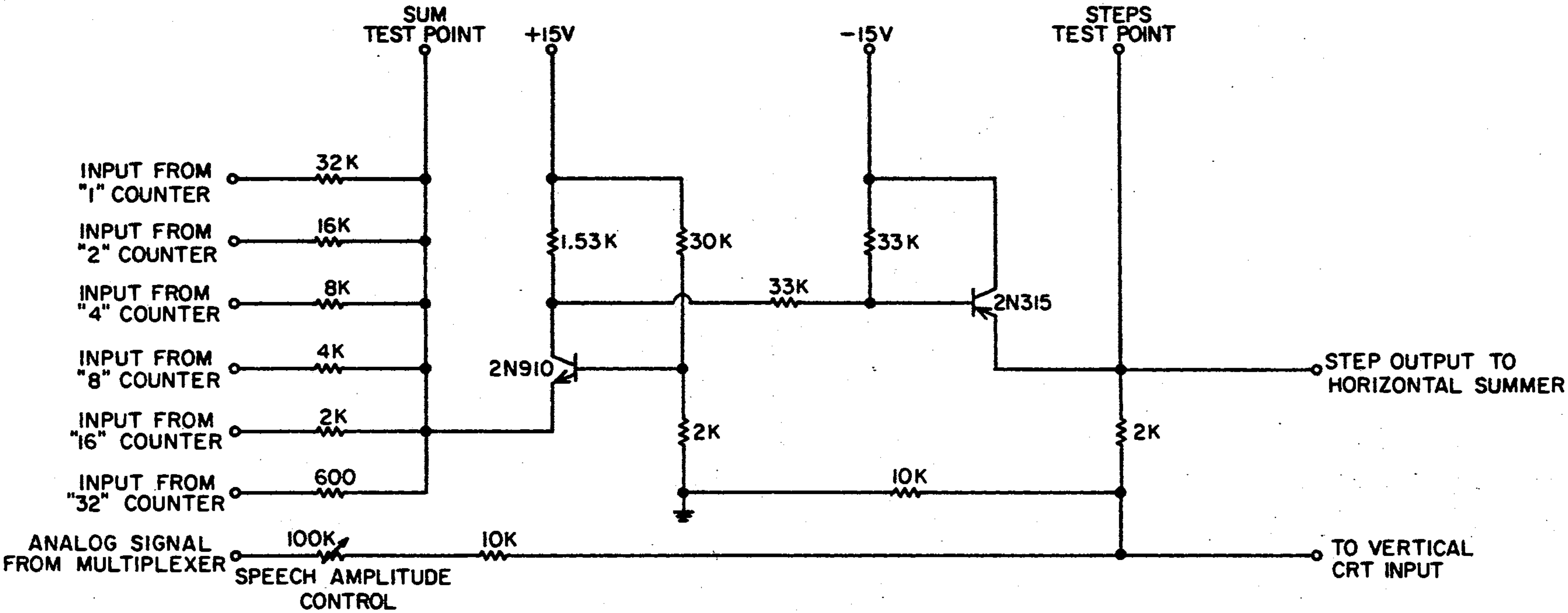


Figure 3.10. Counter Amplifier and Vertical Summer

collector voltage is a series of negative going steps as the Counter counts the first 32 pulses on the "Scan Address End Ring" line from the multiplexer. The 32^{nd} pulse switches the sixth counter (the "32" counter) stage to the "1" state. In this condition the 600 ohm summing resistor draws a larger step increase in current than was drawn by the other summing resistors. This larger current step is converted to a larger negative voltage step which is used to remove the electron beam from the viewing area of the CRT. The 2N315 transistor acts as an isolation emitter follower for driving the vertical and horizontal summing networks. The output terminal which is labeled "Steps Test Point" provides an access terminal on the Control Panel, as shown in Figure 3.11, for viewing the voltage steps which are used to provide the sweep offset. The voltage waveform at this point is shown in Figure 3.12. The 31 uniformly increasing negative voltage steps are followed by the larger 32^{nd} step and then several small steps which have no effect because the beam is off the CRT face at this time. The 64^{th} "End Frame" pulse zeros the counters when S1 is in the "CONTINUOUS" position permitting the sequence to repeat. If S1 is in the "SAMPLE" position the sweep stops after the 32^{nd} sweep because the clock pulse is interrupted by the gate action.

The Vertical Summer which is also shown in Figure 3.10 is a "Tee" resistance divider network. A 100K pot is used to attenuate the analog speech amplitude. This "Speech Amplitude Control" is located on the Control Panel in Figure 3.11. This method of summing would not be acceptable in most applications because the base voltage level from which the steps vary is dependent upon the position of the pot setting and thus a DC offset occurs each time the Speech Amplitude Control is

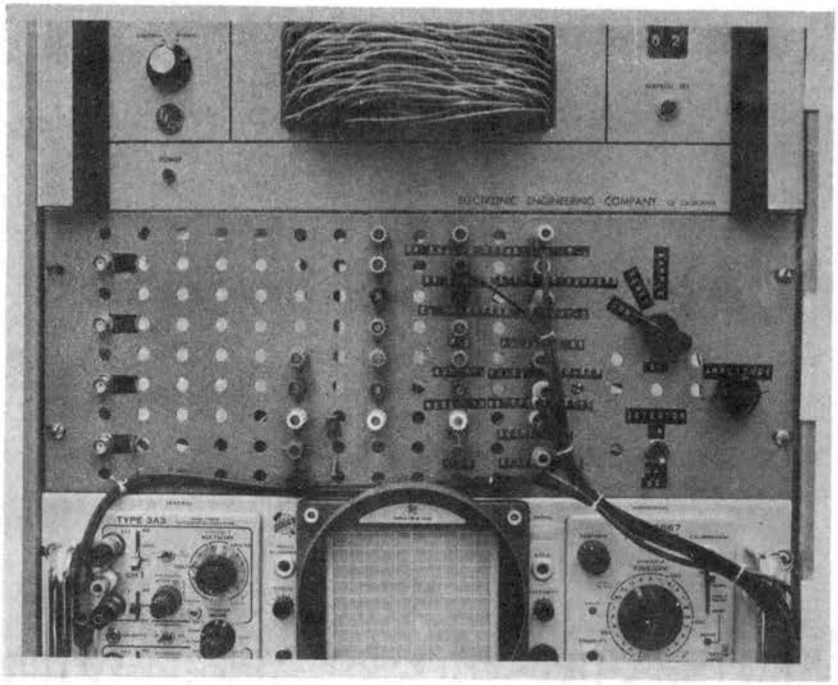


Figure 3.11. Speech Sampler Control Panel

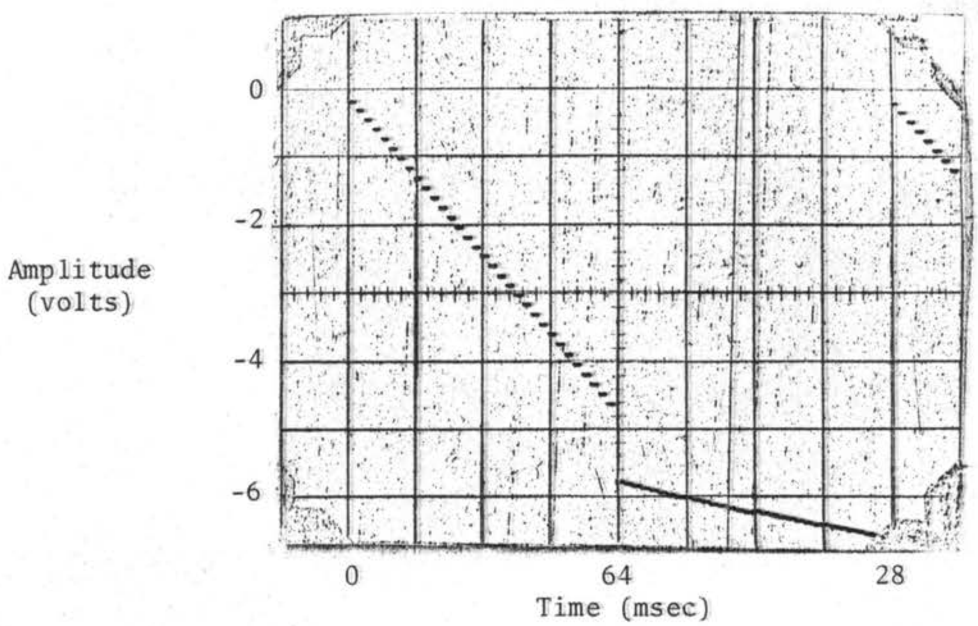


Figure 3.12. Voltage Waveform at "Steps Test Point"

varied. This disadvantage is acceptable here because the Tektronix RM 564 CRT has a vertical amplitude vernier control. If isolation were desired it could be achieved with a summing OA circuit.

3.2.8 Sweep Generator Control. The Burr-Brown Model 1663 Integrator which is used as a sweep generator was found to drift badly any time the reset voltage was not applied, even though its' input was grounded. Repeated and careful calibrations failed to improve its' performance. The Sweep Generator Control was conceived to provide the required positive control of the Sweep Generator and to provide a horizontal electron beam offset during counter reset. The circuit which provides this function is a bistable multivibrator with three control voltages and three output voltages as shown in Figure 3.13. The right side, T2, of the multivibrator is "OFF" only during the sweep. In this condition the collector voltage of -10 volts is applied through a resistance divider to provide a -2.4 volt input to the Sweep Generator. When no sweep is desired this output is at ground potential.

The left side, T1, of the multivibrator is "ON" only during the sweep. In this condition the zero volt level on the collector results in a ground potential being applied to the Sweep Generator Reset and to the Horizontal Summer Sweep Offset. When no sweep is desired T1 is "OFF", and the -10 volts on the collector is used as a signal, through a resistance divider and separate emitter followers, T5 and T4, provide a -6 volt level to the Sweep Generator Reset and the Horizontal Summer.

Three control voltages are used in the Sweep Generator Control and are applied through 1N4381 Germanium diodes at the bottom of Figure 3.13. The Mode Control Switch, when in the "RESET" position, supplies -15 volts from switch level S1(D) to insure that the Sweep Generator is

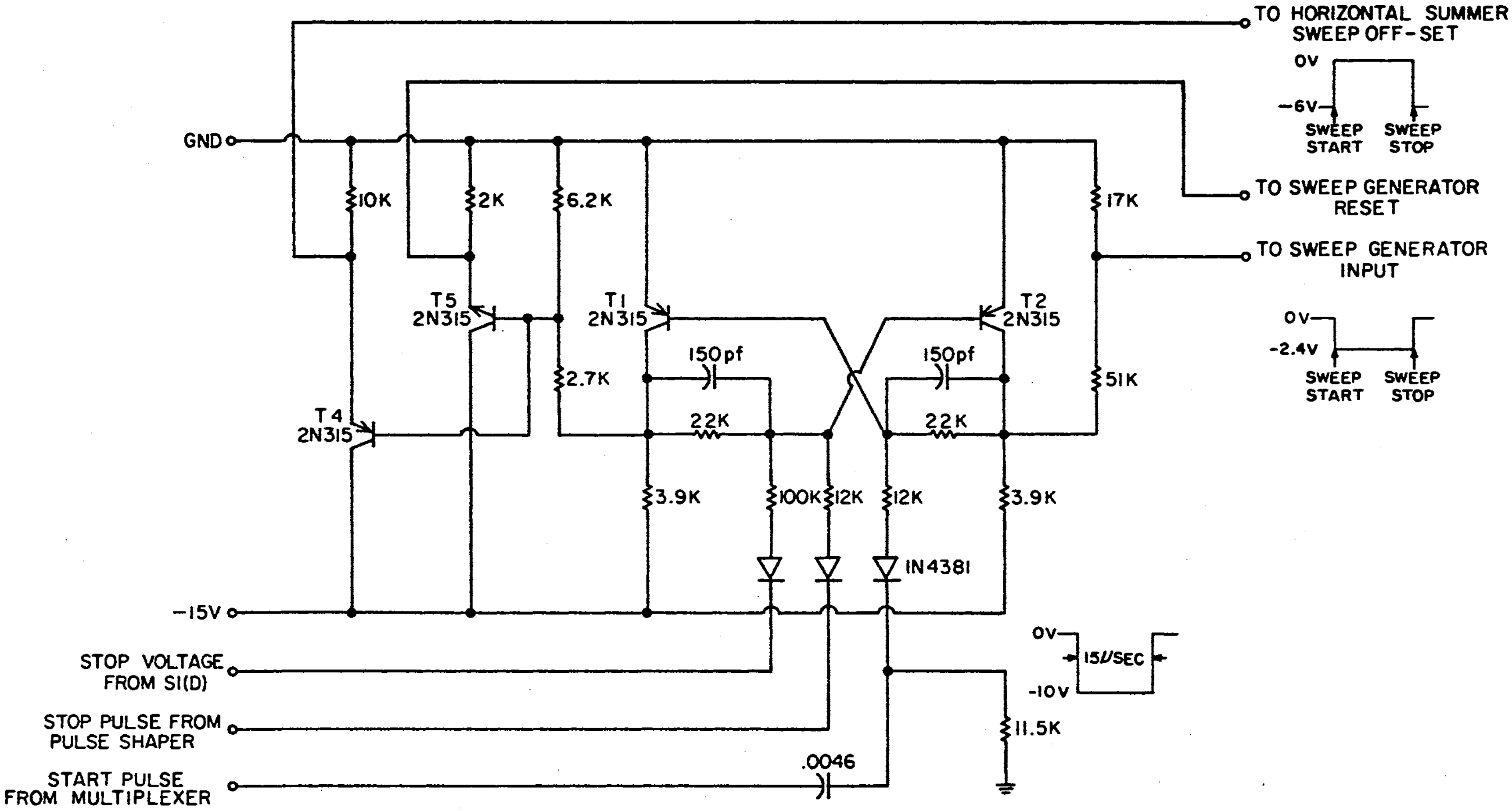


Figure 3.13. Sweep Generator Control

reset in preparation for the sweep condition. Otherwise it is possible to interrupt the clock pulses such that the Sweep Generator would go into saturation, from which it takes typically 2 msec to recover, according to the manufacturers' specification sheet (Appendix C). The remaining two control lines are used to start and stop the sweep in both the "CONTINUOUS" and "SAMPLE" positions of S1. The start pulse is taken from the multiplexer output line which is called the "-A Scan Sync". This line provides a negative going 10 volt pulse, from a zero volt level, of 15 μsec duration. The time of occurrence can be set to correspond to the sampling time of any of the 100 multiplexer input lines, by a thumbwheel switch on the front of the multiplexer. For this system the first two input channels are not used. The thumbwheel switch is set on channel number 2 in order to allow about 38 μsec for resetting the Counter after S1 is placed in the "SAMPLE" position. Channels 3 through 99 are used for input from the TAT channels. The 100th input line is not used but this position on the multiplexer program panel is used to provide the "End Frame" (stop) pulse which is used as an input to the Sweep Generator Control to stop the sweep and prevent saturation of the Sweep Generator integrator.

3.2.9 Sweep Generator. The input wave forms and terminal connections for the Burr-Brown Model 1663 Integrator which generates the sweep, are shown in Figure 3.14. The input is determined by the position of the Mode Control Switch, S1(C), which grounds the input in the "RESET" position and passes the -2.4 volt level from the Sweep Generator Control in the "CONTINUOUS" and "SAMPLE" positions. The -6 volt reset pulse must be a maximum of 40 μsec in duration to reset the integrator when saturation is prevented, as discussed in the previous paragraph.

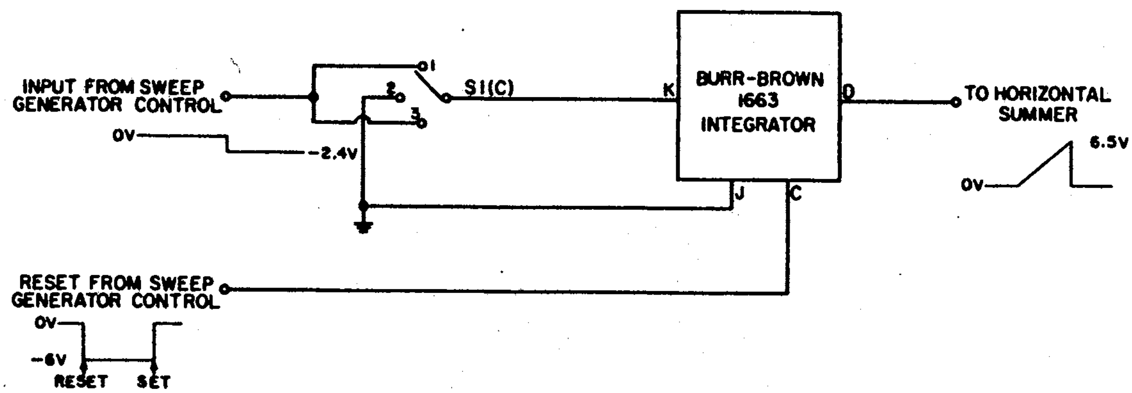


Figure 3.14. Sweep Generator

3.2.10 Horizontal Summer. The ramp voltage from the Sweep Generator is combined with three other voltages in a summing amplifier and inverted as shown in Figure 3.15. The sweep voltage is amplified with a gain of 0.75 and is used to move the electron beam horizontally across the CRT display. The steps from the Counter are amplified with a gain of 0.25 and are used to position each successive sweep to the left of the previous sweep in a manner similar to the vertical movement of the sweep. The horizontal starting point for the first sweep is set by the +15 volt input and the amplifier gain of about 0.17. As described in paragraph 3.2.7, the sweep offset moves the electron beam off the CRT display any time a sweep is not in progress.

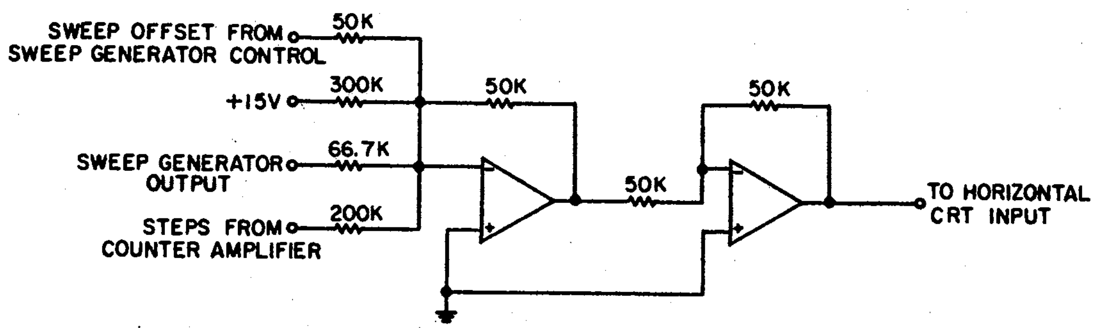


Figure 3.15. Horizontal Summer

3.2.11 Pulse Reset Generator. In order to insure that all counters are in the "zero" state at the beginning of the first 64 msec speech sample, when the Mode Selector Switch is moved to the "SAMPLE" position a 27 μsec pulse of +12 volts is applied to the "RESET" input of each counter. This pulse is provided by the circuit shown in Figure 3.16. When S1(B) is placed in the "SAMPLE" (3^{rd}) position the voltage on the base of T2 drops to -15 volts because T1 turns "OFF" after a delay determined by the discharge of the 10 μfd capacitor which eliminates the effect of any switch chatter. This delay has no effect on the display because the Gate is not opened until after the counters are reset. The transistors T2 and T3 make up a Schmitt trigger circuit in which a

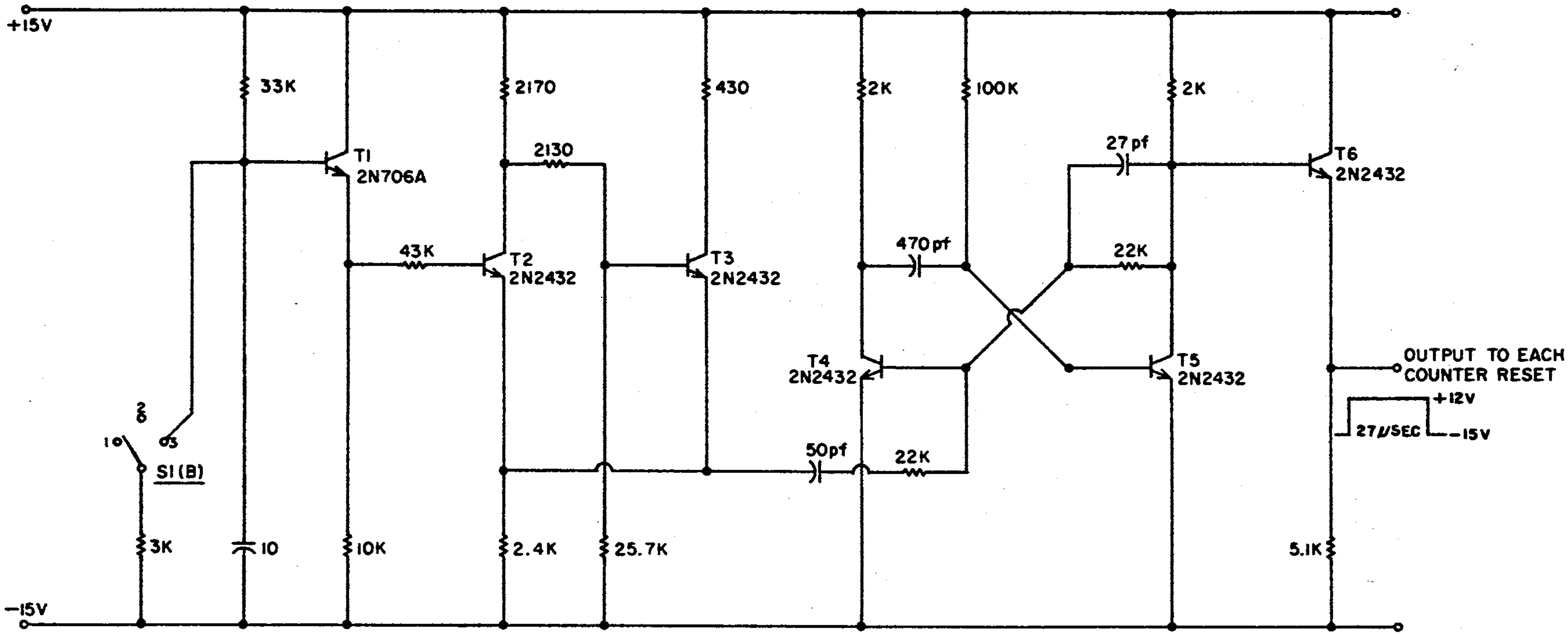


Figure 3.16. Reset Pulse Generator

positive going pulse is passed through the 50 pfd capacitor to trigger the monostable multivibrator. The pulse width of the output, which is taken from the isolation emitter follower, T6, is determined by the discharge of the 470 pfd capacitor through the 100K resistor. With a smaller RC product, this circuit was demonstrated to be capable of providing a 1 μsec pulse with less than 0.25 μsec rise time. It was later found that with positive control of the multiplexer and Sweep Generator, the reset pulse width was not critical as long as it did not exceed the 2 msec time for the first sweep.

3.2.12 Summary of Speech Sampler Controls. The position of the Mode Selector Switch (S1) and the Detector Switch (S2) determine the display which is provided by driving the Vertical input to the CRT with the Vertical Summer output and the Horizontal input to the CRT with the Horizontal Summer output. When S1 is in the "CONTINUOUS" position S2 has no effect and the sequence of 64 msec of data display followed by 64 msec of blank display is continually repeated. When S1 is placed in the "SAMPLE" position the single sample of 64 msec of data (32 sweeps) occurs immediately after the next clock pulse from the Gate if S2 is "OFF" but occurs only after the Speech Tracker output reaches +6 volts when S2 is "ON".

CHAPTER IV

THE FORMANT THEORY OF SPEECH PERCEPTION

4.1 Introduction. Research in speech synthesis and recognition over the past several decades has been heavily influenced by the desire to characterize all phonemes by a minimum number of parameters. Early speech spectrograms tended to show that phonemes could be categorized by grouping the harmonics in such a way that each phoneme could be represented by no more than four parameters, called formants. The use of formants has continued even to current research in spite of the inability of this system of classification to effectively specify phonemes with sufficient precision to permit high accuracy recognition with normal speech variations of a given speaker, or to permit the use of multiple speakers. The purpose of this chapter is to describe the formant classification system, to show some shortcomings of the method, and to illustrate the additional information which is provided by the high resolution filter bank described herein.

4.2 Formant Definition. The speech properties of frequency and intensity as a function of time can be determined either in real-time by a parallel processing system or in non-real-time by recording the speech and later playing it through a tunable bandpass filter. Early spectrograms used the latter method, while more recent efforts to build a usable recognition system employ the parallel processing method. Either method reveals the fact that speech contains a fundamental frequency and many harmonics of varying amplitude as a result of the shape and resonances of the vocal tract.

The fact that the harmonic amplitudes tend to peak in certain frequency bands led to the grouping of several harmonics by the construction of an envelope or curve enclosing the harmonic peaks. The name "formant" applies to this grouping. The "formant frequency" is defined by Fant (33) to be "... the position on the frequency scale of the peak of the spectrum envelope drawn to enclose the peaks of the harmonics". The artistic license implied by the method of measurement is suggested by the further guidance supplied by the same reference that:

> When two formants come close or when the formant to be measured is very low in frequency only one side of the formant 'mountain' may be visible and the estimate has to be based solely on this information. In such cases it pays to go to the broad band spectrogram and determine the center of the formant band.

Other formant characteristics which have been used are formant level (amplitude) and formant bandwidth, but Fant considers these to be redundant measurements because they are well correlated statistically with formant frequency which has a natural range of variation for non-nasal voiced sounds uttered by an average male speaker as follows:

Formant Number	Range of Variation (Hz)
F1	150 to 850
F2	500 to 2500
F3	1700 to 3500
F4	2500 to 4500

4.3 Vowel Specification by Formants. The most extensive cataloging of data on vowel formants which was found in the literature search was conducted at Bell Telephone Labs almost 20 years ago and reported by Peterson and Barney (34). In this research 20 words were spoken by 76 speakers and the frequency of each formant was computed by estimating a weighted average of the principal frequency components of each formant. The authors pointed out that in cases where the frequency distribution, within a formant, is asymmetrical the difference between the estimated formant frequency and that which would be assigned by the human ear, may be appreciable. Further evidence of the non-uniqueness of formant representation of phonemes is provided by a plot of the frequency of the first formant, F1, versus the second formant, F2, for the Peterson and Barney data, which is reproduced in Figure 4.1. The curves were drawn to include about 90 percent of the data points for each phoneme and both male and female speakers were used. The authors suggested that some overlap might be removed by considering male and female data separately but Gerstman (35) used the same data and found that no significant differences in F1 or F2 were revealed when separate averages were plotted for male, female and child talkers. Gerstman concluded that, <u>for a single speaker</u>, the first and second formants are sufficient for vowel classification. Support for this observation was provided by numerous others including Fairbanks and Grubb (36); Flanagan (9); Campanella and Phyfe (37); Forgie and Forgie (38); and Pols, <u>et al</u>. (39), who found that their error rate in recognizing vowels was reduced from 13 percent when only F1 and F2 were used, to 9 percent when F3 was added.

Table IV shows the average frequencies of F1, F2 and F3 separated by men, women and children for the data which is plotted in Figure 4.1. It is useful for comparison with the spectral data which can be obtained from a high resolution spectrum analyzer.

The greatest accuracy of measurement necessary to specify formants was estimated by Flanagan (40) to be ±20 Hz for F1, ±50 Hz for F2 and

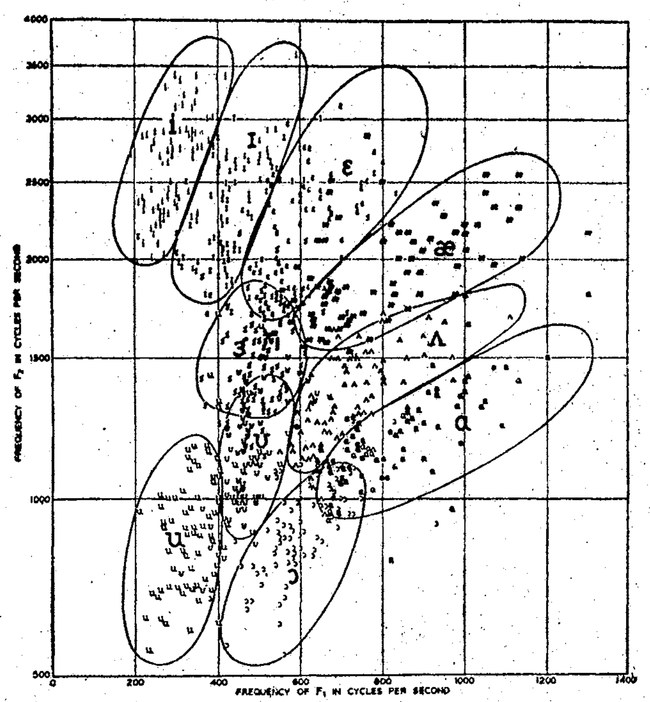


Figure 4.1. F1 Versus F2 for 10 Vowels Spoken by 76 Speakers (34)

±70 Hz for F3. These numbers were arrived at through psychoacoustic experiments and represent the smallest change which could be detected by a listener. The assumption is implied that human perception is based upon formant tracking.

4.4 High Resolution Spectral Measurements. Research in recent years has produced considerable evidence that greater frequency resolution than was available, would be useful in improving vowel recognition techniques. In each case the recommendations have avoided a direct assault upon the formant theory and have sought other parameters to supplement formant measurements. Martin (21), for example, used the rate of change of energy as a function of frequency to resolve

TABLE IV

AVERAGE PITCH AND FORMANT FREQUENCIES FOR 76 SPEAKERS[1]

		Phoneme									
		i	I	ε	ae	a	∂	U	u	Λ	ɜ
Pitch (Hz)											
Men		136	135	130	126	124	129	137	141	130	133
Women		235	232	223	210	212	216	232	231	221	218
Children		272	269	260	251	256	263	276	274	261	261
Formants (Hz)											
	Men	270	390	530	660	730	590	440	300	640	490
F1	Women	310	430	610	860	850	590	470	370	760	500
	Children	370	530	690	1010	1030	680	560	430	850	560
	Men	2290	1990	1840	1720	1090	840	1020	870	1190	1350
F2	Women	2790	2480	2330	2050	1220	920	1160	950	1400	1640
	Children	3200	2730	2610	2320	1370	1060	1410	1170	1590	1820
	Men	3010	2550	2480	2410	2440	2410	2240	2240	2390	1690
F3	Women	3310	3070	2990	2850	2810	2710	2680	2670	2780	1960
	Children	3730	3600	3570	3320	3170	3180	3310	3260	3360	2160

[1]From Peterson and Barney (34).

"allophones": which he defines as "speaker dependent variations of phonemes". In this work Martin used a bank of 19 filters each of which had a Q of less than 8, in order to preserve temporal characteristics, in his study of connected digits. He used 155 speakers and collected data on 34,000 digits in his research claiming 88 to 94 percent recognition accuracy.

Lecours and Sparkes (14) used two filter banks and compared the advantages of each. One filter bank consisted of 32 active filters, each with a bandwidth of 100 Hz. The second used 16 filters, each having a bandwidth of 200 Hz. The conclusions were that narrow-band filters are superior for studying the "formant structure" of vowels but that the temporal transitions of the start and stop consonants require the response of wide bandwidth filters. They suggest the use of a few spectrum analyzers with different filtering characteristics.

An effort to include both the fine frequency resolution of a narrow band system at the low frequencies (below 3000 Hz) where most vowel information is contained, and rapid transient response at high frequencies led Hughes and Hemdal (12) to the use of a roughly constant Q filter bank which had 35 filters, each with a Q of about 10. Their recommendation, along with that of other recent work (Bobrow and Klatt, 13) encouraged the choice of the constant selectivity (Q = 25) system which was constructed as a part of this thesis.

The frequency resolution which is possible with the TAT system is illustrated in Figures 4.2, 4.3 and 4.4. Figure 4.2 shows the system display for the phoneme /i/ as in <u>beet</u> as voiced by a male speaker at 8 different pitch frequencies which are spaced over one octave. The microphone input was used and, for these figures, the gain

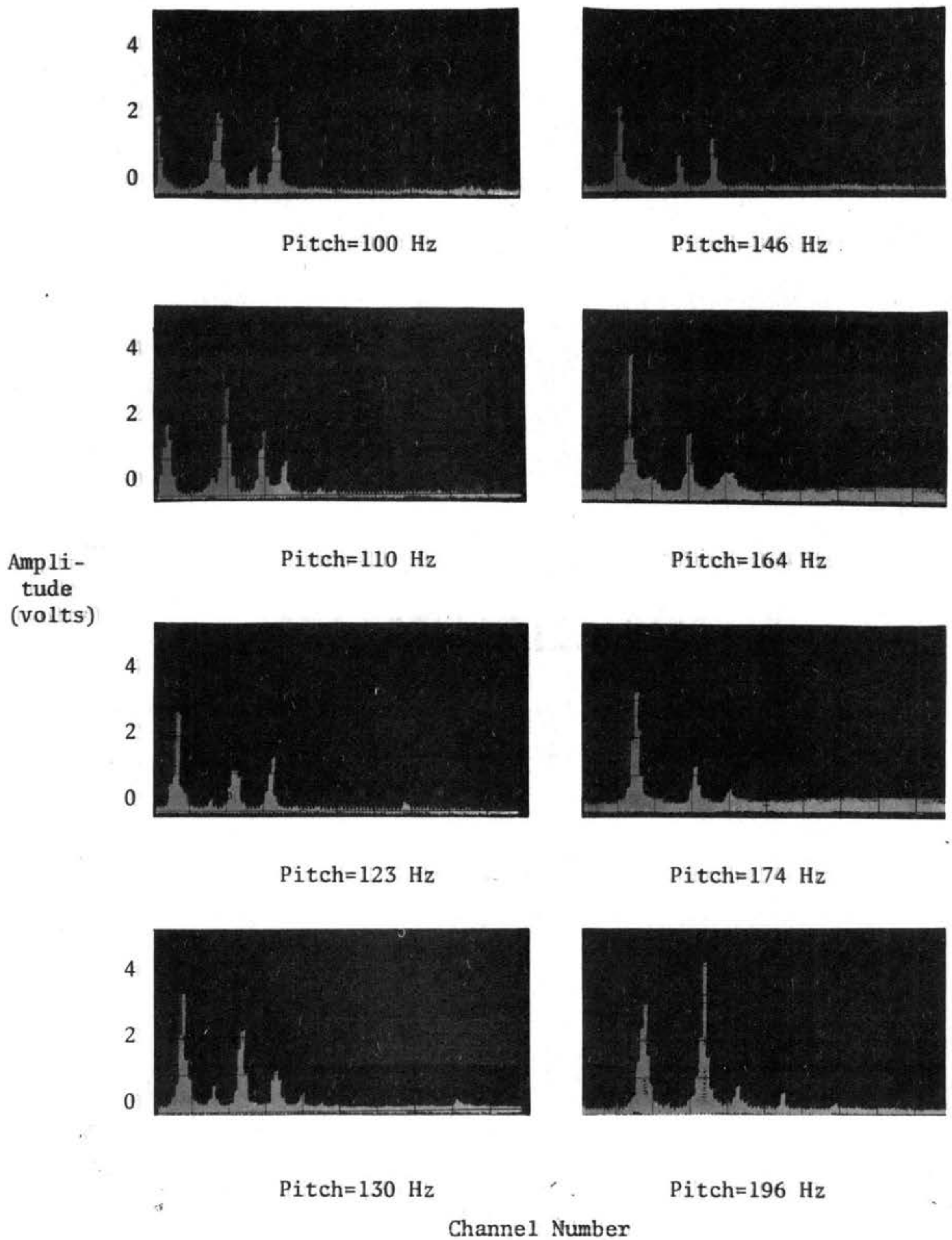


Figure 4.2. TAT System Response to the Phoneme /i/ for Pitch Frequencies Varying Over One Octave

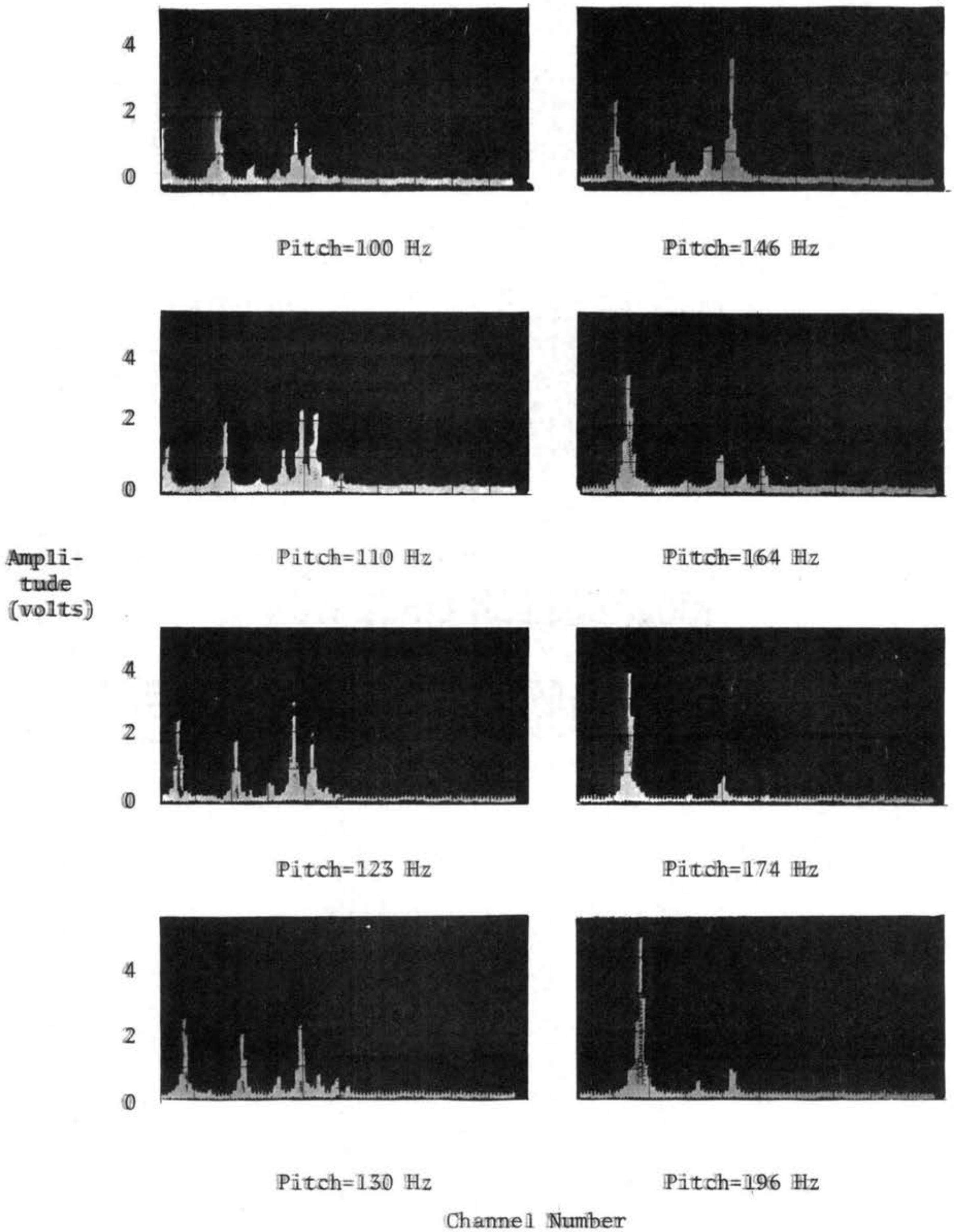


Figure 4.3. TAT System Response to the Phoneme /a/ for Pitch Frequencies Varying Over One Octave

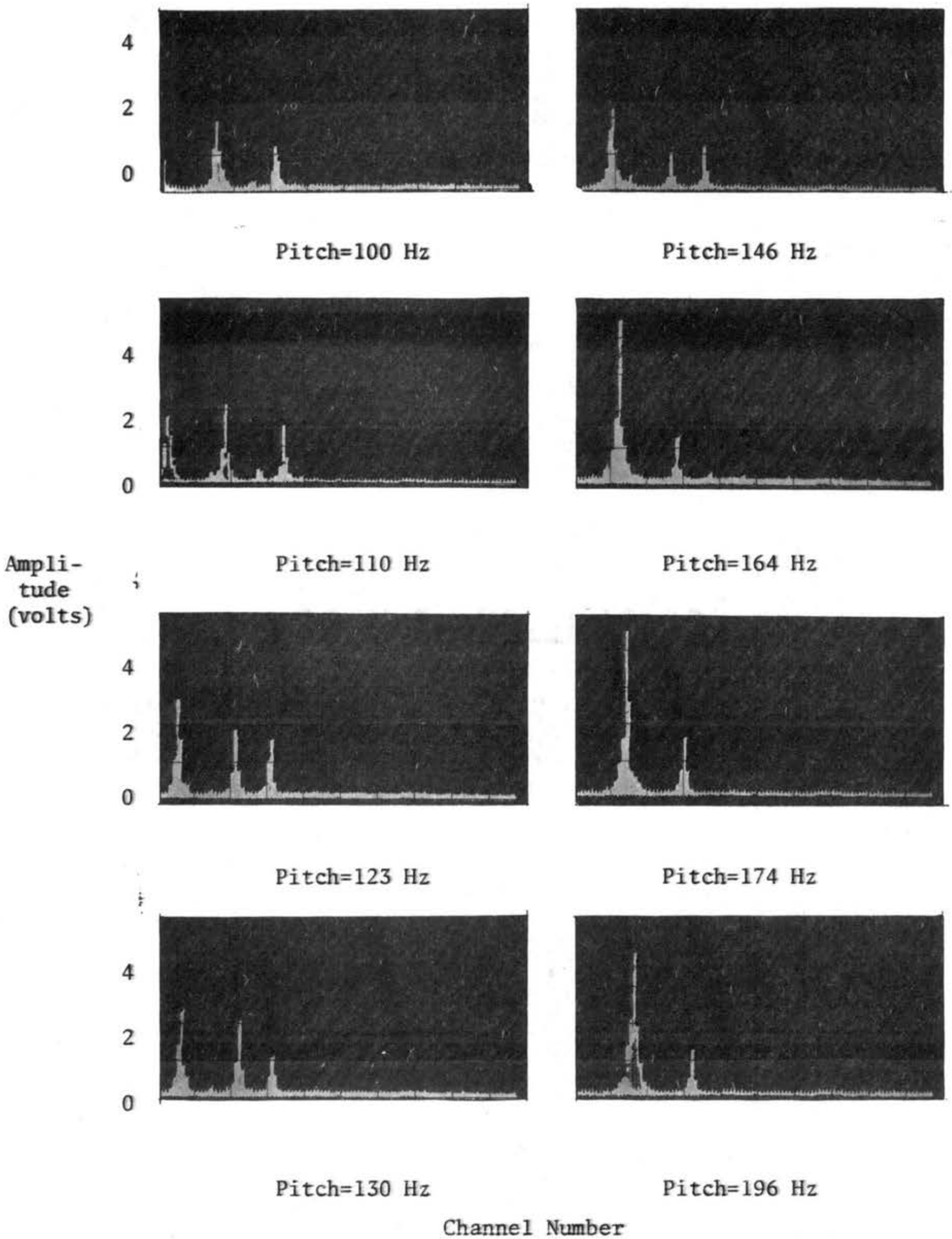


Figure 4.4. TAT System Response to the Phoneme /u/ for Pitch Frequencies Varying Over One Octave

characteristic of the driver amplifier was flat over the speech spectrum. Thus amplitude comparisons can be made between the pitch and the harmonics which make up the first formant, all of which are found to the left of the center (F(50) = 832 Hz) of each photo. Martin (21) reported the energy content of speech to be flat within ± 3 db up to about 800 Hz above which it decreases at about 6 db/octave.

In Figure 4.2 it can be seen that for a pitch of 100 Hz the energies of the 2^{nd} and 4^{th} harmonics are about equal to that of the fundamental but the 3^{rd} harmonic energy is much lower. Use of formant theory would dictate the sketching of an imaginary envelope over these peaks (excluding the fundamental) and arriving at an F1 of about 300 Hz. Increasing the pitch to 110 Hz however results in an asymetric envelope of the 2^{nd}, 3^{rd} and 4^{th} harmonics which is caused by an increase in the 2^{nd} and 3^{rd} harmonic amplitudes and a decrease in the 4^{th} harmonic. The speech analyst is thus forced into a decision on whether to use the envelope technique which would yield an F1 of about 270 Hz or to use the center of the frequencies present which would yield an F1 of about 330 Hz.

The dramatic shifts in energy peaks as a result of small pitch changes are shown by further increasing the pitch to 123 Hz. The 2^{nd} and 4^{th} harmonics drop sharply while the 3^{rd} remains about the same resulting in an F1 of perhaps 320 Hz. Further increases in pitch show similar changes and shifts in F1 which would understandably account for the large areas of overlap and uncertainty in formant plots similar to that of Figure 4.1. The higher pitch frequencies force the voice into a near falsetto condition which is probably of limited interest in speech recognition but the effect on the energy structure is dramatic as is shown at the 196 Hz pitch in Figure 4.2 where a vocal tract resonance at the 2nd harmonic suddenly appears. Figure 4.3 shows the same data for the phoneme /α/ as in got but for this sound the strong 2nd harmonic resonance does not occur at the pitch of 196 Hz. In Figure 4.4 the resonance shifts are shown for the phoneme /u/ as in boot.

The shifts in resonances which are shown in these figures are representative of the variations which can be expected in normal speech either by a single speaker who is permitted to speak without controlling his pitch, or in comparing the same phonemes as spoken by several people. Flanagan (40) reported that a male voice reading factual material varied in pitch usually no more than about one octave. The data which has been presented show that formant coding does a rather poor job of representing the information which is available from a high resolution filter bank system for various pitch frequencies. It is suggested that a better coding system may be necessary in order to normalize speech for pitch variations.

CHAPTER V

DOCUMENTATION OF PITCH EFFECTS ON PHONEME SPECTRA

5.1 Introduction. The high resolution spectral analysis system which was described in Chapter II was used to prepare a data set of eleven phonemes, each voiced over a one octave range of pitch frequencies. Verification of the amplitude normalized data set is made by comparisons with oscilloscope photographs of the system output. A description of the documentation technique and computer interface equipment is included.

5.2 The Data Set. In order to provide a data base for the future development of pitch normalization techniques, the eleven phonemes which are shown in Table V were selected from the International Phonetic Alphabet of the English Language. The first 8 are classified as vowels, the /ɚr/ and /l/ are semi-vowels and /n/ is a nasal consonant. Those pairs of phonemes which are difficult for the human listener to distinguish, were intentionally avoided because they would probably not be used in the selection of a language for automatic recognition. Three male and two female speakers with typical mid-western speech characteristics were used. The male speakers voiced each phoneme at pitch frequencies of 100, 110, 123, 130, 146, 164 and 174 Hz. The female speakers voiced each phoneme at 164, 174, 196, 220, 246, 261 and 294 Hz. These frequencies were selected from Grays' American Institute of Physics Handbook (41) as the scale over one octave within the range of the male and female voices.

TABLE V

LIST OF SELECTED PHONEMES

Phoneme Number	Phoneme	Phoneme Number	Phoneme
1.	/ae/ as in bat	7.	/ə/ as in bawl
2.	/e/ as in ate	8.	/u/ as in boot
3.	/ε/ as in bet	9.	/ər/ as in burr
4.	/i/ as in beet	10.	/l/ as in let
5.	/α/ as in got	11.	/n/ as in net
6.	/o/ as in go		

5.3 Method of Data Collection. The high resolution spectral data was collected, normalized and punched on computer cards by the IBM 1620 Digital Computer with the associated analog-to-digital conversion equipment. The TAT System interface with the computer is shown in Figure 5.1 and control of the data collection is provided by the Computer Interface Switch, S3, which is shown in Figure 5.2. The 97 TAT channel outputs plus the Speech Tracker output are time multiplexed on the analog output line and then converted by the Interface Terminal to a digital signal for input upon command by the IBM 1620 Computer. A one second time constant integrating capacitor was added to each multiplexer input line to reduce amplitude fluctuations while that channel output is being sampled by the computer. This could be used only when taking data on

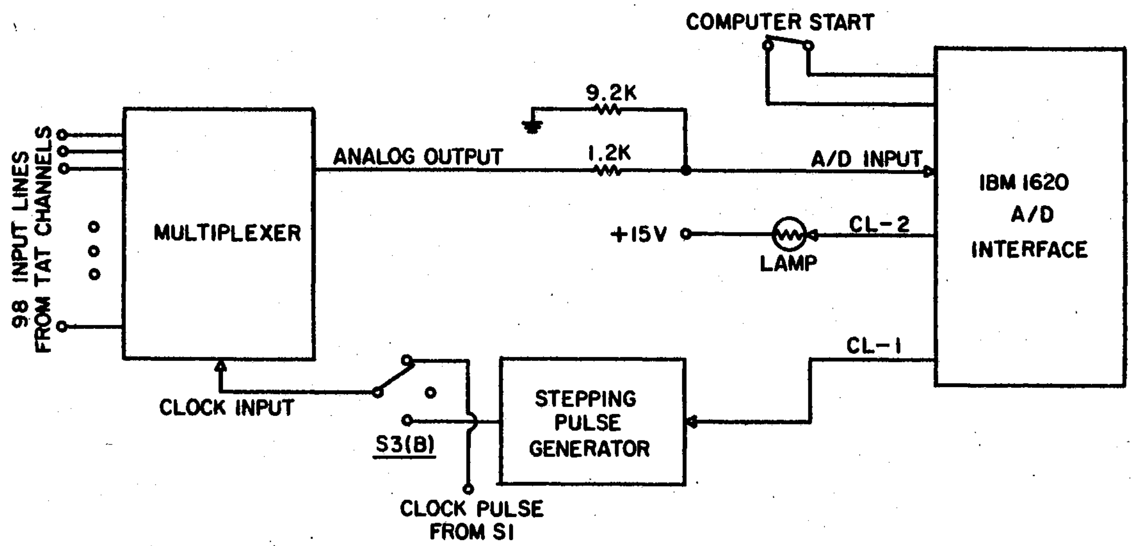
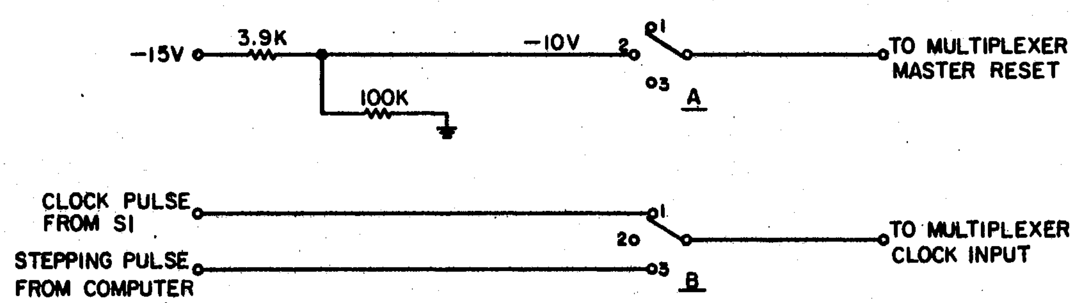


Figure 5.1. Computer Interface Diagram



POSITION	FUNCTION
1	SCOPE
2	RESET MULTIPLEXER
3	COMPUTER INPUT

Figure 5.2. S3, Computer Interface Switch

sustained phonemes in an interval where speech transient characteristics were minimal.

The data sampling technique and operation of the computer interface equipment is outlined with reference to Figures 5.1 and 5.2 and the Phoneme Reading Computer Program which is listed in Appendix E. In preparation for data taking the computer goes into a hold condition indicated by the "PAUSE" statement. S3 is initially in position 1 and the multiplexer therefore steps at a 50 KHz rate as determined by the clock pulse from S1. A continuous spectral display, on the oscilloscope, is provided of the phoneme being voiced into the microphone. This is observed by the speaker who then adjusts his pitch to the desired frequency. When the pitch is correct and the desired phoneme is being clearly voiced, the operator moves S3 to position 2, which provides a -10 volt signal to reset the multiplexer to Channel 00. The operator then moves S3 to position 3 and starts the computer by pressing the remote "Start" button. S3 switches the multiplexer clock input to accept the pulse from the Stepping Pulse Generator which is triggered by computer command on Control Line 1 (CL-1). The circuit is shown in Figure 5.3. CL-2 controls an indicator light which, when illuminated, signals the fact that the computer is ready to accept data and when not illuminated indicates that the computer has completed the sampling of all 98 inputs and is processing the data. The analog signal from the multiplexer which varies between 0v and +5.5v is attenuated by the voltage divider, shown in Figure 5.1, to be compatible with the +4.99v maximum permissible input to the analog-to-digital converter.

The data set consists of eleven phonemes, each of which is voiced by each speaker at seven pitch frequenices. A total of 385 phonetic

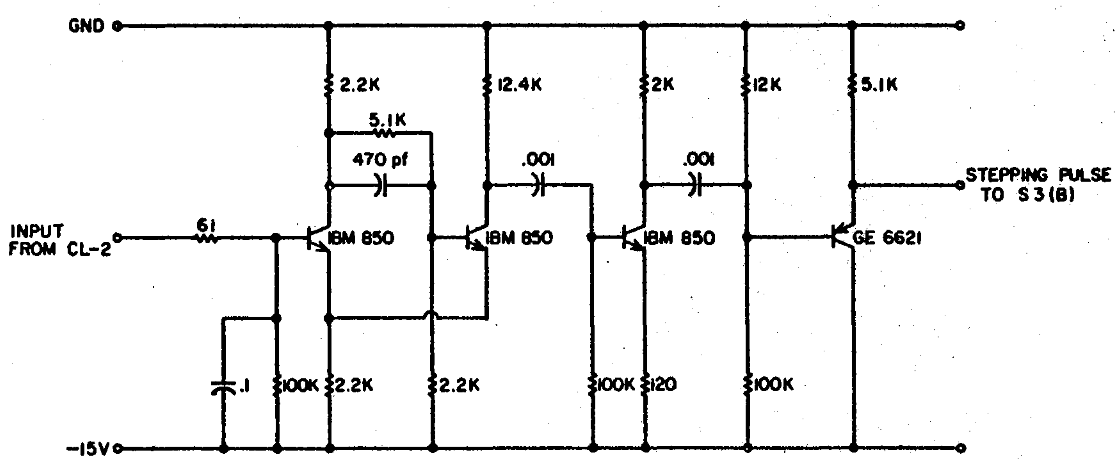


Figure 5.3. Stepping Pulse Generator

spectral records are tabulated in normalized form and plotted as continuous curves showing spectral energy distribution. Table VI shows the identification of speakers and the fact that the first three speakers were male and the last two were female.

The Phoneme Reading Computer Program which is listed in Appendix E normalizes each TAT channel output to a number between 99 and zero by the equation:

TABLE VI

LIST OF SPEAKERS

Speaker Number	Initials	Sex
1.	J.M.M.	M
2.	J.W.M.	M
3.	J.E.B.	M
4.	T.H.B.	F
5.	C.C.S.	F

$$PHONEME(K) = (PHN(K)/TM) 99$$

Where: PHONEME(K) is the normalized amplitude for TAT Channel K which is tabulated in Appendix D;

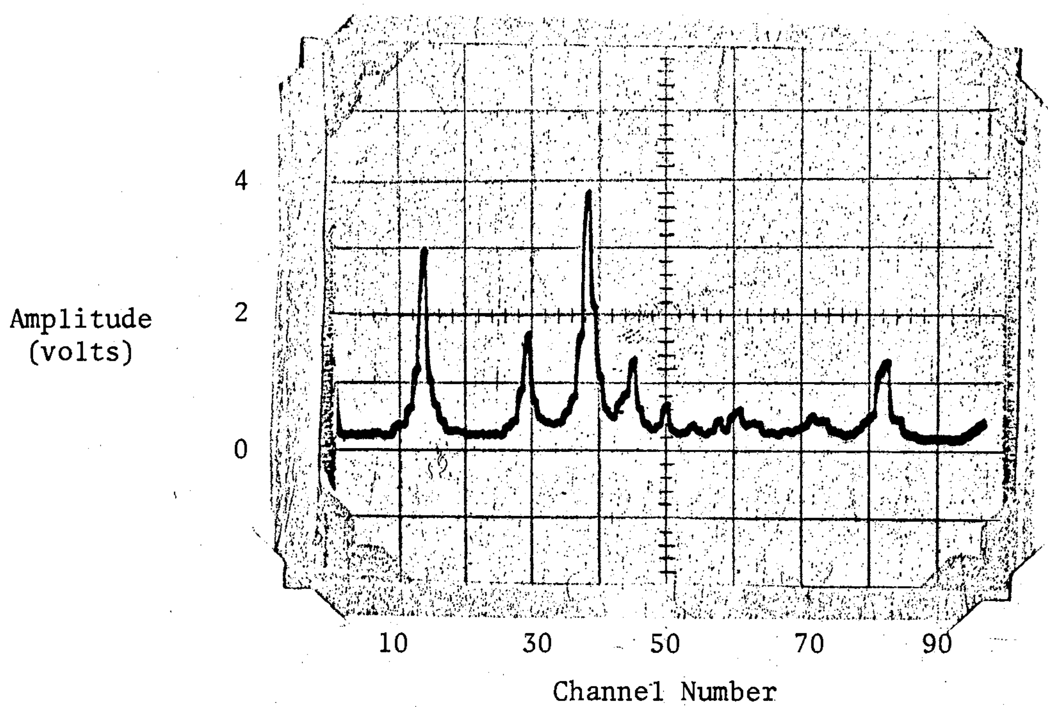
PHN(K) is a number in computer storage which is 20 times the output voltage of TAT Channel K;

TM is a number equal to 20 times the largest output voltage of the 97 TAT channels; and

$1 \leq K \leq 97$.

The data tabulation in Appendix D shows the normalized output with 20 TAT channels per line starting with Channel 1 in the top left corner and continuing to Channel 97 at the bottom. Also shown at the top of each listing are TM (defined above) and ST which is a number equal to 20 times the output voltage of the Speech Tracker circuit. ST is thus a measure of phoneme loudness.

Figures 5.4, 5.5, and 5.6 show comparisons of the CRT display with the computer sampled data which has been plotted with the aid of the



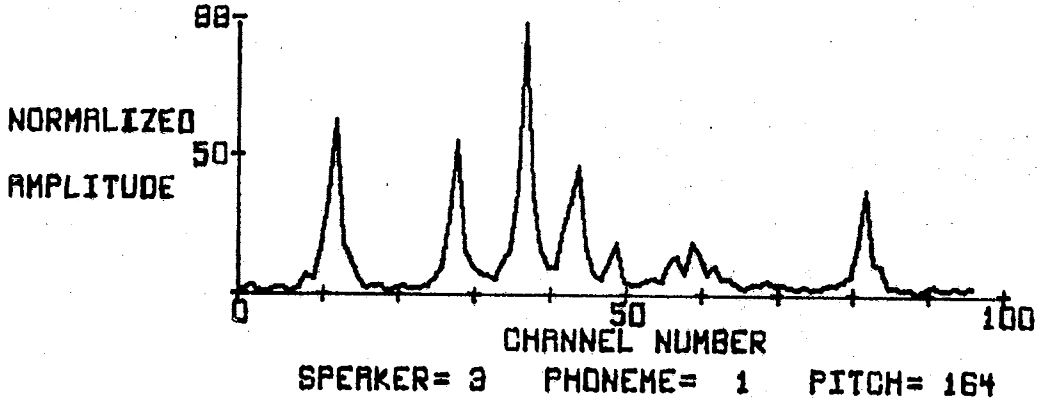


Figure 5.4. Comparison of CRT Display With the Computer Sampled Data for the Phoneme /ae/ as in B<u>a</u>t

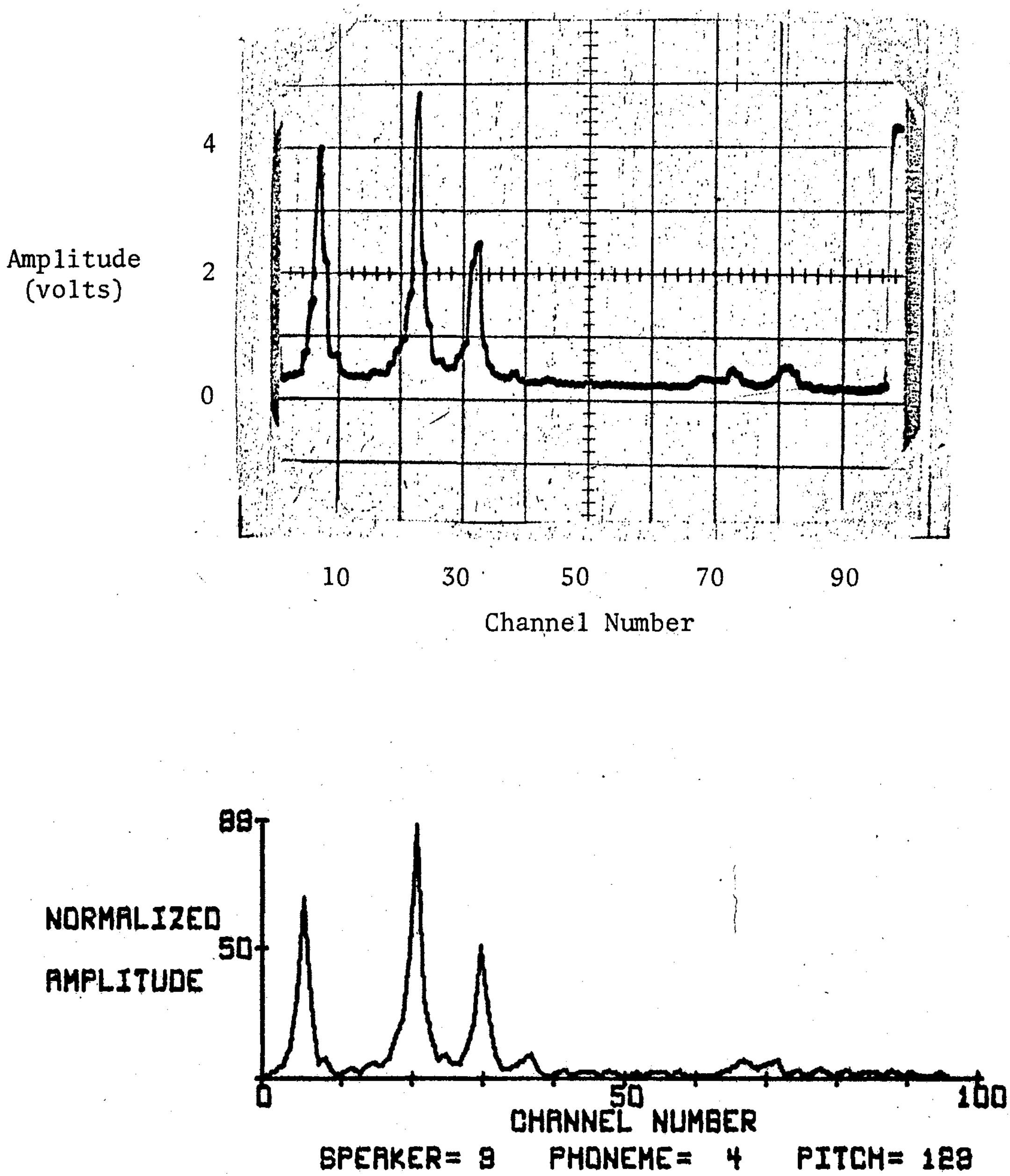
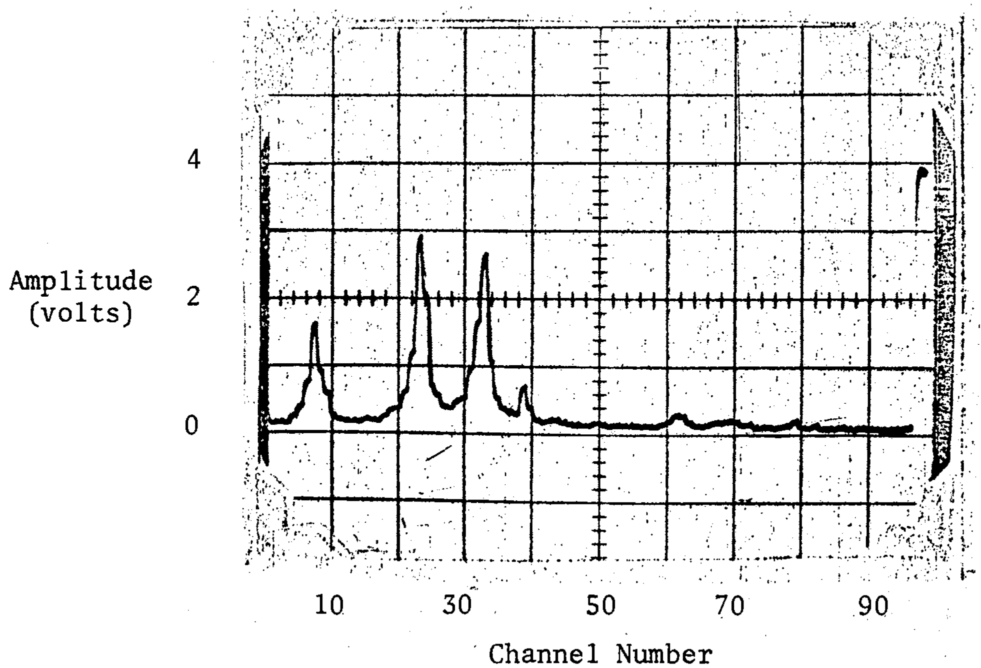


Figure 5.5. Comparison of CRT Display With the Computer Sampled Data for the Phoneme /i/ as in Beet



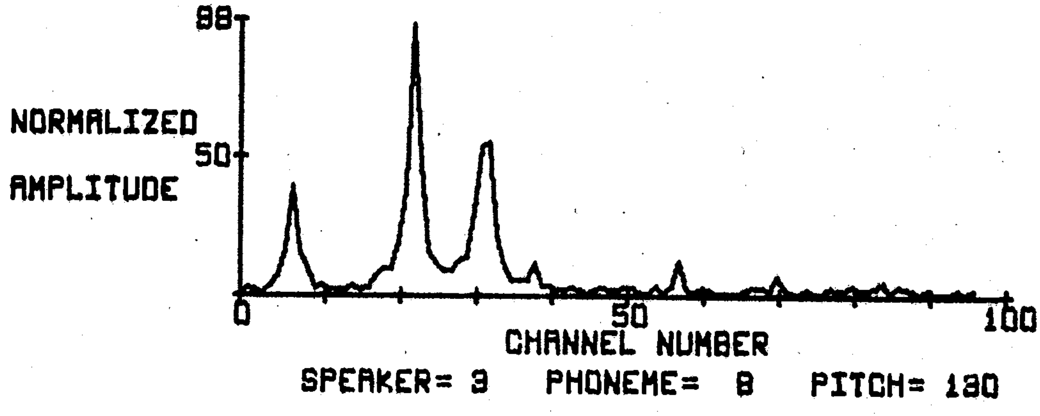


Figure 5.6. Comparison of CRT Display With the Computer Sampled Data for the Phoneme /u/ as in Boot

Calcomp 565 Plotter. The CRT display is smoothed with a 3.5 μsec time constant integrating RC network in series with the vertical input. In Figure 5.4 the phoneme /ae/ is recorded and the computer generated plot can be seen to be an accurate representative of the spectral pattern even though the computer sampled the data at a much slower rate. Some differences are to be expected because the CRT image could not be photographed during the computer sampling. The 50,000 Hz clock was used for the CRT display whereas during the time the computer was sampling the data the channels were sampled at a rate of about 30 channels per second. The second and fourth harmonics are slightly weaker in the photograph than in the plot.

In Figure 5.5 the phoneme /i/ is compared and the spectral displays are virtually identical. In Figure 5.6 the phoneme /u/ is different only in the second harmonic which is slightly lower in the photograph. This is partially due to the smoothing effect of the integrating circuit.

In addition to the tabulation of normalized channel output voltage, maximum channel output voltage level and Speech Tracker output voltage level, Appendix D also contains a computer generated plot of each of the phonemes. The plots are organized by speaker and phoneme so that by scanning vertically down the page, the effect of increasing the voice pitch can be observed. It is clear that dramatic changes in spectral patterns take place as the pitch is varied. These changes very likely account for the poor results of previous research with multiple speakers.

CHAPTER VI

PHONEME TRANSITIONS

6.1 Introduction. The spectral transitions which occur in natural speech have been studied in much less detail than have the properties of sustained phonemes. The sonagram or spectrogram recordings, which have been made by numerous research groups, used comparatively wide bandwidth filters (150 Hz) and time resolution was typically 20 msec (Halle, *et al.*, 42). Recently numerous authors have recognized the importance of these transitions in automatic speech recognition because phonemes in connected speech are heavily dependent upon the muscular tensions and vocal tract shape due to previous and anticipated phonemes. Any recognition scheme must be able to distinguish phonemes either individually or in connected groups.

The three dimensional display system which was functionally described in Chapter III was used to make several recordings by Speaker 3 of spectral transitions which are discussed in this chapter. The purpose of this chapter is to demonstrate the capability of this recording method for the study of phoneme transitions.

6.2 The Sampling Technique. In the study of spectral transitions of phonemes it is necessary to first determine the important parameters for discrimination and then to optimize the recording and display methods to minimize the irrelevant data and to emphasize the useful features. A pilot effort to identify these features could use the display system which was described in Chapter III. It provides a sample of each TAT channel output with a time resolution of 2 msec over a period of 64 msec. The sample is synchronized with the initiation of voicing by means of a Speech Detector which has a response time constant of 1.24 msec which is the same as the time constant of the fastest TAT channel. The first trace starts about 1 msec after initiation of voicing. The 64 msec duration was a limit imposed by the CRT resolution and available photographic recording instrumentation. Thomas, et al. (43) reported that the average duration of all phonemes is between 70 and 80 msec but that vowels are typically 200 to 300 msec in length. He also found in human perception experiments that vowel recognition decreased significantly for durations less than about 125 msec.

Further study may show that less resolution than 2 msec and a larger sample than 64 msec would be desirable. It may also reveal that only certain frequencies need be considered.

6.3 Examples of Phoneme Transitions. In order to illustrate the capabilities of this display, recordings were made of the transitions between all combinations of the vowels /ae/, /e/, /ε/ and /i/ as preceded by the stop-consonants (sometimes called plosives) /b/, /g/, /d/ and /t/. These combinations are illustrated in Table VII.

The recording of each transition was preceded by observing each of the phonemes individually and together by means of the oscilloscope display which was described in Chapter II. The use of a Hewlett Packard Model 1300 XY Display which provided a real time 10 inch by 10 inch CRT image revealed (within the limitation of the TAT channel response times) the channels in which the spectral changes could be observed over a longer time interval than is possible with the Speech Sampler Display.

The three dimensional display can then be used with a transparent overlay to determine the rate of spectral changes.

TABLE VII

Phoneme Transitions

1. /b/+/ae/ as in bat	9. /d/+/ae/ as in dad
2. /b/+/e/ as in bait	10. /d/+/e/ as in day
3. /b/+/ɛ/ as in bet	11. /d/+/ɛ/ as in debt
4. /b/+/i/ as in beet	12. /d/+/i/ as in deep
5. /g/+/ae/ as in gab	13. /t/+/ae/ as in tab
6. /g/+/e/ as in gay	14. /t/+/e/ as in taste
7. /g/+/ɛ/ as in get	15. /t/+/ɛ/ as in tent
8. /g/+/i/ as in geese	16. /t/+/i/ as in teeth

Figure 6.1 shows the energy burst of the phoneme /b/ and the transition to the phonemes /ae/ as in bat, /e/ as in bait, /ɛ/ as in bet and /i/ as in beet. The /b/ is a weak phoneme and therefore causes a slow build-up of voltage from even the highest frequency channels which respond fastest. The response time constant (τ) of each bandpass filter is given by $\tau = 25/\pi F_0$, where F_0 is the resonant frequency. It is seen in Figure 6.1 that the transition to the second phoneme takes place in about 26 msec.

The transition of the tongue controlled /g/ into each of the vowel

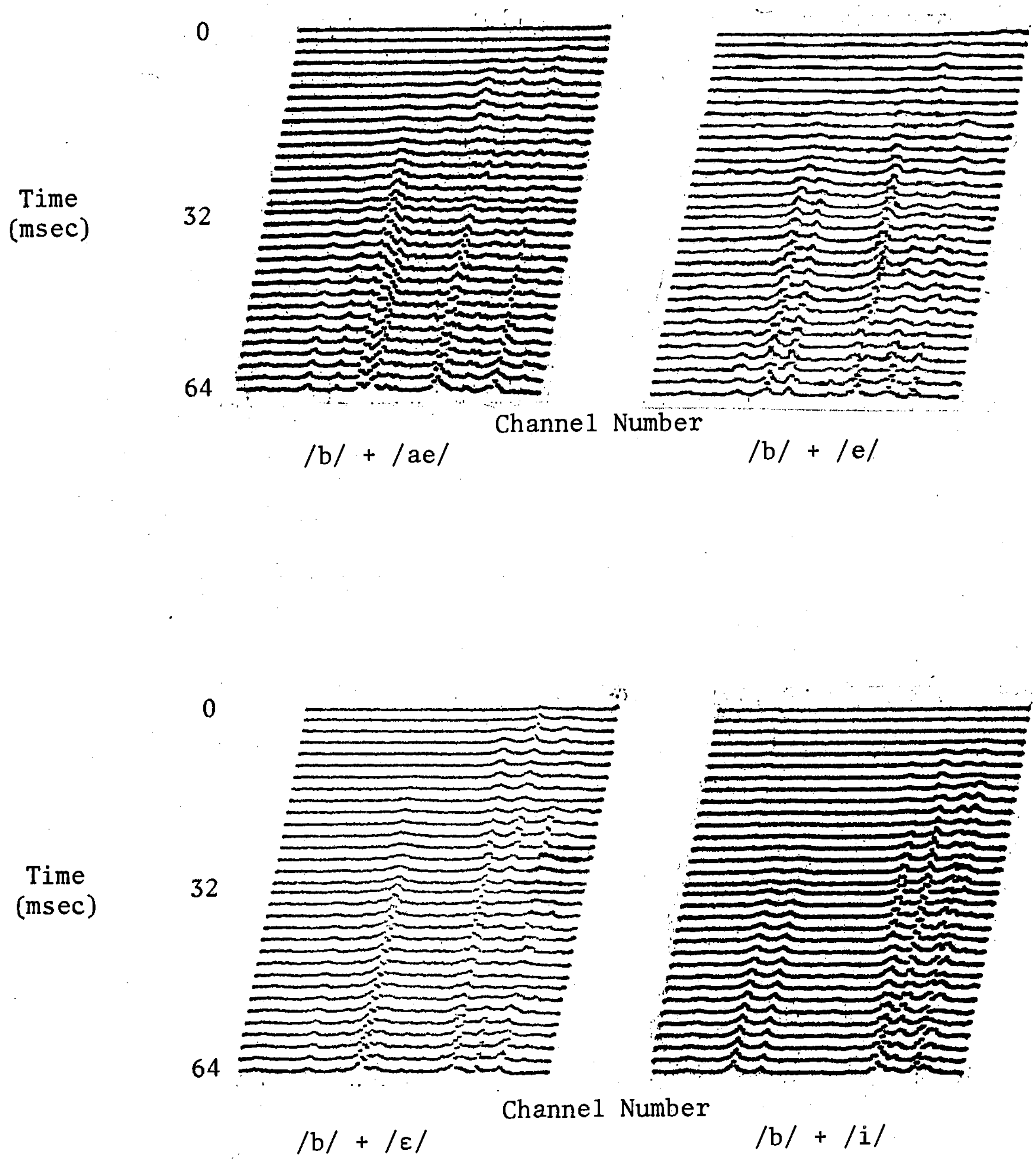


Figure 6.1. Speech Sampler Display of the Phoneme Transitions /b/+/ae/, /e/, /ɛ/ and /i/, as in Bat, Bait, Bet and Beet

phonemes takes place in about 45 msec as shown in Figure 6.2. The /g/ begins with a broad spectral burst in the 2,000 to 5,000 Hz band which is separated by several msec from the following phoneme except in the case of the /ε/ sound.

In Figure 6.3 the energy burst of the phoneme /d/ can be seen to be concentrated in the region of about 4,000 to 5,000 Hz. The second phoneme begins about 20 to 40 msec after initiation of voicing. The high frequency components appear about 20 msec earlier than the low frequency components. Only about half of this difference could be due to the response of the filters.

The phoneme /t/ is shown in Figure 6.4 to consist of a strong burst of energy in the 3400 to 4900 Hz band and, for the limited data which was taken, there was always a period of silence of about 15 msec between the /t/ and the sound which followed it. If this characteristic is confirmed by a more complete study, it may be useful as an identification parameter.

6.4 Other Uses of the Speech Sampler Display. The display can be started after some fixed interval from the start of voicing by any method which controls the delay of the pulse from the Speech Detector. A simple and effective method was found to be the use of the 100 μfd holding capacitor on the Speech Tracker, as was used in Chapter V for reading the voltage by the computer. This delays the start of the sweep to about 200 msec from the initiation of voicing. The sampling interval and sampling duration remain 2 msec and 64 msec, respectively. Figure 6.5 uses this delay and shows four phoneme transitions which were voiced at a much slower rate than normal. In each case all filter channels have had sufficient time to respond to the first phoneme and

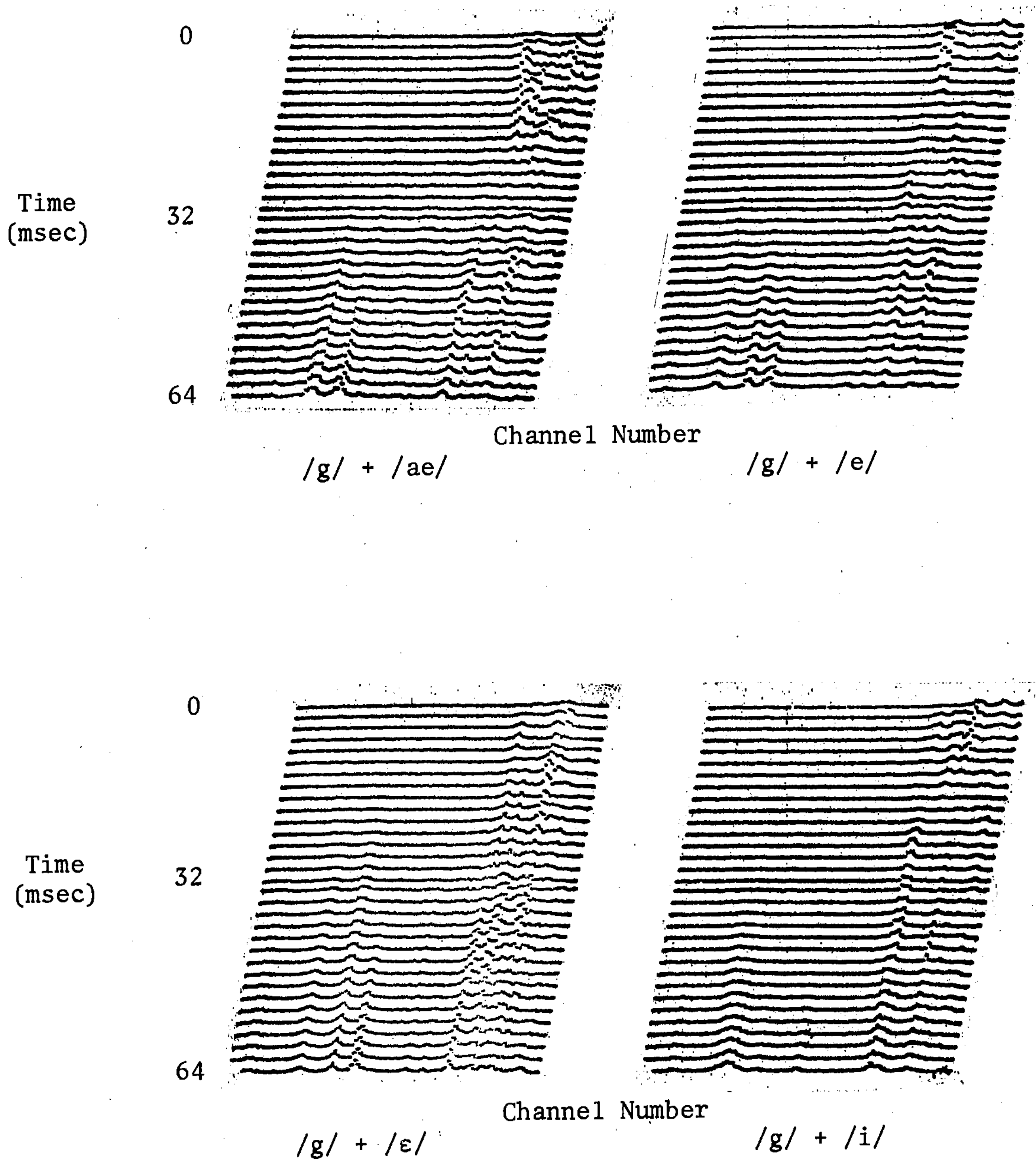


Figure 6.2. Speech Sampler Display of the Phoneme Transitions /g/+/ae/, /e/, /ε/ and /i/, as in Gab, Gate, Get and Geese

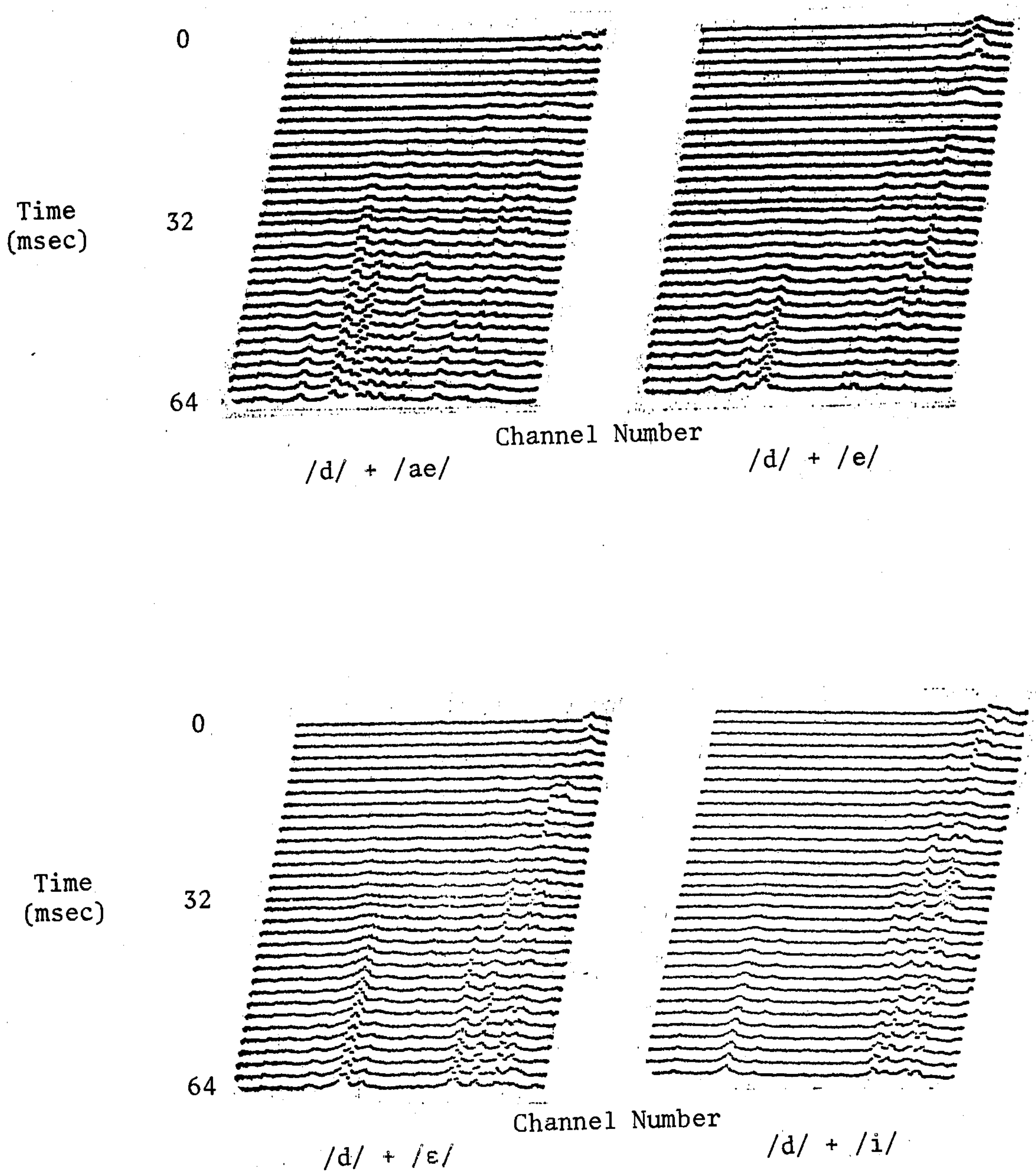


Figure 6.3. Speech Sampler Display of the Phoneme Transitions /d/+/ae/, /e/, /ε/ and /i/, as in Dad, Day, Debt and Deep

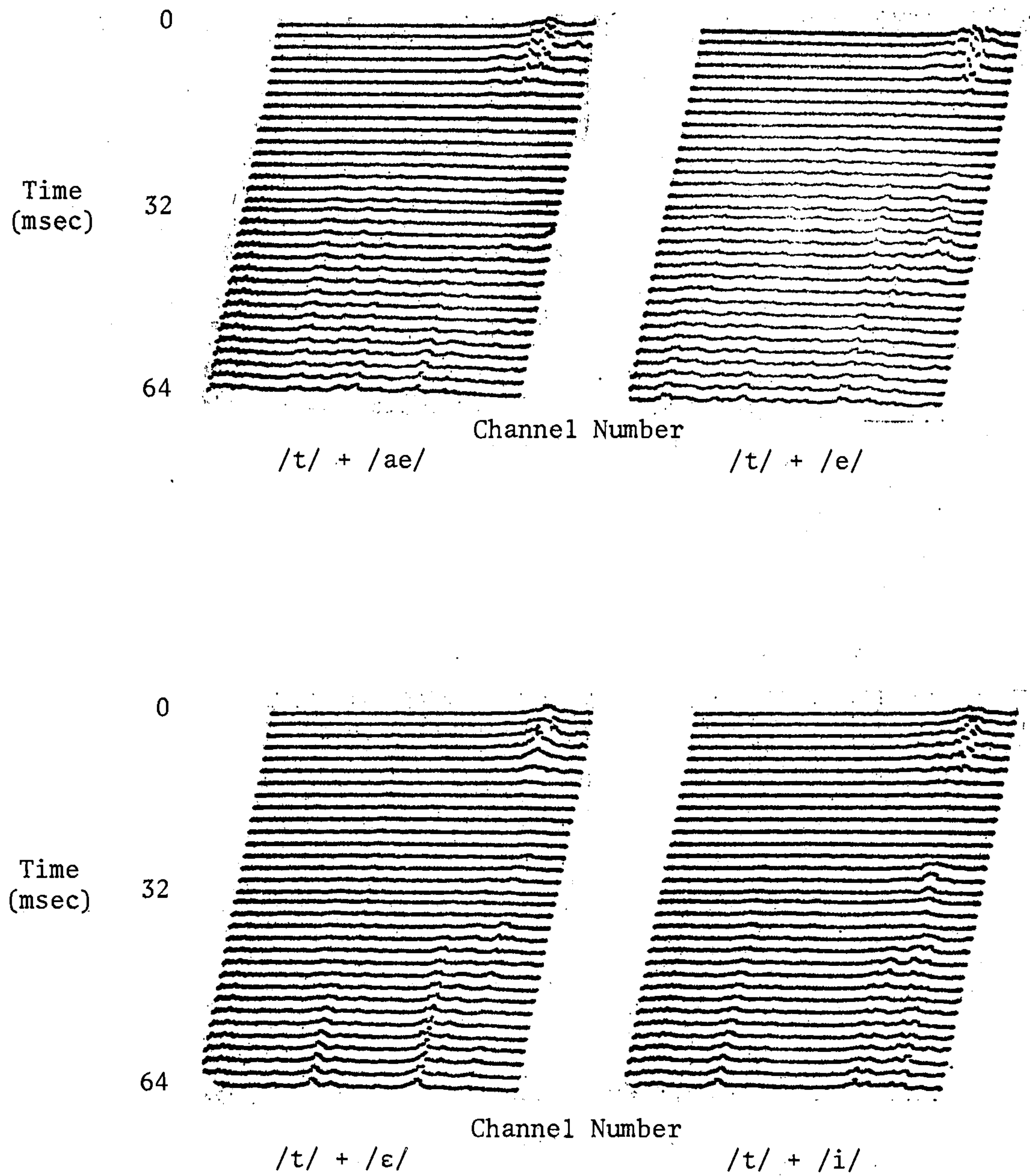


Figure 6.4. Speech Sampler Display of the Phoneme Transitions /t/+/ae/, /e/, /ɛ/ and /i/, as in <u>Tab</u>, <u>Tas</u>te, <u>Ten</u>t and <u>Tee</u>th

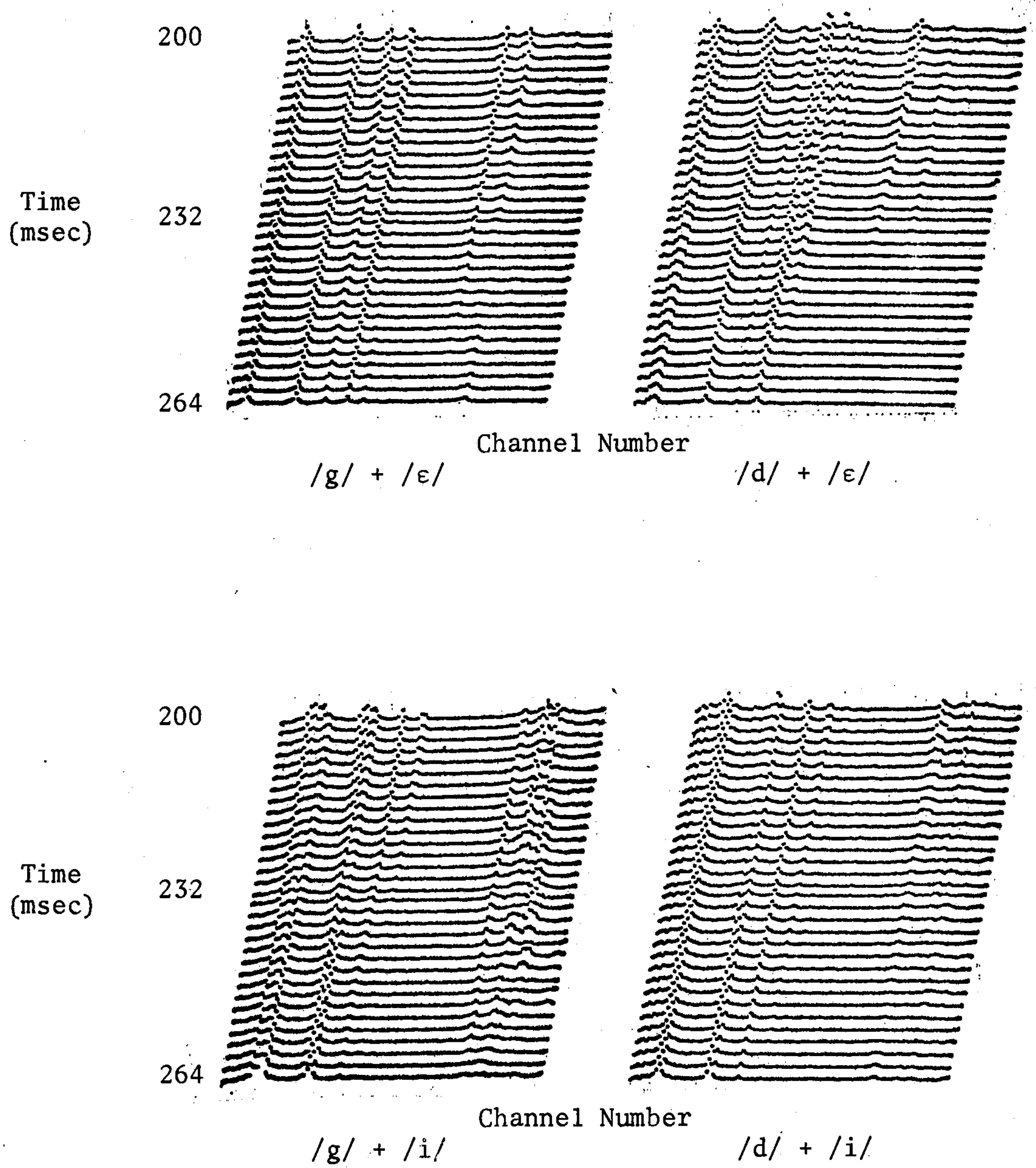


Figure 6.5. Speech Sampler Display of Phoneme Transitions Delayed by 200 msec

are in varying degrees of transition into the second phoneme. The /g/ in the /g/+/i/ transition is seen to be stronger and held longer than in the /g/+/ε/ transition.

Figure 6.6 shows three samples of the /b/+/ae/, as in bat, transition in which an energy shift is moving upward from about channel 58. This is converted by Table I to a frequency shift from about 1200 Hz to about 1500 Hz, during this time interval. In like manner, delayed samples of /b/+/e/, /ε/ and /i/ are shown in Figure 6.7. Slow voicing of the /b/+/e/ as in bait results in a narrowing of the energy in the 1100 to 1700 Hz region. In the case of /b/+/ε/ an increasing peak is seen at about 2500 Hz. A decrease in energy at about 350 Hz occurs in the /b/+/i/ sound.

Not only can the start of the display be delayed but also the duration of the display could be increased by changing the rate of clock pulses which step the multiplexer from channel to channel. If more accurate frequency discrimination is desired the programming panel of the multiplexer can be changed to display only those channels of interest.

A confirmation of the speed of response of each of the high frequency filters can be seen in most of the voiced recordings in this chapter. A careful inspection of the high frequency responses reveals that they have a period of fluctuation of about 8 msec. The output of the high frequency group goes to zero about every fourth trace. This corresponds to the rate of pitch excitation and is caused by the decay of the filters between glottal pulse excitations. The harmonics to which these channels respond are created by resonances of the vocal tract which are excited by bursts of air released by the opening and

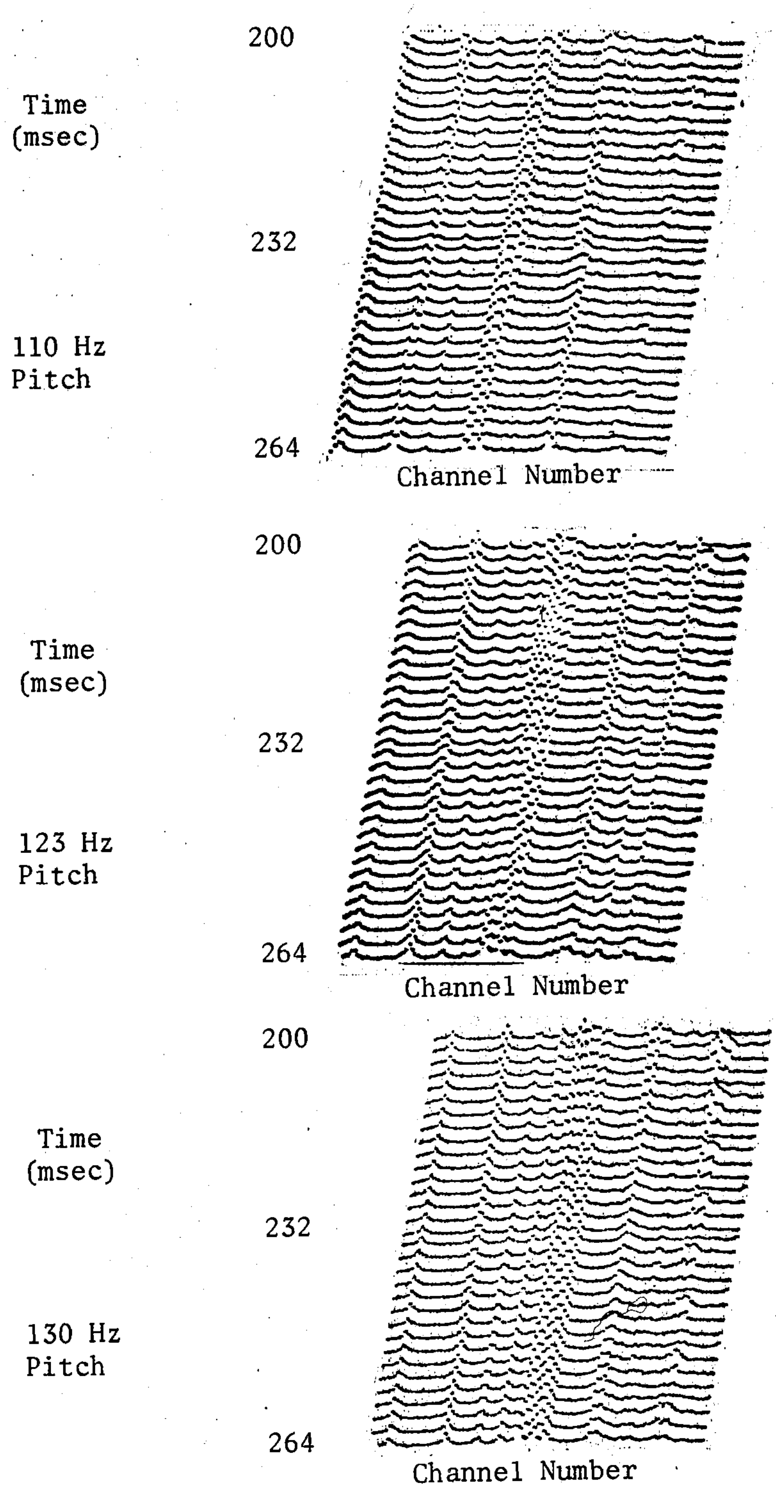


Figure 6.6. Delayed Speech Sampler Display of the Phoneme Transition /b/+/ae/ as in <u>Bat</u> at Three Pitch Rates

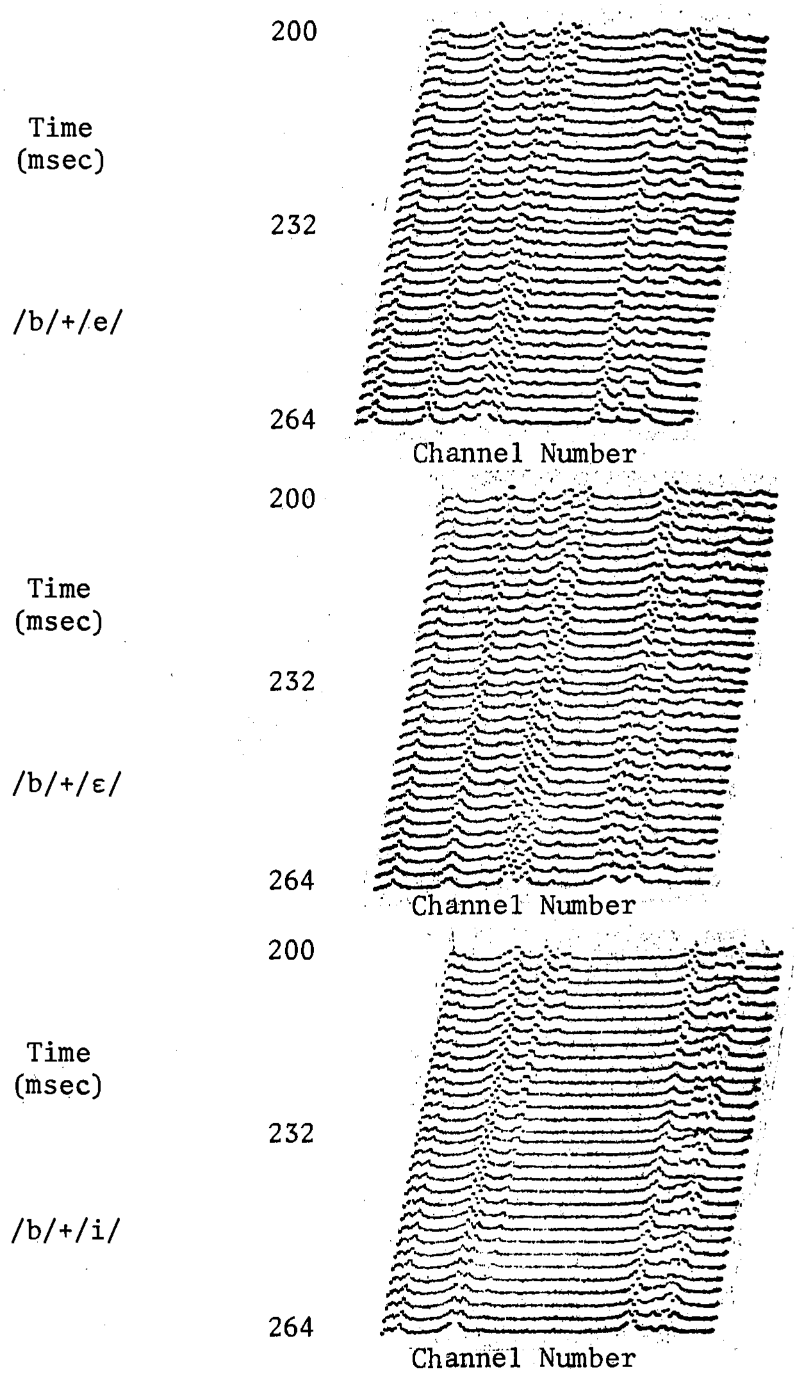


Figure 6.7. Delayed Speech Sampler Display of the Phoneme Transitions /b/+/e/, /ε/ and /i/ at a Pitch of 123 Hz

closing of the glottis at the pitch rate. At a pitch of 125 Hz the glottis releases a burst of air into the vocal tract every 8 msec. A Laplace transform analysis of the TAT channels shows that the one-time-constant decay times of channel numbers 55 and above are less than the 8 msec spacing between the glottal pulse excitations.

CHAPTER VII

PHONEME RECOGNITION

7.1 Introduction. The variation of phoneme spectral characteristics with pitch changes, which can be seen by scanning the data in Appendix D, underlines the importance of finding the key features for identification. Visual inspection of the plots, however, provides more discouragement than ideas for simple detection of unique characteristics. This chapter is devoted to the application of a machine-learning algorithm to the problem of automatically identifying phonemes presented in the form of multi-element vectors of the type produced by the system described herein.

7.2 A Recognition Algorithm. The development of a machine recognition scheme which can consider a large number of descriptive elements and provide the most likely identification of an unknown input is based upon a paper titled "Experiments in Machine Learning" by V. W. Bolie (44). Numerous simplifications were made to tailor the machine learning process to speech inputs and to limit the reference vector (phoneme) adaptation process to a training period after which the reference vectors were not changed. Furthermore, the reference vectors for each phoneme contained only 40 elements in order to reduce computational time and reduce the memory required of the IBM 1620 Computer. A listing of the program is contained in Appendix E.

A 40 element reference vector which is an average of three input speech samples, for each of three phonemes, was calculated by the formula,

$$R(I,J) = \frac{A(1,J) + A(2,J) + A(3,J)}{3}$$

where: I is the phoneme number;

A(M,J) is the normalized amplitude of phoneme sample M, element J and;

$1 \leq J \leq 40$, representing TAT channel numbers 48 through 87.

The selection of only channels 48 through 87, covering the spectral band from 763 to 4150 Hz, was made in an attempt to avoid the radical energy shifts which occur at lower frequencies as the pitch is changed. It was hoped that for some phonemes, a pitch independent recognition parameter might be discovered.

Also calculated during the learning phase is a 40 element tolerance vector for each reference phoneme. The tolerance vector is a measure of the amplitude variation of each of the filter bank channels for the pitch values which were selected for machine training. It is calculated by the relation,

$$T(I,J) = \frac{AMAX1\ (J) - AMIN\ (J)}{2}$$

where: AMAX1(J) is the largest amplitude of element J among the 3 training vectors chosen to represent phoneme I and;

AMIN(J) is smallest amplitude of element J among the same training vectors.

At the conclusion of the learning phase the machine has in storage a reference vector and a tolerance vector for each phoneme which was chosen for identification. An unknown phoneme (stimulus vector) is then read into the machine and a discrepancy vector is calculated by use of each of the reference and tolerance vectors, according to

$$D(I) = \frac{1}{40} \sum_{J=1}^{40} \left[\frac{S(J) - R(I,J)}{1 + T(I,J)} \right]^2$$

where: S(J) is element J of the stimulus factor.

The discrepancy vector is used to calculate a single element output for each of the reference phonemes, the size of which indicates the degree of closeness of fit between the unknown and the stored reference vector. The output vector is a number between 0 and 99 calculated by,

$$W(I) = \frac{99}{1 + D(I)} \quad .$$

The purpose of the tolerance vector as used here is to weight each element of the discrepancy vector according to the repeatability of the voltage from the corresponding filter bank channel for the different pitches used in the calculation of the reference vector.

The calculation of each of these vectors is somewhat different from the methods used by Bolie (44). In calculation of the discrepancy vector numerous combinations were tried to find the one which best identifies the correct phoneme. For example it was discovered that just reversing the sequence of (1) summing the terms and then squaring, or (2) squaring each term and then summing, resulted in the improvement of the ratio of correct W to incorrect W by from 20 to 60 percent and increased the number of correct recognitions by 11 percent. In a like manner the squaring operation increased the W ratio by 25 percent.

In the calculation of the output vector it was found that division, with the addition of a constant to prevent division by zero, resulted in a significant savings in computer time over the use of an exponential function.

The time required for these algorithms was surprisingly short. For example the learning process using three phonemes at three pitches each, was only about 5 minutes. An unknown sound which was already normalized could be read in from IBM cards and recognized in less than one minute.

7.3 Identification Accuracy. The phonemes /ae/ as in bat, /e/ as in ate and /i/ as in beet were used to test the recognition algorithm. The data which is tabulated and plotted in Appendix D for speakers 1, 2 and 3 at pitches of 110, 130 and 164 Hz, was read into the computer and used for calculation of the reference and tolerance vectors. The success of the algorithm in correct identification of the phoneme /ae/ at all of the recorded pitch frequencies for the three male speakers is shown in Tables VIII, IX and X. The response number is the output vector, rounded off to the nearest whole number, and varies from 99, for a perfect fit between the stimulus vector and the reference vector, to 0 which represents a very poor fit. In each case the missed identifications are labeled with an asterisk.

In compiling the results which are shown in Table VIII the reference and tolerance vectors were calculated from the data of speaker 1. The use of the phoneme /ae/ as a stimulus vector resulted in perfect identification of all of the speaker 1 data even at pitches different from those for which the machine was trained. It can be seen that the

TABLE VIII

RESPONSES OF THE SPEAKER-1-TRAINED MACHINE TO THE PHONEME /ae/

Stimulus		Response			Miss
Speaker	Pitch (Hz)	/ae/	/e/	/i/	
1	100	68	5	5	
	110	66	5	1	
	123	33	2	1	
	130	66	2	2	
	146	72	5	4	
	164	69	3	4	
	174	15	3	3	
2	100	11	3	2	
	110	19	6	4	
	123	4	5	2	*
	130	12	3	2	
	146	13	18	8	*
	164	32	34	17	*
	174	41	38	19	
3	100	47	30	11	
	110	41	48	31	*
	123	48	43	32	
	130	41	43	47	*
	146	42	45	25	*
	164	27	15	9	
	174	14	8	2	

TABLE IX

RESPONSES OF THE SPEAKER-2-TRAINED MACHINE TO THE PHONEME /ae/

Stimulus		Response			Miss
Speaker	Pitch (Hz)	/ae/	/e/	/i/	
1	100	26	3	1	
	110	5	93	75	*
	123	6	58	37	*
	130	17	1	42	*
	146	32	3	97	*
	164	18	6	2	
	174	33	2	2	
2	100	30	41	56	*
	110	71	4	77	*
	123	15	4	54	*
	130	75	3	41	
	146	61	2	2	
	164	64	43	3	
	174	34	28	6	
3	100	17	9	5	
	110	21	2	2	
	123	23	40	18	*
	130	21	29	18	*
	146	18	12	14	
	164	14	12	7	
	174	7	2	3	

TABLE X

RESPONSES OF THE SPEAKER-3-TRAINED MACHINE TO THE PHONEME /ae/

Stimulus		Response			Miss
Speaker	Pitch (Hz)	/ae/	/e/	/i/	
1	100	3	11	6	*
	110	3	72	1	*
	123	1	2	1	*
	130	1	4	2	*
	146	2	9	4	*
	164	11	8	5	
	174	2	7	2	*
2	100	61	5	2	
	110	1	8	2	*
	123	58	4	1	
	130	96	4	1	
	146	2	16	6	*
	164	3	33	18	*
	174	7	32	17	*
3	100	28	7	9	
	110	84	24	29	
	123	54	26	44	
	130	77	30	44	
	146	58	20	20	
	164	67	11	6	
	174	8	1	3	

machine failed in 3 out of 7 pitches when each of the other speaker data sets were used as the stimulus vectors. The confidence that a correct choice was made is also seen to be significantly lower even for the correct identifications of the speaker 2 and 3 data as indicated by the much lower ratio of response number for /ae/ to that of the other sounds.

Table IX shows the results when the same stimulus vectors are used but the machine was trained using speaker 2 data. In this case it had difficulty in distinguishing between the phonemes /ae/ and /i/ at the lower pitch frequencies when voiced by speaker 2 and confused /ae/ with both /e/ and /i/ for the speaker 1 stimulus vectors. The speaker 3 input data resulted in correct identifications in 5 of 7 attempts but a low confidence is indicated by the small numbers.

The use of speaker 3 as a reference resulted in perfect identification of stimulus vectors of that speaker and the reflection of good confidence in the decisions as revealed by Table X. However, the algorithm failed in 10 of the 14 attempts to identify the data of other speakers and showed good confidence in only 3 of the decisions.

The results of tests with this algorithm and the use of only data between 763 and 4150 Hz are typical of the multiple speaker results of other researchers, although the success with single speakers and multiple pitches was encouraging. Before extension of the spectral band to the lower frequencies or modification of the recognition algorithm, it might prove instructive to tabulate the results of training the machine with many phoneme combinations and the use of several male and female voices.

On the basis of this preliminary information it appears that additional research will need to be focused on the problem of transforming each stimulus vector into a form which is invariant to speaker differences. The coding techniques for such a transformation are yet to be discovered. The high resolution spectral filter bank system which was built as a part of this thesis should provide an adequate data base for such additional research.

7.4 Computer Operation. The assistance in computer programming and equipment operation, in taking the data for this chapter, provided by Electrical Engineering students J. W. Mabray and R. B. Johnson is gratefully acknowledged.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

8.1 Summary. This thesis is devoted to high resolution studies of speech sounds, including the development of the necessary hardware. The included audio filter bank has 97 channels each of which has a Q of 25 thus providing more than a doubling of selectivity of the best previous system. The 100-6400 Hz breadth of the spectrum covered is believed to contain all of the speech information which is significant in recognition for pitches above 100 Hz.

The formant theory of speech perception which has been the basis for most previous speech research, for both identification and synthesis, is examined with the aid of the high resolution spectral data. The loss of detail in the smoothing process of calculating formants is illustrated and offered as a possible reason for the failure of speech recognition systems to accommodate different speakers and pitch variations.

A method has been developed whereby the output of all spectral filters, when responding to sustained phonemes, can be read into the IBM 1620 Computer. This was used to compile a data set of 11 phonemes, spoken by 3 males and 2 females, each at pitch frequencies varying over one octave. The data is tabulated in normalized form for use by others in future speech recognition research. Plots of all 385 phoneme samples, which were made with the use of the Calcomp 565 plotter, are also included for visual inspection.

Phoneme transitions in connected speech are examined by means of the three dimensional oscilloscope display system which was designed and built as a part of the research. Photographs are included which show the capability of the system to accurately measure energy shifts as a function of time. The differences in several stop-consonant-vowel transitions are also illustrated. Methods for varying the display starting time, duration of the sample, sampling rate and spectral band to be displayed, are outlined.

A machine-learning algorithm was modified and applied to the problem of automatically identifying selected phonemes presented in the form of multi-element vectors of the type produced by the high resolution filter bank. A preliminary test of the algorithm using spectral data restricted to the 763-4150 Hz frequency range resulted in near-perfect recognition of phonemes spoken by the person whose data had been used to train the machine. Rather poor recognition scores for other speakers were obtained.

8.2 Significant Conclusions. The failure of the formant theory of speech perception to provide orthogonal descriptions for speech of different speakers has only recently been recognized. Martin (21), for example, added to his formant data, the rate of change of energy with frequency in his attempts to develop a digit recognition scheme. Mundie and Moore (24) are using temporal shifts. This thesis has provided a basic research tool which preserves the detailed spectral characteristics of speech. Data has been collected which argues for a more descriptive measure of phoneme spectral detail than has previously been thought necessary. The data set which is provided in Appendix D could be used for a preliminary effort. The three dimensional display further provides a means for study of phonemes in connected speech which may vary considerably from the same phonemes spoken in isolation. The recognition algorithm is shown to be very effective in recognition over a wide range of pitch frequencies for a single speaker. Significant improvements in automatic speech recognition should be possible through additional research using a number of both male and female speakers, particularly if a transformation can be developed which will make the filter bank outputs invariant with respect to pitch and speaker differences.

BIBLIOGRAPHY

(1) Lindgren, Nilo. "Machine Recognition of Human Language: Part I - Automatic Speech Recognition." IEEE Spectrum, (March, 1965), 114-135.

(2) Lindgren, Nilo. "Machine Recognition of Human Language: Part II - Theoretical Models of Speech Perception and Language." IEEE Spectrum, (April, 1965), 45-69.

(3) Teacher, Charles F., Henry G. Kellet, and Louis R. Focht. "Experimental Limited Vocabulary Speech Recognizer." IEEE Transactions on Audio and Electroacoustics, AU-15(3), (September, 1967), 127-130.

(4) Ewing, G. D. and John F. Taylor. "Computer Recognition of Speech Using Zero-Crossing Information." IEEE Transactions on Audio and Electroacoustics, AU-17(1), (March, 1969), 37-40.

(5) Sitton, G. A. "Acoustic Segmentation of Speech." International Journal of Man-Machine Studies, 2(1), (January, 1970), 61-102.

(6) Pierce, J. R. "Whither Speech Recognition?" Journal of the Acoustical Society of America, 46(4), (March, 1969), 1049-1051.

(7) Bekesy, George von. "Hearing Theories and Complex Sounds." Journal of the Acoustical Society of America, 35(4), (April, 1963), 588-601.

(8) Flanagan, James L. "Models for Approximating Basilar Membrane Displacement - Parts I and II." Bell System Technical Journal, 39, (September, 1960), 1163-1191 and 41, (May, 1962), 959-1009.

(9) Flanagan, J. L. Speech Analysis Synthesis and Perception. New York: Academic Press, Inc., 1965.

(10) Fletcher, Harvey. Speech and Hearing in Communication. New York: Van Nortand Co., 1953.

(11) Campanella, S. Joseph and Douglas Phyfe. "Application of a Model of the Analog Ear to Speech Signal Analysis." IEEE Transactions on Audio and Electroacoustics, AU-16(1), (March, 1968), 26-35.

(12) Hughes, G. W. and J. F. Hemdal. "Speech Analysis." Final Report, Task 562802, Project 5628, AF 19(628)-305, Purdue University School of Electrical Engineering, TR-ee65-9, AFCRL-65-681, (July, 1965).

(13) Bobrow, Daniel G. and Dennis H. Klatt. "A Limited Speech Recognition System." Final Report, Contract No. NAS 12-138 (Bolt, Beranek and Newman, Inc., Cambridge, Mass.), NASA, (May 15, 1968).

(14) Lecours, M. and J. J. Sparkes. "Adaptive Spectral Analysis for Speech Sound Recognition." IEEE Transactions on Audio and Electroacoustics, AU-16(4), (December, 1968), 523-525.

(15) Hyde, S. R. "Recognition of Speech by Machine." Post Office Electrical Engineers Journal, (July, 1969), 100-104.

(16) Gilmour, W. D. "Electronic Speech Recognition." Wireless World, 75, No. 1400, (1969), 76-80.

(17) Reddy, D. R. "Segmentation of Speech Sounds." Journal of the Acoustical Society of America, 40, (August, 1966), 307-312.

(18) Reddy, D. R. "Phoneme Grouping for Speech Recognition." Journal of the Acoustical Society of America, 41, (May, 1967), 1295-1300.

(19) Stewart, J. L. "Perceptual Mechanisms in Hearing as Explained With an Electronic Analog Simulation for the Ear." Paper presented at a meeting of the American Larynological, Rhinological and Otological Society, Inc., Pebble Beach, California, (January 31, 1965).

(20) Stewart, J. L. "Signal Processing in the Auditory Pathways and Evaluation of the Mammalian Ear." Report on research sponsored by AMRL, (February, 1968).

(21) Martin, T. B. "Acoustic Recognition of a Limited Vocabulary in Continuous Speech." (unpub. Ph.D. thesis, University of Pennsylvania, 1970).

(22) Glaesser, E., W. F. Caldwell, and J. L. Stewart. "An Electronic Analog of the Ear." Technical Documentary Report No. AMRL-TOR-63-60, Biophysics Laboratory, Wright Patterson AFB, Ohio, (June, 1963).

(23) Caldwell, William F. "Recognition of Sounds by Cochlear Patterns." IEEE Transactions on Military Electronics, MIL-7(2, 3), (April-July, 1963), 179-185.

(24) Mundie, J. R. and T. J. Moore. "Speech Analysis as the Ear Sees It: Aural Topography." AMRL Technical Report 67-176, (1967).

(25) Cannon, Mark W. "A Method of Analysis and Recognition for Voiced Vowels." IEEE Transactions on Audio and Electroacoustics, AU-16(2), (June, 1968), 154-158.

(26) Bolie, Victor W. "Computer-Aided Refinements in the Design of an Artificial Cochlea." SWIEEECO Record of Technical Papers, Dallas, Texas, (April, 1970), 308-311.

(27) Lake, O. L. "Analysis of Speech Sounds Through Digital Computer Simulation of the Cochlea." (unpub. Ph.D. thesis, Oklahoma State University, 1970).

(28) Oppenheim, Alan V. "Speech Analysis-Synthesis System Based on Homomorphic Filtering." Journal of the Acoustical Society of America, 45(2), (February, 1969), 458-465.

(29) Huelsman, O. P. Handbook of Operational Amplifier Active RC Networks. Tucson, Arizona: Burr-Brown Research Corp., 1966.

(30) Bolie, Victor W. "Active Filter Circuits." (Personal communication, Oklahoma State University, January, 1971).

(31) Tarmz, R. and M. S. Ghausi. "Very High-Q Insensitive Active RC Networks." IEEE Transactions on Circuit Theory, Vol. 17, No. 3, (August, 1970), 358-366.

(32) Fristoe, H. T. The Use of Q Equations to Solve Complex Electrical Networks. Oklahoma State University Engineering Experiment Station Publication, No. 105, (June, 1961), 105.

(33) Fant, C. M. Gunnar. "On the Predictability of Formant Levels and Spectrum Envelopes From Formant Frequencies." In Readings in Acoustic Phonetics. Ed. Ilse Lehiste. Cambridge, Massachusetts: MIT Press, 1967, 44-56.

(34) Peterson, G. and H. Barney. "Control Methods Used in a Study of the Vowels." Journal of the Acoustical Society of America, 24(2), (March, 1952), 175-184.

(35) Gerstman, L. J. "Classification of Self-Normalized Vowels." IEEE Transactions on Audio and Electroacoustics, AU-16(1), (March, 1968), 78-80.

(36) Fairbanks, G. and P. Grubb. "A Psychophysical Investigation of Vowel Formants." In Experimental Phonetics: Selected Articles. Ed. Grant Fairbanks. Urbana, Illinois: University of Illinois Press, 1966, 99-115.

(37) Campanella, S. J. and D. Phyfe. "Application of a Model of the Analog Ear to Speech Signal Analysis." IEEE Transactions on Audio and Electroacoustics, AU-16(1), (March, 1968), 26-35.

(38) Forgie, J. W. and C. D. Forgie. "Results Obtained From a Vowel Recognition Computer Program." The Journal of the Acoustical Society of America, 31(11), (November, 1959), 1480-1489.

(39) Pols, L. C. W., L. J. T. Van DerKamp, and R. Plomp. "Perceptual and Physical Space of Vowel Sounds." The Journal of the Acoustical Society of America, 42, (August, 1969), 458-467.

(40) Flanagan, J. L. "Estimates of the Maximum Precision Necessary in Quantizing Certain 'Dimensions' of Vowels." The Journal of the Acoustical Society of America, 29, (April, 1957), 533-534.

(41) Gray, D. E. American Institute of Physics Handbook, 2nd Edition. New York: McGraw Hill Book Co., 1963, 115.

(42) Halle, M., G. W. Hughes and J. P. A. Radley. "Acoustic Properties of Stop Consonants." The Journal of the Acoustical Society of America, 29, (January, 1957), 107-116.

(43) Thomas, I. B., Peter B. Hill, Francis S. Carroll, and B. Garcia. "Temporal Order in Perception of Vowels." The Journal of the Acoustical Society of America, 48, (April, 1970), 1010-1013.

(44) Bolie, Victor W. "Experiments in Machine Learning." (A copyrighted Paper, 1969).

APPENDIX A

SUMMARY OF ABBREVIATIONS

This appendix includes a listing in Table XI of the abbreviations which have been used in this thesis. In addition, resistance is in ohms and capacitance is in micro-farads in all of the figures unless otherwise indicated.

TABLE XI

SUMMARY OF ABBREVIATIONS

Abbreviation	Definition
SEF	Single Equivalent Formant Frequency
SEFA	Single Equivalent Formant Amplitude
TAT	Tuned Amplitude Tracker
E	Input voltage
V	Output voltage
Y_z	Admittance of z^{th} branch
A	Gain
R	Resistance
C	Capacitance
Q	Bandpass filter selectivity
Q_1	Feedback parallel LC selectivity
F	Frequency
F_0	Resonant frequency
F(K)	Resonant frequency of channel K
T	$1/2\pi F_0$
k	Feedback ratio
r_p	Equivalent resistance in parallel with ideal L and C
r_s	Equivalent resistance in series with ideal L and C
X_p	Parallel LC reactance
X_s	Equivalent series reactance of a parallel LC
OA	SN72 709N Operational Amplifier
LHPP	Lower half power point frequency
UHPP	Upper half power point frequency

TABLE XI (Continued)

Abbreviation	Definition
msec	milli-seconds
nsec	nano-seconds
μsec	micro-seconds
pfd	pico-farads
μfd	micro-farads (all capacitances are in μfd unless otherwise noted)
CRT	Cathode ray tube oscilloscope
LPF	Low-Pass Filter
ON	Transistor is conducting
OFF	Transistor is not conducting
GPR	Glottal Pulse Rate

APPENDIX B

ACTIVE NETWORK REALIZATION OF A 97-CHANNEL AUDIO FILTER BANK

The manuscript which comprises this appendix is a brief description of the audio filter bank which was constructed as a part of this thesis. It has recently been accepted for publication in the Journal of the Audio Engineering Society. It includes the results of performance measurements for the bandpass and low-pass filters which do not appear elsewhere in the thesis. In addition, an explanation of the SN72 709N OA frequency compensation which was developed by Dr. H. T. Fristoe, is provided.

ACTIVE-NETWORK REALIZATION OF A 97-CHANNEL AUDIO FILTER BANK

by

Victor W. Bolie[1], Harold T. Fristoe[2], and James E. Baker[3]

Introduction

Problems related to speech bandwidth reduction and automatic speech recognition have been investigated in various laboratories for more than twenty years with increasing vigor but with decreasing hope for simple solutions. A good summary of the state of the art as of 1965 was given in the two papers by Lindgren (1,2). The basic problem, of course, is essentially one of efficient coding. It is well-known, for example, that English text can be transmitted at a rate of 120 words per minute by means of a fifty-bit-per-second synchronous binary teletype code. Hence, the utilization factor of an ordinary telephone line can be increased by about two orders of magnitude if something better than the present "vocoder" technique can be devised as a real-time coding scheme. The same coding scheme would naturally facilitate real-time automatic speech recognition for verbal control of computers and other machines. However, investigations of speech signals thus far have not yet even resolved the relative merits of an aritificial cochlea _vs_ an audio filter bank as a preprocessor.

---

[1]Professor and Holder of the Albrecht Naeter Chair, College of Engineering, Oklahoma State University

[2]Professor of Electrical Engineering, Oklahoma State University

[3]Doctoral Candidate in Electrical Engineering, Oklahoma State University

Flanagan (3) made one of the first comprehensive attempts to electronically simulate the functions of the basilar membrane; his model accurately reproduced the impulse responses of the successive adjacent segments of the membrane, but did not simulate the biological property of stripping off more and more of the high frequency components of the acoustic signal as it travels around the cochlear spiral toward the helicotrema. A more realistic but somewhat crude simulation of the basilar membrane functions was achieved by Glaesser, Caldwell, and Stewart (4). Bolie (5) made refinements in the design of the Glaesser-Caldwell-Stewart model of the basilar membrane, by use of a digital computer to search through all the various network parameters. Caldwell (6) studied the responses of the analog ear to sustained speech sounds by extracting a 19-point sample of each cochlear pattern; cross-correlation techniques were then applied in order to develop a confusion matrix for eight voiced vowels and eight whispered vowels; the procedure was not successful in separating the various speech sounds. Guttman and Flanagan (7) simulated the equations of the basilar membrane on a digital computer and produced three-dimensional plots of the transient propagation of high-pass-filtered pulse trains in order to study pitch cues; the results tended to agree with the well-known beat-frequency phenomenon. Mundie and Moore (8) used the Glaesser-Caldwell-Stewart model of the cochlea to generate three-dimensional terrain maps (vibration amplitude vs basilar-membrane distance and time) for various speech phonemes. Campanella and Phyfe (9) used a non-propagating 40-section (Flanagan) model of the cochlea and a high-frequency pre-emphasis (6 db-per-octave rise to 4,000 Hz) of speech sounds to examine the effects of high-order (up to sixth) spatial differencing in processing the resulting cochlear patterns; the results indicated that "high order differencing networks appear unnecessary and even detrimental." Cannon (10) used autocorrelation techniques together with a 24-section analog basilar membrane with 6 db-per-octave pre-emphasis in the 0-4 kHz region to implement automatic recognition of seven vowel sounds, and found that "very high recognition scores were obtained" when only half a glottal-pulse period and only four membrane sections (3000, 1470, 660, 311 Hz) were used.

The audio filter bank has historically been the predominant means used to isolate speech characteristics for classification. The filter design characteristics (including the number of filters, their center frequencies, and their bandwidths) have been heavily influenced by the low selectivity (Q = 2) of the cochlea, the formant theory of speech perception, and a desire to minimize the number of filters to limit data reduction requirements. According to Lindgren (1), one of the first spectral processes was developed at Bell Telephone Laboratories to recognize the spoken digits "zero" through "nine". The first model, which was reported in 1952, used an 800 Hz low-pass filter with an axis-crossing counter to locate the first formant, and a 1000 Hz high-pass filter with an axis-crossing counter to locate the second formant. By 1958 the system was modified to handle 16 different linguistic elements. Recognition accuracy of the order of 95 percent was achieved with a single speaker, but this fell to the 50 percent range when the decision criteria were not adjusted for the particular speaker. Hyde (11) found that the use of a wideband filter and an axis-crossing counter provides an accurate frequency measure for simple waveforms but when two or more components have comparable amplitudes serious errors can arise. However, Gilmour (12) used this technique by selecting a 700 Hz low-pass filter to isolate the first formant, a 700 to 2100 Hz bandpass filter for the second formant, and a 2100 Hz high-pass filter for the third formant; a single-speaker recognition score of 90 percent for seven words and the ten digits was achieved. Efforts based on this wideband filter approach appear to have in common the failure to identify stops and fricatives, and the inability to eliminate speaker differences.

In further attempts to identify the important characteristics of speech for classification and recognition, recent research efforts have turned to narrow-band discrete filters. A comprehensive investigation of speech characteristics was conducted by Hughes and Hemdal (13), who made use of a bank of 35 bandpass filters having center frequencies ranging from 286 Hz to 9500 Hz. The bandwidths varied from 46 Hz for the 286 Hz filter to 963 Hz for the 9500 Hz filter, with all filters having a selectivity (Q) of 10 or less. No pre-emphasis was used. Each tuned filter consisted of a cascade of two series LC resonators the output of which was applied through a cathode follower to a rectifying-and-smoothing network. By means of a stepping switch, the spectral output was then digitized and stored in an IBM 7090 computer for development of phoneme recognition criteria. Nine distinctive features were used to identify 34 phonemes. Three distinctive features were found to reasonably separate the vowels with some overlap which could be removed with the additional dimension of time. The investigators noted that a single speaker was used for all of the research because "a useful method of normalization has not been determined", and concluded that a set of filters with good precision below

500 Hz was needed for accurate first-formant tracking and for measurement of the low-frequency stress cues. Bobrow and Klatt (14) developed a speech recognition system for a limited vocabulary of spacecraft-related words spoken by several speakers. A 19-channel spectrum analyzer was designed to process male voices which varied in glottal pulse rate from 80 to 150 Hz. Bandwidths (BW) were selected to insure that between two and four harmonics were present in every filter. A 360 Hz BW was used for center frequencies varying from 260 Hz to 2780 Hz, a 600 Hz BW at 3260 Hz, a 840 Hz BW at 3980 Hz, and a 1080 Hz BW at 4940 and 6020 Hz. Data were collected in the form of a 19 x 200 matrix by taking 2.0-second word samples at a 100 Hz sampling rate. High-frequency pre-emphasis of 6 db-per-octave was used. Each element of the data matrix was a number ranging from 0 to 63, in logarithmic units, covering a 45 db range of intensity. By use of word-pattern matching rather than phoneme matching, and with the restriction of a small vocabulary, a classification accuracy of 90 percent was attained. However, it was concluded that a finer frequency resolution is required between 300 and 1200 Hz to detect sudden movements in energy concentrations.

Many other reports (15-55) confirm the apparent fact that a much more sophisticated level of understanding of speech and hearing functions must be attained before major breakthroughs in automatic speech recognition can be brought about. From this viewpoint, it may be inferred that an audio filter bank having advanced capability in both the frequency and time domains would be highly desirable for future studies. The purpose of this paper is to report the structure and performance of a bank of 97 tuned amplitude trackers which cover the

100-to-6400 Hz frequency range with uniform-Q resolution and maximum response speed.

The Composite System

The basic scheme of the phoneme analyzer system is shown in Figure 1. The input is a speech phoneme and the output is a visual display of its amplitude spectrum. The microphone (Shure Model 545S) has a cardiod directional pattern and a transfer characteristic which in the frequency range of 50 to 15,000 Hz is flat within $\pm$ 3 db. A sensitivity level of -151 db relative to a zero-db level of one volt per microbar permits this microphone to deliver into a 100-kilohm load a peak-to-peak signal voltage of about 4 millivolts, for speech of normal amplitude from lips positioned ten inches away. The speech amplifier is conventionally designed to have a flat passband characteristic in the 0-20 kHz frequency range, a voltage gain of 66 db, a 1.0 megohm input impedance, and a capability of producing an output signal of $\pm$ 10 volts across a 200-ohm resistive load.

As further shown in Figure 1, the output signal from the speech amplifier is used as the driving signal input to a parallel bank of 97 tuned amplitude trackers. Every tuned amplitude tracker is a series combination of a tuned audio filter, a full-wave rectifier, and a low-pass filter, each constructed by use of operational amplifiers. The tuned audio filter in the K^{th} tracker has a two-pole resonant gain of unity, a tuning sensitivity of Q = 25, a resistive input impedance in the range of 20 to 30 kilohms, and a resonant frequency F(K) in Hz given by the formula.

$$F(K) = (100) \cdot 2^{[(K - 1)/16]} .$$

The full-wave rectifier has a conversion gain of 0.233, i.e., with a sinusoidal input signal of 21.4-volt peak-to-peak amplitude the output signal is a full-wave-rectified sinusoid having a 5-volt peak amplitude and 3.18-volt average amplitude. The low-pass filter has a dc gain of 3.0 and has a flat passband characteristic which extends up to a (6 db) cutoff frequency which is equal to the bandwidth [F(K)/Q] of the corresponding tuned audio filter. Thus, for the $K^{\underline{th}}$ tuned amplitude tracker the overall conversion gain (dc output voltage divided by the peak-to-peak input voltage) is 0.45 at the resonant frequency F(K), and at this frequency F(K) a 20-volt peak-to-peak sinusoid from the speech amplifier will produce a 9-volt dc output signal from the $K^{\underline{th}}$ tracker. Further, each of the trackers will respond proportionally and almost instantaneously to any amplitude fluctuations in its resonating input sinusoid, if the frequency of these input amplitude fluctuations is not appreciably greater than the bandwidth [F(K)/Q] of the particular tracker. Finally, it is seen that the "best" frequencies (F1) and F(97) of the two trackers occupying the extreme end positions in the tracker bank are respectively 100 Hz and 6400 Hz, and that the best frequency of tracker (K + 16) is precisely twice the best frequency of tracker (K). The choice of a uniform tuning-sensitivity factor of Q = 25 makes the characteristic resonance curves of neighboring trackers overlap approximately at their -3 db points.

Referring still further to Figure 1, it is seen that the 97 output terminals of the tracker bank are connected to a corresponding linear array of input terminals of an analog multiplexer. This multiplexer (EECO 765-1 Multiplexer) has a 10-kilohm input impedance and a maximum input voltage of $\pm$ 5 volts for each analog input channel, an analog output voltage of $\pm$ 5 volts across a minimum load resistance of 250 ohms, and a maximum channel-stepping rate of 50,000 channels per second. A simple 50 kHz multivibrator clock is used to advance the multiplexer channel address continuously through repetitive sweeps of all 97 channels. At the end of each sweep an "end frame" pulse appears at a separate sync-output terminal of the multiplexer, thereby enabling the short-term 100-to-6400 Hz spectrum of the microphone signal to be displayed continuously on the CRT oscilloscope.

With Q = 25 for every tuned filter, and with

$$\frac{1}{2}\ \frac{F(K+1)+F(K)}{F(K+1)-F(K)} = \frac{1}{2}\,\mathrm{Coth}\,\frac{\mathrm{Ln}\,2}{32} = 23$$

for every K in the range of $1 \leq K \leq 96$, it is evident that the 97 outputs of the tracker bank comprise, in the frequency domain, a set of 97 essentially independent measures of the audio signal. It is also evident that with Q = 25 the damped oscillatory response of each tuned filter to a pulse input decays to $(100/e^4) = 1.8$ percent of its initial amplitude with the passage of $4Q/\pi = 32$ oscillatory cycles. With the $K^{\underline{th}}$ tracker thus being capable of producing successive independent measures in the time domain at a rate of at least $[\pi \cdot F(K)/4Q]$ samples per second, the total data rate capability of the complete tracker bank is found by summation to be equal to at least $(25\pi/Q)\cdot(1516) = 4775$ independent samples per second. Consequently, if the output voltage of each tracker is sufficiently accurate to justify 4-bit coding, the output data rate capability of the overall system shown in Figure 1 is about 20 kilobits per second.

Tuned Amplitude Tracker Design

The circuit actually used for each of the tuned amplitude trackers is shown in Figure 2. The tuned audio filter is comprised of the operational amplifier U1, the five resistances R_1 through R_5, and the tank circuit formed by the capacitance C_K' connected in parallel with the inductance L_K', which has an internal resistance r_K'. The full-wave rectifier is comprised of the operational amplifier U2, the four resistances R_6 through R_9, and the two diodes D1 and D2. The low-pass smoothing filter is comprised of the operational amplifier U3, the three resistances R_K and $0.75\ R_K$, and the two capacitances C_K and $4C_K$. The sinusoidal input voltage E_K is assumed to reside within the range of $-10 \leq E_K \leq +10$ volts. The smoothed dc voltage V_K resides within the range of $0 \leq V_K \leq 10$ volts.

The structure of the tuned audio filter is a form of Q-multiplier having a resonant frequency determined by the tuning of the r_K'-L_K'-C_K' tank circuit, a Q-enhancement established by positive feedback through the R_4-R_5 voltage divider, and a resonant gain established by the R_1-R_2-R_3 tee network. The voltage transfer function $G(\omega)$ of the K^{th} tuned audio filter is

$$G(\omega) = \frac{-A}{1 + jQ_K\left(\dfrac{\omega}{\omega_K} - \dfrac{\omega_K}{\omega}\right)} ,$$

in which ω_K is the resonant angular frequency, A is the resonant gain, and Q_K is the overall tuning sensitivity. For the case in which $R_4 = R_5$, and in which $R_p = R_1R_2/(R_1 + R_2)$, the resonant gain A is given by the formula

$$A = \frac{2R_p}{R_1 \cdot \left[\dfrac{R_3 + R_p}{(Q_K')^2 \cdot r_K'} - 1\right]},$$

and the overall tuning sensitivity Q_K is given by the formula

$$Q_K = \frac{Q_K'}{1 - \left[\dfrac{(Q_K')^2 \cdot r_K'}{R_3 + R_p}\right]},$$

in which Q_K' is the non-enhanced tuning sensitivity of the r_K'-L_K'-C_K' tank circuit, i.e., $Q_K' = \omega_K L_K'/r_K'$. If the positive input terminal of amplifier U1 is temporarily grounded, and if the resistance R_1 is temporarily made equal to zero, the resultant filter can be tuned to the resonant angular frequency by (slug) adjustment of the inductance L_K', after which the temporarily modified filter can be adjusted to have a resonant gain of unity by making $R_3 = (Q_K')^2 \cdot r_K'$. The unmodified filter will then have a resonant gain of $A = 2R_3/R_1$, which requires that $R_1 = 2R_3$ for unity gain at resonance. Thus, with the tank circuit tuned to ω_K, with $R_4 = R_5 = 33$ kilohms, with $R_3 = (Q_K')^2 \cdot r_K'$, and with $R_1 = 2R_3$, the voltage gain at resonance will be A = 1.0, and the selectivity-multiplication ratio Q_K/Q_K' is determined by the formula

$$\frac{Q_K}{Q_K'} = \frac{3}{2} + \frac{R_3}{R_2}.$$

The value of the resistance R_2 was selected so as to make $Q_K = 25$ for all $1 \leq K \leq 97$. The inductors (manufactured by the United Transformer Company) used in the tank circuits were as follows.

Frequency Range	Inductor Type	Inductance Range
100-200 Hz	HVC-7	0.50-5.0 H
200-476 Hz	HVC-6	0.20-2.0 H
476-1131 Hz	HVC-5	0.07-0.70 H
1131-2691 Hz	HVC-4	0.03-0.30 H
2691-6400 Hz	HVC-3	0.01-0.11 H

The capacitance values used in the tank circuits ranged from 0.01 to 0.50 microfarads, giving a value of about 10 kilohms for $(Q_K')^2 \cdot r_K'$. In cases where the tank selectivity was too large ($Q_K' > 12$) a 10-kilohm shunt resistance was used to desensitize the tank circuit.

In the rectifying portion of the tracker circuit, the diodes D1 and D2, the resistance $R_6 = 22$ kilohms, and the resistances $R_7 = R_8 = R_9 = 11$ kilohms, were arranged so that the output voltage of the amplifier U2 would be a negative-going full-wave rectified sinusoid of -5.0 volts peak amplitude when the input voltage from the amplifier U1 is a sine wave of 20-volt peak-to-peak amplitude. The low-pass smoothing filter has two identical poles, and has its three resistances and two capacitances chosen so that its (6 db) cutoff frequency is equal to the bandwidth $[F(K)/Q_K]$ of the associated tuned audio filter. The dc voltage gain of the low-pass filter is 3.0, and the time constant $R_K C_K$ is chosen so that $R_K C_K = Q/[6\pi \cdot F(K)]$ in order to achieve the desired bandwidth relationship.

Operational Amplifier Implementation

The operational amplifiers U1, U2, and U3 used in the circuit of Figure 2 were of the SN72-709N type, which are manufactured by Texas Instruments, Inc. This type of operational amplifier has an input current requirement of 0.2 microamp, a common-mode rejection ratio of 90 db, an input resistance of 100 to 400 kilohms, an output resistance of 150 ohms, and a power dissipation rating of 80 milliwatts when operated with a $\pm$ 15 volt power supply. However, special attention is required to prevent latch-up and to achieve adequate frequency compensation. The electrical diagram of the SN72-709N integrated circuit is shown in Figure 3. The parenthesized terminal numbers correspond to the dual-in-line pin numbers, which are arranged to increase counterclockwise from the lower left-hand corner (if the dual-in-line is viewed from the top with the indexing notch on the left).

The condition known as latch-up can occur when the common mode input voltage limit is exceeded. This is especially bothersome with low gain (substantial feedback) conditions. If the signal or a transient overloads the input, the inverting transistor saturates and effectively applies the signal to the base of the second transistor stage. Without the inverting action of the first stage, feedback is positive and latch-up occurs. The diode (D3) connected between pins 12 and 10 will limit the output so that it will rise no higher than the common mode voltage limit at the base of transistor T9, thus holding the input transistor out of saturation so that latch-up will not occur.

The type 709 operational amplifier is not frequency compensated internally. It unfortunately has several break-points in its frequency response characteristic. Its highest frequency break-point is at about 10 MHz, and a second break-point is found in the neighborhood of 1.8 to 2 MHz. This frequency characteristic produces an ultimate roll-off of 12 db/octave and a 180-degree phase shift through the uncompensated amplifier. Consequently, a feedback loop connected around the amplifier may produce regenerative oscillations, unless adequate compensation is applied to the open-loop gain characteristic. For moderate closed-loop gain, the roll-off of the open-loop gain must begin at a low enough frequency that the closed-loop gain reduces to less than unity before the amplifier produces 180 degree internal phase shift. If an open-loop roll-off characteristic approximating that of a single-section RC low-pass network is imposed on the amplifier, it will roll off at 6 db/octave and approach a maximum phase shift of 90 degrees in the frequency range of interest. Therefore, as closed-loop gain is decreased the open-loop roll-off frequency must decrease. This is accomplished by use of compensating lag networks, i.e., two external networks must be added to the basic 709 for correct compensation. One network connects between pins 3 and 12 and the other connects between pins 9 and 10.

Pins 3 and 12 are the "input frequency compensation" points. An R_i-C_i step-type of lag network is connected between these pins. It is generally desirable that this network maintain a phase lag for about three decades in frequency. The roll-off or corner frequency of this lag network must be chosen to be low enough to meet the desired closed-loop gain requirement. The internal amplifier impedance between pins 3 and 12 is of the order of 1.0 to 2.0 megohms. The resistor in the lag loop should be about 0.001 times this value for the three-decade step. Once the break-point frequency (f_i) is established in terms of the desired closed-loop gain, the compensating network capacitance is determined by the formula $C_i = [2\pi f_i R_i]^{-1}$. The "output frequency compensation" network is connected between pins 9 and 10. The internal amplifier impedance at these pins is of the order of R_x = 35 to 45 kilohms. A capacitor C_0 connected between these pins introduces a second lag in the amplifier response. The break-point frequency $f_x = 1000\ f_i$ for this lag network is chosen so that it supplements the first correction network after about three decades in frequency. The capacitor connected between pins 9 and 10 can then be evaluated from the formula $C_0 = [2\pi f_x R_x]^{-1}$. Thus, the values of the compensation elements actually used were R_i = 1500 ohms in every circuit, C_i = 0.01 microfarad and C_0 = 40 picofarads in the circuits operating at frequencies below 3 kHz, and C_i = 0.001 microfarad and C_0 = 10 picofarads in the circuits operating at frequencies above 3 kHz. Use of the newer (and more expensive) operational amplifiers containing internal frequency compensation networks would, of course, simplify the system construction somewhat.

Hardware Geometry

Photographs of the completed bank of 97 tuned amplitude trackers are shown in Figure 4. One chassis contains all of the 97 tuned audio filters, plus their common audio driving amplifier. A second chassis contains the 97 corresponding rectification-and-smoothing networks (plus a speech amplitude detector, needed when threshold triggering of external circuits is desired). Each chassis is of conventional 17" x 17" x 3" size to fit a 19-inch equipment rack. The two assemblies are interconnected by means of two pairs of 100-contact plugs and sockets (Cinch-Jones type 37-M100L and 37-F100L) and a corresponding 100-wire cable. The first chassis also includes receptacles for the audio signal input, the $\pm$ 15 volt power inputs, and a ground wire. The second chassis includes the input receptacles for the ground wire and the $\pm$ 15 volt power inputs, and the output fittings for the 100-wire outputs to the sampling switch, and the output of the speech amplitude detector.

The individual components are mounted on a standard 12-pin (Vector) 3" x 4.5" circuit boards. Photographs showing the detailed arrangement of the components on the circuit boards are shown in Figure 5. In general, three tuned audio filters having neighboring resonant frequencies are mounted on one circuit board, and their three corresponding rectification-and-smoothing networks are mounted on another circuit board. For tuning purposes the three slug-adjustable inductors are mounted at the top of the "filter board." On the "rectification board" the three rectification-and-smoothing networks are horizontally arranged along the top, middle, and lower thirds of the board. Thus, a total of 68 circuit boards were used to mount all of the components of the complete (two-chassis) bank of 97 tuned amplitude trackers. The total number of components used included 2720 resistors, 850 capacitors, 400 diodes, 291 operational amplifiers, and 97 inductors. The total current drain required of the (Hewlett-Packard 6523A) power supply was found to be $\pm$ 255 mA at $\pm$ 15 volts for the "filter chassis," and $\pm$ 470 mA at $\pm$ 15 volts for the "rectifier chassis," giving a total power dissipation of 22 watts for the complete tracker bank.

Component Performance

From the previously described design considerations, it can be seen that some rather severe tolerance requirements are imposed on the design of each tuned amplitude tracker, particularly with respect to resonant frequency and bandwidth. It was found that very careful adjustments were necessary to achieve the correct resistance values in each input tee network.

As an example of the resultant performance of each tuned amplitude tracker, Figure 6 shows the measured output voltage of the 800-Hz tuned audio filter (K = 49) as a function of excitation frequency. Figure 7 shows the output voltage of the corresponding smoothing filter as a function of its input ripple frequency. By use of the CRT Lissajous technique, the tuning error in the resonant frequency of each amplitude tracker was made to be less than $\pm$ 2.0 Hz. The error in the resonant gain was found to be less than $\pm$ 2.0 percent, and the error in the bandwidth was found to be less than $\pm$ 5.0 percent.

The response of the 800-Hz tuned audio filter to a suddenly applied 800-Hz sinusoidal input of 20-volt peak-to-peak amplitude is shown in Figure 8. The exponentially controlled build-up of the resonator output voltage is clearly seen. A Laplace transform analysis of the resonator shows that the envelope of the increasing peak-to-peak output amplitude should rise to 63 percent of its final value with the passage of $n = Q/\pi = 8$ cycles of the input waveform, which corresponds to a time lapse of 10.0 millisec in the case of the 800 Hz resonator.

Figures 9a and 9b show the response of the 800-Hz tuned amplitude tracker to an input 800 Hz sine wave which is 100-percent amplitude modulated by a 32-Hz sine wave. The maximum peak-to-peak amplitude of the input signal is 20 volts. The 64 Hz spacing between the sidebands is twice the bandwidth of the 800 Hz filter. The attenuation and phase shift of the sidebands caused by the filter selectivity is shown in the lower trace of Figure 9a. The upper trace in Figure 9b shows the output of the full-wave rectifier, with its zero-volt ground reference being the straight-line upper-limit level. The bottom horizontal trace in Figure 9b shows the zero-volt ground reference level for the smoothing filter output. The center trace in Figure 9b shows the output of the smoothing filter, which introduces some further attenuation and phase shift in the modulation. Thus, it is seen that the smoothing filter performs its intended function of filtering out the rectification ripples while still retaining the capability of responding to fluctuations in the input signal amplitude.

System Performance

When the complete phoneme analyzer system is assembled as previously described with reference to Figure 1, its performance can be tested with a variety of audio input signals. For the results reported here, the sweep time of the CRT oscilloscope was first made approxmately equal to 2.0 msec, and then further adjusted so that the time for one complete sweep through all 97 channels of the multiplexer corresponded to a horizontal distance of 9.7 cm on the CRT face.

Figure 10a shows a photograph of the CRT display when the audio input signal is a steady 800 Hz sine wave of 16-volt peak-to-peak amplitude. As expected, the strongest response (4.1 volts output) is produced by the 800 Hz tracker. Since K = 49 for this tracker, its sampled output appears near the center of the CRT screen. The responses of the two neighboring trackers (K = 48 and K = 50) are each seen to be of about 1.7-volt output amplitude. This corresponds to an adjacent-channel rejection of 7.6 db. The responses of the next pair of adjacent filters (K = 47 and K = 51) are seen to be about 14.2 db below the 800 Hz filter response. The more distant neighbors respond with even less amplitude, and practically no response occurs for those trackers more removed than the K = 42 and K = 56 pair.

For pure tone inputs the general shape of the spectral profile was found to be almost independent of the excitation frequency; the exception being that when the excitation frequency is mid-way between the resonant frequencies of two adjacent trackers the equal responses of the two adjacent trackers are somewhat (about 3 db) less than the peak response of a resonant tracker. This is illustrated in Figure 10b, which shows the spectral profile when the audio input signal is a steady 816 Hz sine wave of 16-volt peak-to-peak amplitude. Thus, as the amplitude of the input sine wave is held constant and the excitation frequency is gradually increased from the lower extreme of 100 Hz to the upper extreme of 6400 Hz, the spectral profile illustrated by Figures 10a and 10b will move rightward from the left edge to the right edge of the CRT face.

Typical responses of the complete speech analyzer system for prolonged phoneme inputs are shown in Figures 11 and 12. Figure 11a shows the speech waveform for the sound [i], voiced by a 30-year-old male person having a natural glottal pulse rate of 141 Hz. Figure 11b shows the resultant spectral profile produced by the system. The fundamental 141 Hz frequency component is correctly displayed by the K = 9 tracker. Prominent harmonic peaks are seen at 283 Hz and 436 Hz, and a distribution of higher harmonics is found between 2263 Hz and 6400 Hz. Figure 12a shows the speech waveform for the sound [ae], voiced by the same 30-year-old male person having a natural glottal pulse rate of

141 Hz. Figure 12b shows the resultant spectral profile produced by the system. The fundamental 141 Hz frequency component is again correctly displayed by the K = 9 tracker. Prominent harmonic peaks are seen at 283 Hz and 423 Hz, and a wide distribution of higher harmonics extends from 519 Hz to 6400 Hz.

Numerous other experiments using various carefully recorded speech sounds as inputs have been conducted with the above described phoneme analyzer system. The preliminary results have already shown that the harmonic structure of any prolonged phoneme changes rather drastically when the glottal pluse rate is changed. It appears possible that the historically accepted "formant" method of representing speech spectra may be open to question.

Acknowledgements

In pursuing the work reported here the authors wish to acknowledge the financial support of the College of Engineering at Oklahoma State University. Sincere thanks are expressed to Dr. W. L. Hughes, Head, School of Electrical Engineering, and to Dr. M. R. Lohmann, Dean, College of Engineering, for making available the necessary experimental facilities and equipment. The diligent secretarial help of Mrs. Cynthia Schroer is also greatly appreciated.

References

1. Lindgren, Nilo. "Machine Recognition of Human Language: Part I--Automatic Speech Recognition", IEEE Spectrum, pp. 114-135, March 1965).

2. Lindgren, Nilo. "Machine Recognition of Human Language: Part II--Theoretical Models of Speech Perception and Language", IEEE Spectrum, pp. 45-69, April 1965.

3. Flanagan, James L. "Models for Approximating Basilar Membrane Displacement - Parts I and II", Bell System Technical Journal, 39: 1163-1191, (September 1960) and 41: 959-1009, (May 1962).

4. Glaesser, E.; Caldwell, W. F.; and Stewart, J. L. "An Electronic Analog of the Ear", Technical Documentary Report No. AMRL-TDR-63-60, Biophysics Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, June 1963 - prepared under contract No. AF33(657)-113 by Santa Rita Technology, Inc., Menlo Park, California.

5. Bolie, Victor W. "Computer Aided Refinements in the Design of an Artificial Cochlea", SWIEEECO Record of Technical Papers, (IEEE Catalog No. 70C5-SWIECO), pp. 308-311, Dallas, Texas, April 1970.

6. Caldwell, William F. "Recognition of Sounds by Cochlear Patterns", IEEE Transactions on Military Electronics, MIL-7(2,3): 179-185, (April-July 1963).

7. Guttman, Newman, and Flanagan, J. L. "Pitch of High-Pass Filtered Pulse Trains", Journal of the Acoustical Society of America, 36 (4): 757-765, April 1964.

8. Mundie, J. Ryland and Moore, Thomas J. "Speech Analysis as the Ear Sees It: Oral Topography", Technical Report AMRL-TR-67-176, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, 1967.

9. Campanella, S. Joseph and Phyfe, Douglas. "Application of a Model of the Analog Ear to Speech Signal Analysis", IEEE Transactions on Audio and Electroacoustics, AU-16(1): 26-35, March 1968.

10. Cannon, Mark W. "A Method of Analysis and Recognition for Voiced Vowels", IEEE Transactions on Audio and Electroacoustics, AU-16 (2): 154-158, June 1968.

11. Hyde, S. R. "Recognition of Speech by Machine", Post Office Electrical Engineers Journal, pp. 100-104, July 1969.

12. Gilmour, W. D. "Electronic Speech Recognition", Wireless World, 75(1400): 76-80, 1969.

13. Hughes, George W. and Hemdal, John F. "Speech Analysis", Final Report, Task 562802, Project 5628, AF 19(628)-305, Purdue University, School of Electrical Engineering, TR-EE65-9, AFCRL-65-681, July 1, 1965.

14. Bobrow, Daniel G. and Klatt, Dennis H. "A Limited Speech Recognition System", Final Report, Contract No. NAS 12-138 (Bolt Beranek and Newman, Inc., Cambridge, Mass.), National Aeronautics and Space Administration, May 15, 1968.

15. Tease, Donald C.; Eldredge, Donald H.; and Davis, Hallowell. "Cochlear Response to Acoustic Transients: An Interpretation of Whole-Nerve Action Potentials", Journal of the Acoustical Society of America, 34(8): 1438-1459, September 1962.

16. Bekesy, George von. "Hearing Theories and Complex Sounds", Journal of the Acoustical Society of America, 35(4): 588-601, April 1963.

17. Moller, Aagi R. "Transfer Functions of the Middle Ear", Journal of the Acoustical Society of America, 35(10): 1526-1534, October 1963.

18. Weiss, Thomas F. "A Model of the Peripheral Auditory System", Kybernetik, Band III/Heft 4, pp. 153-175, November 1966.

19. Glattke, Theodore J. "Unit Responses of the Cat Cochlear Nucleus to Amplitude-Modulated Stimuli", Journal of the Acoustical Society of America, 45(2): 419-425, February 1969.

20. Carterette, Edward D.; Friedman, Morton P.; and Lovell, John D. "Mach Bands in Hearing", Journal of the Acoustical Society of America, 45(4): 986-998, April 1969.

21. Sachs, Murray B. "Stimulus-Response Relation for Auditory-Nerve Fibers: Two-Tone Stimuli", Journal of the Acoustical Society of America, 45(4): 1025-1036, April 1969.

22. Dallos, Peter. "Comments on the Differential-Electrode Technique", Journal of the Acoustical Society of America, 45(4): 999-1007, April 1969.

23. Dallos, P.; Schoeny, Z. G.; Worthington, D. W.; and Cheatham, M. A. "Some Problems in the Measurement of Cochlear Distortion", Journal of the Acoustical Society of America, 46(2): 356-361, August 1969.

24. Honrubia, Vicente and Ward, Paul H. "Dependence of the Cochlear Microphonics and the Summating Potential on the Endocochlear Potential", Journal of the Acoustical Society of America, 46(2): 388-392, August 1969.

25. Plomp, R. and Steeneken, H. J. M. "Effect of Phase on the Timbre of Complex Tones", Journal of the Acoustical Society of America, 46(2): 409-421, August 1969.

26. Goblick, Thomas J. and Pfeiffer, Russell R. "Time-Domain Measurements of Cochlear Nonlinearities Using Combination Click Stimuli", Journal of the Acoustical Society of America, 46(4): 924-938, October 1969.

27. Watson, Charles S. and Gengel, Roy W. "Signal Duration and Signal Frequency in Relation to Auditory Sensitivity", Journal of the Acoustical Society of America, 46(4): 989-996, October 1969.

28. Zerlin, Stanley. "Travelling Wave Velocity in the Human Cochlea", Journal of the Acoustical Society of America, 46(4): 1011-1015, October 1969.

29. Lynn, P. A. and Sayers, B. McA. "Cochlear Innervation, Signal Processing, and their Relation to Auditory Time-Intensity Effects", Journal of the Acoustical Society of America, 47(2-2): 525-533, Febraury 1970.

30. Teacher, Charles F.; Kellett, Henry G.; and Focht, Louis R. "Experimental Limited Vocabulary Speech Recognizer", IEEE Transactions on Audio and Electroacoustics, AU-15(3): 127-130, September 1967.

31. Ross, Paul W. "A Limited-Vocabulary Adaptive Speech-Recognition System", Audio Engineering Society, 15(4): 414-418, October 1967.

32. Purton, Raymond F. "Speech Recognition Using Autocorrelation Analysis", IEEE Transactions on Audio and Electroacoustics, AU-16(2): 235-239, June 1968.

33. Shearme, John N. and Leach, Peter F. "Some Experiments with a Simple Word Recognition System", IEEE Transactions on Audio and Electroacoustics, AU-16(2): 256-261, June 1968.

34. Ewing, G. D. and Taylor, John F. "Computer Recognition of Speech Using Zero-Crossing Information", IEEE Transactions on Audio and Electroacoustics, AU-17(1): 37-40, March 1969.

35. Pierce, J. R. "Whither Speech Recognition", Journal of the Acoustical Society of America, 46(4): 1049-1051, October 1969.

36. Olson, Harry F.; Belar, Herbert; and Rogers, Edward S. "Speech Processing Techniques and Applications", IEEE Transactions on Audio and Electroacoustics, AU-15(3): 120-126, September 1967.

37. Lea, Wayne A. "Establishing the Value of Voice Communication with Computers", IEEE Transactions on Audio and Electroacoustics, AU-16(2): 184-197, June 1968.

38. Oppenheim, Alan V. "Speech Analysis-Synthesis System Based on Homomorphic Filtering", Journal of the Acoustical Society of America, 45(2): 458-465, February 1969.

39. Flanagan, James L. "Use of an Interactive Laboratory Computer to Study an Acoustic-Oscillator Model of the Vocal Cords", IEEE Transactions on Audio and Electroacoustics, AU-17(1): 2-6, March 1969.

40. Flanagan, J. L. and Cherry, L. "Excitation of Vocal-Tract Synthesizers", Journal of the Acoustical Society of America, 45(3): 764-769, March 1969.

41. O'Neill, Lawrence A. "The Representation of Continuous Speech with a Periodically Sampled Orthogonal Basis", IEEE Transactions on Audio and Electroacoustics, AU-17(1): 14-21, March 1969.

42. Kent, R. D. and Moll, K. L. "Vocal-Tract Characteristics of the Stop Cognates", Journal of the Acoustical Society of America, 6(2): 1549-1555, December 1969.

43. Smith, Caldwell P. "Perception of Vocoder Speech Processed by Pattern Matching", Journal of the Acoustical Society of America, 46(6): 1562-1571, December 1969.

44. Miller, J. D.; Ross, P. W.; and Wine, C. M. "An Adaptive Speech Recognition System Operating in a Remote Time-Shared Computer Environment", IEEE Transactions on Audio and Electroacoustics, AU-18(1): 26-32, March 1970.

45. Li, K. P.; Hughes, G. W.; and House, A. S. "Correlation Characteristics and Dimensionality of Speech Spectra", Journal of the Acoustical Society of America, 46(4): 1019-1025, October 1969.

46. Pollack, Irwin. "Depth of Sequential Auditory Information Processing", Journal of the Acoustical Society of America, 46(4): 952-964, October 1969.

47. Beninghof, William J. and Ross, Myron J. "Investigation of an Efficient Representation of Speech Spectra for Segmentation and Classification of Speech Sounds", IEEE Transactions on Audio and Electroacoustics, AU-18(1): 33-42, March 1970.

48. Scarr, R. W. A. "Normalization and Adaption of Speech Data for Automatic Speech Recognition", International Journal of Man-Machine Studies, 2(1): 41-59, January 1970.

49. Sitton, G. A. "Acoustic Segmentation of Speech", International Journal of Man-Machine Studies, 2(1): 61-102, January 1970.

50. Lea, Wayne A. "Towards Versatile Speech Communication with Computers", International Journal of Man-Machine Studies, 2(2): 107-155, April 1970.

51. Bezdel, W. "Some Problems in Man-Machine Communication Using Speech", International Journal of Man-Machine Studies, 2(2): 157-168, April 1970.

52. Martin, T. B. "Acoustic Recognition of a Limited Vocabulary in Continuous Speech", Ph.D. Thesis, University of Pennsylvania, 1970.

53. Velichko, V. M. and Zagoruyko, N. G. "Automatic Recognition of 200 Words", International Journal of Man-Machine Studies, 2(3): 223-234, July 1970.

54. Ainsworth, W. A. "Estimation of Speech Synthesizer Parameters from Acoustic Waveforms", International Journal of Man-Machine Studies, 2(3): 291-302, July 1970.

55. Tatham, Marcel A. A. "Speech Synthesis--A Critical Review of the State of the Art", International Journal of Man-Machine Studies, 2(3): 303-308, July 1970.

LIST OF CAPTIONS FOR PHOTOGRAPHS

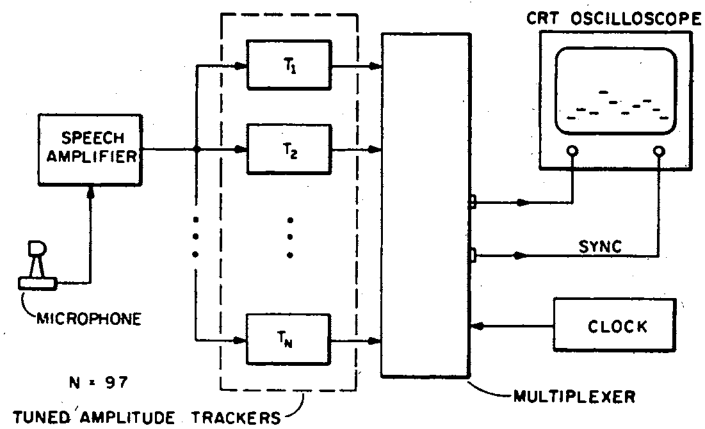


Figure 1. Block Diagram of Speech Analyzer

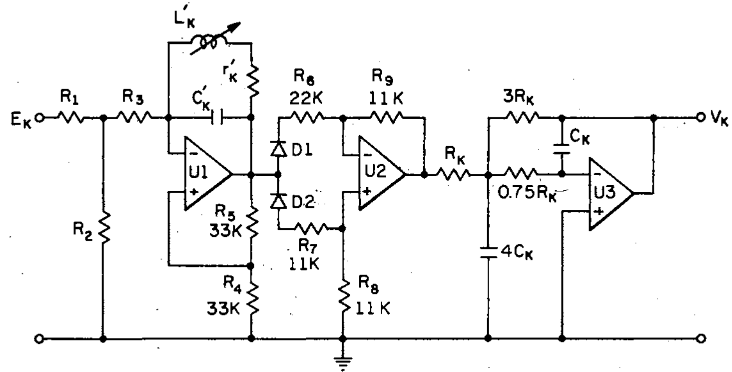


Figure 2. Tuned Amplitude Tracker Circuit

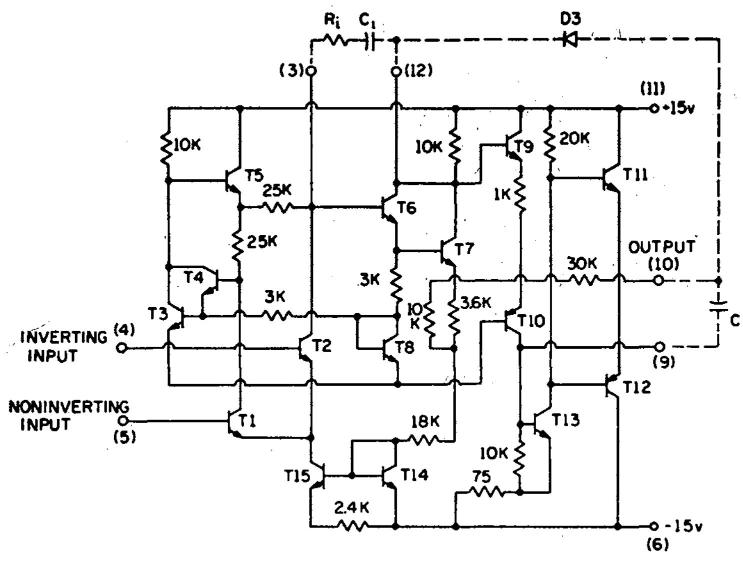


Figure 3. The SN72-709N Operational Amplifier

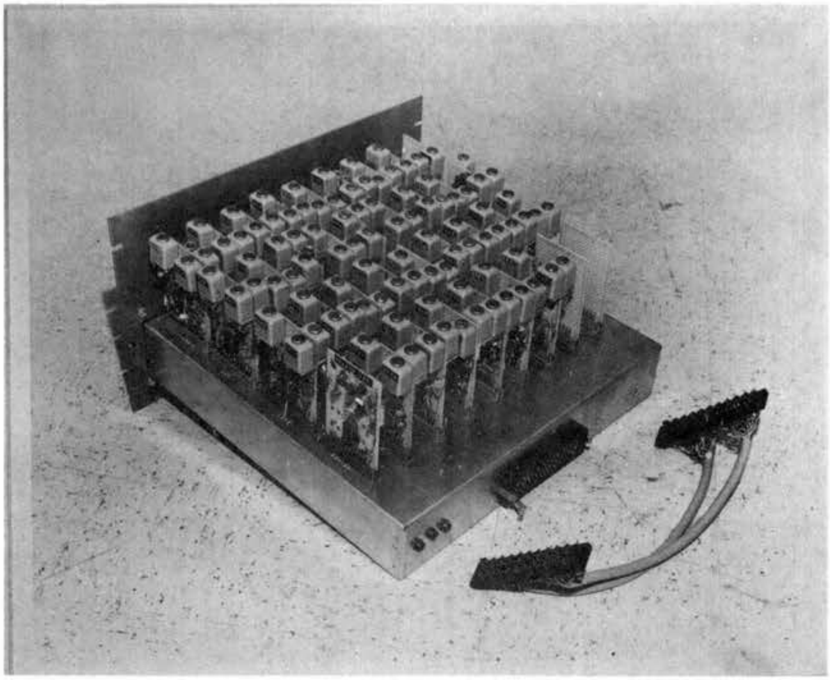


Fig. 4a. Filter Chassis Layout

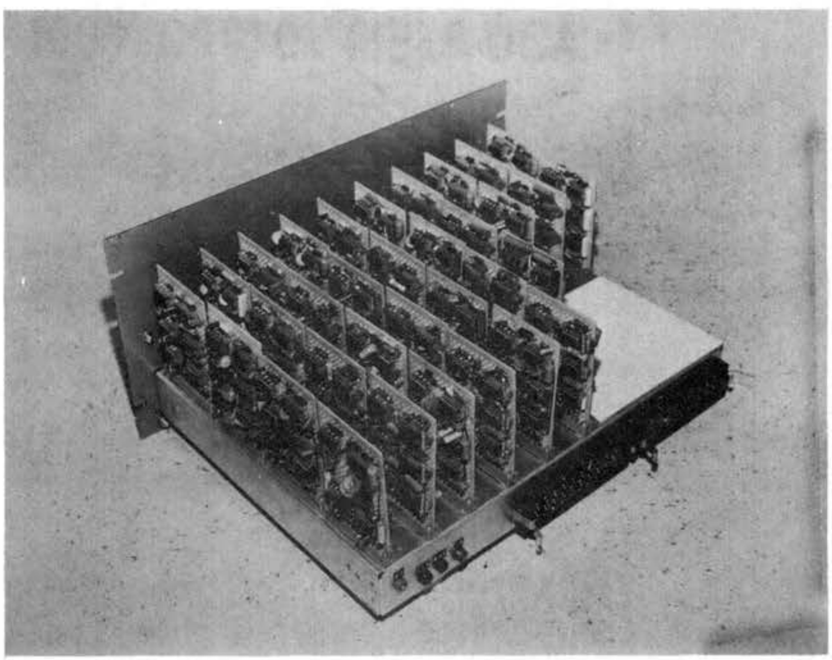


Fig. 4b. Rectifier Chassis Layout

Figure 4. Chassis Arrangement of Circuit Components

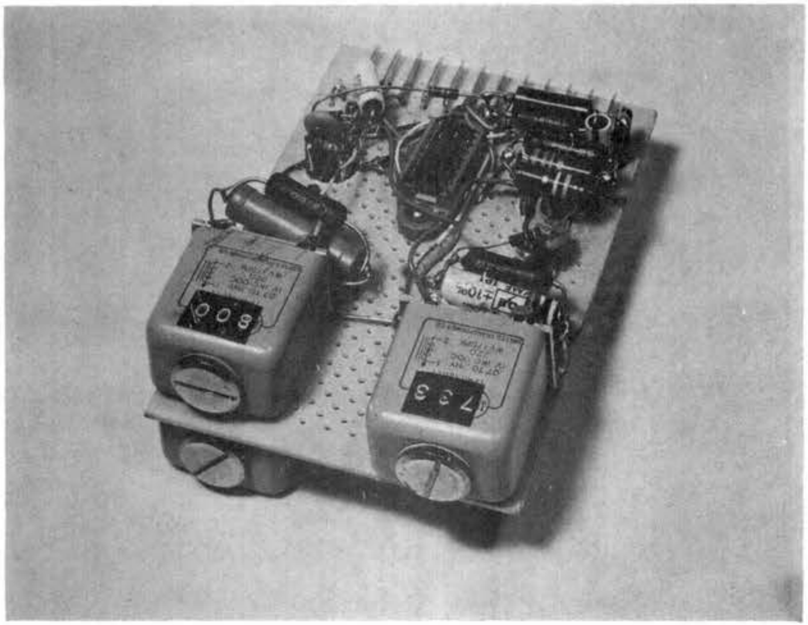


Fig. 5a. Bandpass Filter Circuit Board

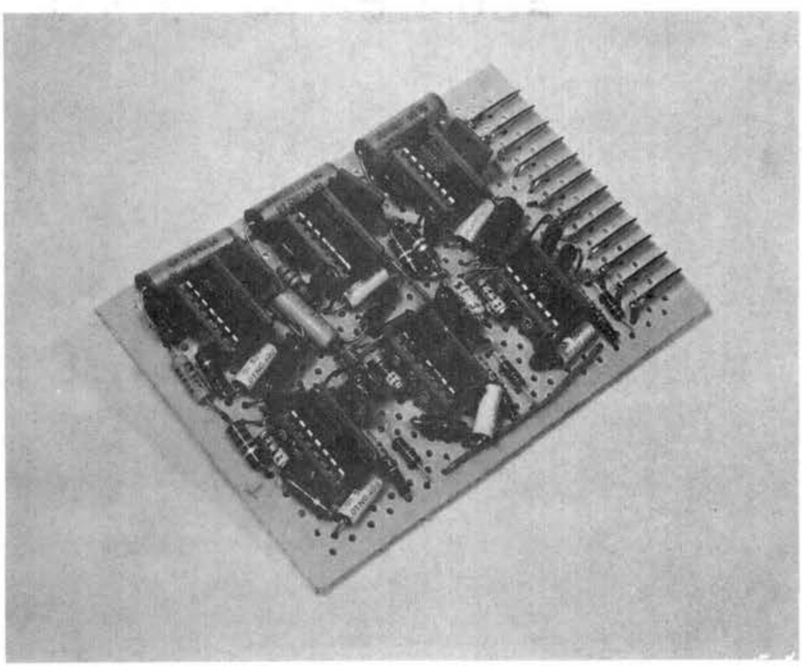


Fig. 5b. Rectifier Circuit Board

Figure 5. Component Layouts on Circuit Boards

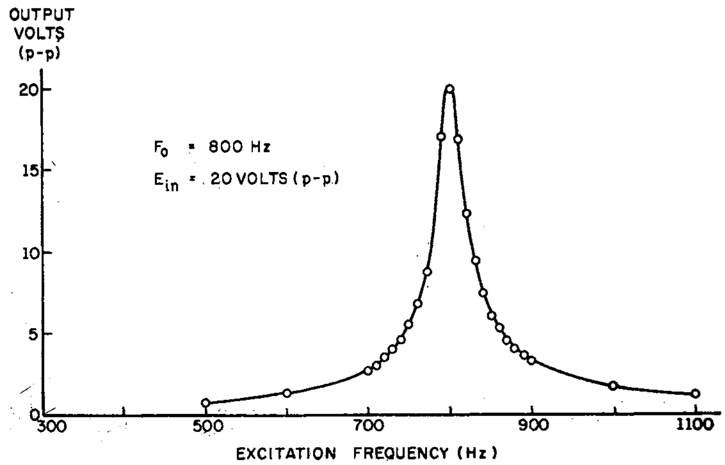


Figure 6. Tuned Filter Frequency Response

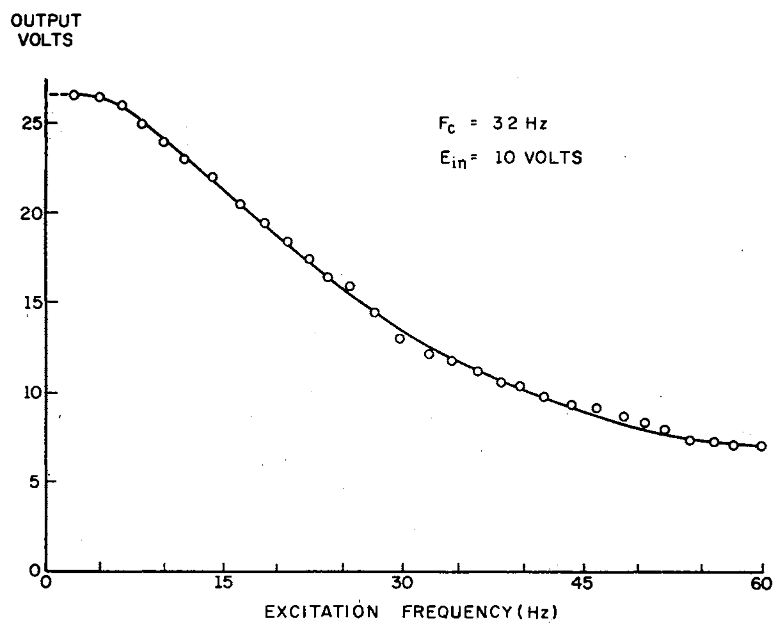
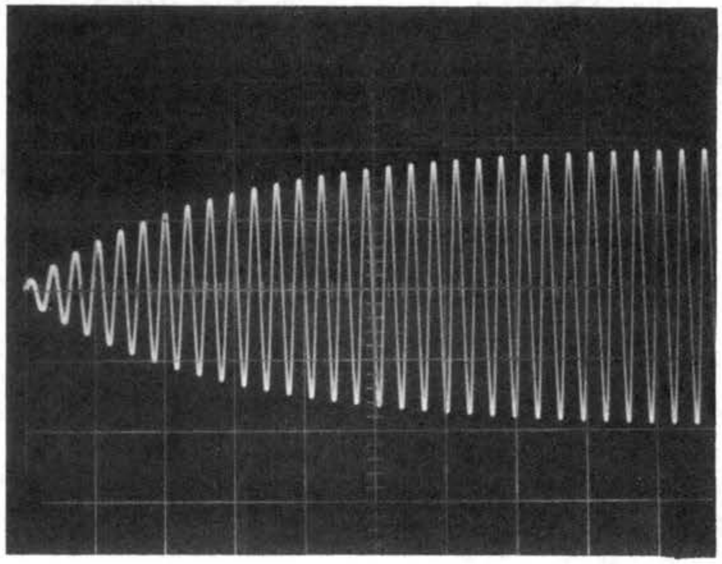


Figure 7. Low-Pass Filter Frequency Response



Horizontal scale: Time, 3.9 msec cm^{-1}.
Vertical scale: Resonator Output, 5.0 volt cm^{-1}.
Resonator input: 20 V p-p, 800 Hz, suddenly applied

Figure 8. Transient Ring-Up of Tuned Audio Filter

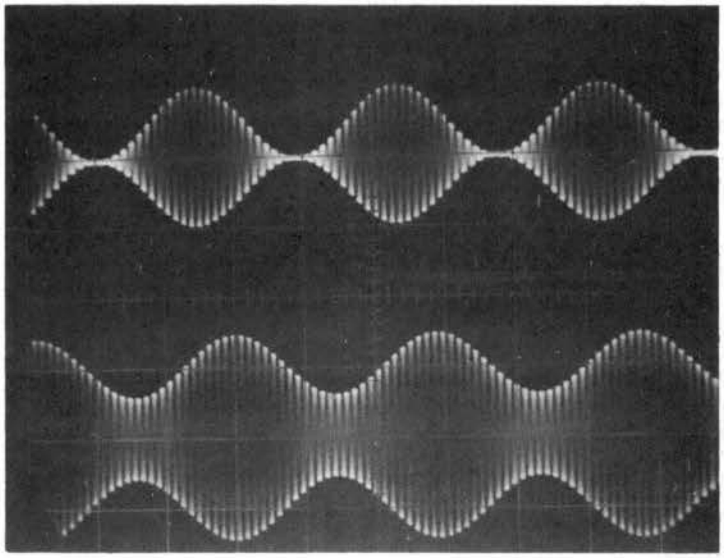


Fig. 9a. Signal Input and Filter Output
Horizontal scale: 10 msec cm^{-1}.
Vertical scale: 10 and 5 volt cm^{-1}.

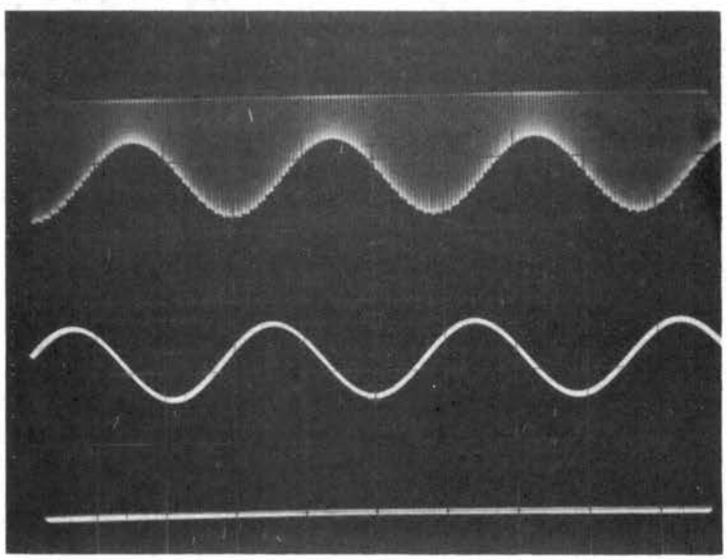


Fig. 9b. Outputs of Rectifier and Smoothing Filter
Horizontal scale: 10 msec cm^{-1}.
Vertical scale: 2 and 1 volt cm^{-1}.

Figure 9. Response of 800-Hz Tracker to AM Input

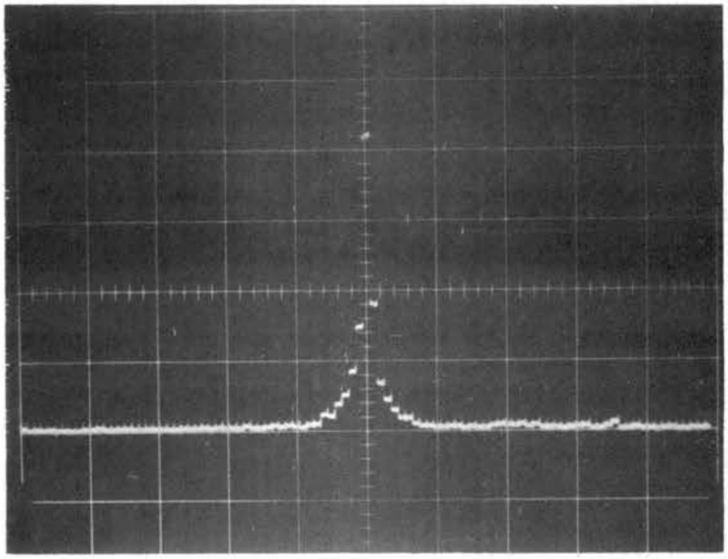


Fig. 10a. Spectro Profile for 800 Hz Input

Horizontal scale: 10 channels/cm.
Vertical scale: 1.0 volt/cm.
Audio input: Sine wave, 800 Hz, 16 V p-p.

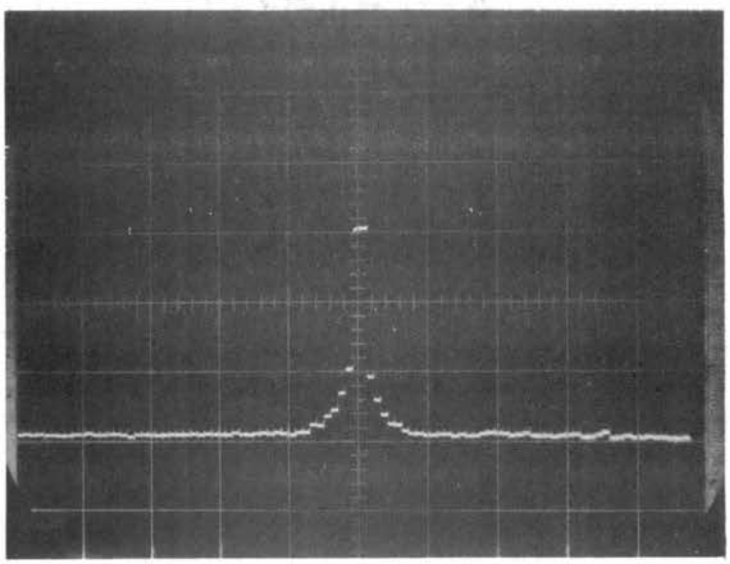


Fig. 10b. Spectro Profile for 816 Hz Input

Horizontal scale: 10 channels/cm.
Vertical scale: 1.0 volt/cm.
Audio input: Sine wave, 816 Hz, 16 V p-p.

Figure 10. System Display for 800 Hz and 816 Hz **Inputs**

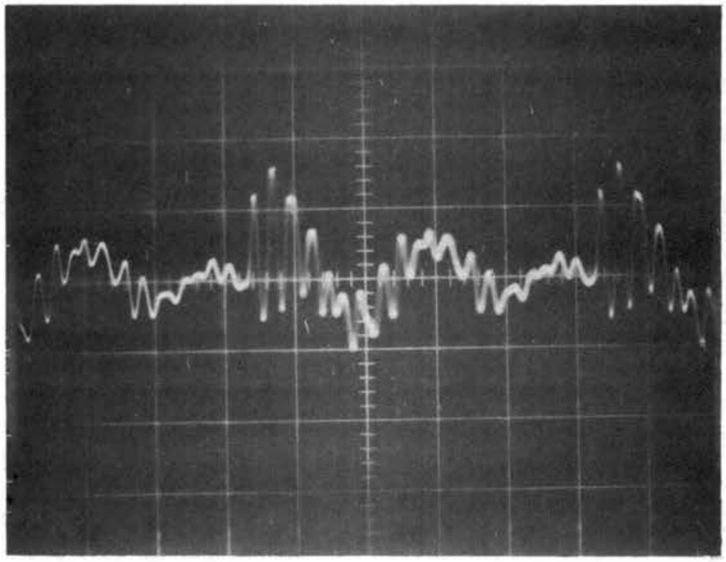


Fig. 11a. Speech Waveform for "ee" as in "beet"
Horizontal scale: 1.4 msec cm^{-1}.
Vertical scale: 1.0 volt cm^{-1}.

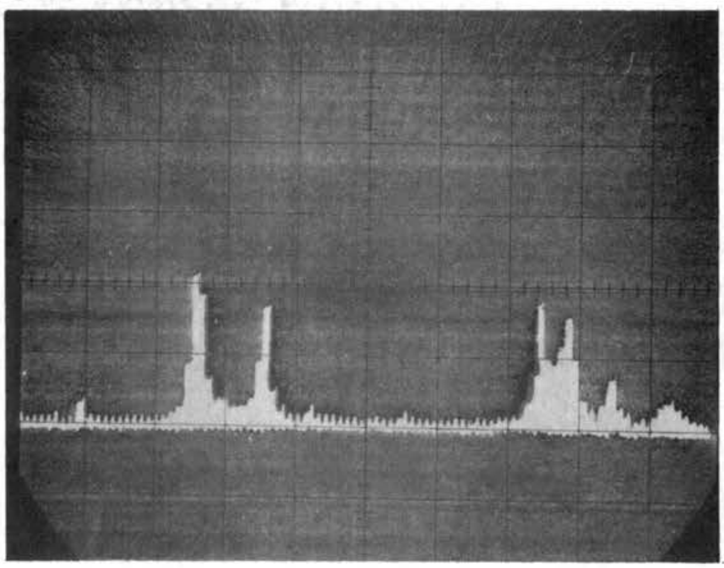


Fig. 11b. Spectral Profile for "ee" as in "beet"
Horizontal scale: 10 channels cm^{-1}.
Vertical scale: 1.0 volt cm^{-1}.

Figure 11. System Input and Output for First Phoneme

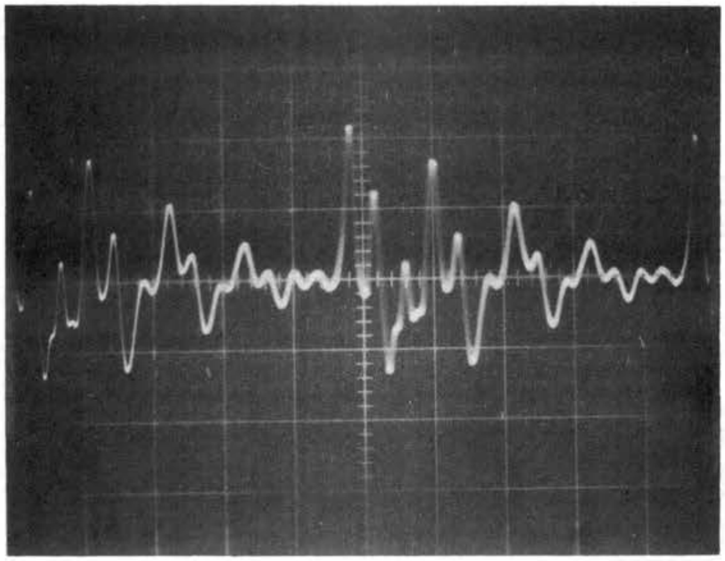


Fig. 12a. Speech Waveform for "aa" as in "bat"
Horizontal scale: 1.4 msec cm^{-1}.
Vertical scale: 1.0 volt cm^{-1}.

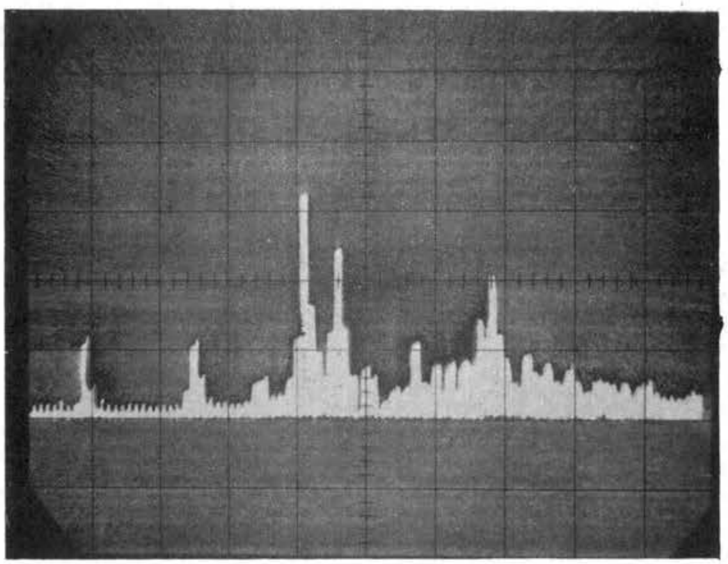


Fig. 12b. Spectral Profile for "aa" as in "bat"
Horizontal scale: 10 channels cm^{-1}.
Vertical scale: 1.0 volt cm^{-1}.

Figure 12. System Input and Output for Second Phoneme

APPENDIX C

EQUIPMENT SPECIFICATIONS

This appendix contains the specifications for the commerical equipment and major components used in this research. These are as follows:

1. Shure 545S Microphone;
2. Operational Amplifier SN72 709N;
3. United Transformer Company HVC Inductors;
4. EECO 765 Multiplexer;
5. RM 564 Tektronix Storage Oscilloscope; and
6. Burr-Brown Sample and Hold/Switched Integrator Model 1663.

SHURE

222 HARTREY AVE • EVANSTON, ILL. (60204) U.S.A.

MICROPHONES AND ELECTRONIC COMPONENTS

AREA CODE 312/328-9000 • CABLE SHUREMICRO

MODELS 545 AND 545S UNIDYNE III

UNIDIRECTIONAL DYNAMIC MICROPHONES

The Model 545 Series Unidyne III Microphones are slender dynamic microphones built to provide wide range reproduction of music and voice, and have an exceptionally uniform and effective unidirectional pickup pattern.

The Models 545-Gold and 545S-Gold are identical to Models 545 and 545S respectively except Models 545-Gold and 545S-Gold have gold finish.

These microphones are particularly suitable for high quality theatre-stage sound systems, recording, cathedrals and churches, and other critical public address systems such as those used in political conventions and legislatures, hotels, stadiums, and public auditoriums.

The microphones feature:

- Unusually effective cardioid pickup pattern. Eliminates feedback (annoying loudspeaker "squeals"). In addition, they prevent echoing (boominess) that sometimes occurs in partially-filled halls. These microphones can also be used closer to loudspeakers than usual, without creating feedback problems.
- Response especially effective for announcing, narration, vocal music, and combo groups.
- Cartridge shock mounted for quiet operation.
- A strong detachable cable especially selected for good shielding from "hum" pickup.
- Dependability and ruggedness under all operating conditions.

The Model 545 Series Microphones are dual impedance for connection into a 50 to 250 ohm line or a high impedance input.

The low impedance connection is recommended where long cable lengths are required or under conditions of severe hum disturbance. The permissible cable length is practically unlimited, since neither response nor level is appreciably affected. For use with high impedance amplifiers, Shure Model A95A Line Matching Transformer is available for coupling the low impedance line to the amplifier input. The Shure Model A95A transformer permits coupling a 50-250 ohm line to the high impedance input.

Furnished Accessories

Swivel Adapter
(for Model 545)....................A25B

Optional Accessories

Line Matching Transformer......A95A
Vibration-Isolation Stand............S39A
Desk Stand
(for Model 545) S33B
(for Model 545S) S36A
Quick Disconnect Isolation Unit
(for Model 545) A45
(for Model 545S) A47
Windscreen Assembly................A2WS

GUARANTEE:

Each microphone is guaranteed to be free from electrical and mechanical defects for a period of one year from date of shipment from factory, provided all instructions are complied with fully. In case of damage, return the microphone to the factory for repairs. Our guarantee is voided if the microphone is subjected to accident or abuse.

Replacement Components

Model R45 Dynamic Replacement CARTRIDGE
Model C56 Cable and Plug Assembly
Model 55A38 Replacement SWITCH (545S only)

Important: Shure Microphone Cables are selected after exhaustive tests to insure superior performance in microphones because of low capacities, superior shielding properties and unusually long life under severe use.

Cables with plastic insulation should not be subjected to excessive soldering-iron heat. Carefully clean and tin the conductors and the connections to which the conductors are to be soldered. The soldering operation can then be done with a minimum of heat, thereby avoiding any possibility of damage to the cable.

INSTALLATION AND CONNECTIONS OF SWITCH

A. To install the 55A38 Replacement Switch in the Shure Model 545S (see Figure A) proceed as follows:

1. Remove the two No. 2-56 screws holding the nameplate and cover to the connector assembly.
2. Remove the nameplate and take the switch out of the switch cover on the connector assembly.
3. Unsolder leads from old switch terminals.
4. Connect the leads to the new replacement switch. Observe lead color and terminal arrangement as in Figure A.
5. Re-assemble switch and nameplate back into the connector assembly and fasten the No. 2-56 screws securely.

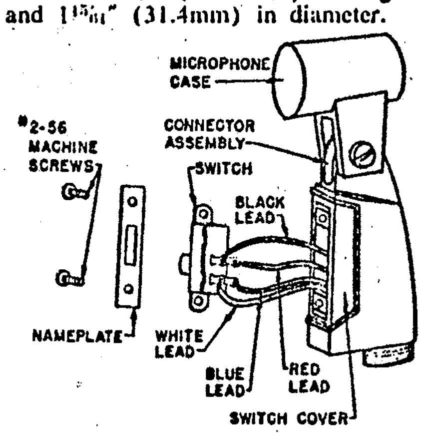
MODEL 545

Architect's Specifications

The microphone shall be the Shure Model 545 or equivalent. The microphone shall be a moving coil type microphone with a frequency range of 50 to 15,000 Hz. This unit shall have a "cardioid" polar characteristic. The cancellation at the sides shall be approximately 6 db and the cancellation at the rear shall be 15 to 20 db. The microphone shall be a dual-impedance microphone with rated impedance of 150 ohms and 40,000 ohms. The microphone rating GM (sensitivity) at 1000 Hz shall be within ±3 db of the following levels.

Low impedance............—149 db
High impedance ....—151 db

The microphone shall be provided with a swivel adapter adjustable through 90° from vertical to horizontal and a receptacle equivalent to the Amphenol 91-MC4F capable of connecting to a three-conductor shielded cable plug. The microphone swivel adapter will mount on a stand having 5/8"-27 thread. The overall dimensions shall be 5 13/16" (147.6mm) in length and 1 15/64" (31.4mm) in diameter.



MODEL 545S
REPLACEMENT OF SWITCH
FIGURE A

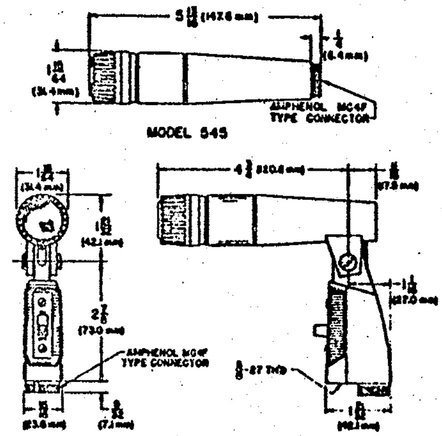
MODEL 545S
Architect's Specifications

The microphone shall be Shure Model 545S or equivalent. The microphone shall be a moving coil type microphone with a frequency range of 50 to 15,000 Hz. This unit shall have a "cardioid" polar characteristic. The cancellation at the sides shall be approximately 6 db and the cancellation at the rear shall be 15 to 20 db. The microphone shall be a dual-impedance microphone with rated impedance of 150 ohms and 40,000 ohms. The microphone rating GM (sensitivity) at 1000 Hz shall be within ±3 db of the following levels.

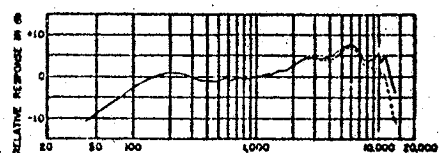
Low impedance ....—149 db

High impedance ....—151 db

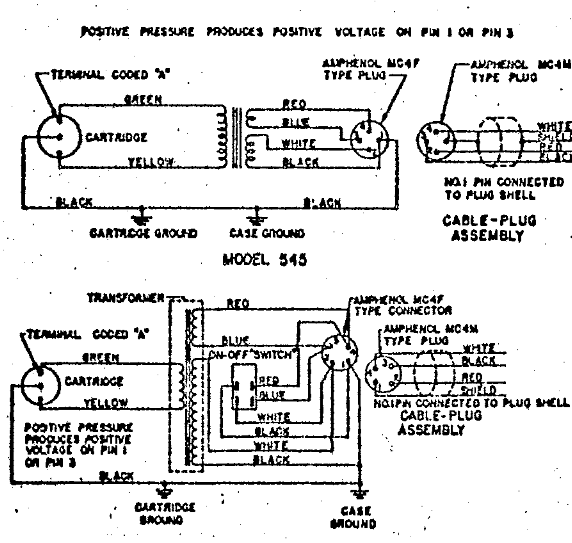
The microphone shall be provided with a swivel, a built in on-off switch and a receptacle equivalent to the Amphenol 91-MC4F capable of connecting to a three-conductor shielded cable plug. The microphone shall mount on a stand having $\frac{5}{8}$"-27 thread. The overall dimensions shall be 4$\frac{15}{16}$" (122.2mm) in height, 5$\frac{13}{16}$" (147.6mm) in depth and 1$\frac{15}{64}$" (31.4mm) in width.



MODEL 545S
OVERALL DIMENSIONS
FIGURE B



FREQUENCY RESPONSE IN HERTZ
FIGURE C



MODEL 545S
INTERNAL CONNECTIONS
FIGURE D

SPECIFICATIONS

Type: Dynamic

Frequency Response: 50 to 15,000 Hz (See Figure C)

Polar Pattern: Cardioid (Unidirectional) pattern—Effective rejection of sound at the rear of the microphone is uniform at all frequencies, while front pickup characteristics are uniform about the axis. (See Figure E)

Impedance: Dual. Connect to cable shield and red conductor for high impedance amplifier inputs. Connect to black and white conductors for balanced line low or medium impedance amplifier input. The shield is connected to the metal parts of the microphone. (See Figure D).

Output Level: 1,000 Hz response.

Model 545 Series Low Impedance
Open Circuit Voltage........................—78 db* (.125 mv)
Power Level........................—57 db**

EIA Microphone Rating
Gm (sensitivity)........................—149 db***

Model 545 Series High Impedance
Open Circuit Voltage........................—55 db* (1.76 mv)
EIA Microphone Rating
Gm (sensitivity)........................—151 db***

*0 db = 1 volt per microbar.
**0 db = 1 milliwatt with 10 microbars
***0 db = EIA Standard SE-105, August 1949.

Cable: 15-foot (4.6 mm) three-conductor shielded with Amphenol MC4M type microphone plug on the microphone end.

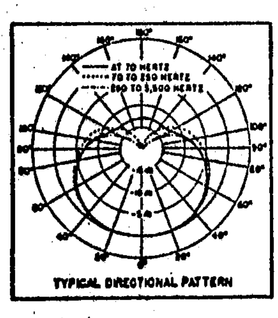
Case: Chrome-plated die-cast case and "Armo-Dur."

Dimensions: See Figure B

Switch: Model 545 None
Model 545S Built in "ON-OFF" switch to control microphone circuit. The switch is an integral part of the receptacle assembly and is a slide-to-talk locking type switch.

Net Weight: Model 545 - 9 ounces (255 grams)
Model 545S - 15 ounces (425 grams)

Packaged Weight: Model 545 - 2 pounds, 3 ounces (992 grams)
Model 545S - 2 pounds, 5 ounces (1049 grams)



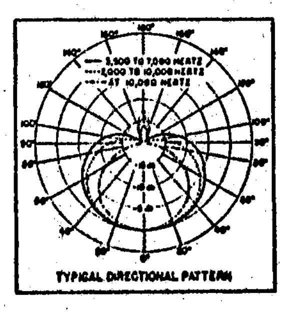


FIGURE E

27A443 (11-68)

U.S. Patents 3,132,713 and D190,864
Printed in U.S.A.

INTEGRATED CIRCUITS

NEW-PRODUCT BULLETIN

This announcement provides preliminary engineering information on new Texas Instruments products. Definitive specifications are now being prepared for publication.

SOLID CIRCUIT® SEMICONDUCTOR NETWORKS†

TYPES SN52 709, SN52 709L, SN52 709N, SN72 709, SN72 709L, SN72 709N HIGH-PERFORMANCE OPERATIONAL AMPLIFIERS

SERIES 52/72 OPERATIONAL AMPLIFIERS

featuring

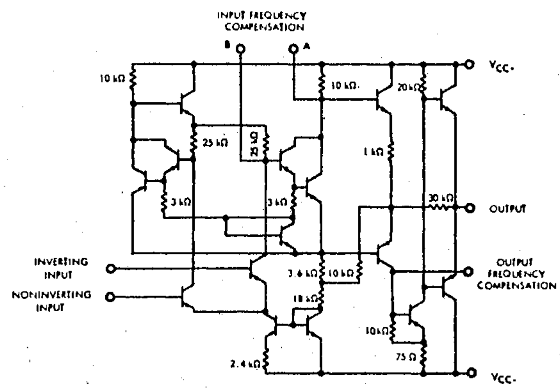
Common-Mode Input Range — ±10 V

Peak-to-Peak Output Voltage Swing — 28 V

description

The SN52 709 circuit is a high-performance operational amplifier with high-impedance differential inputs and a low-impedance output. Component matching, inherent with silicon monolithic circuit fabrication techniques, produces an amplifier with low drift and offset characteristics. Provisions are incorporated within the circuit whereby external components may be used to compensate the amplifier for stable operation under various feedback or load conditions. These amplifiers are particularly attractive for applications requiring transfer or generation of linear or non-linear functions. The SN52 709, SN52 709L and SN52 709N are characterized for operation over the temperature range of −55°C to 125°C. The SN72 709, SN72 709L and SN72 709N are characterized for operation from 0°C to 70°C.

Texas Instruments Series 52/72 catalog lines of linear integrated circuits offer higher reliability, lower cost, smaller size, and less weight than equivalent discrete component circuits.

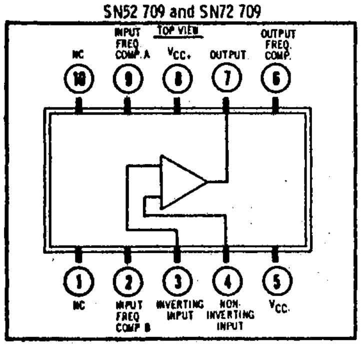


Component values shown are nominal.

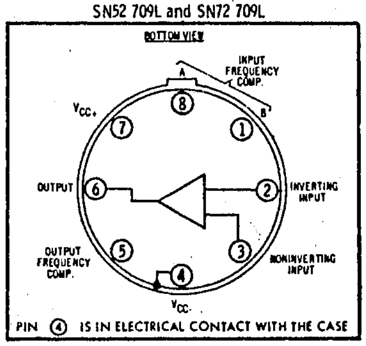
SCHEMATIC DIAGRAM

TERMINAL ASSIGNMENTS

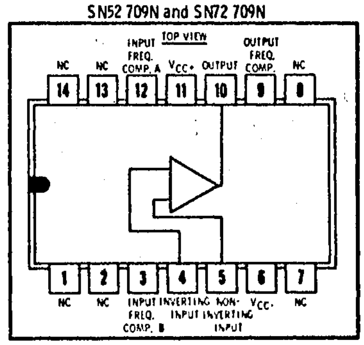
SN52 709 and SN72 709



SN52 709L and SN72 709L



SN52 709N and SN72 709N



NC — No internal connection

†Patented by Texas Instruments.

SC10319
JULY 1967
REPLACES SC9420A DECEMBER 1966

TYPES SN72 709, SN72 709L, SN72 709N
HIGH-PERFORMANCE OPERATIONAL AMPLIFIERS

absolute maximum ratings

Supply Voltages (See Note 1): V_{CC+}	+18 V
V_{CC-}	−18 V
Differential Input Voltage	±5 V
Input Voltage (Either Input, See Note 1)	±10 V
Duration of Short-Circuit Output Current ($T_A = 25°C$)	5 s
Continuous Total Power Dissipation: SN72 709L (See Note 4)	250 mW
SN72 709 (See Note 5)	250 mW
SN72 709N	300 mW
Operating Free-Air Temperature Range (See Notes 4 and 5)	0°C to 70°C
Storage Temperature Range: SN72 709 and SN72 709L	−65°C to 150°C
SN72 709N	−55°C to 150°C

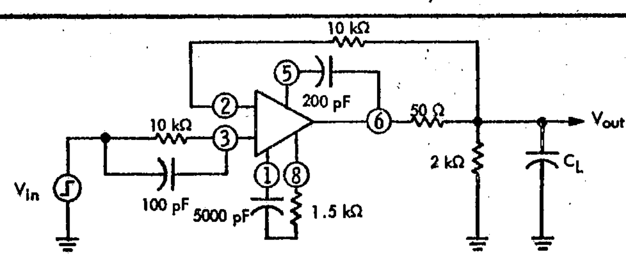
NOTES: 1. These voltage values are with respect to the zero reference level of the supply voltage.
4. At free-air temperature (T_A) above 55°C derate power dissipation linearly at 5.6 mW/deg.
5. At case temperature (T_C) above 100°C derate power dissipation linearly at 5 mW/deg.

electrical characteristics $V_{CC+} = 15$ V and $V_{CC-} = -15$ V, (unless otherwise noted, $T_A = 25°C$)

PARAMETER		TEST CONDITIONS	MIN	TYP	MAX	UNIT
V_{DI}	Differential-input offset voltage	$R_S \le 10$ kΩ, $T_A = 0°C$ to 70°C (See Note 6)			10	mV
		$R_S \le 10$ kΩ (See Note 6)		2	7.5	mV
I_{DI}	Differential-input offset current	$T_A = 0°C$ to 70°C			750	nA
				100	500	nA
I_{in}	Input current	$T_A = 0°C$			2	μA
				0.3	1.5	μA
V_{OM}	Maximum peak-to-peak output voltage	$R_L \ge 10$ kΩ	24	28		V
		$R_L \ge 2$ kΩ	20	26		V
V_{CMI}	Common-mode input voltage range		±8	±10		V
A_V	Large-signal voltage gain	$R_L \ge 2$ kΩ, $V_{out} = \pm 10$ V, $T_A = 0°C$ to 70°C	12,000			
		$R_L \ge 2$ kΩ, $V_{out} = \pm 10$ V	15,000	45,000		
CMRR	Common-mode rejection ratio	$R_S \le 10$ kΩ	65	90		dB
SVRR	Supply voltage rejection ratio	$R_S \le 10$ kΩ		25	200	μV/V
r_{in}	Input resistance		50	250		kΩ
r_{out}	Output resistance			150		Ω
P_T	Total power dissipation			80	200	mW

NOTE 6: $V_{CC+} = 15$ V and $V_{CC-} = -15$ V to $V_{CC+} = 9$ V and $V_{CC-} = -9$ V.

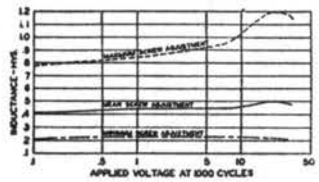
TRANSIENT RESPONSE TEST CIRCUIT FOR SN52 709, SN52 709L AND SN52 709N

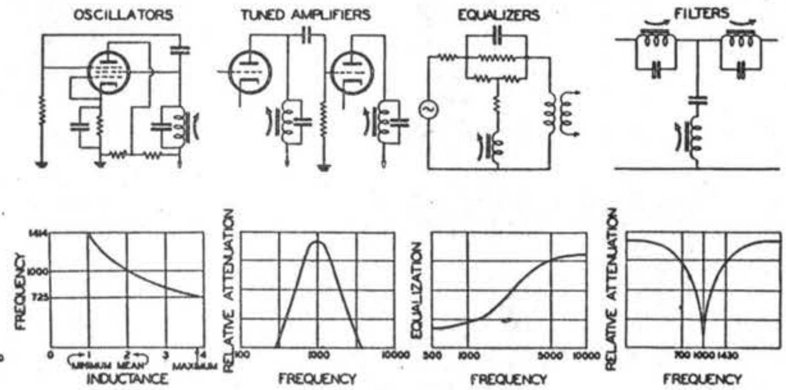


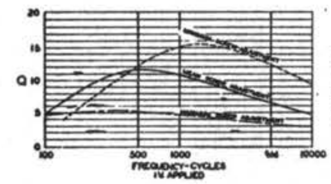
NOTE: Pin numbers shown are for SN52 709L.

FIGURE 1

UTC VARIDUCTORS
VARIABLE INDUCTORS







Type	Mean Hys.	DC MA*	Type	Mean Hys.	DC MA*
VIC-1	.0085	75	VIC-12	1.3	10
VIC-2	.013	60	VIC-13	2.2	8
VIC-3	.021	50	VIC-14	3.4	7
VIC-4	.034	40	VIC-15	5.4	6
VIC-5	.053	35	VIC-16	8.5	5
VIC-6	.084	30	VIC-17	13.	4
VIC-7	.13	25	VIC-18	21.	3.5
VIC-8	.21	21	VIC-19	33.	3
VIC-9	.34	18	VIC-20	52.	2
VIC-10	.54	15	VIC-21	83.	1.5
VIC-11	.85	12	VIC-22	130.	1

UTC type VIC variable inductors offer a revolutionary approach to the problem of tuned audio circuits. By adjusting a set screw in the side of the case, an inductance value of +85%, −45% from mean value is obtainable. Setting is positive. Effective Q for a wide frequency range and variation of inductance with applied AC voltage are shown on the illustrated curves, for a typical VIC unit.

The VIC inductor is housed in a rugged die cast case 1$^{11}/_{32}$" long, 1¼" wide and 1$^{7}/_{16}$" high with mounting centers on terminal board side $^{13}/_{16}$" by $^{29}/_{32}$", tapped for 4-40 screw. Weight is 5½ oz.

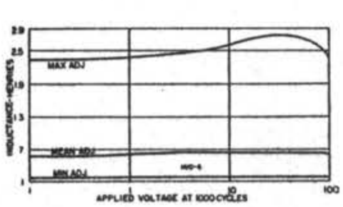
*DC MA shown is maximum recommended . . . will effect some reduction in inductance and Q.

UTC HERMETIC VARIDUCTORS

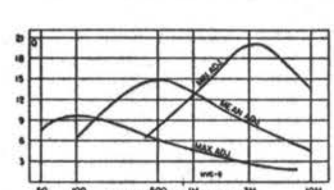
Variable inductors hermetically sealed to MIL-T-27 spec.

UTC variable inductors have served as the ideal solution to many filter, oscillator, equalizer, and tuned amplifier problems—for over a decade. Extended development has now made possible the new HVC series of inductors with improved characteristics. They are hermetically sealed to MIL-T-27 specs . . . extremely compact . . . wider inductance range . . . higher Q's . . . lower and higher frequencies . . . superior voltage and temperature stability.

Adjustment of set screw in top of case permits changing inductance + 200%, −70% of nominal value shown. Setting is positive. Effective Q for a wide frequency range and variation of inductance with applied AC voltage are shown on the illustrated curves, for a typical HVC unit. Case dimensions are 1⅛" long, $^{25}/_{32}$" wide, 1$^{3}/_{32}$" high. The two terminals and two 4/40 mounting studs are on opposite diagonals — $^{15}/_{16}$" spacing.



Type No.	Min. Hys.	Mean Hys.	Max. Hys.	DC* Ma
HVC-1	.002	.006	.02	100
HVC-2	.005	.015	.05	60
HVC-3	.011	.040	.11	40
HVC-4	.03	.1	.3	30
HVC-5	.07	.25	.7	20
HVC-6	.2	.6	2	15
HVC-7	.5	1.5	5	10
HVC-8	1.1	4.0	11	7
HVC-9	3.0	10	30	5
HVC-10	7.0	25	70	3.5
HVC-11	20	60	200	2
HVC-12	50	150	500	1.5



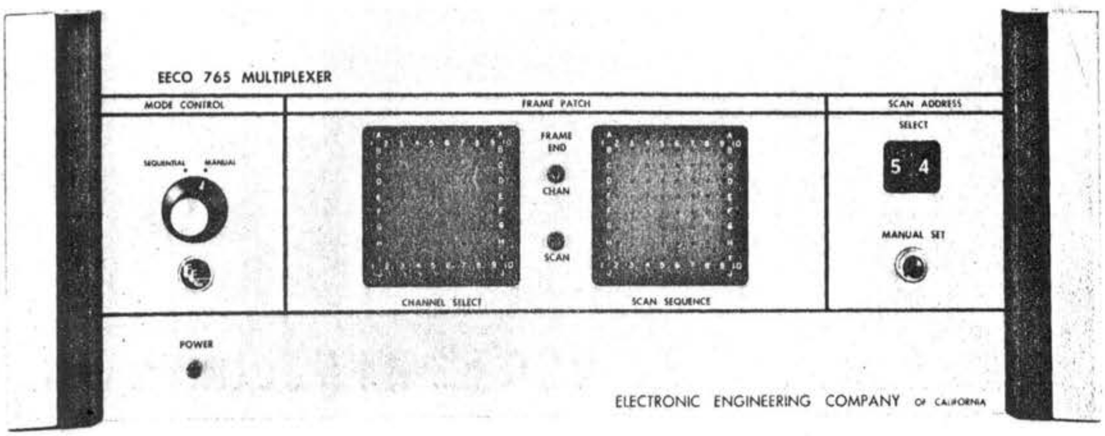
*DC MA shown is maximum recommended . . . will effect some reduction in inductance and Q.

EECO 765
REPRINTED JANUARY, 1967

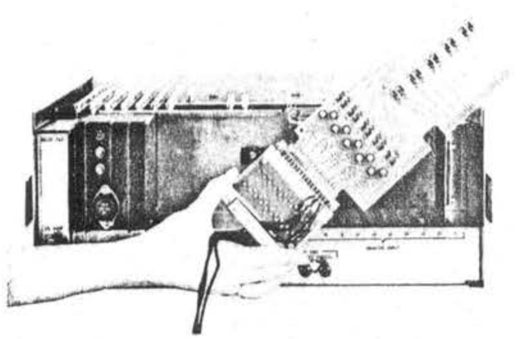
ANALOG MULTIPLEXER

EECO 765-1 With frame patch • **EECO 765-2** Without frame patch

Up to 100 channels
50,000 samples/second

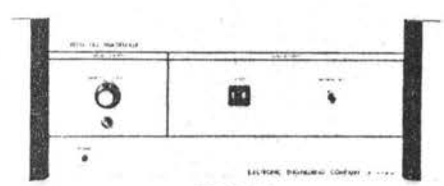


EECO 765-1



10 FET channel switches with adapter for channel attenuators

... **100 megohms minimum input impedance**
... **FET switches — 100% duty cycle**
... **Super-commutation**
... **Display gate for selected channel**



EECO 765-2

ELECTRONIC ENGINEERING COMPANY *of California*

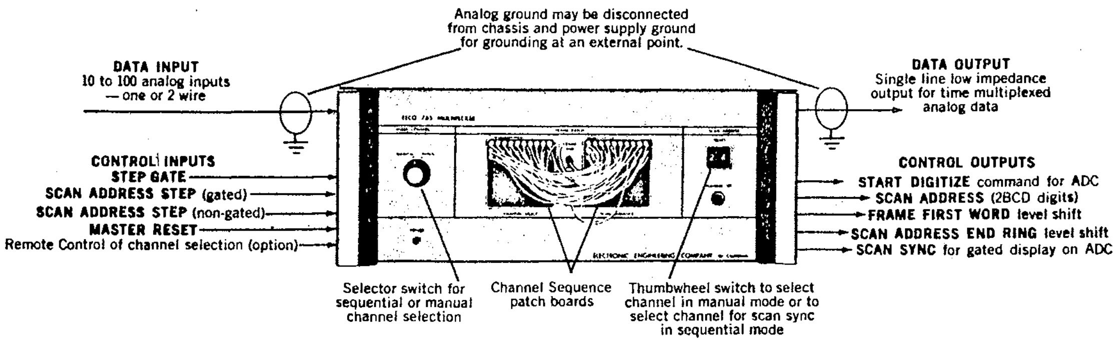
GENERAL . . .

The EECO 765 Analog Multiplexer uses Field Effect Transistor channel switches and a Field Effect input transistor in the buffer amplifier.

10 FET channel switches are mounted on each input circuit card. Multiplexers can therefore be supplied in any multiple of 10 channels.

The FET switches have a very high off impedance ... over 20,000 megohms and the input FET in the buffer amplifier has an input of over 100 megs so the lowest input impedance, with a full 100 channels, is over 100 megohms.

Other features are: super-commutation, sequential or addressable channel selection, very low offset and cross-talk, and selectable SCAN SYNC output for synchronizing external data displays.



SPECIFICATIONS
ANALOG MULTIPLEXER
EECO 765-1— with frame patch
EECO 765-2— without frame patch

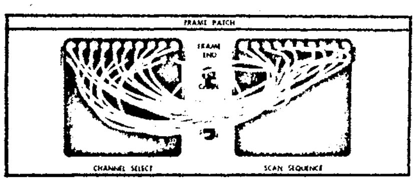
OPERATING MODES

Sequential............ Steps sequentially through scan on external command. Triggers external Analog to Digital Converter after a delay from **SCAN ADDRESS STEP** input command.

Manual............ Steps to channel selected by thumbwheel switches when **MANUAL SET** button is operated. Will continue to trigger external ADC after delay from **SCAN ADDRESS STEP** input command.

Addressable............ Circuits can be provided for random selection of channels by remote addressing. Channel address should preferably be in 2 BCD digit form.

CHANNEL SELECTION............ The patch boards on the EECO 765-1 provide a quick method of changing channel sequencing or of super-commutating any one or several selected channels.
(EECO 765-1)



For super-commutation, it is only necessary to connect one position on the **CHANNEL SELECT** patch to two or more positions on the **SCAN SEQUENCE** patch. The selected channel will then be sampled more than once for each scan. Any number of channels can be super-commutated.

Changing the sequence of channels to be sampled or the number to be sampled in each scan is done by patch connection between the **SCAN SEQUENCE** and **CHANNEL SELECT** patches. Channels can be patched in any arbitrary sequence.

Patching the last **SCAN SEQUENCE** point to the **FRAME END/SCAN** jack and the last **CHANNEL SELECT** point to the **FRAME END/CHAN** jack causes the first channel switch to be closed when the next **SCAN ADDRESS STEP** command is received.

INPUT CHANNELS.... From 10 to 100 single wire channels in multiples of 10.

Up to 50 two wire differential inputs with separate shield grounds or up to 100 two wire differential inputs with a common shield ground.

Analog data ground may be disconnected from the chassis and power supply ground for grounding to an external point. (Voltage differential between analog data ground and chassis must not exceed $\pm$.25v)

SWITCHING RATE..... Up to 50,000 switch closures per second.

Switch settling time is approximately 10 μsec. for an accuracy of .01% of FS when the source impedance is 1,000 ohms or less.

100% duty cycle — no switch closure time limitation as the switch drive is not AC coupled.

DATA INPUTS........... Input lines to the EECO 765 should be on coax cable. RG 174/U or other sub-miniature cable is recommended. Pins are provided for grounding the coax shields.

The inputs run directly to connectors for the channel switch cards or to an input attenuator terminal boards. The attenuator terminal boards accept two RN 60 size resistors per channel for fixed attenuation of the input signals.

Level......................... Input for unity gain is $\pm$5 volts.

Input levels down to $\pm$50 mv can be accepted.

When the input is less than $\pm$5 volts, the buffer amplifier gain is adjusted so that the output is $\pm$5 volts. Accuracy at less than $\pm$5 volt input level depends on source impedance and number of channels.

Input Impedance........ 20,000 megohms and 3 pf per channel in "off" state. Greater than 100 megohms in parallel with [1.5 (N-1) +10] pf where N = number of channels (multiples of 10 up to 100).

Source Impedance..... Up to 10,000 ohms.

Offset Currents.......... Offset current — switch closed — 7.0 (N-1) na maximum at 50°C.

Leakage current — switch open — 7.0 na maximum at 50°C.

Over-voltage rating (no cross-talk) — Up to +9v and −15v.
Maximum input without circuit damage — $\pm$35v.

DATA OUTPUTS

Level......................... Equal to input. Gain is nominally 1.000 with $\pm$.01% gain linearity.

Impedance................ Less than .1 ohm.

Loading...................... Load value should not be less than 250 ohms or more than 5,000 pf. Maximum current is 60 ma into a short circuit.

ACCURACY................. Voltage output is the sum of: (for 0 to 50°C)
— Input voltage X gain (nominally 1.000)
plus — offset voltage —
$\leqq \pm 1$ mv $+7 \times 10^{-4}$(N-1) (R source +1500) mv
plus — AC cross-talk —
$\leqq$ .01% of full scale or
$\leqq 10^{-10}$ (N-1) f(1500 +R source)% of full scale whichever is greater
Where f = frequency of interfering signal
plus — DC cross-talk —
$\leqq$ .01% of full scale or
$\leqq 2.5 \times 10^{-9}$ (N-1) (R source +1500)% of full scale whichever is greater
plus — noise —
~ .2 mv typical wide band peak-to-peak amplitude

Typical Example............ 100 channels with 1k ohm source impedance at 50°C worst case

Voltage output (for 5 volts input) = 5 x .9999′ = 4.9995 volts
offset voltage $\leqq \pm$.0027 volts (.027% of FS)
AC cross talk $\leqq$ .01% of FS up to 400 Hz
or $\leqq .01\% \times \frac{f}{400}$ of FS above 400 Hz

DC cross-talk $\leqq$ .01% of FS

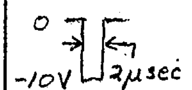
Channel to channel gain accuracy — approx. $\pm$.01%

CONTROL INPUTS..... The external command pulse to advance the internal counter one step and close the switch to the corresponding input channel may be supplied either to a gated input or non-gated input.

Gate Level................. Gate closed — −6 to −12 vdc
Gate open — 0 to −0.5 vdc (or open circuit)
Impedance — 8.2 k ohms
Open/close delay — 1 μsec maximum

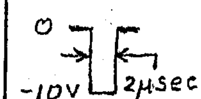
Scan Address Step Through Gated Input......................... Positive going pulse or level shift from a −6 to −12 vdc level to 0 vdc steps the counter and closes the input channel switch.
Rise time — 2 μsec maximum
Duration — 2 μsec minimum
Impedance — 8.2 k ohms
Maximum repetition rate — 50,000/sec.

Scan Address Step Through Non-Gated Input......................... Negative going pulse or level shift from a 0 vdc level to a negative value of from −6 to −12 vdc steps the counter and closes the input channel switch.



Rise time −2 μsec maximum
Duration — 2 μsec minimum
Impedance — 8.2 k ohms
Maximum repetition rate — 50,000/sec.

Master Reset............ A negative going pulse or level shift from a 0 vdc level to a negative value of from −6 to −12 vdc resets the counter to the 00 position.



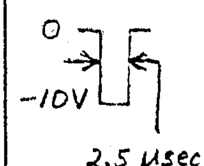
All channel switches are open in this position (except if 100 channels are used in which case channel 100 corresponds to 00 counter position). Channel number one switch is closed and the counter is advanced to position 01 on receipt of the next **SCAN ADDRESS STEP** command pulse.

Rise time — 2 μsec maximum
Duration — 2 μsec minimum
Impedance — 8.2 k ohms

A delay of 5 μsec is required from the end of the reset pulse to the next **SCAN ADDRESS STEP** command pulses.

OUTPUT CONTROL SIGNALS

Start Digitize Signal......................... An output signal is provided 10 μsec after the **SCAN ADDRESS STEP.** This delay is generally sufficient for the switch to settle to less than .01% of FS. This output signal is normally supplied to the A-D Converter to initiate conversion.



Two **START DIGITIZE** outputs are provided — one a positive going pulse from the −6 to −12 v level to 0 vdc and a negative going pulse from 0 level to −6 to −12 v level.
Rise time — 1 μsec maximum
Duration — 2.5 μsec
Loading — 1.2 k ohm and 100 pf maximum
The delay between receipt of the **SCAN ADDRESS STEP** command pulse and the **START DIGITIZE** output can be varied by altering the value of one capacitor.

Scan Address............ A 2 digit output in BCD form is available to indicate the channel being scanned.
Sixteen output lines are used for this **SCAN ADDRESS** signal as both true and complements are brought out.
"One" level — 0 to $\pm$0.5 vdc
"Zero" level — −6 to −10 vdc
Rise and fall time — .5 μsec

Delay from **ADVANCE COUNTER** pulse — 2 μsec max.
Loading — 1.2 k ohms +100 pf max.

Frame First Word....... A dc level shift output is produced when the first channel switch in each frame is closed. This level is maintained during the time the first channel switch is closed.

Level....................... Quiescent — 6v to −12 vdc
First switch closed — 0v to ±0.5 vdc

Delay....................... The interval between receipt of the **SCAN ADDRESS STEP** command pulse to this **SCAN ADDRESS STEP** level shift is approximately 7 μsec.

Loading.................. 1.2 k ohms +100 pf maximum

Scan Address End Ring............. A dc level shift output is produced when the last channel switch in each frame is closed. This level is maintained during the time the last channel switch is closed.

0V
-10V

Level....................... Quiescent — −6v to −12 vdc
Last switch closed — 0v to ±0.5 vdc

Delay....................... The interval between receipt of the **SCAN ADDRESS STEP** command pulse to this **FRAME FIRST WORD** level shift is approximately 7.5 μsec.

Loading.................. 1.2 k ohms +100 pf maximum

Scan Sync Output for Gated Display...... A dc level shift output is produced when the channel switch is closed corresponding to the channel number selected by the front panel thumbwheel switches.
This level shift is available to gate on the output display on the ADC so as to display the digitized output of the selected channel.
Both a positive and a negative dc level shift signal is available.

0V
-10V

Positive Level Shift............................ Quiescent — 6v to −12 vdc
Selected switch closed — 0v to ±0.5 vdc

Negative Level Shift............................ Quiescent — 0v to ±0.5 vdc
Selected switch closed — −6v to −12 vdc

Delay....................... The interval between receipt of the **SCAN ADDRESS STEP** command pulse to this **SCAN SYNC** level shift is approximately 4.5 μsec.

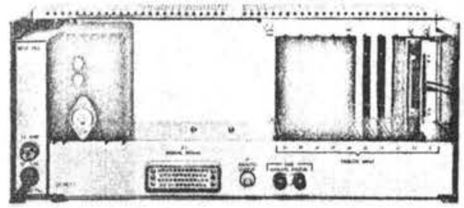
Loading.................. 1.2 k ohms +100 pf maximum

INPUT POWER.......... 117 vac ±10% — 50 to 400 Hz — 25 watts

PHYSICAL.................. 19" wide — 7" high — 19½" deep including mating connectors.
Weight — 35 pounds
Chassis sides drilled for both Chassis Trak CTD 120 tilt slides and Chassis Trak CTN 116 non-tilt slides.
Finish — FED-STD 595-26231, Grey

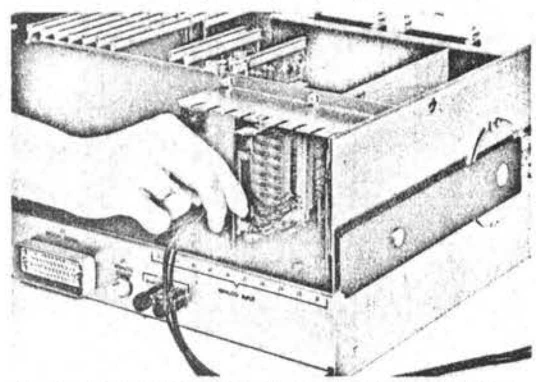
ENVIRONMENT......... Operating temperature range — 0°C to +50°C

CONNECTORS............ Analog Input —
J1 through J10 — wired to attenuator terminal board or to furnished card connectors.
Analog Output —
J11 — BNC UG 1094A/U
Control input and output —
J12 Winchester 50 pin XAC 50SF2006
Grounds
Separate binding posts, isolated from the chassis, are provided for analog and digital (power supply) grounds.
Mating connectors supplied.



ORDERING INFORMATION

When ordering specify —
- Model 765-1 with frame patch
- Model 765-2 without frame patch
- Number of input channels and type (i.e., single wire, etc.)
- If attenuating resistors are to be supplied
- Full scale level of input
- If channel selection by remote address is required



Input leads to channel attenuator board and card connector.

Electronic Engineering Company *of California*

HOME OFFICE:
1601 E. Chestnut Avenue, P.O. Box 58
Santa Ana, California 92702
Phone: (714) 547-5501

WASHINGTON D.C. OFFICE:
Parkway Building, Suite 314, 5801 Annapolis Road
Bladensberg, Maryland 20710
Phone (301) 779-3636

Printed in U.S.A.

SECTION 1
CHARACTERISTICS

General Description

The Type RM564 Storage Oscilloscope is a special-purpose oscilloscope designed to store cathode-ray tube displays for viewing or photographing up to an hour after application of the input signal. In addition, the instrument can be operated as a conventional oscilloscope. The physical dimensions of the Type RM564 permit it to be mounted in a standard 19-inch rack, using only 7 inches of vertical space. Since it is compatible with all Tektronix 2-Series and 3-Series plug-in units, the Type RM564 can be operated in a variety of modes including differential, multi-trace, wide-band, sampling and delayed sweep.

The two separate storage screens of the cathode-ray tube provide convenience and versatility for comparison and analysis of waveforms. Display storage is also convenient for photography and detailed viewing of single-sweep displays and extremely low-frequency waveforms (< 60 cps). You may use either the upper or lower storage screen in storage mode while the other screen is operated in non-store, or you may operate both screens together in either storage or non-store mode.

Cathode-Ray Storage Tube (Patents Pending)

Type—Tektronix-manufactured T5640-200 for the standard instrument; T5640-201 for the Type RM564 (Mod 08) Oscilloscope. The crt is warranted under the Tektronix instrument warranty given in the front of this manual.

Envelope—Rectangular, 4-inch x 5-inch ceramic and glass; flat glass faceplate.

Screen—8-cm × 10-cm viewing area divided into 4-cm × 10-cm upper and lower storage screens. Each screen has its own operating-level and erase circuitry for storage operation.

Accelerating Voltage—3.5 kv.

Deflection—Electrostatic.

Deflection-Plate Sensitivity—Horizontal: approximately 18.5 volts/cm. Vertical: approximately 19.5 volts/cm.

Unblanking—Deflection-type, dc-coupled; ±88 volts approximate cutoff.

Intensity Modulation—Internal: grid. External: cathode. Typically a 3-volt external signal will produce visible intensity modulation.

Focus—Electrostatic.

Storage Duration—At least one hour.

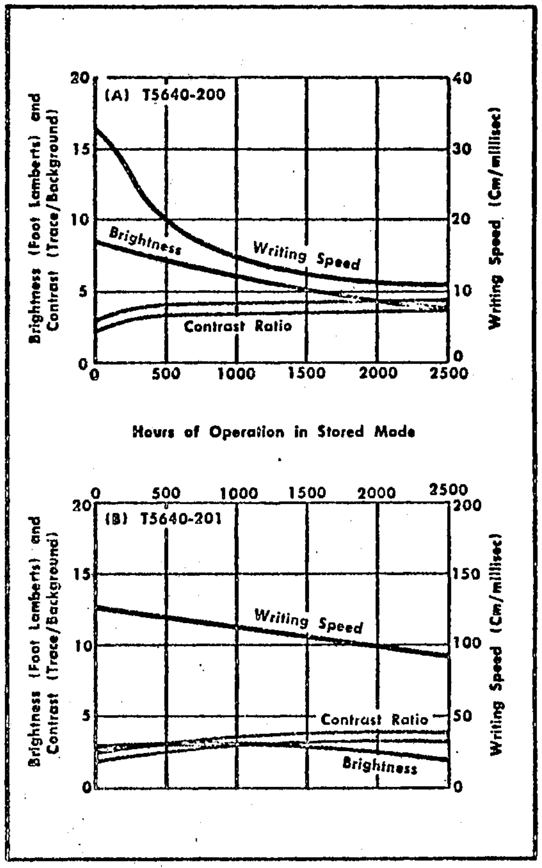


Fig. 1-1. Typical changes in writing speed, brightness and controls for the T5640-200 and T5640-201 crts.

Erase Modes—Manual: front-panel DISPLAY switches. Remote: rear-panel remote-erase jack. A 9-pin plug that mates with the remote-erase jack may be ordered through your Tektronix Field Office. Tektronix part number for the plug is 134-049.

Erase Time—Approximately 250 msec.

Writing Speed, Brightness and Contrast—See Table 1-1 for minimum performance specifications and Fig. 1-1 for expected crt life characteristics. The hours of operation shown are the hours the crt is used in the storage mode with repetitive writing, storing and erasing.

Writing Speed Increase (using trace enhancement)—Approximately 10:1 for T5640-200 crt; approximately 5:1 for T5640-201 crt.

Characteristics—Type RM564

TABLE 1-1

Storage Screen Specifications
(Center 6 x 8 cm area)

Characteristic	Type T5640-200	Type T5640-201
Minimum Initial Writing Speed (60 μamp beam current)	25 Cm/mSec	100 Cm/mSec
Minimum Initial Brightness	6-Foot-Lamberts	2-Foot-Lamberts
Minimum Display-to-Background Contrast Ratio	2:1	2:1

Graticule

Type—External.

Markings—8 vertical and 10 horizontal 1-cm divisions; 2-mm marks along the vertical and horizontal centerlines.

Illumination—Variable edge lighting adjusted with front-panel SCALE ILLUM control.

Calibrator

Waveform—Square-wave signal at power-line frequency.

Output Voltage—1 millivolt to 100 volts peak-to-peak in 6 calibrated decade steps; within 3% of indicated voltage into high-impedance load. Output amplitude at 1 V position is 0.1 volt into a 50-ohm load.

Risetime—5 μsec or less.

Instrument Power

Line Voltage—105 to 125 volts (117 volts nominal), or 210 to 250 volts (234 volts nominal), single-phase ac.

Line Frequency—50 to 60 cps.

Power Consumption (with plug-in units)—Maximum of 250 watts.

Fuse—3-amp slow-blowing type for 117-volt operation; 1.6-amp slow-blowing type for 234-volt operation.

Ventilation

Type—Forced-air cooling.

Overheat Protection—Thermal cutout interrupts instrument power if temperature exceeds safe operating level (ambient air 122° F maximum); restores power automatically when temperature drops below reset level of cutout (approximately 115° F).

Mechanical Characteristics

Construction—Aluminum-alloy chassis and panels. Top and bottom panels separately removable. Photo-etched anodized front panel.

Dimensions—Height: 7 inches. Width: 19 inches. Depth: 20½ inches, excluding right-angle power cord. Additional dimensional information may be found on the Dimension Drawing foldout sheet at the rear of this manual.

Weight—31 pounds without plug-in units.

Standard Accessories

1—Polarized viewer, 016-039.

1—3-wire power cord, 161-013.

2—Instruction manuals, 070-415.

1—3- to 2-wire adapter, 103-013.

2—BNC to binding-post adapters, 103-033.

1—Red test lead, 012-031.

Optional Accessories Available

Slideout track mounting assembly (Mod 171), 351-050.

Cradle mounting assembly, 040-344.

Blank plug-in unit, 040-245.

See current Tektronix catalog for 2-Series and 3-Series plug-in units.

BURR-BROWN

BURR-BROWN MODEL

1663/16

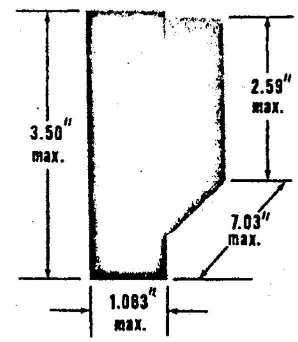
SAMPLE AND HOLD/ SWITCHED INTEGRATOR

FUNCTION MODULE

SPECIFICATIONS - Performance at 25°C and with rated supply.

	Min.	Typ.	Max.	Units
SAMPLE & HOLD				
Inputs	Single ended; 1 or 2 signals may be summed			
Input Impedance		2		kΩ
Input Signal Level - Operating			± 10	V
- Absolute Maximum			± 15	V
Overload Recovery Time		2		ms
Voltage Gain		- 1		ratio
Gain Accuracy at dc		±0.1		%
Small Signal Frequency Response, ±3 dB	1.5	2		MHz
Output		Single ended		
Output Impedance		0.1		Ω
Rated Output - Voltage	± 10			V
- Current	± 20			mA
Maximum Frequency for Rated Output	40			kHz
Output Voltage Drift in HOLD (for 10 ms)		adj to ±1		mV
vs. Temperature		± 1		mV/°C
vs. Supply		± 3		mV/%
in TRACK vs. Temperature		±100		μV/°C
vs. Supply			± 20	mV/%
Aperture Time		50	100	ns
Acquisition Time (to 0.05% for full-scale input voltage step)		10	15	μs
SWITCHED INTEGRATOR				
Integrator Inputs	Single ended; 1 or 2 signals may be summed			
Input Impedance - 1st input		10		kΩ
- 2nd input		100		kΩ
Integrator Input Feedthrough in RESET (for ±10V)			± 2	mV
Initial Condition Inputs	Single ended; 1 or 2 signals may be summed			
Voltage Gain - 1st input		10^4		1/s
- 2nd input		10^3		1/s
Gain Accuracy		±0.25		%
Output		Single ended		
Output Impedance		0.1		Ω
Rated Output - Voltage	± 10			V
- Current	± 20			mA
Output Voltage Drift in COMPUTE (for 10 ms)		adj to ±1		mV
vs. Temperature		±3.5		mV/°C
vs. Supply		±4.5		mV/%
Switching Times - Compute		50	100	ns
-Reset (to within 0.1% of initial condition inputs)			10	μs
SWITCH CONTROL SIGNAL REQUIREMENTS				
TRACK (RESET) 0		- 6	- 4	V
HOLD (COMPUTE)	- 2	0		V
Rise Time		25	40	ns
Noise Immunity -6V 15μsec	1			V
TEMPERATURE RANGE				
Specification	0		+ 60	°C
Operating	- 40		+ 85	°C
Storage	- 55		+100	°C
POWER SUPPLY REQUIREMENTS				
Rated Supply Voltage		± 15		Vdc
Voltage Range	± 12		± 18	Vdc
Supply Drain - Quiescent			± 35	mA
- Rated Output			± 60	mA
Supply Regulation			0.1	%
Noise and Ripple			1	mV, rms

Specifications subject to change without notice. PDS-168 7/67



Module Weight 20 oz. maximum

- **10 μs ACQUISITION TIME**
- **50 ns APERTURE TIME**
- **DUAL SUMMING INPUT**
- **0.1 % f.s. FUNCTION ACCURACY**

THE MODEL 1663/16............. is a sample and hold or switched integrator module designed for ±10V service. A high-gain current amplifier is used in conjunction with a high-gain wide bandwidth operational amplifier to achieve very rapid integrator reset rates. Digital transistor logic is included to provide noise immunity and uniform switching rates. The switching time is essentially independent of the control signal rise time.

Epoxy encapsulated submodules employing silicon semiconductors for operation over a wide temperature range are installed in this Burr-Brown 1600 Series unit. The unit may be accurately balanced by means of three self-contained zero controls. Circuit protection is incorporated to prevent damage to the input stage due to input overvoltage and to the output stage due to short circuit. Operating power is obtained from external, regulated power supplies.

EXTERNAL CONNECTIONS

A Burndy EC4206P5 mating connector, along with hardware for securing the connector to the accessory rack adapter, is furnished with the unit. External connections are made to the connector pins as follows:

Pin A	Case Ground
Pin B	Common
Pin C	Switch Control Signal
Pin D	Output
Pin E	Positive Power, +15 Vdc
Pin F	Negative Power, -15 Vdc
Pin H	Input 1 - Unity Gain, Track or Reset
Pin J	Input 2 - Unity Gain, Track or Reset
Pin K	Input 3 - Integrator Gain of 10^4
Pin L	Input 4 - Integrator Gain of 10^3

APPENDIX D

PHONEME DATA SET

The data set of phonetic sounds which was compiled as a part of this thesis, is included in this appendix. Eleven phonemes were voiced by 3 males and 2 females, each over an octave of vocal pitch. Table XII includes a tabular listing of this data set. At the top of each sample the phoneme and speaker are identified according to Tables V and VI. The parameter identified as "TM" is a number equal to 20 times the maximum voltage applied to the Analog-to-Digital Converter during the sampling of the 97 filter bank channels. This parameter is used to normalize all of these inputs to whole numbers between 99 and 0. The channel from which TM was derived has a value of 99 in Table XII.

The Speech Tracker output was also recorded at the top of each sample as the parameter "ST". The number which is recorded is 20 times the voltage read by the Analog-to-Digital Converter.

All of the data samples of Table XII are plotted in Figures D.1 through D.28 with the aid of the IBM 1620 Computer and associated Calcomp 565 Plotter. The effect of pitch changes can be observed by vertically scanning each figure.

TABLE XII

THE PHONEME DATA SET

```
SPEAKER=1    PHONEME= 1    PITCH=100    TM= 62.50    ST= 92.70
28  5  1  4  0  0  0  0  0  0  2  1  0  3  2  3 18  7  0  0
 1  0  0  2  0  4  1  0  0  0  2  2  8  5  2  2  6 14  7  9
16 47 52 26 50 99 35 15 10  8  6  5  4  4  4  3  4  3  2  5
 7  9 21 35 56 35 19 15  7  6  7  4 10 17 24 12  6  3  1  1
 0  0  3  2  2  1  3  0  0  1  0  0  0  1  0  1  4

SPEAKER=1    PHONEME= 1    PITCH=110    TM= 59.00    ST= 89.00
 3 11 29 11  2  2  0  0  0  1  0  1  2  1  0  2  0  1 10  1
 0  1  1  1  0  3  0  1  1  1  0  2  2  2  4  6  4  5 14 44
14 17 57 99 57 30 29 36  9  9  9  5  5  4  5  9  5  8 12 33
30 67 41 60 14 12  7  7  7  5  6  9 20 37 21 11  4  2  4  2
 0  1  4  0  3  2  2  1  3  1  0  0  0  0  1  1  2

SPEAKER=1    PHONEME= 1    PITCH=123    TM= 72.60    ST= 96.00
 3  1  0  5 20 37 12  2  3  1  3  0  0  0  0  0  0  1  1  2
 8 19  3  1  0  0  0  1  0  2  7  3  0  2  1  0  3  6  3  6
10 23 75 32 23 33 90 49 14 12  8  6  5  7  5  5  6  6 11 11
23 27 75 67 99 46 19 16  8  9 10  9 14 24 11 10  8  6  9 11
 9  8 11 14 13  6  4  2  2  3  3  2  2  2  1  1  3

SPEAKER=1    PHONEME= 1    PITCH=130    TM= 83.00    ST= 95.60
 0  0  0  1  5 16 33  9  7  1  0  0  0  0  0  0  0  1  2  2
 4 14 31 10  2  2  0  1  2  2  4 14  6  1  1  2  1  5 16  7
 5 10 34 59 26 16 22 52 24 13  7 11  4  2  4  2  1  4  4  8
 9 19 20 50 46 99 31 24 13  9 11 10 13 13 12  6  8  5  5  6
 8 12  9  3  8  7  4  1  1  0  2  2  2  1  3  0  1

SPEAKER=1    PHONEME= 1    PITCH=146    TM= 78.00    ST= 95.60
 2  0  0  1  1  0  2  5 25 45 13  6  2  1  1  0  0  1  1  1
 2  1  1  5  9 30  8  4  3  2  3  2  2  3 12  4  5  6  9 18
41 93 37 19 27 46 99 36 13 11  9  6  4  5  4  2  2  5  1  7
 7 10 22 25 56 31 39 20  9  7  9  7  4  5  5  4  2  4  3  5
 2  1  3  3  3  2  2  1  0  1  2  1  1  2  1  0  1

SPEAKER=1    PHONEME= 1    PITCH=164    TM= 55.50    ST= 87.10
 0  1  0  2  2  2  0  2  3  3 10 26 79 21 13  6  1  4  1  0
 0  1  0  1  2  5  7 24 49 14  6  4  0  3  6  9 22 62 29 13
 8 11 31 46 99 36 17 15 23 51 12  6  3  7  4  1  2  5  1  1
 3  1  2  6 12 21 12 20  7  6  6  3  2  2  8  5  3  3  3  2
 3  0  1  0  0  0  2  1  1  1  3  0  0  0  1  0  1

SPEAKER=1    PHONEME= 1    PITCH=174    TM= 57.40    ST= 92.40
 3  1  0  2  0  0  0  0  1  2  6  8 32 63 22  7  2  3  0  1
 2  1  0  2  0  1  2  7 24 48 16  8  4  6  4  7 14 26 99 52
17 11 16 16 43 62 19 10  7 19 52 21  7  7  8  3  1  1  0  2
 4  5  6 22 16 28 46 24  8  5  6  3  7  6  8 26 11  7  1  3
 1  0  5  5  6  3  4  0  0  1  1  0  1  1  2  1  5
```

TABLE XII (Continued)

SPEAKER=1 PHONEME= 2 PITCH=100 TM= 70.90 ST= 92.60

24	2	0	1	0	1	0	0	1	2	1	0	0	0	0	5	11	3	1	2
0	0	1	0	0	10	2	0	3	4	6	11	25	8	8	14	41	99	28	16
8	14	10	4	6	6	2	0	1	3	2	2	3	2	0	1	0	0	1	2
2	2	3	1	1	3	7	10	17	11	7	5	9	10	16	11	5	3	2	3
0	1	4	0	0	1	0	0	1	1	0	1	1	0	0	1	0			

SPEAKER=1 PHONEME= 2 PITCH=110 TM= 64.80 ST= 94.90

4	18	41	17	6	4	0	0	1	0	0	0	1	0	1	2	2	7	24	3
0	0	0	0	1	2	1	2	3	1	1	3	6	12	46	16	10	9	28	71
22	9	8	2	3	3	3	3	3	2	1	0	4	3	4	9	3	4	5	4
3	8	5	8	6	7	10	19	45	64	99	45	22	25	37	33	14	15	15	12
6	6	6	1	4	2	2	4	7	6	6	7	2	0	0	1	0			

SPEAKER=1 PHONEME= 2 PITCH=123 TM= 64.90 ST= 95.20

1	1	2	1	6	22	5	0	1	1	1	0	2	0	0	0	0	0	1	2
7	22	6	1	0	0	0	0	3	5	19	11	6	4	8	14	42	99	29	16
9	13	44	16	7	4	5	0	1	3	2	2	5	6	2	3	2	0	2	2
2	3	5	4	4	7	11	21	48	90	64	39	29	30	49	41	14	6	4	6
3	10	20	9	7	5	3	1	3	4	1	3	2	2	1	3	2			

SPEAKER=1 PHONEME= 2 PITCH=130 TM= 95.20 ST= 97.80

1	1	1	3	8	15	36	8	5	0	1	1	1	1	2	0	0	0	1	0
3	10	27	9	3	2	1	2	3	2	7	16	11	7	9	11	22	42	99	39
16	11	19	34	15	7	5	8	3	2	2	2	0	2	4	2	3	8	2	7
1	2	0	3	4	7	14	31	57	38	14	9	6	10	13	11	3	4	4	4
3	2	1	1	3	1	2	3	1	1	1	0	0	0	1	0	1			

SPEAKER=1 PHONEME= 2 PITCH=146 TM= 75.50 ST= 96.60

0	0	0	2	2	3	2	6	23	52	16	7	2	2	3	1	1	3	1	0
0	1	0	5	8	33	10	7	6	5	8	9	15	28	99	58	25	14	10	13
27	74	32	12	8	5	8	5	2	4	11	6	2	4	13	3	4	12	3	6
6	3	7	4	5	4	13	21	59	36	52	20	15	22	28	14	6	5	4	4
5	3	5	2	7	3	4	3	4	7	8	4	1	1	1	0	1			

SPEAKER=1 PHONEME= 2 PITCH=164 TM= 95.60 ST= 96.70

0	1	1	1	0	1	0	0	3	4	8	19	65	19	10	6	1	1	2	1
0	2	2	2	1	6	8	20	51	18	10	9	10	12	18	26	57	99	83	39
18	14	19	16	30	14	6	4	3	4	1	3	5	11	7	4	4	9	3	4
7	4	9	10	3	4	4	8	12	26	14	10	9	10	6	10	12	6	6	6
7	14	30	15	15	7	4	2	4	2	4	4	2	0	2	2	4			

SPEAKER=1 PHONEME= 2 PITCH=174 TM= 52.30 ST= 73.90

2	1	0	4	0	1	0	1	2	5	11	20	65	99	38	16	4	5	1	2
2	2	1	4	0	3	4	13	47	61	22	10	3	4	1	1	3	8	22	6
3	1	1	4	3	1	0	0	0	2	2	0	1	3	0	0	0	0	0	0
2	0	0	3	0	0	2	2	0	2	3	1	3	3	1	1	1	0	0	1
1	0	2	2	1	0	2	0	0	1	1	0	1	2	1	0	3			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 3 PITCH=100 TM= 54.90 ST= 93.40

7	1	1	1	1	3	0	0	2	0	0	0	0	0	6	14	2	1	3	0
0	0	0	0	4	3	1	1	4	1	4	5	2	0	6	14	27	12	21	41
99	38	33	36	36	15	10	9	6	5	2	2	2	0	3	3	2	3	6	5
6	11	14	25	53	81	51	34	17	11	10	11	13	21	40	42	12	13	8	6
2	4	4	1	3	3	2	4	5	6	3	5	0	0	0	1	0			

SPEAKER=1 PHONEME= 3 PITCH=110 TM= 65.20 ST= 96.70

9	24	47	18	8	3	0	1	0	0	0	0	0	1	2	1	1	7	12	1
0	0	0	1	1	1	0	6	2	0	0	1	0	2	6	5	6	10	24	54
18	17	56	99	45	23	14	13	5	5	5	3	2	3	5	5	3	6	4	6
9	15	24	43	85	61	33	27	14	10	11	8	12	23	43	34	12	7	6	6
5	2	6	3	5	1	3	2	3	4	5	1	0	0	1	0	2			

SPEAKER=1 PHONEME= 3 PITCH=123 TM= 48.80 ST= 94.30

0	2	0	5	14	46	19	5	2	0	1	1	0	1	2	1	0	3	2	2
8	40	12	6	2	3	1	4	3	2	12	18	1	3	6	8	20	56	35	17
14	25	99	88	32	18	22	28	6	5	12	6	0	3	4	1	2	5	2	2
6	7	8	20	54	50	62	36	18	10	11	7	5	8	10	7	7	8	4	3
3	5	15	13	8	1	3	3	2	1	2	1	0	1	3	1	4			

SPEAKER=1 PHONEME= 3 PITCH=130 TM= 71.60 ST= 95.40

2	0	2	9	17	31	69	36	22	7	4	2	3	2	3	2	3	2	2	3
6	13	36	33	7	7	5	4	4	5	6	9	23	10	8	11	20	27	88	99
28	19	46	65	90	39	20	14	20	13	5	5	4	4	12	11	5	6	5	11
9	11	22	32	47	62	65	37	17	14	12	10	13	21	23	19	12	12	10	9
3	4	10	6	9	9	6	2	2	4	3	2	2	1	0	1	3			

SPEAKER=1 PHONEME= 3 PITCH=146 TM= 95.90 ST= 96.70

0	0	1	1	3	5	4	5	28	64	17	10	6	2	3	3	1	1	2	1
0	2	2	6	10	47	17	8	8	4	4	4	5	6	26	22	12	10	18	27
50	99	75	29	23	21	43	24	11	9	16	15	7	6	23	10	4	8	5	6
7	5	11	7	15	17	45	40	18	8	6	6	5	5	6	9	9	11	11	7
5	12	32	26	16	6	3	4	7	9	4	5	2	2	4	5	4			

SPEAKER=1 PHONEME= 3 PITCH=164 TM= 95.60 ST= 96.20

1	0	1	1	0	1	1	0	6	7	14	30	63	16	10	5	2	1	2	2
2	1	3	2	1	6	11	28	50	18	10	7	7	6	8	14	36	90	29	17
12	18	53	72	99	47	23	13	10	11	4	5	6	9	4	4	2	2	3	10
6	6	16	11	12	18	21	38	9	6	3	3	4	4	7	15	10	6	8	5
0	3	9	8	8	4	2	0	1	2	0	1	1	0	1	3	2			

SPEAKER=1 PHONEME= 3 PITCH=174 TM= 56.60 ST= 95.30

0	1	0	2	1	1	0	2	1	0	3	7	29	48	17	6	1	4	1	0
0	1	0	1	1	2	2	9	29	41	13	6	1	3	5	6	15	30	99	37
12	7	8	8	28	26	7	5	2	4	9	2	0	1	5	1	1	4	2	1
4	4	4	19	18	32	54	33	23	9	12	5	4	3	4	5	2	4	2	1
1	0	0	0	1	0	2	2	1	0	2	0	0	0	1	0	2			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 4 PITCH=100 TM= 31.60 ST= 94.90

99	21	7	7	5	2	0	1	2	0	4	5	2	3	15	24	31	15	6	0
4	5	4	12	22	94	28	14	13	7	25	51	75	25	14	5	3	6	2	0
3	4	3	4	7	1	0	0	1	0	3	2	2	2	6	1	0	1	0	0
2	3	0	4	7	3	5	10	10	13	39	53	47	45	40	49	23	16	5	3
2	3	3	3	5	2	3	5	0	0	1	2	0	2	3	2	3			

SPEAKER=1 PHONEME= 4 PITCH=110 TM= 55.50 ST= 96.00

7	16	37	23	9	2	0	0	1	0	2	1	1	1	4	1	1	8	34	6
4	2	2	1	3	6	12	41	37	8	5	4	2	2	11	2	0	1	0	0
1	1	2	3	4	1	1	4	0	0	1	1	0	2	3	1	0	3	0	0
1	2	0	4	6	5	7	11	10	11	22	37	70	72	89	99	78	38	17	12
6	6	7	3	7	3	3	4	3	4	5	6	1	2	2	2	3			

SPEAKER=1 PHONEME= 4 PITCH=123 TM= 45.20 ST= 95.20

0	1	1	3	17	40	11	1	4	0	0	0	0	0	2	2	1	2	7	7
20	58	14	3	2	5	4	7	19	31	99	54	15	5	4	4	3	7	4	0
1	1	1	0	1	1	1	1	3	3	1	1	1	1	1	1	0	1	3	1
1	3	0	0	1	2	2	7	7	8	12	19	40	65	37	36	56	83	48	32
21	24	19	5	8	6	3	3	3	6	8	15	11	7	5	4	7			

SPEAKER=1 PHONEME= 4 PITCH=130 TM= 71.10 ST= 96.10

1	0	0	2	3	6	14	16	7	1	2	0	0	0	1	0	1	2	2	1
5	6	17	17	3	2	3	5	10	11	25	38	92	32	14	6	6	4	6	4
4	1	1	1	1	0	1	2	3	2	3	3	2	1	4	1	1	3	0	2
3	2	3	4	3	3	6	9	12	20	48	57	99	41	41	39	29	17	11	20
21	16	11	5	6	2	4	4	2	6	5	5	3	5	2	1	1			

SPEAKER=1 PHONEME= 4 PITCH=146 TM= 51.00 ST= 95.40

0	1	0	0	3	5	3	8	31	70	54	25	11	3	6	4	0	4	5	3
4	8	7	16	20	91	76	32	24	13	10	12	11	15	44	99	29	13	7	5
4	9	6	2	1	1	0	2	2	3	1	4	1	0	1	0	0	1	1	0
2	4	0	0	1	0	0	3	5	4	14	16	28	8	10	16	24	10	7	6
7	10	13	5	7	1	0	0	0	0	1	2	1	0	3	0	0			

SPEAKER=1 PHONEME= 4 PITCH=164 TM= 48.60 ST= 84.70

2	2	1	5	1	2	0	1	3	6	14	26	99	83	34	15	6	5	1	1
2	0	2	4	1	5	9	17	66	56	20	7	4	3	2	2	5	8	19	4
2	0	2	3	6	3	3	0	0	1	1	0	1	2	2	1	4	0	0	0
1	0	0	2	1	1	5	2	1	2	3	3	10	18	7	5	10	5	2	4
0	0	3	1	1	1	4	1	0	3	0	0	0	0	0	1	3			

SPEAKER=1 PHONEME= 4 PITCH=174 TM= 84.20 ST= 90.90

0	1	0	1	2	1	0	1	1	0	3	6	20	49	23	9	2	3	1	0
1	1	0	3	2	7	9	18	48	99	41	18	7	5	4	3	3	5	11	4
0	1	0	1	1	1	0	2	0	0	0	0	0	0	1	0	0	2	0	0
0	0	0	0	0	0	1	3	1	1	2	1	2	4	9	4	6	12	4	3
3	4	5	1	2	0	1	1	0	0	1	0	0	0	0	0	1			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 5 PITCH=100 TM= 90.50 ST= 97.30

34	7	3	2	2	0	0	0	0	0	1	0	0	0	3	4	16	6	1	0
0	0	0	0	1	2	0	1	0	0	1	2	2	1	2	2	2	5	4	4
8	25	30	19	52	99	38	24	42	19	21	35	13	16	22	44	36	42	34	17
7	5	2	3	3	2	2	4	2	1	3	2	2	4	5	5	2	4	3	2
0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1			

SPEAKER=1 PHONEME= 5 PITCH=110 TM= 77.80 ST= 95.40

1	4	15	10	1	1	0	0	0	1	0	0	0	0	0	0	1	4	15	3
0	0	0	0	0	0	0	1	1	1	1	1	1	1	7	3	4	4	8	25
10	11	38	62	46	29	54	99	29	23	40	19	13	10	22	50	26	64	35	17
7	5	4	3	2	2	2	2	3	3	6	5	9	13	9	5	2	2	4	4
1	1	3	1	1	1	2	0	1	1	0	1	1	0	0	0	1			

SPEAKER=1 PHONEME= 5 PITCH=123 TM= 95.50 ST= 96.60

1	1	0	4	11	25	10	1	1	0	1	0	0	1	0	0	0	0	1	3
9	22	5	2	0	1	0	0	0	1	5	2	1	1	1	1	5	10	3	3
4	10	36	16	14	21	56	30	13	16	25	13	26	39	20	27	33	43	99	45
19	12	15	10	5	4	4	4	3	3	4	2	3	2	2	5	5	3	3	4
4	5	6	2	2	1	2	0	0	1	0	0	0	0	0	0	1			

SPEAKER=1 PHONEME= 5 PITCH=130 TM= 57.90 ST= 96.10

1	1	2	4	8	18	40	28	16	5	2	2	2	0	1	0	0	0	2	2
4	10	29	18	3	3	1	1	3	3	5	11	16	5	4	5	8	11	51	29
14	18	56	87	76	39	48	90	99	45	22	37	23	15	24	25	40	91	41	55
20	14	11	8	4	5	3	4	3	4	6	5	5	4	6	8	3	2	3	4
0	2	3	1	2	2	2	0	0	1	0	1	1	0	1	2	2			

SPEAKER=1 PHONEME= 5 PITCH=146 TM= 54.90 ST= 94.80

0	1	1	2	0	2	0	1	18	46	9	4	2	1	1	1	0	0	0	0
0	1	1	5	8	39	10	4	4	2	1	3	3	6	25	22	8	6	9	16
34	99	47	20	27	48	86	41	17	21	44	24	12	16	26	16	23	52	22	42
23	12	7	5	4	3	3	3	3	1	2	2	4	4	15	12	2	2	2	2
0	3	10	9	7	3	2	0	1	1	0	1	0	0	1	1	2			

SPEAKER=1 PHONEME= 5 PITCH=164 TM= 49.70 ST= 96.50

1	0	1	3	1	1	1	0	2	4	11	25	74	19	10	4	2	1	0	0
0	0	1	2	1	6	11	28	68	23	10	5	4	4	5	9	23	59	31	13
6	5	22	30	56	24	17	17	46	99	38	20	26	48	25	19	37	68	24	33
38	15	8	7	2	3	3	4	2	3	2	1	2	1	0	2	3	3	4	6
2	5	12	3	1	1	2	1	1	2	0	0	0	0	0	1	3			

SPEAKER=1 PHONEME= 5 PITCH=174 TM= 51.30 ST= 86.30

0	1	0	2	1	2	0	2	2	1	4	9	34	75	30	12	3	5	2	0
0	0	0	1	1	3	4	11	36	67	23	10	3	4	5	7	15	27	99	49
15	8	11	8	25	33	11	8	10	21	65	25	10	8	26	11	9	24	49	16
22	27	6	4	1	0	2	3	2	2	4	1	1	0	0	0	0	2	2	2
2	1	2	0	0	0	1	1	0	1	2	0	0	0	0	0	2			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 6 PITCH=100 TM= 90.90 ST= 95.80

45	9	3	4	2	1	0	1	0	0	0	0	0	1	2	2	8	4	1	0
0	0	0	1	1	7	4	4	5	4	7	10	30	15	16	21	49	99	70	33
17	22	24	11	23	43	16	11	22	10	5	6	3	4	4	3	2	3	2	4
2	2	2	5	2	1	2	4	1	2	3	1	0	0	1	0	1	2	6	4
1	0	1	0	1	1	1	0	0	0	1	1	0	0	0	0	2			

SPEAKER=1 PHONEME= 6 PITCH=110 TM= 46.30 ST= 91.90

4	17	29	9	2	1	0	0	1	1	2	2	0	0	1	0	0	10	23	3
2	2	0	0	1	1	1	12	6	2	4	6	9	22	55	16	11	16	46	99
26	13	14	14	8	6	16	18	7	15	21	11	30	15	7	2	0	1	0	0
1	2	1	1	2	0	1	2	2	0	3	2	3	5	12	5	1	2	2	2
4	6	6	2	3	1	1	1	0	0	1	1	0	1	2	0	1			

SPEAKER=1 PHONEME= 6 PITCH=123 TM= 95.90 ST= 96.30

1	1	1	4	12	27	10	0	2	1	1	0	1	0	0	0	0	0	2	4
9	27	6	3	0	2	1	2	5	9	30	17	9	8	13	20	52	99	39	23
13	21	65	27	12	10	16	10	6	10	18	10	23	36	12	7	3	2	3	2
2	2	4	4	5	4	2	4	6	8	3	2	3	3	2	3	1	1	1	1
0	3	6	0	1	1	0	0	0	0	0	0	0	0	0	1	0			

SPEAKER=1 PHONEME= 6 PITCH=130 TM= 94.00 ST= 95.60

0	1	1	2	5	16	34	5	4	1	0	1	1	0	1	1	0	0	2	1
2	11	20	6	1	3	1	2	6	7	17	40	21	8	9	11	22	43	99	34
13	10	15	19	7	5	5	11	6	4	11	21	5	3	7	2	0	1	1	1
0	2	0	1	1	1	0	2	2	1	1	1	1	0	0	1	0	0	1	1
0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0			

SPEAKER=1 PHONEME= 6 PITCH=146 TM= 56.40 ST= 93.20

0	1	1	1	1	4	1	3	23	55	11	6	3	1	2	2	1	1	2	0
0	1	1	4	10	47	13	6	7	3	3	5	8	13	61	33	14	9	12	17
37	99	43	16	15	21	41	16	14	22	83	50	20	20	43	16	5	6	4	2
1	1	0	0	2	2	3	4	6	3	5	5	4	3	2	2	0	1	2	1
0	2	4	1	2	1	0	1	1	1	0	2	0	0	0	0	0			

SPEAKER=1 PHONEME= 6 PITCH=164 TM= 59.10 ST= 93.80

1	1	0	2	2	2	1	3	6	5	13	26	99	44	22	9	4	6	3	1
2	1	0	2	2	4	8	20	64	25	13	7	5	6	10	13	36	84	73	29
15	9	13	14	38	11	7	8	19	49	17	10	16	37	23	6	4	5	1	1
3	1	0	1	0	0	1	3	2	1	4	1	1	1	1	0	0	2	1	2
2	0	1	1	0	0	1	1	0	1	2	1	0	1	0	0	1			

SPEAKER=1 PHONEME= 6 PITCH=174 TM= 56.10 ST= 97.30

1	2	1	2	3	1	0	0	3	0	6	11	31	74	56	21	5	6	4	0
2	3	1	2	4	5	6	14	38	88	53	22	9	7	9	9	15	24	79	99
26	14	13	8	10	15	7	7	8	16	69	47	12	8	20	17	3	4	3	0
1	2	0	1	1	1	1	4	1	0	1	1	0	1	2	2	0	3	1	0
0	0	0	0	1	1	1	3	0	0	0	1	0	0	1	1	1			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 7 PITCH=100 TM= 94.60 ST= 96.50

25	14	5	3	1	1	0	0	1	0	0	1	0	0	4	7	28	24	7	2
1	1	0	0	0	3	3	2	2	1	2	4	5	4	4	6	10	24	29	19
25	59	99	40	26	24	14	15	49	29	18	37	12	6	6	3	1	2	1	1
2	4	1	2	1	0	1	6	4	3	5	4	2	1	2	1	0	0	1	0
0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0			

SPEAKER=1 PHONEME= 7 PITCH=110 TM= 95.40 ST= 96.80

5	12	24	10	3	0	0	0	0	0	1	0	0	2	1	1	3	13	27	7
3	1	0	2	0	0	1	7	2	2	2	2	2	4	6	3	5	7	16	40
15	18	59	99	50	29	48	58	22	33	54	28	49	39	22	26	8	8	3	6
3	3	1	3	2	1	2	3	1	3	4	4	8	6	11	8	2	1	0	1
1	1	1	1	1	0	1	0	0	1	1	0	0	1	0	0	1			

SPEAKER=1 PHONEME= 7 PITCH=123 TM= 95.10 ST= 96.60

1	1	1	5	17	43	21	3	4	0	1	1	0	0	2	0	0	1	3	3
10	32	8	4	2	2	1	2	3	3	16	10	3	3	6	7	20	49	16	13
16	32	99	51	28	24	56	38	13	16	41	20	23	46	18	11	10	5	2	1
2	1	0	1	2	1	1	3	2	2	3	3	3	2	2	1	0	1	0	0
0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1			

SPEAKER=1 PHONEME= 7 PITCH=130 TM= 81.90 ST= 95.60

1	0	0	4	7	20	44	14	10	4	1	0	1	0	0	1	1	1	3	4
5	17	40	15	3	4	2	2	4	4	7	15	13	5	5	7	14	24	82	37
16	19	61	99	57	29	24	41	26	15	23	46	18	15	35	17	18	24	7	5
3	5	3	3	1	3	2	3	4	3	3	3	4	2	2	3	0	0	1	1
0	1	3	1	1	1	0	0	0	1	0	0	1	0	0	1	0			

SPEAKER=1 PHONEME= 7 PITCH=146 TM= 71.10 ST= 88.00

0	0	0	1	0	2	2	3	16	41	24	10	4	2	1	1	1	0	1	1
0	0	1	3	4	35	19	7	5	4	3	3	4	5	18	32	13	8	13	20
34	99	84	32	27	34	86	71	23	17	25	36	15	13	46	22	6	4	2	3
2	1	2	1	0	1	1	1	2	2	2	2	4	2	1	3	1	0	1	1
0	0	2	0	0	1	0	0	0	1	0	1	1	0	0	1	0			

SPEAKER=1 PHONEME= 7 PITCH=164 TM= 39.80 ST= 95.40

1	1	3	3	0	2	0	0	3	4	9	24	98	28	13	6	1	0	1	1
0	1	2	1	0	4	4	14	58	20	6	6	5	4	8	15	33	87	65	28
10	11	27	28	80	35	17	15	38	99	37	23	27	50	37	18	10	23	5	3
0	2	2	2	1	4	2	2	3	2	1	3	4	2	5	7	1	1	3	2
0	1	3	0	1	2	2	0	2	1	0	1	0	0	1	2	2			

SPEAKER=1 PHONEME= 7 PITCH=174 TM= 95.70 ST= 96.50

0	0	0	1	1	1	0	1	1	0	1	3	13	31	12	4	1	2	1	0
0	0	0	0	0	1	1	3	10	18	7	4	2	4	6	7	15	25	91	45
17	13	26	25	79	99	41	21	11	10	20	9	6	7	25	10	4	4	4	2
2	1	0	1	0	0	1	2	1	1	3	1	1	1	0	0	0	1	0	1
1	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	1			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 8 PITCH=100 TM= 51.70 ST= 67.90

43	17	6	5	0	2	1	0	1	2	1	0	1	1	3	8	41	25	7	3
1	0	1	1	0	14	18	6	8	10	17	27	99	56	17	10	11	21	12	5
3	5	8	3	2	4	3	2	13	10	20	44	11	3	1	1	0	0	0	1
1	1	2	0	0	0	1	0	2	2	2	1	4	1	0	1	0	0	1	2
0	1	3	0	0	1	1	0	0	1	0	1	1	0	0	2	0			

SPEAKER=1 PHONEME= 8 PITCH=110 TM= 62.10 ST= 92.20

6	14	32	14	4	0	0	1	1	0	2	0	0	0	2	1	6	19	47	12
6	2	0	2	0	3	10	36	22	9	10	10	16	35	99	31	16	9	4	6
3	2	3	6	2	1	2	3	0	5	8	6	21	16	4	2	0	1	0	1
1	1	0	2	0	0	1	1	0	1	2	1	2	2	0	0	0	0	0	1
1	0	1	1	0	0	1	0	0	1	1	0	0	1	0	0	2			

SPEAKER=1 PHONEME= 8 PITCH=123 TM= 82.20 ST= 96.10

2	2	2	8	19	58	24	4	4	2	2	1	1	2	2	1	2	4	6	10
21	69	26	14	4	6	6	8	17	27	74	99	29	16	10	9	18	42	20	9
4	2	8	5	3	2	6	4	3	9	27	9	3	3	2	1	2	0	0	0
1	0	0	1	0	0	2	1	0	1	2	0	1	0	0	0	1	0	1	2
0	0	2	0	0	1	2	0	0	1	0	2	2	0	0	0	2			

SPEAKER=1 PHONEME= 8 PITCH=130 TM= 79.70 ST= 95.90

0	1	1	4	9	24	53	9	8	2	1	1	1	0	2	2	1	2	6	6
9	30	58	19	6	8	6	8	17	22	48	99	57	24	15	11	13	23	54	18
6	4	7	6	3	3	3	5	3	3	8	13	9	16	33	13	3	2	2	1
0	1	0	0	0	1	0	2	2	1	1	1	1	0	0	2	0	1	1	1
0	1	2	0	0	0	0	0	1	1	0	1	0	0	0	0	0			

SPEAKER=1 PHONEME= 8 PITCH=146 TM= 71.40 ST= 95.40

1	1	0	4	1	2	2	3	14	45	28	11	5	4	3	1	1	2	1	3
4	5	6	17	20	89	54	24	16	10	11	9	14	23	64	99	33	14	6	6
6	15	9	6	3	1	2	1	0	2	6	6	2	4	14	3	2	1	0	0
1	0	0	2	0	0	2	1	0	1	2	0	1	2	0	0	1	0	0	1
0	0	2	1	0	0	2	0	0	0	1	0	0	1	0	0	2			

SPEAKER=1 PHONEME= 8 PITCH=164 TM= 70.10 ST= 96.40

1	2	0	1	1	3	1	2	9	10	21	48	93	26	18	9	2	4	5	2
2	4	2	4	4	12	23	59	99	34	19	12	6	4	6	8	20	54	18	8
4	4	5	5	7	2	0	1	3	6	2	3	11	26	9	2	1	2	0	0
1	1	0	0	1	0	0	1	1	0	2	1	1	0	1	0	0	0	1	0
0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0			

SPEAKER=1 PHONEME= 8 PITCH=174 TM= 82.80 ST= 92.70

0	0	1	1	0	0	0	0	1	2	1	1	7	21	8	3	1	0	1	2
1	0	1	2	0	6	9	15	45	99	45	19	9	5	2	3	4	4	20	14
3	1	2	1	2	8	2	0	0	3	1	0	1	1	8	6	1	0	0	2
0	0	0	0	0	0	1	1	1	2	1	0	1	0	0	0	0	0	1	2
0	0	1	0	0	0	1	0	0	2	0	0	0	0	0	0	1			

TABLE XII (Continued)

SPEAKER=1 PHONEME= 9 PITCH=100 TM= 85.80 ST= 96.20

63	19	10	7	5	3	1	0	2	1	1	1	3	3	12	19	50	22	9	3
1	1	1	1	2	14	7	3	5	4	5	9	27	12	14	19	43	99	65	32
19	37	41	14	14	21	8	5	5	4	2	3	3	3	6	11	8	19	19	27
13	12	6	5	5	5	4	5	3	2	2	3	2	1	1	1	0	1	1	1
0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0			

SPEAKER=1 PHONEME= 9 PITCH=110 TM= 90.80 ST= 96.40

8	20	44	19	8	4	0	0	1	0	0	0	1	1	2	3	5	13	37	10
3	2	1	1	0	2	2	7	5	2	2	4	6	11	32	13	12	15	38	99
34	17	22	29	16	9	7	9	4	4	5	4	7	6	15	35	20	50	29	22
7	7	4	2	2	3	4	5	4	1	1	1	3	1	1	2	1	1	3	2
2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0			

SPEAKER=1 PHONEME= 9 PITCH=123 TM= 89.70 ST= 95.80

1	0	1	7	17	52	34	6	6	3	1	0	1	1	1	1	1	2	4	6
11	37	18	8	1	3	3	3	6	7	16	26	10	9	12	18	43	99	71	33
18	24	89	60	22	14	16	14	6	6	8	5	7	9	11	17	37	22	67	43
11	8	6	5	5	4	6	5	5	4	4	5	9	5	2	2	1	1	3	3
0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	1	0			

SPEAKER=1 PHONEME= 9 PITCH=130 TM= 90.00 ST= 95.70

0	0	0	2	3	10	21	19	10	3	1	0	1	1	0	1	1	1	1	2
2	7	20	21	3	4	3	3	4	5	7	11	28	11	8	11	21	31	99	97
29	18	24	28	32	14	9	9	15	8	4	7	8	7	26	21	17	50	21	18
11	8	11	6	4	5	4	4	3	4	4	3	7	3	1	1	0	0	1	1
0	0	2	0	0	0	0	0	0	1	0	0	1	0	0	1	0			

SPEAKER=1 PHONEME= 9 PITCH=146 TM= 94.10 ST= 96.10

0	0	0	2	2	5	5	6	27	73	28	13	8	4	3	3	2	2	3	3
2	3	4	10	12	63	26	12	12	11	13	14	24	34	98	99	40	17	14	17
26	84	44	16	10	9	12	6	3	4	4	3	5	4	12	7	9	29	14	25
33	13	10	7	13	10	13	11	9	4	4	5	6	3	1	3	1	0	2	2
0	1	2	0	1	2	1	0	0	0	0	0	0	0	0	1	0			

SPEAKER=1 PHONEME= 9 PITCH=164 TM= 94.90 ST= 96.20

0	0	0	1	0	1	1	0	3	4	6	14	56	19	9	5	2	1	2	2
1	1	2	2	1	5	7	14	47	20	11	9	11	11	16	23	51	99	97	45
21	15	21	17	44	18	8	4	5	11	4	3	4	3	4	3	6	18	6	10
20	8	5	6	1	2	1	2	1	1	1	2	2	2	1	2	0	0	0	1
0	1	2	0	0	1	0	0	0	0	0	0	0	0	0	1	1			

SPEAKER=1 PHONEME= 9 PITCH=174 TM= 74.70 ST= 77.10

1	1	0	3	0	0	0	0	0	1	3	4	17	33	10	2	1	2	0	1
2	1	0	2	0	2	5	12	41	55	22	9	5	5	5	7	17	29	99	37
15	7	8	6	15	12	4	3	1	3	5	1	1	1	5	1	3	10	13	8
31	24	7	10	3	1	3	3	2	1	3	0	0	0	0	0	0	1	1	1
2	0	1	1	1	0	1	1	0	0	2	0	0	0	1	0	1			

TABLE XII (Continued)

SPEAKER=1 PHONEME=10 PITCH=100 TM= 41.10 ST= 78.50

13	2	2	3	0	3	0	0	1	1	0	1	2	1	4	15	3	0	0	0
0	0	2	1	2	10	1	0	2	1	0	8	5	3	6	13	44	31	19	29
70	99	50	30	78	37	18	25	11	12	20	16	51	28	48	45	21	17	11	5
1	2	1	0	2	3	3	5	6	4	9	10	8	2	4	8	3	2	4	2
0	0	2	0	1	2	1	1	2	1	0	1	0	0	1	1	1			

SPEAKER=1 PHONEME=10 PITCH=110 TM= 68.40 ST= 83.20

7	22	33	11	4	1	0	0	1	0	1	1	0	0	2	1	3	14	28	4
2	2	0	0	1	2	3	15	6	1	3	4	6	13	26	10	12	17	50	99
28	17	29	49	19	9	10	9	4	6	10	10	36	22	30	45	10	7	2	2
1	1	0	1	2	0	1	2	1	0	2	1	1	1	2	1	0	1	1	0
0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	1			

SPEAKER=1 PHONEME=10 PITCH=123 TM= 54.20 ST= 73.80

1	2	0	6	19	54	19	4	3	0	1	1	0	1	3	1	0	2	2	1
9	35	8	4	3	3	1	5	9	14	53	49	13	9	13	16	42	99	42	17
10	12	39	19	10	5	6	5	3	4	19	10	13	31	18	14	23	7	4	0
2	2	1	1	3	0	0	2	2	1	6	6	5	3	6	3	0	0	1	0
0	1	2	0	2	0	0	0	0	0	1	2	1	0	3	0	0			

SPEAKER=1 PHONEME=10 PITCH=130 TM= 64.70 ST= 76.30

3	0	0	5	6	15	36	11	4	2	3	0	0	2	0	0	0	1	0	1
4	9	25	8	1	1	1	3	4	7	20	39	29	13	8	7	17	28	99	41
16	9	17	30	13	4	5	10	6	6	13	29	9	9	31	10	4	5	1	3
3	1	0	1	0	0	1	3	2	2	7	3	3	4	5	0	0	2	1	1
2	0	0	0	0	0	1	1	0	0	2	0	0	0	1	0	1			

SPEAKER=1 PHONEME=10 PITCH=146 TM= 95.80 ST= 95.90

0	0	0	2	1	3	4	4	21	56	25	12	7	3	3	2	1	2	2	2
2	3	4	8	10	56	29	13	13	11	14	15	26	36	98	99	48	21	15	12
10	25	15	6	5	4	6	2	2	3	4	4	5	6	21	10	8	28	11	5
2	1	2	1	1	1	2	2	4	2	2	1	2	0	0	1	1	0	0	1
0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0			

SPEAKER=1 PHONEME=10 PITCH=164 TM= 95.30 ST= 93.50

1	0	0	2	1	1	1	2	5	7	15	31	74	22	13	7	3	4	3	3
4	4	4	8	6	17	32	71	99	45	27	15	10	8	7	8	21	45	15	7
4	3	5	7	7	2	2	1	1	5	1	0	1	2	1	0	3	3	1	6
3	0	0	1	0	0	1	1	0	0	1	0	1	1	0	0	1	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1			

SPEAKER=1 PHONEME=10 PITCH=174 TM= 67.40 ST= 96.30

1	3	0	1	1	1	0	0	3	1	3	8	18	49	29	11	1	3	3	1
1	4	1	1	1	5	4	13	36	84	49	22	9	6	7	9	15	25	84	99
27	17	14	6	6	6	2	3	5	8	32	20	7	6	30	26	6	13	48	14
3	5	1	1	1	1	0	3	5	2	4	4	4	1	6	3	0	1	2	1
0	1	1	0	0	1	0	0	2	1	0	1	0	0	0	1	0			

TABLE XII (Continued)

SPEAKER=1 PHONEME=11 PITCH=100 TM= 33.30 ST= 84.40

91	16	6	9	3	2	1	1	1	2	5	2	4	9	26	49	64	32	11	6
6	5	4	15	19	99	32	13	8	5	7	10	13	6	2	2	10	22	1	1
4	7	4	5	8	3	2	1	0	3	2	0	2	4	2	2	3	1	0	3
11	17	34	29	10	6	5	4	2	2	4	1	8	5	7	15	13	7	2	3
0	0	3	2	0	1	5	1	0	2	0	0	1	1	0	2	4			

SPEAKER=1 PHONEME=11 PITCH=110 TM= 33.20 ST= 56.80

14	30	71	38	11	3	1	1	0	2	4	1	1	5	5	4	12	34	99	39
14	5	2	5	1	4	8	31	26	7	4	2	2	5	16	3	2	2	0	10
3	0	2	6	2	1	3	3	0	3	2	0	2	3	2	3	5	10	7	28
8	4	5	12	4	2	6	4	2	3	6	4	8	7	10	22	36	15	5	5
0	0	3	1	0	1	4	1	0	3	0	0	0	1	0	2	3			

SPEAKER=1 PHONEME=11 PITCH=123 TM= 67.30 ST= 78.90

1	2	3	9	25	67	31	5	6	2	1	2	2	1	3	3	2	5	10	14
30	99	33	16	5	7	2	2	5	6	23	18	4	1	1	3	3	11	2	0
0	1	7	2	1	2	5	1	1	1	0	0	1	0	1	2	2	5	23	5
3	1	1	0	1	1	2	3	4	3	7	7	7	4	9	19	30	19	6	4
1	2	2	0	1	1	0	0	1	1	0	1	0	0	0	0	0			

SPEAKER=1 PHONEME=11 PITCH=130 TM= 65.30 ST= 87.20

2	2	1	8	17	31	76	38	24	7	4	4	2	3	5	4	3	8	10	11
18	40	99	65	20	14	9	8	7	4	7	13	16	5	3	1	2	4	16	5
2	1	2	5	3	0	1	1	1	1	2	0	0	0	2	0	5	19	5	14
5	2	0	1	0	0	1	2	1	1	4	1	1	1	0	1	3	2	0	2
1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	0	1			

SPEAKER=1 PHONEME=11 PITCH=146 TM= 50.10 ST= 86.80

1	2	0	3	5	6	7	9	39	99	53	26	13	6	8	3	1	3	2	0
1	3	2	11	16	84	39	18	11	4	3	3	0	3	14	15	3	4	1	0
1	8	1	2	1	0	1	2	0	1	1	0	0	1	1	0	1	3	0	0
2	0	0	0	0	0	2	3	2	1	4	1	2	1	1	0	1	2	1	1
2	0	1	0	0	0	1	1	0	1	2	0	0	1	0	0	1			

SPEAKER=1 PHONEME=11 PITCH=164 TM= 82.30 ST= 84.00

1	0	0	2	1	1	1	1	6	7	14	28	99	34	18	9	3	4	1	1
1	1	0	2	0	2	4	13	43	15	6	3	1	1	0	1	2	8	3	1
1	0	1	2	2	0	0	0	0	1	0	0	0	1	0	0	1	1	0	0
0	0	0	1	0	0	2	1	0	0	1	0	1	1	0	0	1	0	0	1
0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1			

SPEAKER=1 PHONEME=11 PITCH=174 TM= 45.50 ST= 75.80

0	1	1	2	1	5	1	0	6	4	7	16	44	99	94	38	11	10	9	2
0	1	1	0	1	5	5	10	34	82	48	18	6	1	3	4	3	4	18	21
2	1	1	0	2	6	1	1	3	2	0	1	0	0	1	1	0	1	6	1
0	4	0	0	0	1	0	3	3	2	2	3	2	0	1	2	0	8	5	4
0	3	1	0	0	1	0	1	1	1	1	3	0	0	1	0	0			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 1 PITCH=100 TM= 57.80 ST= 89.70

66	12	7	4	2	1	0	0	1	1	1	1	2	1	4	10	31	12	3	1
0	0	1	0	0	4	1	0	1	1	1	5	14	4	2	3	2	3	3	5
9	33	33	20	58	94	33	19	29	13	6	7	6	9	6	6	3	3	3	5
6	10	11	16	24	37	64	99	47	24	17	13	15	22	46	38	13	9	7	7
6	14	18	6	7	4	5	4	4	4	5	4	2	1	2	2	2			

SPEAKER=2 PHONEME= 1 PITCH=110 TM= 67.10 ST= 85.90

13	34	64	22	9	4	2	1	1	2	1	0	0	1	1	3	7	21	42	10
4	1	1	0	0	1	2	10	5	2	1	1	2	4	9	3	3	2	6	11
8	14	55	99	44	24	39	47	13	14	17	7	9	5	4	5	2	1	2	4
4	7	10	16	25	32	42	59	24	15	12	11	12	19	35	41	14	7	4	5
2	8	19	7	6	4	4	2	2	2	2	3	3	2	3	3	2			

SPEAKER=2 PHONEME= 1 PITCH=123 TM= 58.80 ST= 85.30

1	2	1	8	23	69	55	11	9	2	1	1	0	1	2	1	1	2	3	3
10	39	23	8	2	2	1	2	3	3	12	26	3	2	3	2	5	13	10	7
10	19	82	76	29	19	38	59	15	10	15	8	4	5	5	2	4	3	3	3
4	5	6	15	23	32	58	61	99	37	23	14	9	7	11	19	34	41	13	9
4	4	9	10	17	6	3	2	1	2	3	3	1	2	1	1	2			

SPEAKER=2 PHONEME= 1 PITCH=130 TM= 60.30 ST= 92.50

0	1	2	7	17	43	87	15	12	4	1	1	2	1	1	2	0	1	2	3
6	25	38	11	2	4	1	1	3	2	7	24	6	2	2	3	4	7	15	7
7	14	59	99	36	23	36	78	29	16	11	14	6	5	10	4	3	3	3	3
4	9	7	16	17	34	46	63	61	26	16	10	7	5	9	16	9	4	4	3
1	2	5	2	3	2	2	1	2	1	1	1	0	0	0	0	1			

SPEAKER=2 PHONEME= 1 PITCH=146 TM= 87.00 ST= 97.00

1	2	1	3	6	8	9	12	52	99	27	18	10	5	6	3	1	2	3	1
2	3	4	11	17	53	16	8	7	3	4	4	4	7	32	13	7	6	9	14
33	81	34	15	20	32	67	25	10	7	9	5	4	4	6	2	2	4	2	2
2	4	7	7	14	15	32	35	64	25	14	9	5	3	5	7	11	10	4	3
1	1	4	3	4	2	3	3	2	1	1	1	0	0	1	0	1			

SPEAKER=2 PHONEME= 1 PITCH=164 TM= 67.20 ST= 89.20

1	0	0	3	1	1	2	1	5	6	14	29	99	30	17	8	4	3	1	0
1	0	0	1	1	2	4	10	28	7	3	0	0	1	2	3	9	25	10	4
2	1	12	17	42	15	7	3	3	11	2	0	1	4	1	1	2	1	0	2
1	0	3	6	7	16	21	54	23	13	7	3	3	3	2	5	9	13	4	5
1	1	4	2	1	1	2	1	0	2	0	0	0	0	0	0	2			

SPEAKER=2 PHONEME= 1 PITCH=174 TM= 62.80 ST= 86.50

2	0	0	1	1	0	1	2	5	4	10	15	44	99	42	17	6	7	4	1
3	1	0	1	0	0	1	3	11	28	11	4	0	1	1	0	3	7	29	15
6	5	13	15	50	84	27	13	4	6	7	3	2	3	6	1	0	1	0	1
3	4	2	11	7	8	24	17	26	11	9	5	4	4	2	5	6	10	0	2
1	0	2	2	1	0	1	0	0	1	1	0	0	2	1	0	2			

TABLE XII (Continued)

```
SPEAKER=2    PHONEME= 2    PITCH=100    TM= 32.00    ST= 74.50
99 25 13 11  5  8  1  0  2  1  0  1  4  4 21 40 74 34 13  3
 0  1  2  0  2 28 11  2  5  2  1  8 23  5  4  8 16 45 20 13
21 78 52 16 14 14  6  2  7  4  1  1  2  0  2  3  3  1  5  2
 0  1  4  4 15 29 17 14 10  5  4  4  6  2 12 16  5  4  8  5
 2  5  5  0  1  2  1  1  4  2  0  2  0  0  2  3  3

SPEAKER=2    PHONEME= 2    PITCH=110    TM= 95.20    ST= 95.40
28 57 99 48 26 11  7  2  5  3  3  3  5  4  9 10 17 39 93 28
15  7  4  3  1  3  6 17 11  4  3  3  4  7 23  7  6  6 18 52
18 10 24 42 23 10  6  5  2  3  2  1  3  2  2  6  2  6  3  4
 2  2  2  3  5  8 16 35 25 15 11  6  5  5 10 15 15  8  4  4
 2  2  6  4  3  0  1  0  0  1  2  2  2  1  1  0  2

SPEAKER=2    PHONEME= 2    PITCH=123    TM= 74.50    ST= 92.10
 3  4  4 14 39 99 36 10 10  3  2  2  1  1  3  2  1  3  5  7
17 55 15  7  2  3  2  3  6  9 33 26  6  5  7 11 29 69 21 10
 7 13 41 17  8  4  5  3  1  1  1  1  0  1  2  1  1  1  3  1
 2  2  2  2  3  3  5 10 20 33 26 14  7  7  9 15 15  7  5  4
 3  7 15 11  8  3  2  1  1  0  1  2  1  1  3  1  1

SPEAKER=2    PHONEME= 2    PITCH=130    TM= 95.30    ST= 95.40
 2  2  3  8 17 39 86 16 14  5  3  3  3  2  2  1  1  1  4  4
 8 24 48 16  4  4  3  2  5  6 12 27 17  8  7 10 20 35 99 38
14 10 25 40 15  7  4  3  1  2  1  1  2  1  2  1  2  1  1  3
 1  2  2  3  2  3  6  9 27 19 18  9  6  3  3  7 10 13  7  6
 2  3  8  4  3  1  1  0  0  1  0  1  1  1  0  1  1

SPEAKER=2    PHONEME= 2    PITCH=146    TM= 70.90    ST= 95.20
 0  1  1  3  4  7  7  9 46 83 21 13  8  4  4  3  1  1  2  2
 1  4  5 14 23 66 20 10  9  6  7  9 16 28 99 42 18  7  4  3
 3  6  4  2  2  3  3  0  0  2  0  0  1  1  1  1  0  0  0  1
 0  1  1  2  1  2  3  3  6 10 20 34 22 14 10 12 16 17  7  7
 2  5 14  4  3  2  1  0  0  1  0  0  1  0  0  1  2

SPEAKER=2    PHONEME= 2    PITCH=164    TM= 91.20    ST= 96.40
 1  0  0  2  2  3  4  3  9 11 25 55 68 23 16  9  4  4  3  2
 2  2  2  4  3 10 21 51 60 22 12  7  7  7 10 17 47 99 30 16
 7  4  6  4  3  2  2  1  3  5  1  0  1  1  0  0  1  1  0  2
 1  0  3  2  2  2  4  5  6 11 24 26 13  6  4  9  7  7  4  4
 1  1  4  3  2  1  1  0  0  1  0  0  0  0  0  0  1

SPEAKER=2    PHONEME= 2    PITCH=174    TM= 51.80    ST= 86.80
 0  1  1  2  1  3  1  0  4  5  9 20 71 99 41 19  6  5  4  2
 1  1  2  2  0  4  4 10 39 58 18  9  5  4  5  7 14 24 93 32
11  6  8  5 16 14  3  0  1  3  1  0  2  1  0  0  1  0  0  2
 2  1  2  3  1  2  4  6 14 25 62 30 31 18 21 41 32 35 12 11
 6  9 21 11  7  2  3  0  1  2  1  1  1  1  1  0  3
```

TABLE XII (Continued)

SPEAKER=2 PHONEME= 3 PITCH=100 TM= 25.60 ST= 63.50

92	16	7	5	3	5	1	1	5	2	1	1	3	2	22	44	83	36	15	3
1	1	1	0	3	40	11	5	8	4	6	20	42	10	8	14	35	99	34	20
23	80	50	15	14	13	4	3	6	6	2	2	3	0	2	3	1	3	4	4
3	5	3	4	7	13	28	55	92	43	27	19	15	11	20	39	46	28	15	12
6	10	19	4	5	3	3	2	3	3	1	3	1	0	1	3	4			

SPEAKER=2 PHONEME= 3 PITCH=110 TM= 79.20 ST= 95.60

7	18	40	16	5	2	1	0	0	0	0	0	0	0	1	1	4	13	27	6
2	0	0	1	0	2	4	16	8	3	3	3	5	11	33	12	12	15	36	99
31	15	18	21	10	5	5	4	1	2	2	0	2	2	1	2	1	2	0	2
1	1	2	4	5	7	16	34	29	17	9	4	4	3	5	8	9	4	2	2
1	1	4	1	1	0	2	0	0	0	0	0	0	0	0	0	3			

SPEAKER=2 PHONEME= 3 PITCH=123 TM= 38.90 ST= 85.60

4	2	1	11	36	88	31	7	7	2	4	1	0	1	2	0	1	3	4	6
19	60	12	5	1	2	2	4	7	9	41	31	7	5	9	13	43	99	33	14
10	14	55	25	9	3	4	4	2	3	3	1	0	1	2	1	1	3	2	1
3	2	1	4	7	7	16	30	45	71	41	22	14	15	16	24	38	40	19	13
7	11	26	36	32	16	8	5	2	2	1	2	0	2	4	4	4			

SPEAKER=2 PHONEME= 3 PITCH=130 TM= 79.40 ST= 87.40

1	2	3	7	17	36	86	23	18	6	2	3	3	1	3	3	2	2	4	4
7	21	59	23	6	7	4	4	7	7	16	33	32	11	8	9	17	30	99	37
12	7	13	16	7	4	3	1	2	2	0	0	1	0	2	1	2	1	2	3
1	1	1	1	3	5	11	19	44	32	17	9	6	3	7	13	23	21	12	8
4	6	16	18	13	5	3	1	2	1	0	1	0	0	0	1	2			

SPEAKER=2 PHONEME= 3 PITCH=146 TM= 51.20 ST= 81.90

1	2	1	1	3	5	4	8	43	99	25	15	7	3	4	2	0	1	2	0
1	3	2	7	13	62	19	8	7	2	4	6	8	14	56	45	15	6	6	7
21	68	30	11	8	6	12	5	2	2	2	2	1	1	3	1	0	3	1	0
1	2	1	2	4	4	8	12	25	25	60	31	15	11	10	13	16	22	12	8
3	5	11	10	8	3	2	2	1	1	1	1	0	0	2	0	2			

SPEAKER=2 PHONEME= 3 PITCH=164 TM= 51.50 ST= 66.90

1	1	2	2	0	2	0	0	5	6	14	33	93	26	14	7	1	1	1	1
0	1	2	1	0	4	6	17	42	12	4	4	5	5	8	15	37	99	42	18
6	5	12	15	26	8	3	1	1	3	0	0	1	1	1	2	1	0	1	1
0	0	1	1	2	4	6	13	21	43	19	11	7	4	6	9	15	30	13	9
3	4	10	12	23	15	6	2	2	1	0	1	0	0	0	1	1			

SPEAKER=2 PHONEME= 3 PITCH=174 TM= 87.90 ST= 94.40

0	0	1	1	0	1	0	0	2	2	3	7	28	44	14	6	1	1	1	1
0	1	1	1	0	3	3	7	26	36	12	6	4	4	5	9	17	29	99	47
15	9	11	6	10	13	4	1	1	2	1	1	1	1	1	1	0	0	0	1
0	1	1	2	1	2	4	5	12	12	34	15	10	4	3	6	7	9	4	3
0	1	4	4	2	1	1	0	0	0	0	0	0	0	1	2	2			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 4 PITCH=100 TM= 47.10 ST= 82.80

70	13	7	6	3	4	1	0	1	1	0	1	4	6	20	36	56	25	10	4
2	3	4	9	12	72	29	13	14	13	24	45	99	34	14	10	7	10	3	1
1	1	2	1	1	2	1	0	2	2	0	0	1	0	1	2	1	0	2	1
0	1	0	0	1	2	3	5	11	19	38	18	9	5	9	15	6	4	5	6
6	6	4	0	1	1	0	1	1	2	1	2	1	0	0	1	3			

SPEAKER=2 PHONEME= 4 PITCH=110 TM= 47.60 ST= 90.50

12	30	61	21	7	3	1	0	2	2	2	1	3	3	7	9	18	48	92	25
12	7	5	6	3	17	35	99	55	21	10	7	4	3	4	4	1	1	1	1
0	1	2	3	1	2	1	0	1	2	0	0	1	0	1	2	1	1	2	1
0	1	1	0	1	2	3	4	7	9	22	38	17	8	11	17	23	35	29	22
23	39	24	6	7	3	1	1	2	2	1	2	1	0	1	1	1			

SPEAKER=2 PHONEME= 4 PITCH=123 TM= 69.10 ST= 77.90

2	1	3	11	28	76	37	7	7	2	2	1	2	3	4	3	4	6	10	15
33	99	34	17	6	7	6	7	13	22	67	52	15	7	3	3	3	3	1	1
1	0	2	2	0	0	1	1	0	2	1	0	0	1	0	0	1	0	0	1
1	0	0	1	0	0	2	3	4	8	19	20	16	9	8	14	22	11	7	7
6	11	22	4	4	1	1	0	0	1	1	0	1	0	0	0	1			

SPEAKER=2 PHONEME= 4 PITCH=130 TM= 68.20 ST= 87.10

1	0	1	6	12	40	57	10	8	4	2	1	1	1	1	2	2	4	6	10
17	47	63	22	6	7	6	7	14	20	47	99	38	17	8	5	5	3	4	2
2	1	3	3	0	0	1	0	0	2	2	1	2	1	0	0	1	0	0	2
1	2	2	3	1	2	3	4	10	16	30	37	18	11	11	21	37	41	19	14
14	28	33	8	8	4	3	1	1	2	0	1	1	2	2	3	3			

SPEAKER=2 PHONEME= 4 PITCH=146 TM= 83.90 ST= 95.40

0	1	0	2	3	4	4	6	24	61	34	17	8	4	5	3	2	4	4	3
4	6	7	17	23	99	55	26	18	9	7	5	2	2	5	5	2	2	1	0
1	2	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0
1	0	0	0	0	0	1	3	3	2	6	9	16	18	9	7	9	14	15	12
9	15	18	4	4	0	1	1	1	1	1	1	0	0	0	0	1			

SPEAKER=2 PHONEME= 4 PITCH=164 TM= 79.80 ST= 92.70

0	1	0	1	1	1	0	0	4	3	9	19	64	36	19	9	2	3	3	1
1	2	2	3	3	10	16	32	99	63	27	15	7	4	4	3	1	2	1	0
1	1	1	1	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	0
0	1	0	0	0	1	1	3	4	6	12	22	16	14	7	9	13	17	12	10
8	15	28	8	7	2	1	0	1	1	0	1	0	0	0	1	2			

SPEAKER=2 PHONEME= 4 PITCH=174 TM= 75.70 ST= 92.10

1	0	0	2	1	1	1	1	4	5	10	16	50	95	41	18	6	7	4	2
3	2	1	4	2	6	10	19	59	99	39	18	8	6	3	2	2	3	9	3
2	1	1	2	1	1	0	0	0	1	1	0	0	1	0	0	0	1	0	0
1	1	0	1	0	0	2	3	4	5	13	18	34	22	16	31	20	13	7	7
5	8	19	9	6	0	1	1	0	1	1	1	0	0	1	0	2			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 5 PITCH=100 TM= 53.20 ST= 88.60

61	13	6	8	4	2	0	1	0	1	3	3	3	7	20	35	49	24	8	3
3	2	1	4	3	20	6	3	1	2	3	5	10	5	4	4	9	19	8	11
24	69	47	32	90	99	40	39	49	20	35	32	38	66	37	63	20	10	4	4
2	1	1	2	0	2	4	4	2	4	3	1	1	1	0	4	4	4	2	4
1	0	3	2	1	1	2	0	0	2	1	0	1	0	0	0	2			

SPEAKER=2 PHONEME= 5 PITCH=110 TM= 82.80 ST= 94.70

8	18	36	16	6	2	1	0	0	0	1	0	0	1	2	2	6	15	37	9
4	0	0	1	1	1	3	10	5	2	2	1	2	5	13	5	6	6	13	39
15	17	58	99	52	27	32	39	11	14	29	18	42	55	20	19	7	5	2	4
1	1	1	2	1	1	2	2	2	3	4	3	5	2	1	2	1	0	1	2
1	0	2	0	0	0	1	0	0	1	0	0	1	0	0	0	0			

SPEAKER=2 PHONEME= 5 PITCH=123 TM= 90.60 ST= 95.60

1	2	3	9	25	53	20	3	5	2	0	1	1	1	1	1	0	1	3	5
13	40	10	5	0	3	1	0	1	3	8	6	3	2	2	5	10	23	9	10
12	30	99	45	23	22	47	28	9	8	13	10	21	38	24	43	51	17	9	5
1	2	3	2	2	3	3	4	5	5	4	3	4	3	10	14	7	3	3	2
0	2	6	0	2	1	2	0	1	0	0	0	0	0	0	1	1			

SPEAKER=2 PHONEME= 5 PITCH=130 TM= 95.70 ST= 96.00

1	1	0	4	11	28	56	11	9	2	2	1	0	0	1	0	0	1	2	2
6	16	27	7	2	1	1	2	2	2	6	13	4	2	3	3	8	19	32	13
12	18	65	99	38	20	20	38	12	8	10	14	10	16	28	17	35	33	8	5
3	2	1	4	2	3	3	5	4	3	4	3	3	3	3	5	2	2	0	1
1	1	1	1	1	0	1	0	0	0	0	0	0	1	0	0	1			

SPEAKER=2 PHONEME= 5 PITCH=146 TM= 91.30 ST= 96.80

2	1	0	3	5	6	9	11	43	99	31	18	9	5	5	2	1	3	2	1
3	2	1	7	10	45	16	8	6	3	4	3	1	2	8	7	6	7	10	16
32	93	53	21	23	34	94	50	17	14	33	27	12	12	27	14	21	59	22	11
6	4	6	4	4	2	4	5	5	3	7	3	2	3	4	6	13	13	5	4
3	5	13	6	4	0	2	2	2	2	3	2	0	0	1	0	2			

SPEAKER=2 PHONEME= 5 PITCH=164 TM= 50.30 ST= 94.40

0	2	0	0	1	3	0	2	10	10	19	49	99	28	18	10	2	4	5	1
0	2	1	0	0	6	8	28	49	13	5	4	2	1	4	9	24	65	24	10
5	7	22	32	51	19	10	11	35	53	11	9	14	27	20	23	52	76	22	13
4	4	4	2	2	4	2	5	5	4	3	3	3	0	3	8	10	13	9	8
8	21	44	17	13	4	2	1	3	2	0	2	0	0	0	1	1			

SPEAKER=2 PHONEME= 5 PITCH=174 TM= 65.60 ST= 96.60

1	0	0	2	1	1	2	1	4	4	10	15	53	99	46	20	8	6	4	1
2	0	1	2	1	2	4	7	22	47	18	6	3	4	4	6	12	20	70	36
12	8	21	20	66	98	33	17	14	23	50	20	13	16	52	26	17	35	46	16
5	3	2	4	1	2	5	4	5	4	6	2	3	2	1	6	14	15	5	7
4	11	16	4	2	1	2	1	1	3	1	0	0	0	0	0	2			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 6 PITCH=100 TM= 61.10 ST= 79.20

44	7	1	3	2	1	0	2	1	0	1	0	0	2	11	19	43	21	7	1
0	1	0	1	2	15	5	4	4	3	8	18	37	15	13	17	44	99	31	16
17	49	21	9	15	10	4	7	13	7	25	24	13	18	7	2	0	0	1	0
1	2	1	0	1	0	0	2	2	1	2	3	2	0	4	3	0	1	2	2
0	2	1	0	0	0	0	0	1	1	0	1	0	0	0	1	0			

SPEAKER=2 PHONEME= 6 PITCH=110 TM= 83.10 ST= 96.80

12	21	46	38	15	6	4	1	2	1	1	0	1	2	4	4	8	19	51	15
7	2	2	3	1	4	9	26	23	9	9	9	17	33	99	40	19	11	13	36
14	6	8	12	7	3	4	8	3	10	26	11	11	14	4	1	0	2	0	2
2	3	0	2	1	0	1	2	2	1	3	1	1	1	4	2	0	1	2	1
1	0	1	0	1	0	1	1	0	0	0	0	0	0	1	1	0			

SPEAKER=2 PHONEME= 6 PITCH=123 TM= 40.20 ST= 73.60

3	2	0	11	31	97	57	12	10	2	3	2	0	1	4	1	1	6	7	7
21	76	31	13	4	5	3	8	12	16	57	78	17	11	13	16	41	99	55	22
10	8	20	15	8	6	16	24	5	7	28	14	8	24	14	3	2	4	1	0
1	1	0	1	2	0	2	5	3	1	3	2	1	5	13	7	3	7	10	9
2	1	1	0	2	1	1	3	1	0	0	1	0	0	2	1	1			

SPEAKER=2 PHONEME= 6 PITCH=130 TM= 63.20 ST= 96.00

2	3	2	8	18	43	99	23	18	6	3	4	2	1	2	2	0	3	7	8
13	40	67	22	6	7	4	6	12	16	38	84	42	18	10	9	13	30	58	20
7	7	17	24	7	6	7	20	8	10	31	37	9	4	5	2	0	1	2	2
1	3	0	0	0	1	0	2	2	2	1	2	2	0	4	7	0	1	3	4
4	2	2	0	0	1	0	0	1	1	0	1	0	0	0	1	0			

SPEAKER=2 PHONEME= 6 PITCH=146 TM= 91.00 ST= 95.80

1	1	2	4	5	7	10	12	57	64	18	12	8	4	4	3	1	2	3	3
3	5	7	17	29	56	20	11	10	8	10	12	21	37	99	41	20	9	7	8
14	22	10	5	4	7	8	3	4	12	26	11	6	7	5	3	1	1	0	2
0	1	1	1	0	1	1	1	1	1	0	1	2	0	2	11	3	1	2	2
0	0	2	0	1	1	1	0	0	0	0	0	0	0	0	1	1			

SPEAKER=2 PHONEME= 6 PITCH=164 TM= 53.60 ST= 82.10

0	0	1	0	1	3	2	1	8	7	15	37	99	25	16	8	2	3	4	1
1	3	2	1	2	10	15	43	75	22	10	9	5	4	6	13	33	88	31	14
6	6	7	7	11	4	2	3	14	30	5	5	12	24	8	2	0	0	1	0
1	2	1	0	2	0	0	2	2	0	2	2	1	1	3	2	0	3	2	2
3	7	3	1	2	1	1	2	0	0	0	1	0	0	2	1	1			

SPEAKER=2 PHONEME= 6 PITCH=174 TM= 69.90 ST= 94.10

0	0	0	2	0	2	2	2	6	7	11	19	62	99	38	19	7	6	5	3
3	2	3	2	1	5	8	15	52	79	28	13	8	4	3	5	9	17	65	26
7	5	8	5	16	19	5	2	3	8	13	6	6	8	22	8	1	0	0	0
0	0	0	0	0	1	1	1	2	1	1	0	1	0	1	6	2	1	3	2
0	0	1	0	0	1	0	0	1	1	0	1	0	0	0	0	0			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 7 PITCH=100 TM= 36.40 ST= 94.30

78	16	6	7	3	1	0	1	0	1	4	2	2	7	17	28	94	47	14	5
4	3	1	5	2	18	9	4	1	2	6	10	29	13	5	5	12	26	13	17
34	99	80	38	74	99	35	26	47	21	16	17	34	72	38	45	21	12	4	5
2	0	1	4	1	2	5	5	3	5	4	1	2	2	0	5	7	5	2	6
2	1	3	1	0	1	3	1	1	4	1	0	1	0	0	0	4			

SPEAKER=2 PHONEME= 7 PITCH=110 TM= 78.20 ST= 96.20

23	51	99	39	20	9	4	2	5	2	2	3	3	3	8	9	14	38	84	24
13	7	3	2	2	3	4	17	9	2	3	4	5	11	23	8	7	10	29	56
18	19	58	98	45	24	24	22	9	12	20	13	31	27	43	83	24	21	8	7
3	4	2	5	6	5	5	8	5	3	4	5	4	3	2	3	3	2	1	1
1	1	1	0	1	0	1	1	1	1	2	1	0	0	1	0	1			

SPEAKER=2 PHONEME= 7 PITCH=123 TM= 68.60 ST= 97.50

3	3	5	13	34	94	62	12	14	4	2	2	2	0	2	2	1	2	5	5
12	42	16	5	1	3	0	2	4	4	12	18	3	2	3	6	11	33	23	13
13	26	99	75	28	17	22	23	9	10	30	19	21	43	32	23	43	17	6	3
3	4	4	5	7	5	7	6	7	4	6	6	4	2	7	13	7	4	4	5
7	6	3	1	3	1	2	2	1	0	0	1	0	0	1	1	1			

SPEAKER=2 PHONEME= 7 PITCH=130 TM= 46.80 ST= 85.80

3	1	0	6	12	29	75	19	11	4	3	2	0	3	1	0	0	1	0	1
5	16	45	18	4	2	1	2	0	2	5	11	9	5	4	5	10	18	68	25
12	13	59	99	50	22	21	32	16	12	20	41	24	33	94	41	32	41	9	8
3	2	1	4	1	3	5	6	4	5	5	2	2	2	2	14	27	12	6	10
8	11	13	5	2	1	3	0	1	4	1	0	1	0	0	0	2			

SPEAKER=2 PHONEME= 7 PITCH=146 TM= 72.60 ST= 96.10

2	1	0	3	5	6	8	11	40	99	50	25	13	7	6	2	1	3	2	1
2	2	2	7	8	50	21	10	5	3	4	4	3	8	24	27	10	7	7	14
28	87	50	21	17	20	55	31	8	8	20	21	12	16	54	26	10	15	6	3
3	1	4	5	7	5	9	8	8	5	9	4	5	3	3	8	16	10	3	6
6	9	8	3	2	2	3	2	1	2	0	0	0	0	0	0	2			

SPEAKER=2 PHONEME= 7 PITCH=164 TM= 86.90 ST= 96.10

1	1	1	2	3	3	2	2	10	10	22	49	99	32	22	11	4	5	4	1
2	2	0	2	2	5	7	21	40	12	6	4	1	2	4	6	16	41	18	8
6	7	21	30	61	24	14	13	29	58	16	10	15	30	16	8	17	23	6	3
3	1	1	1	2	1	2	5	2	4	4	4	3	3	2	3	9	24	6	6
6	12	10	3	2	0	1	1	0	1	1	1	0	1	0	0	1			

SPEAKER=2 PHONEME= 7 PITCH=174 TM= 77.10 ST= 97.00

0	0	0	1	0	1	1	0	4	5	8	14	45	90	38	17	6	4	4	2
1	1	2	1	0	3	2	4	17	37	11	6	4	3	5	8	12	21	76	46
13	10	19	16	54	99	31	14	10	11	29	15	12	15	57	33	10	15	26	8
4	7	2	1	1	1	2	3	5	2	3	3	2	1	3	7	11	20	6	4
4	8	5	1	3	0	0	0	0	0	0	1	0	0	1	0	1			

TABLE XII (Continued)

SPEAKER=2 PHONEME= 8 PITCH=100 TM= 85.50 ST= 89.10

57	13	8	7	4	4	1	0	2	1	0	2	5	5	16	26	67	33	14	6
4	4	4	7	10	51	35	16	15	12	21	34	99	39	16	10	7	7	4	2
1	2	2	0	1	2	0	1	3	2	1	2	1	0	0	1	0	0	1	1
2	6	11	10	1	1	0	2	2	1	2	3	5	1	1	0	0	0	1	0
0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	1			

SPEAKER=2 PHONEME= 8 PITCH=110 TM= 75.60 ST= 95.10

19	40	83	37	19	7	6	2	3	1	2	1	3	4	10	11	20	43	99	29
17	9	6	8	5	16	34	88	62	26	17	11	14	25	70	24	12	6	3	5
2	1	3	5	1	1	2	2	1	3	2	0	1	1	0	1	1	0	1	3
5	12	8	10	2	1	3	3	4	4	6	8	16	6	3	4	3	1	5	8
1	1	4	1	0	1	1	0	0	1	0	0	1	0	0	1	1			

SPEAKER=2 PHONEME= 8 PITCH=123 TM= 84.10 ST= 96.30

2	2	4	10	29	62	23	4	6	3	1	2	3	2	4	4	3	6	12	18
36	95	28	16	6	9	7	9	21	37	99	60	21	11	7	6	5	4	2	1
0	1	3	1	1	2	1	0	1	1	0	0	1	0	0	1	1	0	2	1
1	2	11	6	7	4	3	3	3	2	3	4	6	5	2	2	1	1	3	3
4	2	2	0	0	0	0	0	1	1	0	1	0	0	0	0	0			

SPEAKER=2 PHONEME= 8 PITCH=130 TM= 93.60 ST= 96.40

3	3	2	8	18	32	69	38	26	9	6	5	3	3	5	3	3	6	7	8
14	29	71	45	14	10	8	10	15	18	37	60	99	37	19	10	9	6	7	4
3	2	3	4	3	0	1	1	1	1	2	1	0	1	0	0	0	1	0	1
3	6	7	6	2	1	1	4	1	2	3	2	3	7	2	1	1	1	0	2
3	4	2	1	1	0	1	0	0	0	0	0	0	0	0	0	2			

SPEAKER=2 PHONEME= 8 PITCH=146 TM= 76.50 ST= 95.00

0	1	0	2	3	4	5	7	28	67	33	16	8	4	5	2	2	3	4	2
5	6	6	17	23	99	49	24	16	8	8	5	4	8	28	25	8	4	0	1
2	3	1	2	1	0	1	1	0	1	1	0	0	2	1	0	1	1	0	1
3	5	21	10	7	3	4	2	2	1	3	2	5	13	5	3	3	2	2	6
10	8	5	1	0	0	1	1	0	2	0	0	0	0	0	0	1			

SPEAKER=2 PHONEME= 8 PITCH=164 TM= 76.10 ST= 93.70

0	1	0	2	3	4	2	4	11	12	25	58	90	29	20	11	4	7	5	2
3	4	2	5	5	13	27	64	99	33	20	10	5	3	3	2	6	14	3	1
2	1	1	2	1	0	0	1	0	2	1	0	0	1	0	0	0	1	0	2
4	8	24	10	3	1	1	2	0	1	2	1	2	2	2	1	1	0	0	1
1	1	2	1	1	0	2	0	0	0	0	0	0	1	0	0	2			

SPEAKER=2 PHONEME= 8 PITCH=174 TM= 94.80 ST= 95.50

0	1	0	1	2	2	1	1	7	5	11	19	53	99	45	21	7	8	6	2
2	2	2	2	3	5	7	14	42	80	32	15	6	4	4	2	1	3	8	2
1	1	0	1	3	4	0	1	0	0	0	0	0	0	1	0	0	1	0	0
0	1	0	6	4	1	1	2	1	0	2	0	0	0	2	3	0	1	1	1
1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0			

TABLE XII (Continued)

```
SPEAKER=2    PHONEME= 9    PITCH=100    TM= 72.80    ST= 92.90
58 12  6  7  4  2  1  0  0  1  2  2  5  7 21 36 44 22  9  4
 3  1  2  5  7 33 15  7  8  8 17 28 58 23 15 18 49 99 30 15
 8  6  7  5  8  7  4  3  3  3  5  3  2  5  2  3  3  5  7 22
13  9  4  4  3  7  4  3  1  2  2  0  1  1  0  0  1  1  1  2
 0  0  0  0  0  0  1  0  0  2  0  0  0  0  0  0  1

SPEAKER=2    PHONEME= 9    PITCH=110    TM= 76.60    ST= 92.30
 7 17 39 19  8  3  0  0  1  0  1  1  1  1  4  3  5 16 46 11
 5  3  1  2  2  5  8 29 27  9  8 10 15 29 99 36 17 10  6 10
 3  2  2  4  4  2  2  6  1  2  5  2  1  3  4  4  5 13 17 22
14  8  4  9 11 10  5  4  2  1  3  1  1  0  0  0  0  1  1  1
 1  0  1  0  0  0  0  0  0  1  1  1  0  1  0  0  0

SPEAKER=2    PHONEME= 9    PITCH=123    TM= 77.90    ST= 96.10
 2  2  4 10 25 74 41  7  8  3  1  2  2  2  2  3  2  3  6  8
17 59 24 10  3  6  4  5 11 16 46 57 16  9 11 16 41 99 41 19
 8  6  9  4  4  4  8  5  3  3  4  3  3  5  5  8 13 15 54 26
12  8  4  4  6  8 13 11  5  2  3  2  2  1  2  1  0  0  1  0
 0  2  2  0  2  0  0  0  0  0  0  1  0  0  1  1  0

SPEAKER=2    PHONEME= 9    PITCH=130    TM= 43.10    ST= 81.20
 3  2  0  8 20 47 99 27 20  5  6  2  1  1  3  0  1  4  5  6
15 39 84 31  9  7  4  6  8 13 35 71 42 19  8  6 10 21 46 15
 6  3  9 16  5  2  4  7  0  2  3  2  3  4  9  5 11 16 23 67
20 26 11 14  7 12 19 29  9  5  4  1  2  1  0  1  2  2  2  5
 2  1  4  1  0  0  2  0  1  3  1  0  2  0  0  0  2

SPEAKER=2    PHONEME= 9    PITCH=146    TM= 95.10    ST= 96.10
 1  0  0  3  4  6  7  9 38 75 20 12  6  4  3  3  2  3  2  3
 3  3  4 13 20 58 21 11 10  9 11 12 21 36 99 45 22  9  7  6
 6  6  5  2  2  3  4  1  2  4 10  4  4  6  7  3  2  3  3  7
 9 19 37 15  8  8 12 12 15  6  3  2  3  2  2  3  1  0  0  1
 0  2  4  1  1  1  1  0  0  0  0  0  0  0  0  1  1

SPEAKER=2    PHONEME= 9    PITCH=164    TM= 95.00    ST= 96.30
 1  0  0  1  2  1  2  2  8  8 17 33 99 31 20 10  4  5  3  2
 2  2  1  3  1  5 10 23 63 23 12  7  4  5  6  9 22 58 29 13
 6  4  7  7 10  3  1  1  2  8  2  1  1  4  2  0  2  3  1  5
 8  4 11 12  4  5  4  8  2  2  2  0  1  1  1  0  1  0  0  0
 0  0  1  0  0  0  1  0  0  0  0  0  0  0  0  0  1

SPEAKER=2    PHONEME= 9    PITCH=174    TM= 90.80    ST= 96.10
 0  0  0  2  1  2  2  2  5  6  9 16 49 94 41 19  8  8  5  4
 3  2  2  3  1  7 11 19 57 99 37 17  9  6  5  6 11 20 69 35
11  6  8  4  6  9  3  1  2  5 17  7  4  3 12  7  2  2  5  7
21 31 18 37 24 14 11  9 18  7  4  4  6  3  2  3  1  0  0  1
 0  1  4  2  1  1  1  0  0  0  0  0  0  0  0  1  1
```

TABLE XII (Continued)

SPEAKER=2 PHONEME=10 PITCH=100 TM= 42.80 ST= 83.90

99	20	12	9	7	6	3	1	6	2	1	3	7	8	35	64	74	38	20	6
3	3	2	1	4	29	7	4	7	4	7	19	36	12	7	10	23	50	18	10
8	25	17	6	9	8	1	6	10	7	4	7	10	17	12	23	9	27	11	8
2	4	1	0	1	0	0	2	3	1	3	3	3	0	3	3	0	1	3	0
1	2	3	0	2	0	0	0	1	0	1	2	1	0	3	1	3			

SPEAKER=2 PHONEME=10 PITCH=110 TM= 54.80 ST= 84.70

23	57	99	39	21	8	2	0	3	0	1	2	2	2	8	7	13	37	69	15
8	5	1	2	3	7	13	45	20	5	5	7	10	26	48	16	9	10	26	53
12	6	8	13	6	3	7	7	2	3	3	3	11	7	13	22	5	16	3	1
0	2	0	1	1	1	1	4	2	0	1	2	0	2	6	14	12	11	3	1
0	1	0	0	1	1	1	3	1	0	0	1	0	0	1	0	1			

SPEAKER=2 PHONEME=10 PITCH=123 TM= 44.80 ST= 73.00

3	1	4	15	41	99	37	7	7	1	3	0	1	2	3	1	3	2	3	5
18	58	17	9	3	3	4	3	7	13	47	46	12	9	10	15	41	90	41	19
8	4	18	11	5	5	15	11	3	9	27	10	13	31	12	4	4	2	0	2
1	0	0	2	0	1	3	3	2	4	3	0	1	1	0	5	9	8	8	11
3	0	2	1	0	1	3	1	0	4	1	0	0	0	0	0	2			

SPEAKER=2 PHONEME=10 PITCH=130 TM= 82.70 ST= 95.40

2	2	2	8	14	29	67	21	13	5	3	3	2	3	2	0	0	2	1	3
7	19	45	18	4	4	2	3	4	6	15	30	28	12	8	8	17	28	99	40
14	9	14	19	8	4	5	8	4	5	11	24	9	11	30	10	3	2	0	1
1	1	1	4	2	2	3	4	2	2	4	1	3	3	8	22	16	15	9	6
2	1	5	2	3	1	3	1	0	0	0	0	0	1	1	0	2			

SPEAKER=2 PHONEME=10 PITCH=146 TM= 69.60 ST= 88.40

2	2	0	5	4	5	6	9	36	99	35	19	10	7	5	2	1	2	1	1
3	3	4	12	16	74	28	13	10	6	9	9	17	29	92	78	32	14	8	7
10	31	15	7	4	4	11	5	1	3	8	5	4	7	23	8	3	5	1	2
1	0	0	1	0	0	2	2	2	3	2	0	2	1	1	6	12	20	13	8
1	1	3	0	0	0	1	0	0	1	1	0	2	0	0	0	1			

SPEAKER=2 PHONEME=10 PITCH=164 TM= 52.70 ST= 83.40

0	1	0	0	1	3	1	2	11	11	23	56	99	27	18	10	2	4	4	2
1	3	1	1	1	8	13	42	69	21	10	8	3	3	4	10	27	79	22	10
4	5	6	8	10	3	0	1	3	6	1	3	2	6	2	1	4	9	1	1
1	3	0	0	0	0	0	2	2	1	2	3	2	0	1	0	0	1	2	1
1	2	3	1	3	1	0	0	1	0	1	2	1	0	2	0	0			

SPEAKER=2 PHONEME=10 PITCH=174 TM= 64.90 ST= 86.50

1	1	1	2	3	2	0	0	5	2	8	16	44	99	52	22	6	7	5	0
2	3	2	2	3	5	6	14	43	91	38	17	8	4	6	4	6	10	45	28
7	5	5	2	6	7	1	0	1	1	12	5	2	1	10	4	0	1	6	0
1	2	1	0	2	0	0	1	1	0	2	2	2	1	3	4	7	14	5	1
1	3	7	5	7	1	0	0	0	0	0	1	0	0	2	0	0			

TABLE XII (Continued)

SPEAKER=2 PHONEME=11 PITCH=100 TM= 83.40 ST= 94.80

57	14	8	6	5	3	1	1	4	2	4	6	11	12	36	53	99	48	25	10
9	8	8	14	20	79	33	16	12	5	6	9	12	4	3	2	3	9	1	0
0	3	0	1	3	2	0	2	0	0	0	0	0	0	1	0	0	2	1	0
2	1	0	5	4	2	2	3	2	2	4	1	1	0	0	0	0	1	1	1
2	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0			

SPEAKER=2 PHONEME=11 PITCH=110 TM= 72.60 ST= 85.60

15	30	57	32	14	6	2	0	3	2	1	3	6	6	14	17	29	70	99	30
17	9	5	5	2	10	20	41	18	6	2	3	5	11	13	5	2	1	2	1
0	0	2	2	1	1	1	0	2	1	0	0	1	0	0	1	1	0	2	3
3	3	2	1	1	2	1	4	4	2	2	3	2	0	0	1	0	2	2	2
0	2	1	0	0	1	0	0	1	1	0	2	0	0	1	0	0			

SPEAKER=2 PHONEME=11 PITCH=123 TM= 73.50 ST= 75.00

0	1	2	6	20	55	24	4	6	1	0	0	1	0	2	3	3	5	12	15
31	99	30	13	5	6	2	3	5	4	12	11	1	0	0	1	0	4	1	1
0	2	1	0	0	0	0	0	1	1	0	2	0	0	1	0	0	0	1	0
4	3	2	1	2	0	0	1	2	0	2	2	2	0	2	1	1	1	1	0
0	1	1	0	2	0	0	1	0	0	0	1	0	0	1	1	0			

SPEAKER=2 PHONEME=11 PITCH=130 TM= 75.00 ST= 91.90

3	1	3	10	19	35	86	37	24	8	5	3	4	4	5	4	5	6	9	10
18	37	99	61	19	14	10	7	6	4	4	6	11	3	1	1	2	1	1	1
0	0	0	2	0	1	1	1	0	3	0	0	0	0	0	0	1	1	0	2
0	0	0	1	0	0	2	2	2	3	2	1	2	1	0	1	1	0	1	2
0	0	2	0	0	0	1	0	0	1	1	0	1	0	0	0	1			

SPEAKER=2 PHONEME=11 PITCH=146 TM= 50.10 ST= 61.70

0	1	2	4	3	7	5	7	38	92	24	15	8	4	3	4	1	1	1	2
0	4	7	18	28	99	32	15	11	4	0	2	2	0	6	5	1	0	1	0
0	1	2	0	1	3	1	0	2	1	0	0	1	0	1	2	1	0	2	0
0	1	7	0	1	2	3	2	5	2	2	1	2	0	1	2	2	3	6	2
0	0	2	0	1	1	1	0	3	1	0	0	0	0	0	1	0			

SPEAKER=2 PHONEME=11 PITCH=164 TM= 68.30 ST= 74.00

0	0	1	1	2	3	2	1	9	7	13	27	99	59	29	15	5	5	5	1
0	0	0	0	0	3	2	8	31	26	7	4	1	0	0	1	0	1	3	1
0	2	0	0	0	1	0	0	1	2	0	2	0	0	0	0	0	0	1	0
1	2	0	4	1	1	0	2	2	1	3	2	2	0	2	0	0	0	1	0
0	1	2	0	2	0	0	0	0	0	0	1	0	0	2	0	0			

SPEAKER=2 PHONEME=11 PITCH=174 TM= 55.40 ST= 63.60

1	1	0	2	1	1	1	2	5	4	10	16	51	99	39	15	5	6	4	0
3	0	0	0	0	0	2	8	28	44	16	6	0	0	1	0	0	2	3	0
3	1	0	2	0	0	0	1	0	2	2	0	0	2	0	0	0	1	0	1
2	1	0	6	1	0	1	2	0	1	6	3	2	3	1	0	0	1	0	3
1	1	1	2	0	0	1	0	0	1	1	0	0	2	1	0	2			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 1 PITCH=100 TM= 40.90 ST= 67.70

11	2	0	0	1	0	0	2	2	0	4	1	4	9	49	62	12	9	4	1
3	1	0	3	15	3	1	3	3	4	19	28	4	4	4	9	25	11	14	37
99	42	37	45	50	19	12	13	2	6	4	3	2	4	2	0	1	4	7	14
17	21	11	13	3	2	3	4	1	4	9	11	7	5	1	0	1	1	0	2
1	0	7	6	4	1	4	0	0	1	1	0	1	2	2	0	5			

SPEAKER=3 PHONEME= 1 PITCH=110 TM= 56.30 ST= 96.00

13	38	57	29	12	7	1	0	2	1	0	2	4	4	10	14	22	54	99	28
15	8	5	4	2	9	14	33	18	6	4	6	10	23	23	14	15	21	70	88
30	17	45	51	20	11	13	8	5	4	1	1	2	0	2	2	2	4	5	4
7	6	2	0	2	2	1	3	4	2	3	6	8	2	1	2	0	2	5	5
5	18	19	3	4	1	0	1	2	2	1	3	0	0	0	1	1			

SPEAKER=3 PHONEME= 1 PITCH=123 TM= 44.30 ST= 63.60

0	2	1	5	28	50	10	4	6	1	1	2	0	0	1	2	0	4	10	16
35	61	18	6	1	4	0	3	7	13	37	19	5	2	5	9	27	54	17	11
15	42	99	35	17	9	5	4	3	3	3	4	2	1	3	2	0	2	7	0
3	3	2	1	4	1	2	4	5	4	5	4	4	1	4	1	0	0	2	0
1	3	6	1	3	1	0	1	0	0	1	2	0	1	3	1	1			

SPEAKER=3 PHONEME= 1 PITCH=130 TM= 73.30 ST= 84.50

1	1	1	0	4	9	22	8	5	0	1	1	1	0	2	0	0	0	1	0
2	8	22	8	3	2	0	1	3	1	6	15	11	5	7	8	17	32	99	37
15	11	25	44	23	8	5	7	4	1	2	4	1	1	4	1	1	3	1	0
1	1	0	1	2	0	0	1	2	0	4	5	4	2	3	1	0	1	2	2
4	12	16	4	5	1	0	0	0	0	0	1	0	0	2	2	3			

SPEAKER=3 PHONEME= 1 PITCH=146 TM= 80.50 ST= 95.40

0	1	0	0	1	3	1	4	21	54	18	10	4	1	2	1	0	1	2	1
1	3	2	6	9	47	16	8	6	3	3	5	4	7	32	20	8	7	12	21
42	99	41	16	12	8	8	4	3	3	4	3	1	1	3	1	1	7	3	6
5	8	18	5	3	1	1	2	3	2	6	6	5	5	2	1	0	2	4	6
10	20	18	4	6	1	0	0	1	0	1	2	1	0	2	1	1			

SPEAKER=3 PHONEME= 1 PITCH=164 TM= 77.10 ST= 96.50

1	2	0	0	1	1	0	1	6	5	15	34	63	18	12	5	1	2	2	0
1	2	1	1	2	5	9	27	55	16	10	7	6	5	11	15	38	99	36	16
9	9	24	33	46	17	7	5	11	18	4	3	3	4	5	4	11	13	7	18
14	7	10	5	5	3	1	3	3	4	3	3	2	1	2	1	1	3	3	4
6	18	38	11	10	2	2	1	1	0	1	3	1	1	2	1	2			

SPEAKER=3 PHONEME= 1 PITCH=174 TM= 46.90 ST= 96.30

1	0	1	2	1	1	3	0	0	1	2	0	14	43	21	7	4	2	0	1
0	0	1	2	0	2	4	5	17	51	25	8	4	5	5	8	17	23	78	99
26	12	18	14	35	80	38	14	6	9	18	10	4	3	4	5	4	7	35	14
14	39	15	9	3	3	7	6	9	9	16	14	20	12	5	5	4	4	5	11
9	24	49	25	14	7	6	2	1	5	3	4	7	5	1	4	5			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 2 PITCH=100 TM= 38.20 ST= 69.70

34	3	3	4	3	2	3	0	0	1	3	0	9	16	59	99	50	29	12	5
4	0	2	8	13	30	12	7	10	15	37	82	43	18	6	12	41	24	5	5
3	3	3	3	0	2	2	1	1	6	1	0	2	1	0	1	2	0	1	3
2	1	3	3	4	11	33	33	17	14	17	22	23	7	2	2	1	0	3	6
5	13	15	4	3	3	2	0	1	2	0	2	2	1	1	3	2			

SPEAKER=3 PHONEME= 2 PITCH=110 TM= 69.80 ST= 92.70

7	19	36	22	10	3	0	0	1	0	1	2	2	3	10	10	20	45	81	24
13	7	2	4	3	7	12	33	21	9	11	14	26	52	99	39	20	12	12	19
6	3	3	5	3	1	1	2	1	1	2	0	0	0	1	0	0	1	1	0
3	2	4	13	26	14	6	7	5	6	14	16	7	2	1	0	1	2	1	3
5	5	5	2	1	0	1	0	0	1	1	1	0	2	0	0	0			

SPEAKER=3 PHONEME= 2 PITCH=123 TM= 50.90 ST= 95.70

4	2	5	17	52	79	24	7	7	3	2	0	2	3	3	3	5	8	15	26
58	80	30	15	3	8	8	9	19	36	84	44	18	12	15	28	91	99	33	19
10	9	15	6	4	3	4	0	1	4	1	1	3	1	1	2	0	0	1	3
2	9	17	22	11	10	5	6	7	8	16	26	15	5	4	6	3	3	6	8
8	7	6	0	2	2	3	1	3	1	0	1	0	0	1	1	1			

SPEAKER=3 PHONEME= 2 PITCH=130 TM= 46.50 ST= 95.90

2	0	2	8	19	51	81	13	10	3	3	0	2	3	3	2	4	5	8	13
21	71	64	24	5	8	7	8	13	20	50	99	35	17	11	17	38	80	96	37
15	9	19	16	6	4	6	6	2	5	5	3	4	1	1	1	1	0	1	2
0	2	4	9	17	27	34	30	16	13	17	28	16	7	5	8	5	4	7	9
9	30	42	17	18	9	6	3	6	3	1	1	1	0	3	4	5			

SPEAKER=3 PHONEME= 2 PITCH=146 TM= 67.70 ST= 91.70

1	3	1	1	3	5	4	8	42	74	20	13	6	2	4	2	0	1	2	0
2	4	4	10	19	57	16	8	8	3	7	9	16	28	99	43	19	9	5	2
5	12	4	3	3	2	3	2	0	1	1	1	0	1	2	1	0	2	0	0
1	1	0	6	20	9	10	8	6	3	7	8	4	2	2	1	2	3	6	11
22	16	9	2	4	0	2	1	0	1	2	1	0	1	0	0	1			

SPEAKER=3 PHONEME= 2 PITCH=164 TM= 70.50 ST= 94.00

1	0	0	2	1	2	2	1	5	6	15	34	72	21	14	6	3	2	1	0
1	0	1	3	2	7	15	35	58	18	10	5	5	6	9	14	41	99	31	16
7	3	8	9	10	4	3	1	2	7	1	0	1	2	0	1	2	2	1	3
2	2	6	11	30	51	17	28	11	16	17	27	16	12	6	5	4	5	6	11
14	28	20	6	4	2	2	0	1	3	1	0	1	1	0	2	4			

SPEAKER=3 PHONEME= 2 PITCH=174 TM= 62.00 ST= 96.60

0	0	0	1	1	2	2	0	3	3	6	10	34	83	48	20	8	6	4	2
2	0	1	3	1	6	10	15	42	99	59	22	10	6	4	4	6	7	27	25
5	1	3	3	4	11	4	1	0	3	4	0	1	1	1	2	2	2	4	3
2	3	6	14	22	34	92	54	34	26	42	36	24	11	8	7	6	7	9	15
22	37	27	11	9	3	4	2	1	3	1	0	0	1	0	2	4			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 3 PITCH=100 TM= 39.50 ST= 70.50

25	14	3	3	0	2	2	0	1	4	2	0	4	3	14	27	64	44	14	4
4	2	4	2	1	21	16	4	6	5	8	21	49	24	11	17	43	99	39	17
8	26	15	7	6	8	2	1	3	3	0	1	2	0	2	2	2	1	4	3
7	11	22	6	2	3	2	4	6	3	7	10	5	0	1	2	0	2	4	4
3	9	8	3	4	1	0	1	2	0	1	3	1	0	4	1	1			

SPEAKER=3 PHONEME= 3 PITCH=110 TM= 74.40 ST= 94.90

10	20	42	36	15	5	1	1	3	0	2	3	4	5	15	15	26	64	99	30
20	11	4	5	3	5	7	21	13	4	5	7	11	28	52	19	13	15	35	69
24	11	13	17	8	3	3	3	1	2	2	0	0	0	1	0	0	2	1	3
6	7	2	3	1	0	1	2	1	2	5	7	7	3	1	0	0	1	0	3
5	8	8	4	2	0	2	0	0	0	1	0	0	1	1	0	3			

SPEAKER=3 PHONEME= 3 PITCH=123 TM= 60.60 ST= 82.90

1	3	5	14	38	79	31	5	7	3	0	2	3	1	3	5	3	6	12	16
34	99	30	13	5	7	5	5	12	16	53	36	12	7	10	17	44	92	39	18
8	8	23	8	5	4	4	3	2	2	1	2	0	0	0	1	0	1	3	3
9	6	3	1	0	0	0	2	3	2	5	8	12	4	3	1	0	1	3	4
10	13	9	2	4	0	0	0	0	0	1	2	1	0	3	1	1			

SPEAKER=3 PHONEME= 3 PITCH=130 TM= 95.30 ST= 93.40

1	1	1	4	9	19	43	15	9	2	1	1	0	1	2	1	0	2	2	1
5	14	34	15	4	4	2	4	4	4	12	23	27	11	8	8	18	30	99	48
18	11	15	19	11	4	4	4	2	2	2	1	1	1	1	0	0	2	0	3
2	4	1	2	0	0	1	2	1	2	4	4	5	4	1	1	2	1	1	4
6	13	20	6	5	1	2	0	0	0	0	0	0	1	0	1	3			

SPEAKER=3 PHONEME= 3 PITCH=146 TM= 47.60 ST= 90.60

3	1	0	4	3	4	4	9	38	99	30	17	8	7	4	1	1	2	0	1
4	3	5	15	19	88	37	16	10	7	8	6	13	22	70	77	30	12	8	12
25	78	40	16	8	8	19	7	1	5	3	0	1	1	0	1	2	2	1	4
3	6	28	14	6	5	4	3	3	5	4	4	9	11	3	4	4	3	7	13
23	38	48	22	15	4	3	0	1	3	1	2	5	2	2	5	5			

SPEAKER=3 PHONEME= 3 PITCH=164 TM= 42.10 ST= 78.20

0	1	2	2	1	5	0	0	6	6	12	36	99	24	13	7	2	2	4	0
0	0	1	0	1	6	7	25	63	18	7	7	4	3	7	15	35	92	53	23
8	8	7	5	10	3	0	1	3	6	2	3	2	1	3	0	0	2	2	1
6	12	40	46	11	7	3	6	3	0	4	4	6	10	5	3	1	4	3	6
12	33	18	7	7	2	2	3	0	0	1	2	0	1	2	0	3			

SPEAKER=3 PHONEME= 3 PITCH=174 TM= 47.90 ST= 88.60

1	3	1	1	1	2	0	1	5	2	7	15	40	99	51	20	5	5	4	0
2	3	1	2	3	5	5	13	38	91	40	17	5	5	6	7	10	19	67	62
15	8	6	5	10	17	5	4	1	2	5	1	0	1	3	0	1	3	7	3
13	29	8	7	6	0	3	5	5	4	8	5	13	7	4	2	1	3	2	7
10	13	25	11	6	1	3	0	0	1	2	0	2	3	3	3	8			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 4 PITCH=100 TM= 46.20 ST= 93.80

31	4	2	4	1	3	2	1	2	4	2	3	10	15	48	91	58	35	17	8
6	3	3	7	10	30	14	11	15	22	48	99	68	29	11	9	11	6	2	4
2	2	2	2	0	1	2	1	1	4	1	0	0	1	0	1	2	1	0	3
1	0	0	2	2	10	26	40	16	11	10	8	10	5	0	3	3	4	7	18
17	8	7	5	4	3	7	3	2	4	1	0	0	0	0	1	3			

SPEAKER=3 PHONEME= 4 PITCH=110 TM= 51.40 ST= 88.70

8	25	43	14	6	4	1	0	4	1	0	1	3	1	10	13	25	64	99	25
16	8	4	2	2	8	14	42	26	9	8	11	20	43	83	27	12	6	7	7
2	3	3	4	1	1	0	1	2	3	1	3	1	0	0	0	0	1	1	1
1	4	2	5	20	23	20	28	15	12	23	37	17	6	5	3	1	2	4	5
15	13	9	2	7	4	4	2	2	0	1	2	0	0	3	0	1			

SPEAKER=3 PHONEME= 4 PITCH=123 TM= 65.20 ST= 73.00

0	2	4	11	31	70	29	5	7	2	0	2	3	1	4	5	4	7	15	22
48	99	31	15	6	8	5	5	11	20	51	27	9	3	3	4	6	8	3	0
0	1	2	0	1	1	1	0	2	1	0	0	1	0	1	1	1	0	2	0
0	0	0	0	1	2	4	6	5	3	4	5	6	0	1	2	0	1	3	1
0	1	2	0	1	1	1	0	2	1	0	1	0	0	0	1	0			

SPEAKER=3 PHONEME= 4 PITCH=130 TM= 63.40 ST= 75.20

1	2	3	5	11	27	60	18	12	4	1	2	3	2	3	4	2	5	9	12
19	54	99	43	14	14	7	6	10	12	27	59	44	17	7	7	4	6	15	3
0	1	3	2	1	2	0	0	1	1	0	0	1	0	1	1	0	0	1	0
0	0	0	0	1	3	8	16	12	5	7	10	9	4	3	3	1	1	3	1
0	0	1	0	1	1	0	0	2	1	0	1	0	0	0	1	0			

SPEAKER=3 PHONEME= 4 PITCH=146 TM= 42.30 ST= 77.90

0	2	3	2	1	5	2	3	26	72	40	18	9	4	3	4	0	0	1	1
0	4	6	13	18	99	52	20	13	5	1	3	3	3	16	29	7	2	3	0
0	5	4	0	1	2	1	0	3	2	0	0	1	0	1	2	1	0	3	1
0	1	0	0	0	2	3	8	34	28	18	12	15	20	14	13	5	6	8	9
12	22	12	6	14	4	1	3	3	3	1	4	0	0	2	2	2			

SPEAKER=3 PHONEME= 4 PITCH=164 TM= 80.00 ST= 82.70

1	0	0	1	0	2	2	0	3	4	8	17	62	38	18	8	3	4	3	0
2	4	2	3	3	11	16	31	99	58	26	13	8	4	1	3	5	8	9	2
2	1	1	1	4	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
0	0	1	0	0	0	1	1	0	1	3	2	2	3	1	2	1	2	1	0
1	3	6	6	13	3	1	0	1	0	0	0	1	0	0	0	2			

SPEAKER=3 PHONEME= 4 PITCH=174 TM= 59.90 ST= 87.30

0	0	1	3	0	0	0	1	3	2	5	11	33	78	39	16	3	5	5	1
0	1	2	2	0	5	8	15	43	99	44	17	8	6	2	1	3	4	15	8
2	2	2	1	3	7	0	1	2	1	0	0	1	0	0	0	1	1	0	1
0	0	0	2	0	0	4	6	20	9	5	4	3	1	2	2	0	3	6	5
5	12	13	18	13	2	3	1	0	1	0	0	0	2	1	0	3			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 5 PITCH=100 TM= 57.60 ST= 71.60

20	4	1	2	0	1	0	0	2	0	0	1	2	2	13	28	21	11	3	1
1	0	1	3	2	9	3	3	3	1	5	15	11	5	7	9	32	31	20	29
69	99	45	19	14	9	14	28	13	12	16	5	6	4	2	0	0	2	1	0
0	1	0	1	2	0	0	2	3	1	1	1	2	1	0	1	0	0	2	4
2	1	5	3	2	0	1	1	0	1	1	0	0	0	2	0	0			

SPEAKER=3 PHONEME= 5 PITCH=110 TM= 69.90 ST= 79.60

7	19	39	16	6	4	0	0	1	1	1	0	1	2	6	7	15	37	61	16
11	4	1	2	2	4	5	20	12	4	4	5	7	14	37	14	13	15	46	99
30	18	50	76	31	16	30	25	10	21	20	10	17	8	5	3	0	0	1	0
0	1	0	0	0	1	1	0	1	2	2	0	1	0	0	0	1	0	0	3
7	5	5	2	2	1	0	0	0	0	0	1	0	0	1	2	1			

SPEAKER=3 PHONEME= 5 PITCH=123 TM= 71.40 ST= 95.20

3	2	1	11	37	41	12	3	5	2	1	0	1	1	2	3	2	4	10	15
35	73	22	12	4	5	3	5	8	12	37	22	8	8	14	20	59	99	33	19
15	35	67	26	14	12	15	6	5	10	5	5	25	19	6	2	0	1	1	1
2	0	0	1	1	0	0	1	2	1	1	0	2	0	0	1	0	0	1	3
0	1	5	2	1	0	0	0	0	0	1	0	0	0	2	0	0			

SPEAKER=3 PHONEME= 5 PITCH=130 TM= 58.10 ST= 96.30

2	1	2	8	14	34	78	26	15	6	5	2	1	1	4	2	1	4	6	4
12	33	77	34	10	10	4	3	6	8	14	27	43	16	7	10	20	29	99	62
20	14	37	60	43	18	16	31	30	13	13	33	11	3	5	3	1	0	0	1
0	1	2	0	0	1	3	2	0	1	3	2	1	1	1	0	0	3	1	0
1	2	3	0	1	2	2	0	1	1	0	1	3	0	0	1	2			

SPEAKER=3 PHONEME= 5 PITCH=146 TM= 76.80 ST= 83.00

2	0	0	1	1	1	0	0	10	13	2	1	0	0	1	2	0	0	0	1
0	0	0	1	2	10	2	1	0	1	3	1	1	4	21	11	5	5	10	16
38	99	40	14	12	14	23	7	3	6	18	8	3	4	7	2	2	0	0	1
2	0	0	0	1	1	0	1	1	0	2	2	1	0	1	2	1	1	2	6
11	5	5	1	0	0	3	0	0	1	1	0	0	0	1	0	1			

SPEAKER=3 PHONEME= 5 PITCH=164 TM= 51.20 ST= 94.60

1	2	0	1	0	3	1	0	6	6	11	32	99	29	12	7	4	3	1	1
2	2	0	1	1	4	8	26	60	17	10	9	5	4	9	16	38	83	60	27
11	11	29	32	79	32	16	8	13	43	15	10	17	40	24	8	3	5	1	0
1	4	0	0	0	2	1	1	1	2	3	0	3	1	0	2	4	1	0	4
4	2	3	1	2	1	0	0	1	0	0	3	0	0	1	3	1			

SPEAKER=3 PHONEME= 5 PITCH=174 TM= 94.20 ST= 96.70

0	0	2	1	0	1	1	0	2	2	6	10	32	62	24	9	4	5	2	0
1	2	1	0	0	3	2	5	19	32	11	6	6	4	6	9	19	30	99	62
21	13	20	18	51	79	27	15	11	18	54	25	14	15	57	29	8	5	4	2
0	2	2	1	0	1	3	2	3	2	4	2	3	3	3	1	1	3	2	2
5	11	9	7	7	2	1	1	2	1	0	1	1	0	0	0	2			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 6 PITCH=100 TM= 48.90 ST= 94.10

33	6	4	3	3	5	1	0	3	1	0	3	8	10	42	71	44	23	11	3
1	2	3	4	10	23	6	3	6	6	12	33	27	13	16	33	99	67	24	14
13	13	10	8	15	9	4	7	5	6	8	5	3	2	1	3	0	0	0	1
0	1	2	2	0	3	1	1	1	3	4	5	4	1	0	3	0	1	6	5
0	1	5	5	2	3	1	0	1	2	0	1	2	0	0	2	1			

SPEAKER=3 PHONEME= 6 PITCH=110 TM= 51.40 ST= 75.00

7	20	37	16	6	3	2	0	0	0	1	0	2	4	9	9	20	46	99	28
14	4	2	3	1	3	6	15	8	2	3	2	7	17	54	18	10	6	13	38
12	3	11	21	10	4	5	5	2	6	18	4	2	3	2	0	2	0	0	0
1	0	0	2	1	0	3	2	1	2	3	0	2	2	1	1	3	2	1	2
1	0	3	2	3	1	3	1	0	1	2	0	1	2	3	1	5			

SPEAKER=3 PHONEME= 6 PITCH=123 TM= 26.10 ST= 61.90

3	5	2	5	15	46	32	16	10	1	3	4	2	0	3	1	0	1	5	3
17	58	56	17	4	6	0	5	11	12	41	88	32	12	11	15	32	94	99	36
13	13	23	17	9	12	33	59	16	14	34	21	5	1	1	2	0	1	3	3
2	6	1	0	0	1	0	4	6	5	4	6	4	0	0	2	0	3	5	5
0	4	3	0	0	2	0	1	3	3	1	4	0	0	1	2	1			

SPEAKER=3 PHONEME= 6 PITCH=130 TM= 48.70 ST= 53.30

1	0	2	2	0	8	22	2	2	2	2	0	1	0	0	0	1	0	1	2
2	9	32	8	0	2	1	1	3	4	7	19	13	5	3	7	17	31	99	35
11	6	9	8	3	4	9	22	10	7	5	7	2	1	0	0	1	0	0	2
1	0	1	1	0	0	2	2	2	3	3	0	2	0	0	0	1	0	2	4
1	0	2	0	0	0	2	0	1	3	1	0	1	0	0	0	2			

SPEAKER=3 PHONEME= 6 PITCH=146 TM= 28.10 ST= 47.10

2	4	1	4	5	3	0	2	24	47	9	5	2	1	5	1	0	0	1	0
0	4	3	9	22	78	20	8	7	0	4	9	11	20	99	45	14	5	7	4
27	73	23	8	10	7	11	2	3	2	20	8	3	1	5	1	0	0	1	0
2	4	2	1	4	1	1	2	4	0	4	4	5	1	4	1	0	1	3	0
1	3	4	0	3	1	0	0	2	0	2	4	1	0	4	1	0			

SPEAKER=3 PHONEME= 6 PITCH=164 TM= 79.10 ST= 79.60

1	0	2	1	0	1	1	0	4	5	10	22	56	15	7	3	1	0	2	1
1	1	2	2	0	6	11	29	57	18	10	7	7	6	9	15	40	99	39	20
9	6	12	10	15	5	3	0	5	13	2	1	3	3	0	0	0	0	0	1
0	1	1	0	0	0	1	0	1	3	4	8	4	2	0	1	0	0	2	2
0	0	2	0	0	1	0	0	0	1	0	1	1	0	0	1	1			

SPEAKER=3 PHONEME= 6 PITCH=174 TM= 61.70 ST= 88.30

0	0	1	1	1	3	0	0	3	2	2	7	26	63	30	13	3	3	3	1
0	1	2	2	1	7	7	13	43	99	42	18	9	5	6	8	13	19	77	50
13	8	11	8	23	49	17	8	9	13	46	23	7	3	3	3	1	0	1	0
0	1	1	0	0	2	1	1	2	1	4	4	6	1	1	3	0	0	4	3
0	4	15	3	3	2	1	0	0	0	0	0	1	0	1	2	1			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 7 PITCH=100 TM= 96.30 ST= 96.60

20	4	0	2	1	1	0	0	0	0	0	1	1	2	8	12	39	22	8	1
1	1	0	1	2	9	6	2	3	1	3	5	13	9	7	8	16	33	51	24
24	51	99	39	23	23	17	13	34	24	11	20	8	4	4	1	0	1	1	0
1	2	0	0	0	0	0	2	2	2	3	3	1	0	0	0	0	0	1	2
2	4	4	0	1	0	0	0	1	1	0	1	0	0	0	1	0			

SPEAKER=3 PHONEME= 7 PITCH=110 TM= 62.80 ST= 86.00

2	8	20	11	4	2	0	1	0	0	0	0	0	1	4	4	7	23	56	13
6	3	0	1	2	2	2	12	8	1	3	5	5	10	38	14	12	14	37	99
44	23	38	59	45	19	23	47	12	7	10	5	3	2	3	0	0	0	0	0
1	1	0	0	1	0	0	1	2	0	2	2	2	0	1	0	0	0	2	1
3	5	5	0	1	0	0	0	1	0	0	2	0	0	0	0	0			

SPEAKER=3 PHONEME= 7 PITCH=123 TM= 70.30 ST= 95.40

0	0	0	2	3	16	14	1	0	0	1	0	0	1	1	0	2	1	1	2
7	29	14	6	2	2	2	2	3	4	14	20	8	9	12	17	42	99	62	29
16	18	66	52	19	12	22	31	11	18	61	32	10	8	5	1	1	1	0	1
1	1	0	2	0	0	0	1	0	1	2	1	1	1	0	0	0	1	0	1
2	1	1	0	0	0	1	1	0	1	2	0	0	0	0	0	1			

SPEAKER=3 PHONEME= 7 PITCH=130 TM= 53.60 ST= 86.80

1	0	1	3	4	10	27	10	4	1	2	0	1	2	1	0	1	1	1	2
6	14	46	29	6	4	3	2	2	2	5	9	14	7	6	7	16	25	99	57
19	13	32	51	45	17	18	33	40	18	9	17	7	5	8	2	1	1	0	0
2	1	0	3	0	0	1	2	0	1	3	1	2	2	0	0	0	1	0	2
2	1	3	3	1	0	0	0	0	1	2	1	0	2	1	0	1			

SPEAKER=3 PHONEME= 7 PITCH=146 TM= 72.60 ST= 80.10

1	1	0	2	0	0	0	1	9	26	6	2	0	1	0	0	0	0	0	0
1	1	0	5	6	33	8	4	2	2	2	2	2	5	20	10	6	5	8	16
38	99	38	16	10	6	8	5	4	11	44	23	6	4	3	0	0	1	0	0
1	1	0	2	0	0	0	1	0	1	2	1	1	2	0	0	0	1	0	2
1	0	1	2	0	0	0	0	0	1	1	1	0	1	0	0	0			

SPEAKER=3 PHONEME= 7 PITCH=164 TM= 50.30 ST= 75.90

3	1	0	2	0	0	0	1	2	4	11	24	75	19	9	2	1	2	1	1
2	0	0	1	0	3	8	26	54	15	8	5	2	4	5	7	25	73	25	12
8	6	18	27	47	15	12	15	44	99	25	13	9	17	6	1	1	2	0	1
2	1	0	2	0	0	1	2	0	2	4	1	1	2	0	0	2	9	4	5
4	2	4	4	3	0	1	1	0	2	2	1	0	1	1	0	3			

SPEAKER=3 PHONEME= 7 PITCH=174 TM= 95.90 ST= 97.40

0	0	0	1	0	1	0	1	0	0	1	2	9	25	8	3	0	2	0	0
0	0	0	0	0	1	0	2	7	12	3	2	0	1	2	3	7	12	48	23
7	4	6	5	16	20	10	12	21	37	99	47	19	11	12	6	3	4	1	1
0	1	0	1	1	1	1	2	1	0	1	0	0	0	1	1	0	2	1	1
0	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 8 PITCH=100 TM= 52.10 ST= 57.80

21	2	1	4	1	0	0	0	0	0	2	1	1	4	9	15	41	18	4	2
3	3	2	7	7	44	20	11	9	13	28	49	99	39	14	7	5	6	0	1
3	4	2	3	1	0	0	1	0	2	2	1	1	4	1	0	0	1	0	1
2	1	0	1	0	0	1	2	1	2	4	1	1	1	0	0	0	2	1	2
2	0	1	0	0	0	1	1	0	1	3	1	0	0	0	0	1			

SPEAKER=3 PHONEME= 8 PITCH=110 TM= 53.20 ST= 66.00

4	17	27	9	4	2	0	0	1	0	1	2	2	2	8	8	16	43	99	27
13	7	2	1	1	3	3	20	13	4	3	6	8	18	60	18	5	3	3	3
1	2	2	2	0	1	0	1	1	2	0	2	1	0	0	1	0	6	3	2
0	2	0	0	0	1	0	2	3	3	3	2	2	0	0	1	0	1	2	2
0	1	1	0	0	1	1	0	2	1	0	0	0	0	0	1	1			

SPEAKER=3 PHONEME= 8 PITCH=123 TM= 48.00 ST= 61.60

1	2	0	7	24	51	25	5	4	0	1	1	0	1	3	3	2	7	10	14
37	90	25	13	5	8	5	9	18	31	99	60	16	9	6	5	10	22	3	0
1	2	1	2	2	2	2	2	0	1	1	1	0	1	2	1	0	1	0	0
3	2	0	1	2	1	1	3	4	2	2	2	2	1	2	1	0	1	1	0
0	2	2	0	2	1	1	1	0	0	0	2	0	0	3	1	0			

SPEAKER=3 PHONEME= 8 PITCH=130 TM= 68.10 ST= 94.50

2	1	0	3	6	15	39	15	9	2	3	1	1	1	3	1	2	6	9	9
18	38	99	48	15	11	8	8	12	13	29	54	55	21	10	5	5	5	11	3
3	1	1	2	1	0	1	2	1	1	2	2	0	0	3	0	3	12	2	0
2	1	0	0	0	0	1	2	2	1	6	2	0	0	1	0	0	2	1	1
3	1	1	2	4	0	3	2	1	0	1	0	0	0	1	0	1			

SPEAKER=3 PHONEME= 8 PITCH=146 TM= 45.80 ST= 66.90

1	3	1	1	1	3	0	4	29	77	26	14	5	1	2	1	0	1	3	2
2	6	4	12	18	99	40	18	13	6	3	5	2	3	20	27	4	3	3	1
2	11	4	1	0	1	0	1	2	2	0	1	0	0	0	1	0	1	2	1
0	1	1	0	3	3	3	7	22	11	5	3	2	0	0	2	0	1	4	2
2	2	2	0	1	2	2	0	3	1	0	0	0	0	1	2	2			

SPEAKER=3 PHONEME= 8 PITCH=164 TM= 51.10 ST= 84.50

1	1	0	2	2	2	0	1	2	1	8	23	65	17	10	4	0	2	1	0
0	2	0	2	3	9	15	41	99	29	14	8	2	2	4	3	7	27	6	0
1	2	1	3	4	1	0	0	0	0	0	1	0	1	2	0	0	0	0	0
0	2	0	1	2	1	1	5	5	12	5	4	2	1	2	1	0	1	1	0
1	2	3	1	3	1	0	0	0	0	0	1	0	0	3	0	0			

SPEAKER=3 PHONEME= 8 PITCH=174 TM= 67.00 ST= 81.20

0	0	0	2	0	1	1	1	2	3	3	6	23	56	32	13	4	4	2	2
1	0	0	2	0	5	7	14	40	99	42	16	6	4	1	2	3	5	15	7
1	0	1	1	0	2	1	0	0	3	1	0	0	0	0	0	1	0	0	2
1	0	0	0	0	0	1	1	3	3	7	2	2	0	0	0	1	0	1	2
0	0	2	0	0	0	1	0	0	2	1	0	1	0	0	0	1			

TABLE XII (Continued)

SPEAKER=3 PHONEME= 9 PITCH=100 TM= 56.20 ST= 73.00

33	4	2	4	2	1	2	0	0	0	1	0	4	7	23	38	57	26	11	3
2	0	1	4	5	27	9	4	4	4	9	17	40	16	11	16	45	99	31	16
11	23	13	7	6	7	4	1	1	2	1	0	1	2	2	6	5	7	6	14
5	2	8	11	5	3	3	2	1	2	2	0	1	1	0	1	1	1	1	4
0	0	1	1	0	1	2	1	0	2	0	0	0	0	0	1	2			

SPEAKER=3 PHONEME= 9 PITCH=110 TM= 51.80 ST= 84.20

6	20	44	24	9	6	1	0	1	1	0	1	3	3	10	12	21	52	99	29
14	7	3	2	2	6	9	27	19	7	7	9	20	44	67	28	16	12	35	53
13	6	11	12	5	4	6	5	3	4	2	2	4	1	7	12	11	38	14	17
7	11	6	7	4	2	0	2	3	2	2	3	2	0	0	1	0	1	2	2
0	3	2	0	0	0	0	0	1	1	1	2	0	0	1	0	0			

SPEAKER=3 PHONEME= 9 PITCH=123 TM= 74.00 ST= 79.20

0	0	1	1	4	18	6	0	2	0	0	0	0	0	1	1	0	1	3	3
8	33	9	2	1	3	2	2	5	5	19	18	7	5	10	16	42	99	41	18
8	5	8	3	3	3	3	2	2	3	5	2	2	2	3	5	8	8	35	17
9	7	17	16	4	3	2	2	3	1	1	1	1	0	0	1	0	1	2	1
0	1	1	0	0	1	0	0	1	1	0	1	0	0	0	0	0			

SPEAKER=3 PHONEME= 9 PITCH=130 TM= 85.80 ST= 95.10

0	0	0	0	3	11	25	6	5	1	0	0	0	0	1	1	1	1	3	2
4	13	41	16	4	4	2	2	5	3	7	16	18	6	6	8	16	26	99	40
13	7	8	4	4	3	3	5	4	3	2	9	3	1	8	7	14	34	12	18
6	9	7	9	5	5	2	2	2	1	1	0	1	0	0	1	0	0	2	1
0	0	1	0	0	1	1	0	1	0	0	0	0	0	1	1	1			

SPEAKER=3 PHONEME= 9 PITCH=146 TM= 44.00 ST= 66.80

2	0	1	4	3	4	5	5	28	81	32	13	6	4	4	1	3	1	0	0
2	1	4	13	18	89	39	17	11	7	9	8	17	30	99	95	37	14	7	5
7	22	11	6	4	3	7	2	0	2	1	0	1	2	1	1	3	1	0	1
2	0	7	7	16	10	13	9	3	3	3	0	2	2	0	1	2	2	1	2
0	0	1	1	0	1	3	1	0	2	0	0	0	1	0	1	2			

SPEAKER=3 PHONEME= 9 PITCH=164 TM= 69.00 ST= 93.90

1	0	0	3	1	1	1	0	2	4	10	22	60	18	9	4	1	2	0	1
1	1	1	3	1	5	10	27	67	22	12	7	6	8	11	16	41	99	53	26
12	8	11	10	16	5	2	2	1	8	2	1	1	4	1	0	0	3	0	2
3	3	7	11	2	1	1	2	0	1	3	2	1	2	1	1	1	2	1	4
4	3	4	2	2	0	1	1	0	1	2	2	1	2	2	1	3			

SPEAKER=3 PHONEME= 9 PITCH=174 TM= 82.90 ST= 90.40

2	0	0	1	0	0	0	0	1	1	4	4	13	37	27	8	2	3	1	0
2	0	0	0	0	0	1	4	11	28	24	8	3	3	4	5	12	17	49	99
27	13	12	6	7	10	5	3	1	2	6	3	0	0	3	1	1	2	7	2
2	3	0	1	3	0	2	2	1	1	3	0	0	0	0	0	0	2	3	3
2	0	0	0	0	0	1	1	0	0	2	0	0	0	0	0	2			

TABLE XII (Continued)

SPEAKER=3 PHONEME=10 PITCH=100 TM= 39.20 ST= 72.60

56	21	8	5	2	3	0	1	3	2	2	5	4	5	18	27	88	99	34	11
6	6	2	2	2	20	26	10	10	9	15	26	81	70	22	17	27	54	76	26
14	25	41	13	12	21	9	10	41	26	5	6	2	0	1	1	0	1	2	1
1	4	0	0	0	0	0	3	4	2	3	4	3	0	1	1	0	1	3	3
2	5	7	1	4	2	0	1	2	2	1	4	1	0	1	1	0			

SPEAKER=3 PHONEME=10 PITCH=110 TM= 41.70 ST= 48.00

1	7	22	6	2	3	1	0	3	1	0	0	0	0	2	4	7	22	62	18
4	1	1	0	0	4	2	12	13	4	4	7	12	24	99	41	16	9	16	37
12	6	9	13	9	5	3	6	5	8	21	8	5	3	1	2	0	1	2	1
0	2	0	0	0	1	0	3	4	2	2	3	2	0	0	2	0	2	4	4
1	4	4	0	2	3	0	1	2	2	1	3	0	0	0	1	0			

SPEAKER=3 PHONEME=10 PITCH=123 TM= 66.60 ST= 70.40

1	0	0	2	7	26	11	2	1	0	1	0	0	0	1	0	1	3	4	5
16	51	19	8	2	2	2	4	4	4	19	17	5	5	9	13	39	99	35	16
9	6	12	6	5	2	7	6	9	20	39	15	6	6	3	0	1	1	1	0
2	0	0	0	0	0	1	2	1	1	3	1	1	0	0	0	0	1	1	1
2	1	2	1	2	0	2	1	0	1	2	0	0	1	0	0	1			

SPEAKER=3 PHONEME=10 PITCH=130 TM= 47.90 ST= 73.80

1	3	1	2	5	14	28	12	8	1	2	3	1	0	2	0	0	1	3	2
7	22	57	30	7	5	1	3	6	5	14	34	36	11	8	7	13	32	99	36
11	8	13	20	9	3	1	3	3	3	12	32	7	3	8	1	0	1	1	0
1	2	0	0	2	0	0	1	2	0	3	2	2	0	2	0	0	0	2	0
1	2	3	0	3	1	1	0	1	0	1	2	1	0	3	2	5			

SPEAKER=3 PHONEME=10 PITCH=146 TM= 45.00 ST= 57.30

1	0	0	3	0	2	2	1	14	42	8	1	1	2	0	1	2	1	0	1
1	0	2	10	14	71	24	10	7	5	7	7	15	28	99	66	27	10	4	4
7	20	7	5	2	3	6	1	0	3	2	0	4	11	28	7	3	1	0	1
1	0	0	2	0	1	4	2	1	2	2	0	2	2	0	1	3	1	0	1
0	0	2	1	0	1	4	0	0	1	0	0	0	1	1	1	3			

SPEAKER=3 PHONEME=10 PITCH=164 TM= 50.30 ST= 76.00

2	2	0	4	0	1	0	1	1	4	10	22	80	24	11	4	0	2	0	1
2	1	0	3	0	4	8	25	71	23	12	8	6	8	11	17	44	99	65	29
14	9	12	13	28	7	4	4	8	29	7	3	3	9	3	0	0	2	0	1
3	1	0	1	0	0	1	3	1	2	4	2	1	0	0	0	0	2	2	2
2	1	5	5	3	0	1	1	1	1	3	0	0	0	1	0	4			

SPEAKER=3 PHONEME=10 PITCH=174 TM= 56.40 ST= 90.30

3	1	0	2	0	0	0	1	3	2	7	10	31	78	37	13	4	5	3	1
3	1	0	1	1	2	6	13	39	82	34	14	6	5	6	7	17	27	99	69
22	11	11	7	15	25	8	5	4	7	32	13	5	5	33	17	4	3	2	0
2	0	0	0	0	0	1	3	2	1	4	1	1	0	1	0	1	2	1	2
3	5	22	10	9	1	4	8	10	4	2	1	0	0	1	0	3			

TABLE XII (Continued)

```
SPEAKER=3   PHONEME=11   PITCH=100   TM= 44.70   ST= 83.00
28  2  2  5  3  2  3  0  2  2  4  4 14 20 57 99 97 51 24 10
 7  1  2  4  4 14  5  2  0  0  1  1  1  2  1  1  5  2  0  0
 1  0  1  3  1  1  3  3  1  3  5  0  5  9  2  2  4  7  0  1
 1  0  0  2  0  1  4  2  1  2  2  0  2  2  1  2  7  5  1  3
 1  0  7  8  8  6  6  1  0  1  1  0  0  1  1  1  4

SPEAKER=3   PHONEME=11   PITCH=110   TM= 66.90   ST= 69.90
 3 13 31 12  4  3  0  0  1  0  0  1  3  3  9 12 19 46 99 26
15  7  4  2  1  4  2  6  3  1  0  1  1  0  2  2  0  0  2  8
 0  0  1  1  1  2  1  1  0  1  0  0  3  2  1  3  1  5  1  2
 0  1  1  1  0  2  1  1  2  2  0  1  2  1  1  4  3  1  1  2
 0  1  4  2  3  3  1  0  0  1  0  0  1  0  1  2  1

SPEAKER=3   PHONEME=11   PITCH=123   TM= 74.20   ST= 80.80
 2  1  2  7 19 61 27  5  6  3  2  1  1  1  2  2  3  6 11 17
31 99 34 16  4  6  4  3  3  3  4  2  0  0  0  1  2  7  1  2
 1  0  1  1  0  0  2  1  1  4  8  1  2  5  0  1  2  1  6  3
 2  0  0  1  0  0  2  3  2  3  4  2  2  1  1  2  2  1  1  3
 1  0  3  2  3  4  4  0  0  2  1  0  0  0  0  0  1

SPEAKER=3   PHONEME=11   PITCH=130   TM= 37.00   ST= 53.80
 2  3  2  4 11 23 45 24 14  0  2  3  2  1  4  1  0  1  6  4
14 42 99 41 13  9  2  2  3  0  1  6  4  0  4  1  0  1  1  0
 2  3  2  3  4  1  0  3  2  0  5 16  2  2  6  2  3  9  1  0
 2  8  1  2  4  2  2  6  2  0  4  3  2  2  4  5  2  5  4  2
 2  4  9 12 22 12  4  2  1  0  1  2  0  1  4  2  2

SPEAKER=3   PHONEME=11   PITCH=146   TM= 49.40   ST= 66.50
 0  1  2  3  3  7  5  7 41 99 26 15  8  4  3  4  1  0  1  2
 0  4  6 15 23 90 28 12  8  4  0  2  2  2 11  6  1  1  0  1
 3 15  6  3  2  7  9  1  0  5 12  5  5  9 12  7  8 20  5 15
 4  3  3  2  2  3  4  2  2  2  1  2  4  2  2  6  9  8  5  6
 0  5 17 24 22 10  4  0  0  2  0  1  2  1  2  5  4

SPEAKER=3   PHONEME=11   PITCH=164   TM= 57.80   ST= 77.10
 1  1  0  2  2  2  1  3  7  7 18 42 99 28 18  8  3  5  2  0
 0  0  0  0  1  3  5 20 42 10  4  2  0  0  1  1  2 11  2  0
 1  1  0  2  4  1  1  3  5 16  3  2  7 22 10  5 13 26  7 15
16  5  3  4  3  4  3  7  3  5  4  3  1  2  3  4  7 16  6  4
 2  2  3  2  4  2  2  2  1  0  1  0  0  0  3  5  6

SPEAKER=3   PHONEME=11   PITCH=174   TM= 68.70   ST= 70.00
 1  0  0  2  0  2  1  1  5  6  8 14 43 99 44 20  7  7  4  3
 1  0  0  1  0  1  2  4 12 32  9  3  0  1  0  1  2  3 15  8
 1  0  0  1  1  5  1  1  0  3  8  1  0  1  1  1  1  2  7  3
 3  6  0  2  0  1  4  3  2  2  1  0  1  1  0  1  2  3  4  4
 1  1  7 12 13  5  3  1  0  2  0  0  0  0  0  1  2
```

TABLE XII (Continued)

SPEAKER=4 PHONEME= 1 PITCH=164 TM= 19.90 ST= 42.30

3	5	2	4	4	3	0	0	6	4	26	80	99	23	15	4	0	1	1	0
1	3	2	1	4	4	4	31	19	0	1	5	1	1	5	3	8	12	3	0
4	6	16	44	34	13	15	21	84	69	16	17	32	35	13	7	14	12	4	13
7	11	25	17	41	27	10	14	6	3	7	5	4	4	8	12	11	13	5	3
0	2	2	1	3	2	3	5	1	1	1	2	0	1	4	2	3			

SPEAKER=4 PHONEME= 1 PITCH=174 TM= 42.40 ST= 81.10

1	2	1	2	2	2	0	0	3	1	7	14	45	99	51	20	4	6	3	0
0	1	0	1	1	2	1	6	23	47	14	7	1	1	3	3	3	6	24	8
4	5	13	17	64	92	29	18	18	33	94	38	12	7	15	6	3	7	10	3
9	14	5	19	16	16	34	18	10	4	8	4	4	3	6	11	12	20	6	4
2	1	3	0	2	0	2	2	1	1	2	1	0	0	1	0	1			

SPEAKER=4 PHONEME= 1 PITCH=196 TM= 70.10 ST= 96.10

1	1	0	2	0	0	0	0	1	1	3	4	7	11	30	44	99	46	19	8
6	3	1	2	0	1	1	2	2	3	9	17	40	14	5	3	3	2	2	5
12	40	19	9	9	9	18	34	61	25	9	6	12	20	6	3	3	6	1	2
5	5	3	9	11	9	29	23	7	7	4	2	3	4	3	10	13	15	5	4
1	0	3	1	0	1	1	0	0	1	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 1 PITCH=220 TM= 51.90 ST= 82.20

0	0	1	0	0	2	0	0	2	0	0	0	2	1	8	10	17	44	99	26
11	5	2	0	0	3	1	1	4	2	3	5	11	23	56	17	8	4	5	4
5	11	46	88	37	17	13	14	28	57	75	33	14	7	9	11	3	4	6	15
4	6	14	16	17	41	33	22	28	12	11	12	7	11	15	41	13	14	7	6
2	3	3	0	1	1	0	0	1	1	0	2	0	0	0	0	0			

SPEAKER=4 PHONEME= 1 PITCH=246 TM= 70.70 ST= 95.80

0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	1	0	3	8	10
24	75	20	9	3	4	0	2	2	1	1	3	1	0	2	3	9	28	9	3
3	4	10	9	22	37	99	61	20	12	7	7	15	30	12	5	3	7	21	6
4	6	17	12	12	25	27	16	6	3	4	3	2	2	3	4	8	9	18	9
2	2	3	1	3	1	2	1	0	0	0	0	0	0	1	0	1			

SPEAKER=4 PHONEME= 1 PITCH=261 TM= 51.90 ST= 95.20

1	0	0	2	0	1	1	0	0	2	1	0	0	1	0	1	2	4	8	10
16	37	99	56	14	12	6	4	5	3	2	2	3	1	1	3	7	12	49	19
7	4	10	7	14	19	39	81	79	36	14	10	9	7	21	9	3	2	6	24
9	6	8	19	15	10	18	41	12	7	7	5	7	15	5	11	8	6	8	5
1	1	4	1	1	2	1	0	0	1	0	1	1	0	0	2	2			

SPEAKER=4 PHONEME= 1 PITCH=294 TM= 57.60 ST= 66.60

1	0	0	2	0	1	1	0	0	1	1	0	0	1	0	0	1	0	0	1
1	2	3	12	17	76	24	11	7	4	2	1	1	1	0	2	2	2	3	7
17	44	15	7	5	5	7	8	15	28	99	57	20	10	8	7	9	21	6	8
10	20	62	29	16	28	90	53	15	10	9	4	5	5	2	4	4	3	10	5
1	1	3	0	1	1	2	0	0	1	1	0	1	0	0	0	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 2 PITCH=164 TM= 50.30 ST= 65.90

1	2	0	1	1	1	0	0	5	4	11	26	99	32	16	7	1	2	2	0
1	2	1	0	1	3	5	19	43	11	6	5	3	2	6	9	24	64	25	10
6	8	24	34	59	20	7	4	5	10	2	2	2	4	3	1	0	1	0	0
1	2	1	3	4	6	7	19	20	36	23	15	8	10	13	16	24	14	7	5
2	3	6	3	5	2	2	2	0	0	0	1	0	0	1	1	1			

SPEAKER=4 PHONEME= 2 PITCH=174 TM= 72.60 ST= 95.80

1	1	0	1	2	2	0	0	5	4	9	17	52	88	37	16	5	5	4	1
2	2	1	1	2	4	4	10	35	44	16	8	5	4	6	8	16	28	99	40
14	9	12	11	37	40	11	6	3	4	5	2	1	2	5	2	1	3	2	1
3	3	1	5	4	6	9	15	34	24	44	19	14	11	16	29	26	23	9	6
3	4	7	4	6	2	3	2	1	1	1	0	0	0	1	0	1			

SPEAKER=4 PHONEME= 2 PITCH=196 TM= 86.60 ST= 95.80

0	0	0	1	0	1	0	0	2	1	1	3	6	7	21	30	99	51	20	8
5	3	2	1	1	3	2	3	6	6	12	20	68	29	11	8	8	6	11	17
29	83	75	28	18	11	7	5	9	4	1	1	1	0	1	1	1	1	1	0
0	1	0	0	2	2	8	13	10	19	9	5	4	3	4	12	8	12	5	3
1	1	3	0	2	1	1	0	1	0	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 2 PITCH=220 TM= 84.90 ST= 95.80

0	0	1	1	0	1	0	0	1	1	0	1	3	4	9	11	18	42	97	29
15	8	4	3	1	3	2	3	5	5	8	11	21	39	99	33	17	8	7	7
7	13	45	81	45	20	10	5	3	4	4	2	2	1	3	6	1	1	1	7
1	2	3	7	4	8	14	15	49	50	19	20	10	15	13	28	14	16	6	6
3	4	8	4	4	2	2	0	0	1	0	0	1	0	0	0	1			

SPEAKER=4 PHONEME= 2 PITCH=246 TM= 52.20 ST= 80.60

1	0	1	2	0	1	0	0	1	1	1	0	1	1	1	1	3	5	12	18
41	99	23	12	3	4	2	1	3	2	2	2	3	2	4	9	32	71	18	9
4	3	8	6	14	25	74	33	9	8	3	2	7	11	2	1	1	0	5	2
2	2	7	5	5	11	19	17	45	74	24	12	10	6	13	12	37	17	18	10
6	3	8	2	2	1	2	0	0	2	1	1	1	0	0	0	1			

SPEAKER=4 PHONEME= 2 PITCH=261 TM= 47.60 ST= 94.10

0	1	1	1	1	2	0	0	2	1	0	0	1	0	2	3	2	5	11	13
23	67	98	32	10	10	5	4	7	4	3	4	5	4	10	17	37	89	99	36
14	10	14	8	15	21	49	95	27	14	5	4	4	3	6	3	1	1	3	3
1	2	2	4	4	6	18	19	31	60	58	26	28	13	20	27	22	36	13	12
3	5	9	2	5	2	2	1	2	1	0	1	0	0	1	1	2			

SPEAKER=4 PHONEME= 2 PITCH=294 TM= 38.80 ST= 69.30

1	0	0	3	0	2	1	1	1	3	2	0	0	1	0	0	1	1	1	2
1	1	2	8	9	67	37	14	8	5	2	2	1	1	0	3	5	4	7	15
31	99	60	23	12	7	6	5	6	15	49	57	12	5	3	3	4	15	4	4
3	4	20	11	4	9	40	34	14	22	56	20	17	43	16	22	33	14	19	12
6	4	7	2	1	1	3	1	1	3	1	0	0	1	0	1	3			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 3 PITCH=164 TM= 66.80 ST= 85.20

1	1	0	1	1	1	0	1	4	3	10	25	70	20	12	5	1	2	1	0
0	1	0	0	1	3	4	14	33	8	4	3	1	2	3	5	12	37	12	7
8	11	37	52	99	37	18	12	16	33	8	5	5	12	5	1	2	4	1	2
3	2	4	6	9	16	16	42	15	11	8	6	4	4	6	8	13	17	5	4
2	2	4	2	2	0	1	0	0	1	1	1	0	1	0	0	1			

SPEAKER*4 PHONEME= 3 PITCH=174 TM= 95.60 ST= 96.50

0	1	0	1	0	1	0	0	3	2	5	11	33	59	24	10	3	3	2	0
1	1	0	0	1	2	2	5	19	32	11	5	3	2	3	3	5	9	31	12
7	10	27	27	90	99	38	19	11	8	5	5	4	3	7	2	1	1	1	1
4	4	3	10	9	17	22	14	11	5	9	4	4	5	7	12	9	6	2	2
1	1	2	0	1	0	0	0	0	0	0	0	0	0	1	0	0			

SPEAKER=4 PHONEME= 3 PITCH=196 TM= 26.30 ST= 69.30

2	3	1	4	3	3	0	1	1	0	2	2	1	4	20	33	99	42	13	1
1	2	0	1	3	3	0	4	3	1	7	19	68	22	7	6	4	7	6	8
25	95	59	22	15	12	17	36	76	28	7	7	12	35	12	4	3	9	3	1
5	7	0	7	15	11	38	66	31	57	28	21	13	13	16	35	39	69	23	15
7	5	8	3	7	2	4	3	2	2	3	1	0	0	2	0	3			

SPEAKER=4 PHONEME= 3 PITCH=220 TM= 58.30 ST= 86.00

1	1	0	2	0	0	0	0	0	0	2	1	0	2	4	4	9	25	71	21
10	4	1	2	0	1	1	1	1	1	3	3	6	15	52	16	8	4	4	5
9	14	56	99	66	26	14	8	5	7	15	4	2	2	4	8	3	1	0	3
2	1	3	10	5	12	28	15	19	20	8	7	4	5	4	12	9	16	4	5
1	0	2	1	0	1	2	1	0	2	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 3 PITCH=246 TM= 64.00 ST= 72.20

1	1	0	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	4	6
21	40	8	4	2	2	0	1	1	0	1	2	1	1	4	6	21	33	8	3
3	4	10	9	25	39	99	41	14	9	4	4	9	12	5	2	1	2	4	0
1	2	6	3	4	8	9	14	29	38	13	8	5	5	11	9	18	11	12	5
2	2	3	1	2	1	1	1	0	0	0	0	0	0	1	0	1			

SPEAKER=4 PHONEME= 3 PITCH=261 TM= 40.20 ST= 68.10

1	0	0	3	0	1	2	0	0	1	1	0	0	1	0	0	2	2	4	6
14	41	85	27	5	6	3	3	3	3	2	2	1	3	3	8	19	39	99	33
11	6	9	7	10	16	38	92	34	19	6	3	4	6	14	4	2	1	2	7
2	1	4	19	4	6	26	41	10	11	20	7	10	7	8	27	17	35	11	13
3	4	8	2	2	0	2	0	0	2	1	1	2	0	0	0	2			

SPEAKER*4 PHONEME= 3 PITCH=294 TM= 53.60 ST= 82.50

1	2	0	1	0	1	0	0	1	0	1	1	0	0	1	0	0	0	2	1
2	5	7	18	29	74	21	11	9	4	3	4	2	1	2	3	3	4	8	16
45	71	28	11	9	6	6	7	16	31	83	36	13	7	7	5	7	13	4	4
6	16	28	17	24	48	99	53	23	22	20	11	13	11	9	20	12	17	25	12
10	6	9	2	6	1	1	0	1	0	1	1	1	0	2	0	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 4 PITCH=164 TM= 48.20 ST= 84.50

1	2	0	1	1	2	0	1	6	5	11	26	99	47	23	10	1	4	3	1
1	3	1	1	1	7	10	28	96	44	19	11	4	2	4	5	12	44	29	8
3	3	4	3	13	3	0	1	2	3	1	1	1	1	3	1	0	1	1	0
1	2	1	1	3	2	2	6	8	10	25	43	35	29	22	34	54	34	12	8
4	5	10	8	13	8	4	3	1	0	1	1	0	0	2	1	1			

SPEAKER=4 PHONEME= 4 PITCH=174 TM= 88.20 ST= 92.50

0	1	0	1	1	1	0	0	4	3	6	12	36	66	26	12	3	4	4	2
2	3	2	3	2	7	10	20	60	99	38	19	10	6	4	3	3	6	22	7
2	2	2	1	4	5	0	0	1	1	4	1	1	0	2	0	0	0	0	0
0	1	0	1	1	1	1	3	4	4	14	15	26	14	13	22	24	16	7	4
1	2	4	2	4	2	2	1	0	0	0	1	0	0	1	0	0			

SPEAKER=4 PHONEME= 4 PITCH=196 TM= 81.00 ST= 91.10

0	0	0	0	0	1	0	0	1	0	0	1	3	3	16	25	50	23	10	3
2	2	1	0	1	3	2	5	10	12	24	41	99	35	15	9	6	3	3	2
2	11	4	1	1	1	0	2	3	2	0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	1	2	4	6	9	20	10	16	7	11	21	22	14	8
3	4	6	3	5	2	1	1	1	0	0	1	0	0	0	0	0			

SPEAKER=4 PHONEME= 4 PITCH=220 TM= 44.20 ST= 70.80

1	1	0	2	2	2	0	2	1	0	1	1	1	4	12	14	28	77	87	21
13	6	1	3	2	3	2	5	7	6	14	19	40	99	78	30	17	9	4	1
2	2	4	7	3	1	0	2	1	2	1	0	0	1	1	0	0	2	0	0
2	1	0	1	1	0	2	4	3	2	7	6	14	24	11	15	19	16	10	5
3	2	3	1	2	0	2	2	0	1	2	1	0	0	1	0	1			

SPEAKER=4 PHONEME= 4 PITCH=246 TM= 52.80 ST= 66.30

2	1	0	3	0	0	0	0	0	0	2	1	0	2	2	1	2	5	9	15
29	99	42	20	6	5	3	3	3	2	3	3	2	4	6	11	31	79	34	14
6	3	4	4	2	2	7	4	0	2	1	0	1	2	1	0	1	0	0	0
1	0	0	2	0	0	2	2	2	4	8	13	23	11	27	14	27	14	16	8
5	2	7	2	3	1	3	0	0	0	1	0	0	1	1	0	3			

SPEAKER=4 PHONEME= 4 PITCH=261 TM= 42.50 ST= 77.80

1	2	0	3	2	2	0	1	1	0	1	1	0	1	2	1	1	5	7	7
17	49	99	33	10	9	3	5	4	1	1	3	0	2	4	7	16	37	87	27
8	4	4	3	3	2	4	15	3	2	1	0	0	1	3	0	0	2	1	0
2	1	0	1	1	0	3	6	5	6	16	13	38	41	19	30	28	67	22	22
9	11	12	5	6	0	2	2	0	1	2	1	0	0	1	0	2			

SPEAKER=4 PHONEME= 4 PITCH=294 TM= 61.70 ST= 77.30

1	1	0	3	0	0	0	0	0	0	1	0	0	1	1	0	0	2	2	3
5	7	9	26	42	99	33	17	13	7	5	3	1	2	1	1	2	1	1	3
14	23	7	3	2	0	1	0	0	1	4	0	1	1	1	0	1	0	0	0
1	0	0	1	1	0	3	2	2	5	8	6	19	27	11	22	12	11	13	6
3	1	5	2	3	1	3	0	0	1	0	0	0	0	0	0	2			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 5 PITCH=164 TM= 66.90 ST= 80.40

0	1	0	1	0	1	0	0	1	1	3	10	41	10	5	2	0	0	1	0
0	1	0	0	0	1	0	6	24	6	2	3	1	0	1	3	4	14	7	5
6	10	32	41	98	39	21	21	47	99	34	22	21	38	25	13	11	26	9	8
12	5	2	2	1	2	1	3	2	2	2	3	2	1	1	1	0	1	5	7
7	4	4	0	1	0	0	0	1	0	0	1	0	0	1	0	0			

SPEAKER=4 PHONEME= 5 PITCH=174 TM= 72.30 ST= 83.20

0	0	1	1	0	1	0	0	0	1	1	3	17	35	11	4	0	0	0	0
0	0	1	0	0	1	0	1	11	23	6	3	2	1	1	2	2	3	17	9
4	6	18	18	62	99	32	16	13	21	56	23	13	14	53	27	10	11	26	8
3	4	2	2	0	1	1	1	2	2	2	1	2	0	0	1	0	0	3	5
3	3	4	1	0	0	1	0	0	1	0	0	1	0	0	0	1			

SPEAKER=4 PHONEME= 5 PITCH=196 TM= 94.60 ST= 95.70

0	0	0	1	0	0	0	0	0	0	1	0	2	3	10	14	51	32	11	4
3	1	1	1	0	1	1	1	1	2	4	7	28	15	5	3	4	3	4	7
13	35	51	19	15	14	20	30	99	52	18	13	18	34	28	13	9	15	19	6
2	2	1	2	1	1	2	2	1	1	1	0	1	1	1	1	1	1	7	8
3	1	2	1	1	0	1	0	0	0	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 5 PITCH=220 TM= 64.30 ST= 81.70

1	0	1	1	0	0	0	0	0	1	1	0	1	0	1	2	6	19	35	7
3	1	1	0	0	1	1	0	1	2	2	2	5	14	20	6	4	3	4	5
7	14	58	99	35	18	15	15	30	63	67	31	18	13	39	61	16	9	12	33
6	3	3	3	0	1	2	2	2	2	2	1	2	1	0	2	3	6	11	10
2	1	3	0	0	0	1	0	0	1	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 5 PITCH=246 TM= 95.60 ST= 96.90

0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	3	3
12	37	11	4	2	2	1	1	1	0	1	2	1	2	4	5	13	36	11	6
6	7	17	15	34	52	99	86	30	19	12	13	29	48	19	11	9	18	52	16
4	4	8	5	2	2	2	3	2	1	1	2	1	1	3	2	1	2	3	4
9	3	2	0	1	0	0	1	0	0	0	1	0	0	1	0	1			

SPEAKER=4 PHONEME= 5 PITCH=261 TM= 96.70 ST= 96.80

0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2	2
6	21	27	8	2	2	1	2	1	1	1	1	0	1	2	3	7	18	22	7
4	4	10	8	18	25	56	99	36	21	11	10	12	20	27	12	8	12	30	59
12	7	4	7	2	2	3	4	2	2	3	2	2	1	2	3	3	7	11	23
6	6	4	0	2	0	1	0	0	0	1	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 5 PITCH=294 TM= 94.60 ST= 96.50

0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0
1	2	2	7	10	52	20	9	6	3	2	2	1	1	2	3	4	4	8	14
28	86	39	16	12	10	10	10	18	28	99	69	25	14	16	18	32	83	28	14
7	6	9	4	2	2	4	3	2	2	3	2	1	1	0	1	1	1	4	4
10	6	3	0	1	0	0	0	0	1	0	1	0	0	0	0	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 6 PITCH=164 TM= 76.90 ST= 96.10

0	0	1	1	0	2	0	0	3	3	8	20	60	16	9	5	2	2	2	1
0	1	1	1	1	5	7	20	44	14	8	6	8	7	11	17	42	99	47	22
11	10	24	29	52	21	14	15	45	86	23	15	17	30	12	6	3	2	2	1
0	0	0	0	0	1	1	2	2	1	1	1	1	0	0	1	1	1	7	6
1	1	2	0	1	2	1	0	1	1	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 6 PITCH=174 TM= 87.40 ST= 95.40

0	0	0	1	1	1	0	1	1	0	2	4	14	36	21	7	2	3	2	0
1	1	0	1	1	2	2	5	12	30	18	8	5	5	7	9	18	26	82	93
28	16	17	10	20	34	17	13	18	28	99	71	23	13	13	9	4	3	2	2
2	1	0	1	0	0	1	2	1	1	2	2	1	1	1	0	0	1	1	2
1	0	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0			

SPEAKER=4 PHONEME= 6 PITCH=196 TM= 93.20 ST= 95.80

0	0	1	1	0	1	0	0	0	0	0	1	3	4	12	19	51	23	9	3
2	2	1	2	0	3	2	3	7	9	18	30	86	33	14	10	9	8	12	20
34	99	78	30	20	14	11	15	37	16	7	9	25	53	20	10	5	7	3	2
1	1	0	0	0	1	2	2	2	3	1	2	2	1	1	3	2	5	6	4
0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0			

SPEAKER=4 PHONEME= 6 PITCH=220 TM= 75.00 ST= 94.70

0	1	0	1	1	1	0	1	0	0	1	0	0	0	3	2	6	16	50	18
6	3	0	1	1	1	1	2	3	3	5	6	9	17	64	34	14	8	7	9
11	17	53	78	99	40	20	12	8	10	32	18	9	10	27	60	37	14	5	4
3	2	1	2	1	0	1	2	0	1	2	1	1	1	0	0	1	2	1	3
1	0	1	1	1	0	1	0	0	0	0	0	0	0	1	0	1			

SPEAKER=4 PHONEME= 6 PITCH=246 TM= 57.90 ST= 90.30

1	0	0	2	0	1	0	0	0	1	1	0	1	0	0	0	1	1	3	5
12	45	17	6	1	2	1	1	3	3	3	4	5	5	9	15	38	96	49	22
10	8	16	12	24	37	99	97	27	18	10	11	30	66	30	14	6	6	25	12
2	2	2	3	1	2	2	2	1	2	1	1	2	1	1	2	2	2	11	10
1	1	2	0	0	1	1	0	0	1	0	0	0	0	1	1	2			

SPEAKER=4 PHONEME= 6 PITCH=261 TM= 63.30 ST= 82.90

1	1	0	1	0	0	0	0	1	0	2	0	0	0	0	0	0	1	1	1
4	12	34	10	2	2	0	1	1	1	2	3	2	3	5	7	16	33	99	34
12	8	10	8	11	14	31	71	38	20	9	8	14	23	67	25	9	5	5	20
4	1	1	3	1	0	2	2	1	1	2	0	2	1	1	2	2	5	4	9
2	1	2	0	0	0	2	0	0	1	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 6 PITCH=294 TM= 81.70 ST= 95.30

0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1
0	0	0	3	5	28	7	3	2	2	1	1	1	1	1	3	5	5	10	19
42	99	39	16	11	8	8	7	13	24	75	34	14	9	15	20	44	84	24	14
6	7	13	5	2	2	3	1	2	2	1	1	2	1	0	2	1	2	6	4
6	3	5	0	1	1	1	0	0	0	0	0	0	0	0	1	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 7 PITCH=164 TM= 96.70 ST= 97.00

0	1	0	1	1	1	0	0	2	2	5	11	40	35	15	7	2	3	2	0
1	1	1	2	1	4	6	14	49	28	12	7	4	3	4	4	5	6	9	10
11	18	48	53	99	97	43	26	24	40	29	18	16	25	32	13	7	17	7	5
6	3	1	2	2	3	2	3	2	3	4	5	4	7	5	6	4	2	3	5
11	7	5	1	2	1	1	1	1	2	2	1	0	0	1	0	2			

SPEAKER=4 PHONEME= 7 PITCH=174 TM= 96.80 ST= 96.90

0	1	0	1	0	1	0	0	1	1	2	5	16	39	22	8	2	2	2	0
1	1	1	1	1	3	3	6	20	46	21	9	5	4	5	5	7	8	29	28
12	14	33	31	86	99	85	39	24	25	79	54	20	13	24	23	7	6	16	6
4	7	3	2	2	1	3	3	5	4	11	9	12	6	7	4	2	2	2	5
7	12	7	1	3	1	1	1	0	1	1	1	0	0	0	0	1			

SPEAKER=4 PHONEME= 7 PITCH=196 TM= 60.30 ST= 63.00

0	1	0	2	1	1	0	0	0	0	1	0	0	0	2	5	12	4	1	0
0	0	0	0	1	1	0	0	1	0	2	8	19	5	3	2	1	2	3	4
12	43	31	11	13	14	24	43	99	40	14	11	15	30	13	5	4	15	6	1
2	2	1	1	1	0	0	1	2	1	2	2	2	0	1	1	0	1	3	4
7	8	6	0	1	0	0	0	0	0	0	1	0	0	0	0	0			

SPEAKER=4 PHONEME= 7 PITCH=220 TM= 95.80 ST= 96.10

0	0	0	1	0	0	0	0	0	0	0	0	0	1	3	4	7	16	47	28
12	5	2	2	0	2	1	2	3	3	4	4	7	13	48	29	13	7	6	8
9	14	45	63	93	38	21	15	20	32	99	62	25	15	22	41	25	9	5	4
3	2	2	3	1	1	2	2	1	2	2	2	3	3	4	4	3	2	3	7
6	10	6	1	1	0	1	0	0	1	0	1	0	0	0	1	1			

SPEAKER=4 PHONEME= 7 PITCH=246 TM= 85.20 ST= 95.40

0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	3	3
9	33	13	4	1	2	0	1	2	1	1	2	2	2	4	8	23	56	18	9
5	6	12	10	23	37	99	69	23	16	10	12	28	51	22	11	5	6	17	6
2	2	1	1	0	0	0	1	1	1	1	1	1	0	1	0	0	0	2	4
13	9	8	1	3	0	0	0	0	0	0	1	0	0	1	0	1			

SPEAKER=4 PHONEME= 7 PITCH=261 TM= 89.70 ST= 95.80

0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	2	3	3
8	18	48	27	7	5	3	2	2	2	2	2	2	2	3	5	10	17	62	29
10	7	14	11	20	26	48	99	81	39	18	15	17	23	64	30	11	8	7	19
5	3	2	4	1	1	2	4	1	2	2	1	1	1	0	2	1	2	5	15
6	4	4	1	1	0	1	0	0	1	0	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 7 PITCH=294 TM= 95.70 ST= 95.80

1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1	1	2	6	10	40	11	4	3	2	1	1	1	1	1	1	2	2	5	10
25	49	20	9	8	7	9	10	20	34	99	51	21	12	13	14	27	54	14	8
4	4	9	3	1	1	2	2	1	1	2	1	2	1	1	2	1	1	5	9
26	13	8	1	2	1	1	0	0	0	0	0	1	1	0	1	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 8 PITCH=164 TM= 95.70 ST= 96.40

0	0	0	1	1	1	1	1	4	4	9	17	65	26	14	7	2	3	3	2
2	2	2	4	3	9	16	33	99	58	27	14	8	5	5	6	13	33	33	13
5	3	4	3	8	3	2	2	2	8	6	3	6	14	22	8	5	15	5	2
3	1	0	2	0	0	1	2	1	1	2	1	1	1	1	4	14	7	1	2
1	0	1	1	3	1	1	0	0	0	0	0	0	0	0	0	0			

SPEAKER=4 PHONEME= 8 PITCH=174 TM= 96.40 ST= 96.50

1	1	0	2	1	1	0	0	2	2	4	7	19	44	41	16	5	5	4	3
4	4	3	6	4	11	16	26	61	98	99	44	23	15	11	8	10	13	38	54
14	8	8	5	5	6	4	2	1	4	8	8	8	12	42	61	17	11	25	14
5	9	6	5	8	4	6	7	5	6	4	3	2	2	2	4	7	11	3	3
2	2	8	13	13	4	2	1	0	0	1	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 8 PITCH=196 TM= 37.40 ST= 51.40

1	2	1	3	2	2	0	1	1	0	1	1	0	1	2	3	19	17	3	0
0	1	0	0	1	3	1	4	7	7	17	28	99	91	26	13	7	5	3	1
3	7	17	5	3	1	1	2	8	6	2	1	2	6	10	2	4	16	40	8
3	2	0	1	1	0	2	4	2	2	4	2	2	1	1	2	2	3	2	2
2	1	2	2	1	0	1	1	0	1	2	2	0	2	1	0	1			

SPEAKER=4 PHONEME= 8 PITCH=220 TM= 73.30 ST= 80.20

1	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1	3	17	5
1	1	1	0	0	1	1	1	4	4	7	9	18	30	99	60	24	11	6	4
2	2	6	8	8	4	1	0	1	2	8	3	2	1	5	15	8	3	3	9
5	1	1	3	0	1	1	1	2	1	1	1	1	0	0	3	3	2	2	2
0	1	2	0	1	0	0	0	0	1	0	1	0	0	0	0	0			

SPEAKER=4 PHONEME= 8 PITCH=246 TM= 90.00 ST= 91.70

0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
5	21	17	6	2	2	1	1	3	1	3	4	4	5	10	15	35	99	65	27
12	8	9	5	5	4	7	8	2	1	0	1	0	2	2	1	1	4	16	12
1	1	1	3	1	0	1	2	1	0	1	0	0	0	2	2	3	4	1	1
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0			

SPEAKER=4 PHONEME= 8 PITCH=261 TM= 53.70 ST= 59.00

1	1	1	2	0	1	0	0	0	1	1	0	1	1	0	1	0	0	0	1
2	13	25	7	1	3	0	0	1	1	0	2	3	4	6	12	25	51	99	33
11	7	8	4	4	4	4	11	2	2	0	0	2	2	5	3	2	4	18	44
5	2	1	2	0	1	2	2	2	1	1	0	2	0	0	2	1	2	2	2
0	0	1	0	0	1	1	0	1	1	0	1	0	0	0	0	0			

SPEAKER=4 PHONEME= 8 PITCH=294 TM= 44.00 ST= 65.50

1	0	0	3	0	2	1	0	0	1	1	0	0	1	0	0	1	1	1	2
2	3	4	17	32	67	21	11	8	4	2	2	1	1	1	3	5	5	10	22
57	99	38	15	8	5	3	2	2	7	15	4	3	1	1	1	4	7	3	6
16	38	64	22	5	3	5	3	3	5	5	3	7	5	5	21	6	3	4	5
3	2	6	1	2	1	1	0	0	1	0	1	1	1	1	2	1			

TABLE XII (Continued)

SPEAKER=4 PHONEME= 9 PITCH=164 TM= 29.70 ST= 66.60

2	3	2	4	2	2	0	0	2	0	5	15	73	23	11	4	0	1	1	0
0	1	0	1	2	5	5	20	78	29	11	9	5	6	10	16	37	90	86	32
14	11	23	28	99	49	16	12	15	37	22	11	10	22	33	16	22	76	28	24
62	27	26	55	19	36	21	36	24	23	17	8	5	2	2	1	1	3	4	5
6	9	10	1	2	0	2	1	0	2	3	2	0	1	1	0	1			

SPEAKER=4 PHONEME= 9 PITCH=174 TM= 80.80 ST= 95.90

0	0	0	1	0	1	0	0	1	1	3	6	22	42	16	6	1	2	1	0
1	0	0	0	0	1	2	6	24	29	11	6	3	3	6	7	17	30	99	39
15	10	15	14	50	60	19	10	7	13	28	12	7	9	31	15	11	22	36	21
51	75	25	19	12	16	30	23	48	22	11	6	4	3	2	2	2	3	3	4
4	6	6	2	2	1	1	0	0	0	1	0	0	0	0	0	1			

SPEAKER=4 PHONEME= 9 PITCH=196 TM= 70.50 ST= 86.30

1	1	0	1	0	1	0	0	0	0	1	1	0	0	2	4	22	18	4	0
0	1	0	0	1	1	0	0	2	1	3	6	23	17	5	4	4	4	7	10
19	47	99	35	20	11	7	7	27	16	4	4	6	12	13	6	6	15	30	13
23	53	30	16	25	16	16	39	16	11	11	5	3	1	2	1	0	1	1	1
2	2	3	0	1	0	1	1	0	0	0	0	0	0	1	0	1			

SPEAKER=4 PHONEME= 9 PITCH=220 TM= 72.00 ST= 94.80

1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	6	19	3
1	1	0	0	0	0	0	1	1	1	3	4	7	16	47	14	7	5	4	6
9	16	57	99	56	25	13	8	5	7	9	4	4	5	13	30	8	8	15	49
15	12	26	45	11	16	24	17	34	25	9	5	3	2	1	1	1	1	1	2
3	1	5	1	1	0	1	0	0	0	0	0	0	0	1	0	2			

SPEAKER=4 PHONEME= 9 PITCH=246 TM= 57.40 ST= 94.30

0	1	1	2	0	2	0	0	1	0	0	0	0	0	0	1	1	0	2	1
4	22	7	1	0	2	1	1	2	1	1	2	2	1	5	10	26	69	34	15
6	5	9	6	13	23	62	64	17	12	6	7	14	29	13	9	10	22	86	40
21	35	99	69	20	21	31	21	26	44	16	9	5	2	2	2	1	1	3	2
2	2	3	0	1	0	0	0	0	0	0	1	1	0	1	0	1			

SPEAKER=4 PHONEME= 9 PITCH=261 TM= 58.40 ST= 95.60

1	0	0	2	0	1	1	0	0	0	1	0	0	1	0	0	1	1	1	1
3	8	27	21	3	3	2	2	2	2	2	2	2	2	4	7	16	28	99	58
17	9	11	7	9	12	23	46	55	25	10	8	11	15	57	33	12	10	18	54
31	18	25	61	56	28	22	50	22	14	26	13	6	3	1	2	1	1	2	3
1	2	5	1	1	1	1	0	0	1	0	1	1	0	0	1	1			

SPEAKER=4 PHONEME= 9 PITCH=294 TM= 71.80 ST= 86.40

0	1	0	1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0
0	1	0	1	2	19	3	1	1	0	1	2	1	0	2	2	3	4	8	14
34	99	38	15	11	6	4	3	5	8	37	19	5	3	4	4	7	24	6	5
10	22	63	24	9	10	28	16	8	12	26	9	5	7	2	2	1	2	1	1
1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1			

TABLE XII (Continued)

```
SPEAKER=4    PHONEME=10    PITCH=164    TM= 41.90    ST= 69.10
 1  2  0  2  0  2  0  0  3  3  8 22 75 20  9  5  0  2  2  1
 0  2  1  0  0  5  6 20 67 35 14 10  6  5  8 15 34 65 99 37
15 13 27 29 94 44 16 12 25 60 26 16 16 32 21  8  3  9  3  3
 2  3  1  1  0  1  1  3  3  3  3  3  3  0  1  4  6 15 17 16
12 11  8  0  1  1  0  0  1  1  1  2  0  0  0  0  0

SPEAKER=4    PHONEME=10    PITCH=174    TM= 43.90    ST= 74.20
 0  1  0  3  1  2  0  1  1  0  1  3 14 44 28  9  2  3  2  0
 0  1  0  0  1  3  2  8 23 61 31 12  5  4  6  8 16 27 92 99
27 14 14  8 14 25 10  7  8 16 65 46 12  7 18 14  3  3  3  1
 2  2  0  0  0  0  2  3  3  2  4  2  2  0  2  2  4  7 12 15
 3  2  2  0  0  0  1  1  0  1  2  1  0  0  1  0  1

SPEAKER=4    PHONEME=10    PITCH=196    TM= 44.90    ST= 68.90
 0  1  1  3  0  2  0  0  0  1  0  0  1  2  4  9 39 25  5  2
 0  1  1  1  0  3  1  2  5  7 14 25 92 56 17 11  9  6  8 14
22 65 99 35 18 12  9 11 45 25  6  7 19 41 33 13  4  4  5  2
 0  2  1  2  1  2  2  2  1  2  1  1  3  1  2  7  4  5  8 17
 1  1  3  1  0  2  1  0  0  1  0  1  1  0  1  2  1

SPEAKER=4    PHONEME=10    PITCH=220    TM= 66.40    ST= 76.10
 0  0  1  2  0  2  0  0  0  0  0  0  1  0  1  2  2  8 31 12
 3  1  1  1  0  2  1  1  3  3  5  8 15 27 99 44 19  7  5  4
 3  4 20 31 31 12  7  5 10 18 61 25  8  3  5  9  2  1  1  1
 0  0  0  0  0  1  1  1  2  1  1  0  1  0  0  1  1  5  3  5
 0  0  1  0  0  1  1  0  1  1  0  0  0  0  0  1  1

SPEAKER=4    PHONEME=10    PITCH=246    TM= 46.30    ST= 58.30
 1  2  1  2  2  1  0  0  1  0  1  1  0  1  2  0  0  0  1  0
 5 22  2  1  1  1  0  1  1  0  1  2  1  2  6 11 34 67 17  7
 4  4  9  8 21 36 99 46 14  9  5  7 23 38 13  7  6 18 62 15
 3  4 10  4  2  2  2  4  2  1  2  2  1  1  3  4  8 10 25  9
 3  2  3  0  2  1  1  1  0  0  0  1  0  0  1  1  2

SPEAKER=4    PHONEME=10    PITCH=261    TM= 80.30    ST= 86.70
 1  0  1  1  0  0  0  0  0  0  1  0  1  0  0  0  0  0  0  1
 2  7 21  6  0  1  0  0  1  1  1  2  3  3  5  9 18 33 99 33
11  7  8  5  5  6 12 28 10  6  3  5 13 23 46 18  6  3  1  3
 0  0  0  1  0  1  1  1  1  1  1  0  1  0  0  1  0  1  2  5
 0  2  3  1  1  1  1  0  0  0  0  0  0  0  0  0  1

SPEAKER=4    PHONEME=10    PITCH=294    TM= 58.70    ST= 89.60
 1  2  0  1  0  1  0  0  1  0  1  1  0  0  1  0  0  0  0  0
 0  1  1  4 11 25  5  2  3  1  1  3  2  1  3  4  6  8 16 28
75 99 42 18 14  8  5  4  5 10 26 11  8  7 17 25 60 74 20 11
 4  5  3  2  2  1  2  3  2  1  2  1  1  1  1  1  0  2  1  1
 0  0  0  0  1  1  1  1  0  0  0  0  0  0  1  0  1
```

TABLE XII (Continued)

SPEAKER=4 PHONEME=11 PITCH=164 TM= 42.50 ST= 52.30

1	1	0	3	1	2	0	2	4	4	11	26	99	46	22	10	2	4	2	0
0	0	0	0	1	2	3	13	44	15	5	2	0	0	1	1	2	10	6	0
2	1	0	1	3	0	0	1	1	3	2	0	0	1	2	0	2	13	4	5
18	5	2	6	3	6	4	8	5	4	6	3	3	2	1	3	7	10	14	15
5	4	6	3	2	0	1	1	0	1	1	1	0	1	1	0	1			

SPEAKER=4 PHONEME=11 PITCH=174 TM= 51.60 ST= 78.00

0	1	0	3	2	3	1	2	6	4	9	17	44	99	91	37	11	12	7	3
2	2	0	1	1	4	3	9	26	66	40	15	4	3	3	2	3	5	16	24
4	1	1	2	3	7	3	2	1	2	6	5	1	2	18	28	4	4	16	5
2	1	0	0	0	0	2	4	3	2	4	3	3	3	1	1	5	14	9	5
2	1	1	0	0	0	1	1	0	1	2	1	0	0	0	0	1			

SPEAKER=4 PHONEME=11 PITCH=196 TM= 80.70 ST= 87.90

0	1	0	1	0	1	0	0	2	1	2	4	7	9	26	38	99	49	22	9
6	4	2	1	1	2	1	2	4	4	10	17	56	22	8	4	2	1	1	0
1	5	4	1	1	0	0	0	1	0	0	0	1	4	2	0	0	2	0	0
1	2	0	1	2	1	1	3	1	2	2	3	2	2	2	4	5	11	9	5
0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0			

SPEAKER=4 PHONEME=11 PITCH=220 TM= 44.80 ST= 60.50

1	2	1	2	1	1	0	0	1	0	1	2	2	3	9	9	14	33	99	84
31	14	6	4	2	3	0	1	3	1	3	6	9	16	63	54	17	5	3	0
1	2	5	8	13	2	0	0	1	1	3	2	1	1	4	10	5	1	1	0
2	2	1	6	2	1	6	4	3	3	3	5	5	1	3	3	5	15	11	4
1	2	3	0	2	0	1	0	0	0	1	1	1	0	2	0	0			

SPEAKER=4 PHONEME=11 PITCH=246 TM= 52.70 ST= 71.80

1	0	0	3	1	1	1	0	0	0	1	0	1	2	3	3	4	8	14	22
57	99	25	14	4	6	3	3	3	2	2	1	1	3	5	12	39	62	14	7
3	1	2	3	3	7	22	7	1	3	1	0	5	8	1	1	3	11	33	8
2	1	1	1	0	2	3	3	9	9	5	5	5	5	11	8	20	8	15	7
3	1	4	0	0	0	1	0	0	1	0	1	1	0	0	0	1			

SPEAKER=4 PHONEME=11 PITCH=261 TM= 55.70 ST= 72.70

1	2	0	1	0	1	0	0	1	1	1	1	0	0	1	1	0	4	7	8
15	38	99	43	12	10	3	4	4	2	2	3	1	1	2	5	9	20	79	29
8	6	4	2	1	1	1	6	3	2	1	2	1	2	17	5	0	1	3	14
3	3	1	6	2	1	4	16	5	4	14	7	4	5	3	9	5	15	11	9
3	4	5	1	2	1	1	0	1	1	0	2	0	0	0	0	0			

SPEAKER=4 PHONEME=11 PITCH=294 TM= 49.50 ST= 57.00

1	0	0	3	1	1	1	0	0	0	1	0	0	1	1	0	1	0	0	0
2	4	7	17	26	99	33	15	10	4	3	1	1	1	2	1	2	2	2	5
18	56	24	9	4	2	2	0	0	2	4	1	1	1	1	0	2	5	0	0
1	3	19	8	1	1	4	3	1	2	7	1	3	6	2	5	6	4	7	3
1	0	2	1	0	0	2	0	0	1	0	0	0	0	0	0	2			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 1 PITCH=164 TM= 48.80 ST= 90.00

1	0	1	3	2	1	2	0	0	0	2	3	35	13	4	1	3	1	0	0
0	0	0	1	0	2	4	9	30	8	2	0	1	1	0	3	6	17	6	6
5	6	30	42	99	38	20	16	36	87	25	12	9	13	6	4	6	14	3	6
5	2	4	8	4	12	13	27	27	53	45	51	28	33	37	40	33	16	6	8
2	2	9	9	8	6	5	1	2	5	3	2	2	1	0	1	5			

SPEAKER=5 PHONEME= 1 PITCH=174 TM= 66.40 ST= 96.30

1	2	0	1	1	1	0	0	1	0	2	4	15	40	23	7	0	1	1	0
0	1	1	0	2	1	0	2	12	30	16	7	2	1	3	2	2	4	18	14
5	6	16	16	52	99	39	20	15	22	77	42	13	9	18	14	4	6	15	4
4	9	4	8	12	11	25	29	62	69	51	53	22	23	38	33	31	15	6	4
1	2	3	2	6	4	4	4	3	3	4	3	0	1	2	1	2			

SPEAKER=5 PHONEME= 1 PITCH=196 TM= 85.40 ST= 96.20

0	0	1	1	0	1	0	0	0	1	0	1	4	5	16	28	44	20	9	3
1	1	1	1	0	2	0	0	2	3	6	14	29	9	4	5	4	3	5	9
18	63	33	14	17	22	37	66	99	43	17	14	12	9	6	5	3	3	2	3
3	5	6	13	17	22	60	50	21	20	9	14	10	14	5	5	2	2	4	4
3	3	6	3	6	2	3	1	3	2	1	1	1	0	0	1	2			

SPEAKER=5 PHONEME= 1 PITCH=220 TM= 58.00 ST= 93.20

0	1	1	2	0	3	0	0	0	1	0	0	1	1	1	3	3	10	39	13
3	2	2	1	0	3	0	0	0	1	0	2	2	4	20	8	3	1	2	2
2	7	32	52	46	20	12	10	19	37	99	44	18	9	9	12	4	2	2	5
0	2	5	13	6	12	26	19	26	35	14	16	13	17	18	29	17	16	6	4
0	1	4	0	4	4	3	1	5	3	2	3	2	0	1	2	2			

SPEAKER=5 PHONEME= 1 PITCH=246 TM= 62.00 ST= 95.20

0	1	1	2	0	3	0	0	0	0	0	0	1	1	0	2	0	0	1	3
9	34	7	3	0	3	0	0	0	0	0	1	1	0	1	4	9	21	4	2
0	3	10	8	21	38	99	47	16	11	6	9	32	48	14	8	4	4	14	3
1	3	12	8	7	16	24	20	42	62	21	13	15	7	17	11	13	5	4	2
1	1	3	0	3	4	3	0	2	1	1	2	1	0	1	1	2			

SPEAKER=5 PHONEME= 1 PITCH=261 TM= 65.90 ST= 96.10

0	0	0	1	0	1	1	1	0	2	1	0	0	1	0	0	1	2	2	5
6	19	55	20	3	4	3	2	2	3	1	1	0	1	0	4	8	15	52	22
7	4	10	9	15	24	48	99	73	38	16	11	10	10	27	11	5	4	3	11
3	2	4	15	6	9	25	55	19	17	29	12	16	29	14	34	21	23	13	10
2	1	4	2	4	5	10	3	1	4	6	5	6	3	0	1	3			

SPEAKER=5 PHONEME= 1 PITCH=294 TM= 77.70 ST= 88.50

1	0	2	1	0	0	0	0	0	1	1	0	2	0	0	0	0	0	0	1
0	0	2	3	6	25	5	1	1	1	1	1	2	0	0	1	2	1	4	9
23	49	20	7	5	5	7	6	15	30	99	45	18	9	6	4	3	4	1	2
2	2	6	2	1	3	11	6	10	21	30	13	16	22	8	13	9	7	11	7
2	1	4	1	1	1	1	0	0	1	0	1	3	1	0	0	1			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 2 PITCH=164 TM= 52.30 ST= 78.30

1	2	1	1	3	1	0	0	1	0	3	11	28	6	4	1	0	2	0	0
0	0	0	0	1	2	1	10	17	2	0	2	0	0	3	3	6	20	7	4
6	10	34	51	99	37	18	16	30	63	16	8	6	12	6	1	2	5	2	1
4	1	1	3	4	6	9	22	24	35	24	14	7	8	11	12	22	34	22	16
6	4	5	4	5	0	3	4	7	9	7	4	1	3	2	1	5			

SPEAKER=5 PHONEME= 2 PITCH=174 TM= 43.80 ST= 72.20

1	4	0	0	0	2	0	0	2	2	1	6	7	28	15	4	0	1	2	0
0	3	0	0	1	1	0	4	16	42	17	9	3	1	5	4	7	16	63	50
14	10	17	15	54	99	37	16	8	4	14	9	4	2	8	4	0	1	3	0
2	4	1	3	7	3	10	12	19	16	10	9	12	15	25	17	14	13	6	3
1	2	3	4	11	5	3	4	5	13	11	6	0	1	2	1	2			

SPEAKER=5 PHONEME= 2 PITCH=196 TM= 48.70 ST= 73.00

1	2	1	2	3	1	0	0	1	0	1	1	0	1	5	6	26	10	2	0
0	1	0	1	2	1	0	2	1	0	3	9	33	12	6	5	4	8	9	15
31	99	85	33	21	14	11	19	51	19	4	4	1	5	4	2	1	5	1	0
1	2	0	2	5	3	10	22	18	38	24	19	14	14	13	19	15	33	15	13
5	3	4	4	12	11	8	4	2	2	8	5	0	0	1	0	2			

SPEAKER=5 PHONEME= 2 PITCH=220 TM= 46.50 ST= 64.50

1	0	0	2	0	1	1	1	1	3	1	0	0	1	0	1	4	10	40	20
5	1	0	0	0	1	1	0	1	3	1	1	3	5	30	15	7	2	4	6
7	11	47	69	99	37	18	8	7	11	33	17	7	2	2	3	2	0	1	3
3	1	4	6	2	4	17	13	35	65	26	21	16	12	13	20	21	19	13	13
3	3	8	2	5	6	6	0	1	3	2	4	5	1	0	1	2			

SPEAKER=5 PHONEME= 2 PITCH=246 TM= 64.90 ST= 71.40

0	1	0	0	0	1	0	0	2	0	0	0	0	0	0	1	0	2	5	6
17	44	9	2	0	2	0	1	2	1	0	3	1	0	2	6	18	45	12	5
3	5	9	8	22	39	99	46	16	11	4	5	5	5	2	1	0	1	3	2
1	3	3	1	2	5	5	9	26	37	15	16	12	6	14	9	12	10	12	8
3	4	3	0	2	1	2	1	1	0	1	3	1	0	1	0	0			

SPEAKER=5 PHONEME= 2 PITCH=261 TM= 28.80 ST= 57.30

5	2	0	2	2	0	1	3	3	1	6	1	0	0	2	0	1	3	3	1
9	17	57	18	3	0	1	4	4	1	5	2	0	0	3	1	10	25	99	35
14	6	7	6	11	10	32	78	63	27	12	5	1	2	6	0	1	3	2	2
6	2	0	4	3	0	8	22	14	23	85	37	25	45	15	25	14	32	13	12
7	5	7	6	13	7	20	6	1	3	6	4	0	2	1	0	3			

SPEAKER=5 PHONEME= 2 PITCH=294 TM= 58.60 ST= 75.20

1	0	1	1	0	1	1	0	1	2	1	0	1	0	0	0	1	0	0	2
1	0	2	5	8	51	15	5	4	3	2	1	2	0	0	2	4	3	8	17
36	99	40	15	8	4	3	0	2	6	22	9	4	1	0	0	1	3	0	1
1	1	2	1	0	0	4	1	2	5	12	4	7	12	2	3	2	0	5	3
1	1	4	0	0	1	1	0	0	2	1	2	3	0	0	1	1			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 3 PITCH=164 TM= 58.20 ST= 83.70

1	0	0	2	0	1	2	0	1	2	5	13	52	15	6	3	2	1	0	1
0	0	0	1	0	4	7	16	50	20	7	3	3	4	5	10	25	51	50	20
9	8	28	36	99	41	18	9	7	17	4	2	3	7	4	2	3	7	1	3
3	1	2	6	3	7	6	12	9	12	19	28	25	28	16	20	12	5	3	4
1	0	3	1	2	2	3	2	6	9	5	2	0	1	0	1	2			

SPEAKER=5 PHONEME= 3 PITCH=174 TM= 50.80 ST= 94.50

0	1	1	1	0	3	0	0	1	0	0	3	18	45	18	8	1	1	2	0
0	0	1	0	0	3	1	4	16	36	11	6	4	2	5	9	15	23	92	71
20	12	19	15	51	99	39	16	9	9	20	11	5	1	12	11	3	2	9	2
1	4	2	1	4	5	12	11	25	21	36	39	35	20	21	18	21	34	19	12
3	3	7	1	8	6	4	3	8	7	8	13	8	2	2	2	3			

SPEAKER=5 PHONEME= 3 PITCH=196 TM= 73.40 ST= 90.90

0	1	0	1	2	1	0	0	0	0	0	1	1	2	10	16	24	11	4	0
1	1	0	0	1	1	0	2	2	1	5	12	29	11	6	5	7	7	12	18
37	99	79	31	21	12	9	13	29	11	3	3	2	6	4	2	1	5	2	0
2	3	0	2	3	2	5	9	7	14	11	18	13	19	14	23	14	26	5	3
1	1	2	1	6	5	2	3	5	4	3	1	0	0	1	0	1			

SPEAKER=5 PHONEME= 3 PITCH=220 TM= 83.10 ST= 91.20

0	1	0	0	0	1	0	0	1	0	0	1	0	0	1	1	2	10	27	8
2	2	0	0	0	1	0	1	2	1	1	3	4	9	24	8	3	3	5	6
8	16	57	99	46	22	10	7	6	11	11	6	2	1	3	4	0	1	1	3
1	2	2	3	1	2	2	4	12	11	19	34	16	22	10	16	12	26	13	10
4	5	6	3	6	4	4	2	2	2	4	9	6	2	2	1	1			

SPEAKER=5 PHONEME= 3 PITCH=246 TM= 35.80 ST= 54.40

2	0	0	3	0	2	1	1	1	4	2	0	0	1	0	0	2	1	4	9
21	58	13	4	0	2	1	0	2	3	1	2	3	2	3	10	36	79	22	11
5	3	11	7	16	32	99	55	14	12	5	3	8	8	1	0	1	0	5	3
3	2	7	3	1	6	11	6	16	28	13	17	26	9	18	11	16	9	24	17
8	5	10	3	4	4	8	0	1	3	1	3	7	3	1	2	2			

SPEAKER=5 PHONEME= 3 PITCH=261 TM= 55.40 ST= 96.10

1	0	1	3	1	1	1	0	0	0	1	0	0	1	1	0	2	1	1	2
7	21	47	15	4	3	3	1	1	0	1	0	1	2	3	5	14	28	63	20
7	3	8	7	13	18	44	99	34	18	7	3	4	5	6	3	3	2	1	4
1	0	1	7	2	4	12	19	19	35	76	34	37	42	18	33	18	28	8	8
2	3	6	4	8	4	4	2	2	7	9	5	1	2	0	0	3			

SPEAKER=5 PHONEME= 3 PITCH=294 TM= 94.60 ST= 95.90

0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0
0	1	1	4	7	40	17	7	6	3	2	2	1	0	2	3	4	6	10	17
33	99	49	19	13	8	5	3	4	6	18	17	4	1	1	1	0	4	1	1
1	2	6	2	0	1	2	5	3	3	7	8	16	42	16	9	4	3	10	5
1	2	2	0	1	0	0	0	0	0	0	1	0	0	0	0	0			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 4 PITCH=164 TM= 43.10 ST= 63.80

2	2	0	5	1	0	0	0	0	1	5	11	52	19	7	1	1	1	0	0
1	0	1	3	1	6	12	30	99	38	16	7	2	3	2	2	5	15	5	1
2	0	1	3	2	0	2	0	0	1	1	0	1	2	2	0	3	0	0	0
1	0	0	2	1	0	4	3	1	2	4	2	6	10	26	42	50	41	26	22
5	2	7	9	12	5	5	2	0	3	2	0	0	1	0	1	3			

SPEAKER=5 PHONEME= 4 PITCH=174 TM= 42.40 ST= 90.50

1	3	1	3	3	2	0	0	2	0	4	9	30	68	51	18	3	4	3	0
1	2	0	2	3	6	6	16	45	99	77	30	9	6	6	3	1	3	4	0
1	2	0	2	3	3	0	2	0	1	3	2	0	1	4	1	0	3	1	0
1	2	0	2	2	1	2	6	3	2	6	7	8	14	38	55	53	33	42	55
47	36	30	28	62	45	14	9	6	6	3	5	2	2	3	2	3			

SPEAKER=5 PHONEME= 4 PITCH=196 TM= 45.40 ST= 65.50

1	3	1	1	1	1	0	0	1	0	1	2	1	0	6	9	44	27	8	0
1	2	1	0	1	1	0	1	6	5	16	29	99	51	18	8	3	2	2	0
2	7	8	2	3	1	0	0	2	1	1	2	1	0	2	0	0	0	1	0
1	2	1	1	2	1	1	3	4	3	7	10	18	32	53	51	35	54	21	13
6	6	8	3	13	11	5	1	2	0	2	3	1	0	3	0	0			

SPEAKER=5 PHONEME= 4 PITCH=220 TM= 63.90 ST= 74.10

2	1	0	2	0	0	0	0	0	0	2	0	0	1	3	2	7	21	63	22
10	4	1	2	0	0	0	2	2	3	7	8	14	28	99	37	17	7	3	2
3	1	2	4	2	0	0	1	0	2	2	0	0	1	0	0	0	1	0	0
2	1	0	1	0	0	1	2	2	2	5	6	8	22	17	19	12	19	10	9
4	2	3	2	3	1	2	1	0	1	2	1	0	1	0	0	0			

SPEAKER=5 PHONEME= 4 PITCH=246 TM= 35.00 ST= 82.40

1	4	1	1	0	2	0	0	2	2	2	3	1	0	1	1	0	4	11	14
32	99	67	26	7	7	0	2	5	1	2	5	1	0	4	6	14	39	48	14
5	5	4	2	3	1	0	5	1	1	1	2	2	1	3	1	0	0	1	0
1	3	2	1	3	1	1	3	4	1	8	9	17	17	56	96	57	91	59	98
23	25	22	13	34	28	19	7	3	0	2	5	7	3	4	1	2			

SPEAKER=5 PHONEME= 4 PITCH=261 TM= 46.60 ST= 87.80

3	1	0	3	0	0	0	1	0	1	2	1	0	2	1	0	0	3	4	7
15	37	91	49	12	8	3	4	2	2	3	4	2	4	4	5	13	24	98	42
14	7	6	6	3	1	1	3	0	2	2	0	1	3	1	0	1	0	0	0
2	0	0	3	1	0	3	4	2	4	9	8	15	31	41	99	54	44	35	37
18	12	14	9	19	16	9	2	0	1	2	0	1	2	2	0	4			

SPEAKER=5 PHONEME= 4 PITCH=294 TM= 45.10 ST= 77.50

2	1	3	2	0	0	0	0	1	2	2	0	3	1	0	0	0	0	1	2
2	4	10	17	25	99	30	12	10	5	2	3	3	1	1	3	2	1	5	10
27	80	30	10	4	4	1	0	0	3	5	2	2	1	1	2	1	3	0	1
0	1	2	2	1	3	4	3	3	5	8	10	24	54	26	53	43	21	55	28
24	15	16	7	19	18	19	5	3	3	0	1	3	1	1	3	2			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 5 PITCH=164 TM= 91.60 ST= 94.60

0	0	1	1	0	2	0	0	1	2	4	14	30	7	3	2	0	0	0	0
0	0	0	0	0	2	2	10	14	3	0	1	1	1	1	3	5	15	3	2
1	3	12	19	20	11	9	11	37	47	17	22	63	99	39	32	62	59	18	12
5	5	6	4	2	3	2	2	2	3	2	2	3	2	2	3	2	4	3	3
0	1	2	1	0	1	1	0	0	0	0	0	0	0	0	1	0			

SPEAKER=5 PHONEME= 5 PITCH=174 TM= 67.90 ST= 94.50

0	0	0	2	0	1	1	0	1	2	3	6	30	42	12	5	2	2	1	1
0	0	0	1	0	1	2	4	20	23	7	2	1	1	0	2	4	7	24	9
4	3	13	14	51	52	22	17	27	52	99	45	25	24	57	32	27	57	77	30
10	6	4	6	2	3	4	4	3	4	5	2	3	3	2	4	3	3	2	3
0	0	2	1	0	0	2	1	0	2	2	2	0	0	0	0	2			

SPEAKER=5 PHONEME= 5 PITCH=196 TM= 82.60 ST= 96.60

1	0	0	1	0	0	0	1	0	0	1	0	0	1	6	9	35	17	5	1
1	0	0	0	0	0	0	1	1	0	2	4	16	10	3	1	2	3	2	2
6	18	29	10	7	6	9	15	52	30	11	10	19	36	32	19	25	56	99	34
13	11	5	5	3	1	2	4	2	1	3	1	0	0	1	0	0	2	1	1
1	0	0	0	0	0	1	1	0	0	3	2	0	0	0	0	1			

SPEAKER=5 PHONEME= 5 PITCH=220 TM= 47.90 ST= 84.60

0	1	2	3	0	2	0	0	0	1	0	0	2	1	1	3	7	24	42	9
3	2	2	1	0	2	0	0	1	1	1	3	6	14	22	7	4	1	3	5
5	12	56	99	36	19	16	15	35	72	95	42	25	14	25	42	9	5	6	19
3	2	4	2	0	1	2	0	2	3	2	2	4	1	0	2	1	1	3	3
1	1	4	1	3	3	2	0	0	2	1	1	3	1	0	0	2			

SPEAKER=5 PHONEME= 5 PITCH=246 TM= 49.20 ST= 85.70

2	1	0	2	1	0	0	1	1	0	3	0	0	0	2	0	2	6	11	16
39	99	22	12	4	3	2	3	3	1	3	2	0	1	3	4	19	47	11	4
5	3	7	6	18	31	94	52	16	12	12	15	52	86	29	13	12	25	85	27
8	5	5	3	2	0	2	4	3	1	4	1	1	0	1	0	1	2	5	3
2	0	1	0	2	1	2	2	1	1	3	1	0	0	1	0	1			

SPEAKER=5 PHONEME= 5 PITCH=261 TM= 58.80 ST= 95.10

1	0	0	3	1	1	2	0	0	0	1	0	0	1	1	1	2	2	3	4
10	24	67	26	5	5	4	2	2	2	1	0	1	2	1	4	10	19	59	20
6	4	9	8	12	20	45	94	46	26	12	12	18	30	64	28	14	17	36	99
24	13	7	10	2	3	4	5	2	4	4	1	2	1	0	2	2	3	3	5
1	0	2	0	0	0	2	0	0	2	1	0	1	0	0	0	2			

SPEAKER=5 PHONEME= 5 PITCH=294 TM= 94.60 ST= 96.00

0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
0	1	2	3	5	31	9	2	2	1	0	1	1	1	1	2	2	2	3	7
15	50	21	9	7	8	9	9	18	30	99	67	25	14	12	9	7	16	6	4
1	1	4	1	1	1	2	1	2	1	1	0	1	0	0	1	0	1	2	1
0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 6 PITCH=164 TM= 44.10 ST= 78.30

2	1	2	2	0	1	1	0	3	4	9	22	70	16	8	3	2	0	1	2
1	1	3	1	0	4	7	20	53	17	8	6	7	6	8	13	36	90	48	24
15	17	38	43	99	40	21	16	36	82	30	20	37	73	42	18	10	10	3	4
3	2	4	2	0	0	2	1	2	3	4	2	5	1	0	1	0	0	1	3
1	1	5	0	0	0	1	0	0	2	1	1	3	0	0	1	1			

SPEAKER=5 PHONEME= 6 PITCH=174 TM= 45.30 ST= 80.60

1	2	0	3	2	2	0	1	0	0	1	2	7	29	17	5	0	2	0	0
0	1	0	1	1	2	0	5	9	25	17	7	1	3	6	8	13	24	74	99
25	14	17	15	41	87	35	17	9	11	40	27	11	15	64	61	14	10	8	2
1	2	0	1	2	1	2	5	2	0	2	1	0	1	2	2	1	3	2	1
0	0	0	0	1	1	1	3	0	0	1	1	0	0	2	0	1			

SPEAKER=5 PHONEME= 6 PITCH=196 TM= 91.00 ST= 95.60

1	0	0	2	0	0	0	0	0	0	1	0	2	3	10	16	41	16	6	2
1	0	0	1	0	1	2	1	2	3	7	12	32	11	5	5	7	6	10	17
32	99	56	23	16	14	20	33	52	22	9	7	20	38	12	7	6	15	4	3
1	0	0	1	0	1	4	3	2	4	2	2	1	1	0	1	2	2	0	2
0	0	1	0	0	0	1	0	0	2	0	0	0	0	0	0	1			

SPEAKER=5 PHONEME= 6 PITCH=220 TM= 69.50 ST= 95.60

0	2	0	0	0	1	0	0	1	1	1	2	0	0	2	4	6	22	58	16
6	4	1	0	0	1	0	1	2	2	2	5	7	15	51	17	7	5	6	6
8	15	51	87	55	26	14	12	18	32	66	29	16	14	44	99	28	14	8	14
3	3	1	1	0	1	0	2	3	2	2	3	1	0	0	2	0	3	2	1
0	1	1	0	1	1	1	0	1	1	0	1	0	0	0	1	1			

SPEAKER=5 PHONEME= 6 PITCH=246 TM= 53.20 ST= 82.90

0	1	0	2	1	1	0	2	1	0	1	0	0	0	2	0	1	5	8	11
32	67	14	7	2	2	2	3	3	1	4	3	2	3	8	13	47	99	26	13
8	5	10	8	21	33	97	47	14	11	9	9	34	55	15	5	2	3	3	2
2	1	0	2	0	0	1	2	0	1	3	2	2	2	0	0	0	1	0	2
1	1	2	2	1	0	2	0	0	1	1	0	0	2	1	0	2			

SPEAKER=5 PHONEME= 6 PITCH=261 TM= 67.80 ST= 95.60

0	2	0	1	0	1	0	0	1	1	0	1	0	0	1	1	0	2	4	4
7	20	58	23	6	6	2	2	4	2	2	3	2	1	4	8	15	28	99	37
12	9	13	8	18	23	44	94	71	34	14	11	10	11	39	15	5	3	4	7
1	0	0	0	0	1	2	2	3	1	1	0	1	0	0	1	1	1	2	1
0	0	1	0	0	1	1	0	1	1	0	0	0	0	0	1	1			

SPEAKER=5 PHONEME= 6 PITCH=294 TM= 90.10 ST= 96.60

1	0	0	2	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
1	2	2	7	8	46	22	9	5	3	3	3	1	2	2	2	5	6	9	17
34	99	51	21	14	9	7	7	9	17	58	56	18	11	10	9	16	46	20	9
4	3	2	4	1	0	2	4	0	1	4	2	1	2	0	0	0	1	3	2
1	0	1	1	0	0	1	0	0	0	1	1	0	1	0	0	0			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 7 PITCH=164 TM= 54.00 ST= 92.50

2	2	0	4	0	0	0	0	0	0	2	4	26	8	2	0	1	0	0	0
0	0	0	1	0	1	2	4	16	8	2	0	1	2	2	3	6	11	15	6
5	4	21	27	91	45	24	21	44	99	56	29	21	31	32	15	12	30	8	6
3	1	0	2	0	1	2	3	2	3	2	0	2	1	0	1	3	1	1	3
0	0	3	0	0	0	2	0	1	2	1	0	2	0	0	0	1			

SPEAKER=5 PHONEME= 7 PITCH=174 TM= 56.70 ST= 82.50

1	1	0	2	2	1	0	1	0	0	1	3	14	38	13	4	1	3	0	0
0	0	0	0	1	0	1	4	12	25	13	4	0	1	2	0	3	7	29	19
8	5	14	14	47	86	36	18	18	32	99	51	21	17	46	32	10	12	25	8
4	3	1	4	1	0	2	2	0	1	3	0	2	2	1	1	2	2	0	1
0	0	2	1	2	0	3	0	1	2	1	0	0	1	0	1	2			

SPEAKER=5 PHONEME= 7 PITCH=196 TM= 84.80 ST= 92.40

0	0	0	0	0	1	0	0	1	0	0	0	0	0	4	7	31	17	6	1
0	1	0	0	0	1	0	1	2	1	3	7	24	12	3	2	1	2	3	4
9	26	42	15	12	12	17	27	99	55	18	15	17	29	28	14	11	25	45	13
4	6	3	2	3	1	1	2	1	0	1	1	1	0	1	1	0	1	1	0
0	0	1	1	4	1	0	1	0	0	0	0	0	0	1	0	1			

SPEAKER=5 PHONEME= 7 PITCH=220 TM= 49.30 ST= 84.70

0	1	0	2	1	2	0	2	1	0	1	0	0	0	2	0	2	13	20	2
1	0	0	0	1	0	0	2	1	0	2	2	1	7	17	3	3	4	3	3
7	10	47	87	34	16	16	18	35	69	99	43	20	14	37	65	14	8	4	4
4	2	0	1	0	0	1	3	2	2	4	2	1	1	0	0	0	4	2	2
2	1	2	2	2	0	1	1	0	1	2	2	1	4	5	1	1			

SPEAKER=5 PHONEME= 7 PITCH=246 TM= 45.90 ST= 69.30

2	1	3	2	0	0	0	0	1	2	1	0	3	1	0	1	0	0	0	2
5	28	14	4	0	3	0	0	0	1	0	1	2	1	2	6	10	32	17	6
1	3	10	7	18	31	81	96	27	18	12	17	50	99	58	29	15	21	70	58
9	7	4	2	1	2	1	3	4	3	2	3	2	0	0	2	0	2	7	6
0	3	2	0	2	3	0	0	1	0	1	3	1	0	2	0	0			

SPEAKER=5 PHONEME= 7 PITCH=261 TM= 74.00 ST= 96.00

0	1	2	2	0	1	0	0	0	1	0	0	1	0	0	1	0	0	2	2
5	16	46	17	4	4	1	1	1	1	0	2	2	2	2	5	7	13	52	22
6	6	13	10	18	27	48	99	74	36	16	14	14	16	44	21	8	8	15	47
11	7	4	4	2	3	2	4	2	1	1	1	2	1	1	2	1	1	2	3
0	1	3	1	8	6	2	0	1	1	0	0	0	0	0	1	1			

SPEAKER=5 PHONEME= 7 PITCH=294 TM= 85.30 ST= 95.20

0	0	0	2	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	2	4	19	5	2	1	1	0	0	0	0	0	1	2	3	3	8
21	36	14	6	5	6	9	10	19	37	99	44	18	10	7	5	4	5	1	3
1	0	0	0	0	0	1	1	1	2	1	0	1	0	0	0	0	0	1	1
0	0	2	1	2	0	1	0	0	1	1	0	1	0	0	0	0			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 8 PITCH=164 TM= 58.00 ST= 82.90

1	0	0	2	0	1	1	1	3	5	10	25	52	13	6	4	2	2	2	3
2	1	2	5	2	15	29	74	99	32	18	10	7	4	4	8	25	46	11	5
3	2	5	5	2	1	1	0	3	4	2	0	4	3	2	3	16	7	2	6
1	1	2	1	0	1	1	0	1	2	1	1	3	1	2	5	1	0	1	2
0	0	2	1	0	2	1	0	0	1	0	1	1	1	0	2	1			

SPEAKER=5 PHONEME= 8 PITCH=174 TM= 78.20 ST= 96.10

0	0	0	1	1	1	1	2	3	2	6	10	33	62	33	13	4	5	4	2
3	2	1	2	2	6	10	19	50	99	51	21	10	7	6	6	12	21	72	48
15	8	8	5	10	16	6	4	4	9	34	16	5	5	15	8	1	1	0	1
1	1	0	2	0	0	0	1	0	1	2	1	1	2	0	0	0	0	0	1
0	0	1	1	0	0	1	0	0	0	1	0	0	1	1	0	2			

SPEAKER=5 PHONEME= 8 PITCH=196 TM= 84.90 ST= 95.60

1	0	1	1	0	0	0	0	0	1	1	1	4	4	13	20	68	38	15	6
2	2	2	2	0	3	2	3	8	10	19	29	99	63	22	14	9	4	3	2
0	3	5	2	1	2	1	0	3	2	0	0	1	1	1	1	1	3	3	0
0	0	0	0	0	1	1	1	2	1	0	0	1	0	0	1	0	0	2	1
0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0			

SPEAKER=5 PHONEME= 8 PITCH=220 TM= 94.90 ST= 96.60

1	0	0	1	0	0	0	0	0	0	1	0	1	2	4	4	10	28	46	12
6	3	1	3	1	2	2	4	6	8	13	16	31	59	99	36	21	11	6	5
3	2	6	8	3	1	1	0	0	2	2	0	1	2	9	16	3	1	0	1
0	0	0	1	0	0	2	1	0	0	1	0	1	0	0	0	1	0	0	1
0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1			

SPEAKER=5 PHONEME= 8 PITCH=246 TM= 61.70 ST= 95.50

0	0	0	1	1	1	0	2	1	0	1	0	0	0	3	1	2	7	11	13
28	87	50	20	7	6	4	4	6	3	6	5	4	5	11	15	41	99	67	29
14	8	9	5	5	3	11	11	2	3	3	1	2	9	3	0	0	3	17	9
2	1	0	2	0	0	1	2	0	1	3	1	1	2	0	0	0	1	0	1
1	1	1	2	2	3	1	0	0	1	1	1	0	2	0	0	1			

SPEAKER=5 PHONEME= 8 PITCH=261 TM= 57.00 ST= 80.40

1	0	1	3	1	1	1	0	0	0	1	0	0	1	2	1	3	3	5	7
15	34	93	46	13	10	7	4	4	2	2	1	2	3	4	7	16	25	99	46
13	6	7	5	3	3	6	12	10	6	1	0	2	4	17	6	3	3	5	25
4	1	0	1	0	0	2	2	1	3	2	0	1	0	0	1	1	1	1	3
0	0	1	0	0	0	2	0	0	3	1	0	1	0	0	0	1			

SPEAKER=5 PHONEME= 8 PITCH=294 TM= 54.30 ST= 67.10

1	0	0	3	0	1	2	0	0	1	1	0	0	1	0	0	2	1	1	2
2	2	5	14	19	99	40	17	11	6	3	2	0	1	0	2	2	2	2	5
11	41	20	7	1	2	2	2	1	5	9	8	3	3	6	11	26	83	32	16
4	2	1	1	0	0	2	2	2	3	2	0	2	0	0	1	1	1	10	6
1	0	2	0	0	2	11	1	1	3	1	0	1	0	0	0	1			

TABLE XII (Continued)

SPEAKER=5 PHONEME= 9 PITCH=164 TM= 60.40 ST= 96.20

0	0	0	1	0	2	1	1	6	8	16	34	99	29	17	7	3	1	2	2
1	1	2	1	0	6	8	21	61	20	7	6	5	4	7	13	30	77	34	15
6	6	18	21	39	15	7	5	15	28	7	7	15	29	21	20	39	88	27	25
27	14	10	10	3	2	0	3	3	1	3	3	2	1	2	1	0	0	1	0
0	1	1	0	2	0	0	1	0	0	0	1	0	0	1	0	1			

SPEAKER=5 PHONEME= 9 PITCH=174 TM= 83.50 ST= 97.40

0	1	0	1	1	3	1	1	6	6	9	17	43	99	62	27	9	9	7	4
2	3	1	1	0	4	4	9	28	61	27	13	7	5	8	10	17	27	89	89
26	16	16	9	16	28	10	4	3	3	6	4	4	3	13	9	7	18	54	18
16	29	9	6	4	2	2	3	4	1	4	3	3	1	3	2	0	2	1	0
0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0			

SPEAKER=5 PHONEME= 9 PITCH=196 TM= 71.70 ST= 94.20

0	0	0	1	0	1	1	1	1	2	2	2	5	7	20	30	99	55	22	9
6	3	2	2	0	2	2	1	5	7	13	20	73	41	13	8	6	3	4	6
11	33	46	16	8	6	4	4	19	10	1	1	3	5	4	3	2	7	11	6
13	32	14	6	9	5	3	4	3	1	1	0	1	0	0	1	0	0	2	1
0	1	1	0	0	0	0	0	1	1	0	2	0	0	0	0	0			

SPEAKER=5 PHONEME= 9 PITCH=220 TM= 96.30 ST= 97.20

0	1	0	1	1	1	0	0	2	0	2	3	5	5	13	13	19	42	99	45
22	12	5	5	3	4	2	3	5	4	8	10	16	28	93	47	20	10	7	5
5	7	24	39	34	14	7	4	3	4	11	5	2	2	4	6	3	5	7	22
13	8	9	21	6	5	5	4	2	1	1	1	0	0	1	1	0	1	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0			

SPEAKER=5 PHONEME= 9 PITCH=246 TM= 53.30 ST= 96.00

0	2	0	0	0	1	0	0	2	1	1	2	1	0	1	2	0	5	10	12
31	99	47	18	6	7	2	4	6	4	4	9	5	5	10	16	37	87	79	33
13	11	13	7	9	12	31	48	12	8	4	5	4	6	5	2	0	2	6	5
3	6	13	21	4	3	5	4	2	1	2	2	2	0	1	0	0	0	2	0
0	2	2	0	1	0	0	0	1	0	1	2	1	0	2	0	0			

SPEAKER=5 PHONEME= 9 PITCH=261 TM= 57.80 ST= 92.00

0	1	1	2	0	2	0	0	0	1	0	0	1	1	1	2	1	1	4	4
12	29	82	41	11	10	5	3	3	2	1	3	4	3	4	8	15	26	99	53
15	9	11	5	4	5	7	15	16	8	1	1	2	2	3	2	0	0	1	10
2	2	3	6	3	2	1	1	1	2	1	1	3	1	0	2	0	0	1	2
0	0	2	1	0	2	1	0	0	1	0	1	1	1	0	2	1			

SPEAKER=5 PHONEME= 9 PITCH=294 TM= 95.00 ST= 96.40

0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	1	2	2	3	3
5	6	9	23	38	73	27	15	11	6	6	4	4	4	5	6	11	11	18	34
79	99	48	22	16	11	9	6	7	13	32	12	5	4	4	3	7	10	2	3
5	13	19	8	3	5	10	5	1	2	2	0	2	1	0	1	1	0	0	0
0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1			

TABLE XII (Continued)

SPEAKER=5 PHONEME=10 PITCH=164 TM= 64.10 ST= 81.30

2	1	0	1	1	0	1	2	6	6	15	31	99	30	17	6	2	4	2	1
2	1	0	2	1	4	9	26	52	16	9	6	4	6	7	12	33	80	28	14
9	9	26	37	57	22	12	10	26	43	12	7	14	27	11	3	3	2	0	0
1	0	0	2	1	0	3	2	0	0	2	0	1	1	0	1	2	1	0	2
0	0	1	0	0	1	2	1	0	2	0	0	0	0	0	0	2			

SPEAKER=5 PHONEME=10 PITCH=174 TM= 53.00 ST= 94.60

1	1	0	1	1	1	0	1	4	3	8	12	37	89	52	19	6	7	4	2
3	2	0	2	0	2	4	10	28	58	29	12	6	7	6	8	16	27	89	81
24	15	24	22	59	99	60	24	12	14	32	21	8	6	14	10	4	2	0	0
1	0	0	2	0	1	3	2	1	2	2	0	1	1	0	1	1	2	1	3
0	0	1	0	0	0	2	0	0	2	1	0	1	0	0	0	1			

SPEAKER=5 PHONEME=10 PITCH=196 TM= 47.20 ST= 89.80

1	1	2	1	0	1	0	0	1	2	2	2	6	6	19	31	99	50	19	7
3	2	2	2	0	4	1	1	5	6	12	26	87	37	13	10	7	6	10	14
24	77	83	27	17	13	13	18	58	25	6	6	12	24	12	5	0	3	3	2
1	3	0	0	0	0	0	2	3	1	2	3	2	0	2	1	0	0	2	1
0	1	2	0	3	2	1	1	0	0	0	2	0	1	2	1	2			

SPEAKER=5 PHONEME=10 PITCH=220 TM= 41.60 ST= 90.30

2	1	3	2	0	1	0	0	1	2	2	1	4	2	4	6	12	31	99	29
12	5	4	2	0	4	1	2	4	3	3	7	14	30	75	29	12	6	6	12
5	12	50	84	61	28	18	16	9	19	43	18	10	7	10	25	4	1	0	1
0	1	1	1	0	3	2	1	3	2	0	1	3	0	1	3	1	0	3	3
0	0	3	0	0	2	2	0	3	1	0	0	0	0	1	2	3			

SPEAKER=5 PHONEME=10 PITCH=246 TM= 89.00 ST= 96.60

0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	1	2	5	6
13	43	31	11	3	4	1	2	3	2	2	4	3	3	6	10	22	53	58	23
10	9	12	8	16	24	55	99	28	17	9	9	13	26	20	8	2	2	2	1
1	1	0	1	1	0	1	1	1	1	2	2	2	0	1	1	0	0	1	0
0	1	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0			

SPEAKER=5 PHONEME=10 PITCH=261 TM= 56.60 ST= 94.60

0	1	0	0	0	2	0	0	2	1	0	1	0	0	0	1	0	1	3	3
6	20	56	24	5	5	0	2	3	2	1	4	2	1	3	7	12	26	99	46
13	10	11	6	9	13	25	57	48	23	9	9	7	6	21	9	1	1	1	0
1	2	0	0	1	0	0	1	2	0	2	2	2	0	2	0	0	0	1	0
0	1	2	0	2	1	0	0	0	0	0	1	0	0	2	0	0			

SPEAKER=5 PHONEME=10 PITCH=294 TM= 47.60 ST= 75.30

3	1	0	2	0	0	0	1	0	1	3	1	0	1	1	0	0	1	0	0
2	2	2	10	16	58	16	8	3	2	3	3	1	3	3	2	4	7	8	21
54	99	40	18	11	7	7	7	11	29	83	38	13	9	6	3	6	13	0	1
2	0	0	2	1	0	3	2	1	1	3	0	2	2	1	0	2	1	0	1
1	0	2	1	1	0	3	0	0	0	1	0	0	1	1	0	3			

TABLE XII (Continued)

SPEAKER=5 PHONEME=11 PITCH=164 TM= 75.90 ST= 91.40

0	1	1	1	1	3	1	1	8	9	18	42	99	30	19	11	3	3	3	1
0	1	2	2	1	8	11	28	62	21	8	5	2	1	2	4	6	20	7	2
0	1	1	0	2	1	0	0	2	2	0	0	0	0	0	1	0	0	1	1
0	1	2	0	2	3	3	7	4	3	1	2	1	0	0	1	0	1	2	1
0	1	1	0	0	0	0	0	1	1	0	2	0	0	0	0	0			

SPEAKER=5 PHONEME=11 PITCH=174 TM= 44.20 ST= 89.70

1	0	0	2	0	2	1	2	5	6	8	14	41	99	97	38	12	10	6	4
2	0	1	1	0	2	2	4	15	42	30	9	3	1	0	2	2	1	12	23
4	1	4	2	1	4	2	0	1	3	4	2	3	1	1	1	0	0	6	2
2	7	3	2	0	1	2	1	2	2	2	2	4	1	0	2	0	0	1	2
0	1	3	1	0	2	1	0	0	1	0	1	1	1	1	3	2			

SPEAKER=5 PHONEME=11 PITCH=196 TM= 95.20 ST= 95.50

0	0	0	1	0	1	0	0	1	2	2	3	6	8	21	31	99	51	21	10
6	3	1	2	0	1	1	1	1	2	3	5	22	8	1	1	1	1	1	2
3	14	14	5	1	1	1	1	7	3	0	0	0	2	0	0	1	1	1	0
0	1	0	1	0	0	2	1	0	1	1	0	1	0	0	0	1	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1			

SPEAKER=5 PHONEME=11 PITCH=220 TM= 75.20 ST= 89.70

0	0	0	1	0	1	1	0	0	1	1	0	3	4	9	10	17	34	99	55
25	10	5	5	2	3	3	2	2	1	2	1	4	10	39	23	9	3	0	0
1	0	2	5	7	1	1	0	0	1	3	0	0	1	1	0	1	0	0	0
1	0	0	3	0	0	2	1	0	1	2	0	1	1	0	0	1	0	0	1
0	0	1	1	0	0	1	0	0	0	1	0	0	1	0	0	1			

SPEAKER=5 PHONEME=11 PITCH=246 TM= 58.90 ST= 85.60

1	1	0	1	1	0	0	1	1	0	2	0	0	0	2	1	2	7	11	15
38	99	26	13	5	4	2	3	3	1	2	1	0	0	2	3	13	38	8	2
2	1	0	1	1	1	7	3	1	1	1	0	0	1	1	0	0	1	1	0
2	0	0	0	0	0	1	3	2	1	3	1	0	0	1	0	0	2	1	1
1	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0	1			

SPEAKER=5 PHONEME=11 PITCH=261 TM= 84.40 ST= 96.00

0	0	0	1	0	0	1	0	0	0	1	0	1	1	3	2	3	5	8	10
18	38	99	49	15	11	8	5	5	3	3	1	1	2	2	2	5	10	38	15
4	1	2	2	1	0	1	1	0	1	0	0	0	1	0	0	1	0	0	0
1	0	0	1	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0
0	0	1	1	0	0	1	0	0	0	0	0	0	1	0	0	1			

SPEAKER=5 PHONEME=11 PITCH=294 TM= 63.40 ST= 96.00

2	1	0	1	0	0	0	0	0	0	2	0	0	1	1	0	0	2	2	3
6	8	9	27	41	99	34	19	13	7	6	4	1	2	1	0	0	1	0	1
6	10	2	3	0	0	0	0	0	1	1	0	0	2	0	0	0	0	0	0
1	0	1	2	0	0	3	2	0	1	2	0	2	1	0	0	1	0	0	1
0	0	1	1	1	0	3	0	0	0	0	0	0	1	1	0	2			

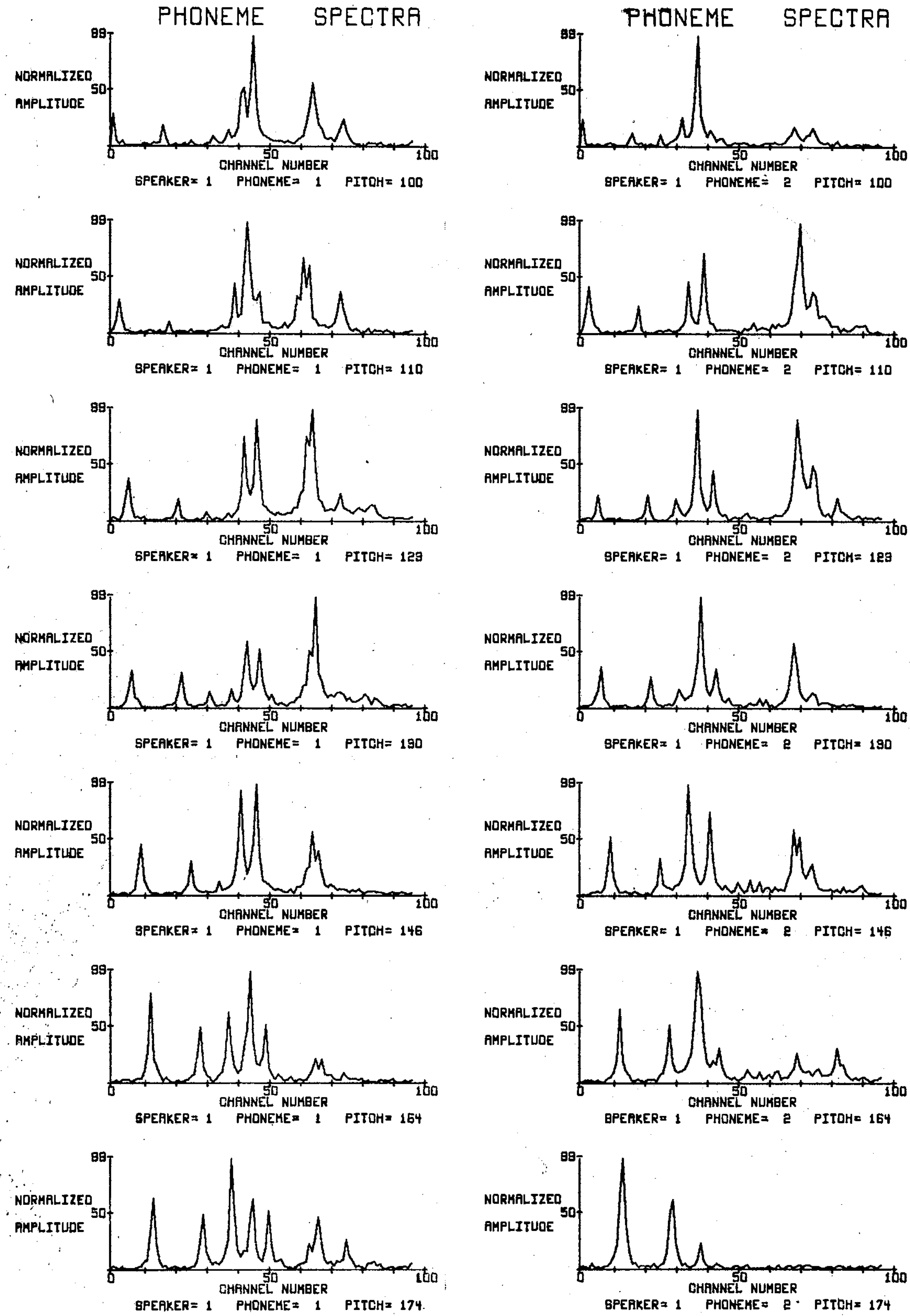


Figure D.1. Spectral Plots of the Phonemes /ae/ and /e/ Voiced by Speaker 1 at Several Pitches

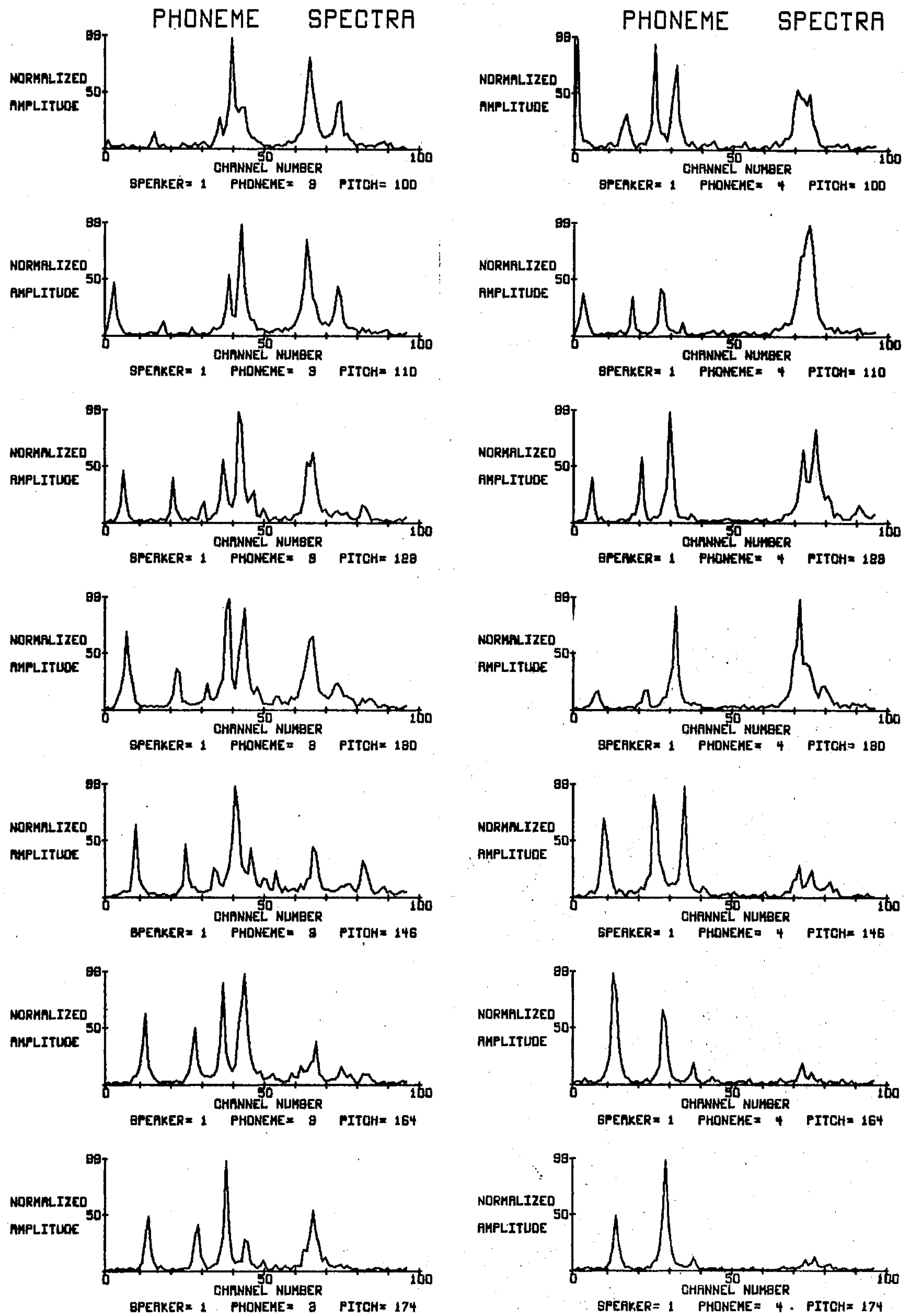


Figure D.2. Spectral Plots of the Phonemes /ε/ and /i/ Voiced by Speaker 1 at Several Pitches

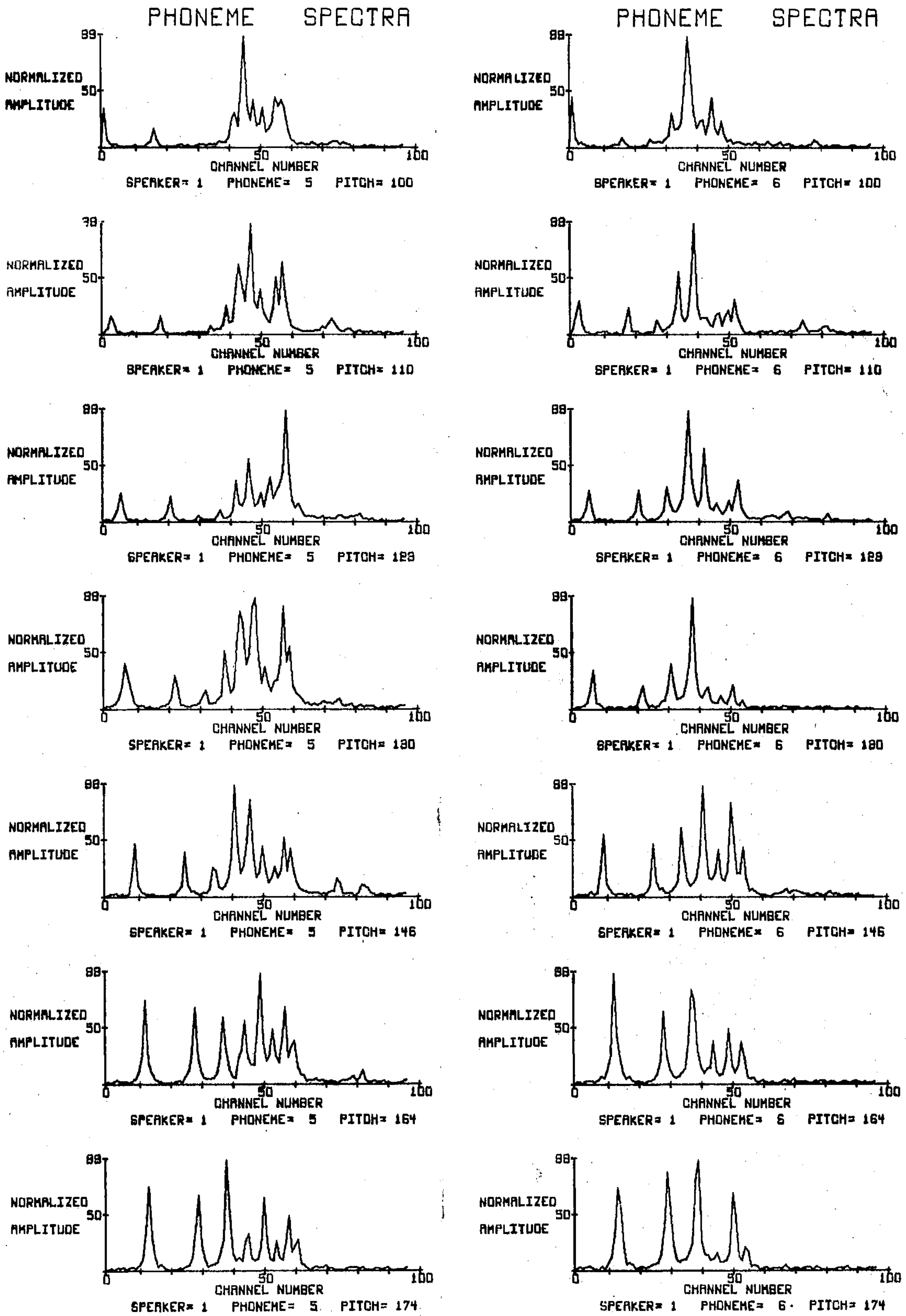


Figure D.3. Spectral Plots of the Phonemes /α/ and /o/ Voiced by Speaker 1 at Several Pitches

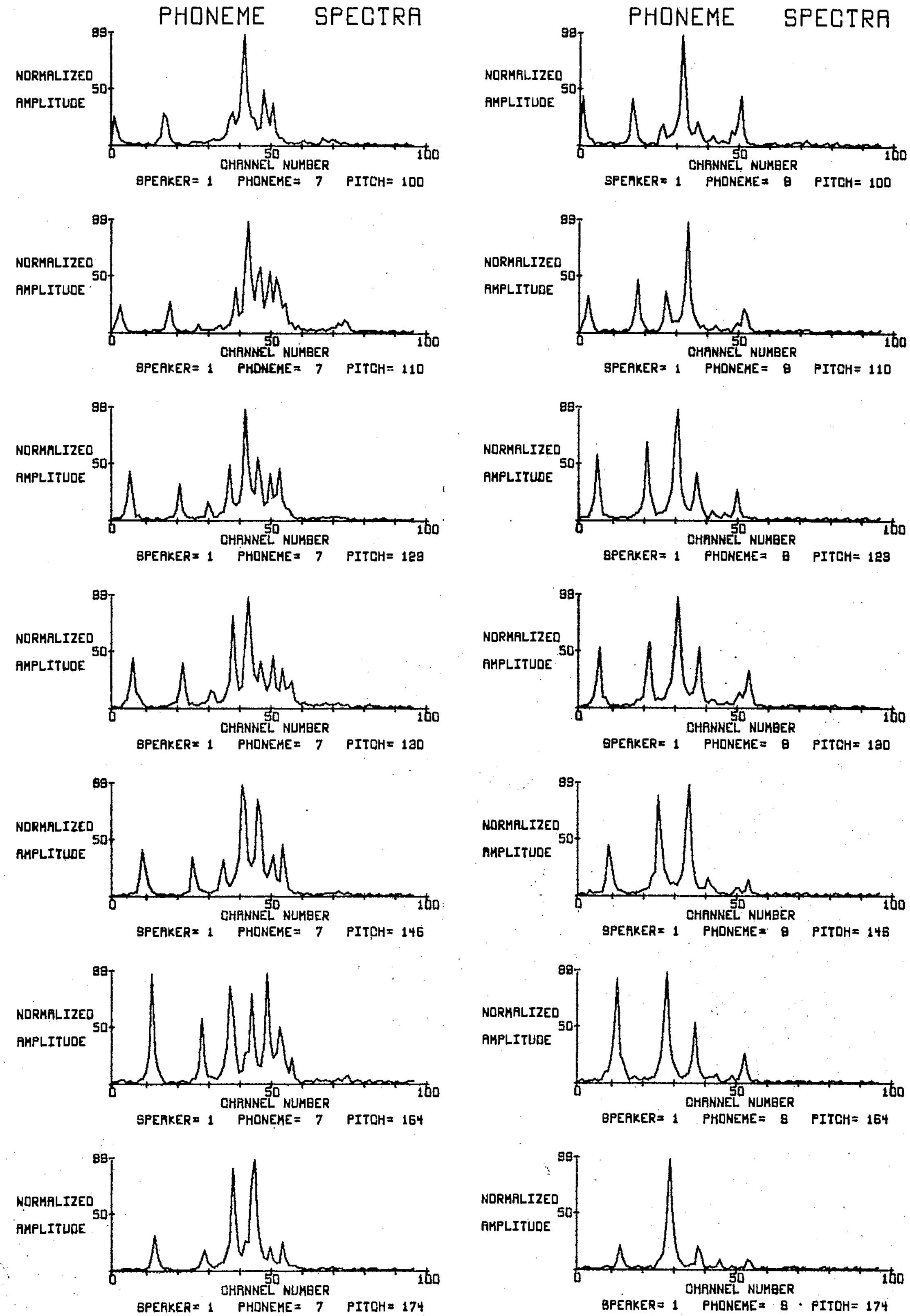


Figure D.4. Spectral Plots of the Phonemes /ɔ/ and /u/ Voiced by Speaker 1 at Several Pitches

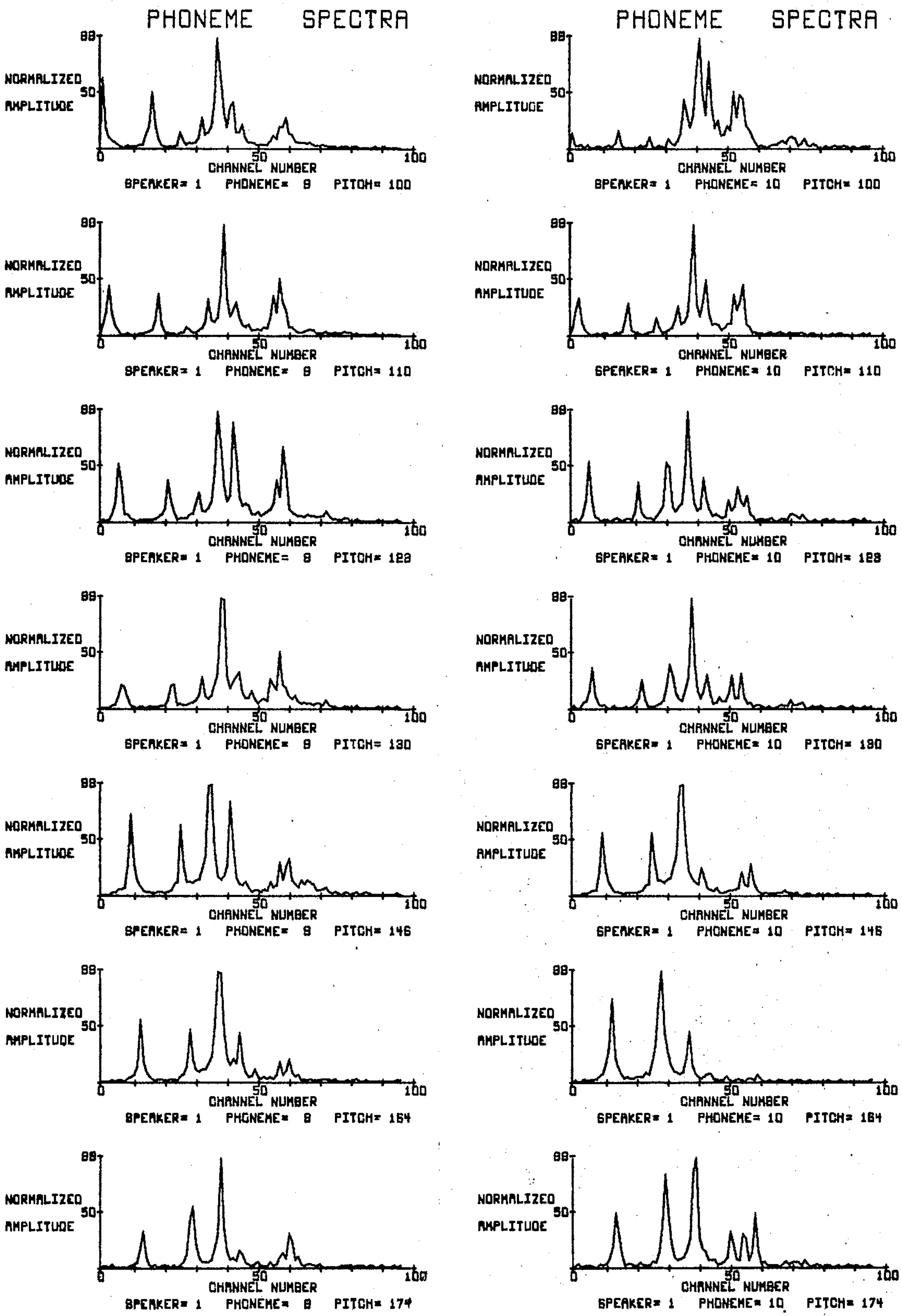


Figure D.5. Spectral Plots of the Phonemes /ər/ and /l/ Voiced by Speaker 1 at Several Pitches

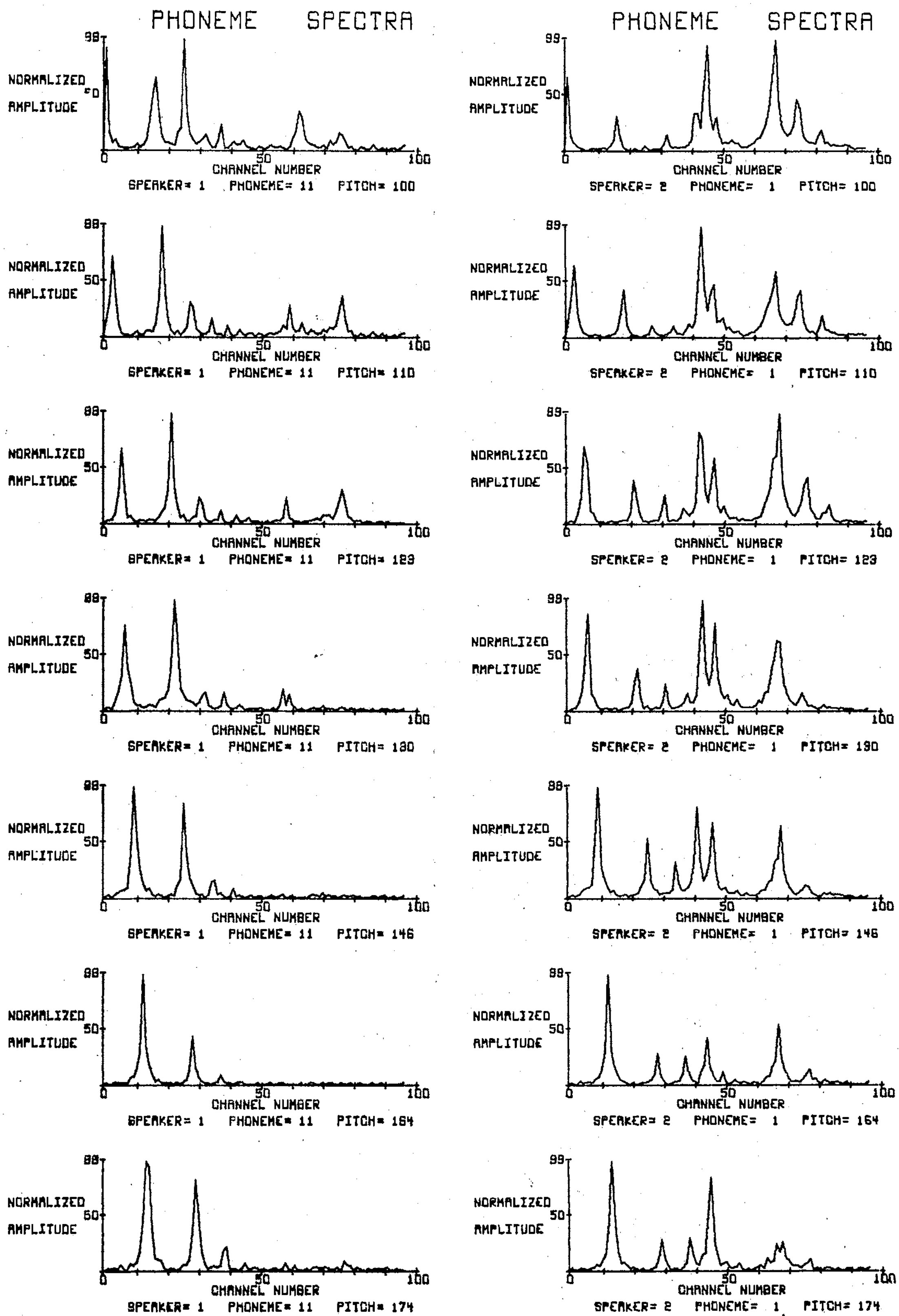


Figure D.6. Spectral Plots of the Phonemes /n/ Voiced by Speaker 1 and /ae/ Voiced by Speaker 2 at Several Pitches

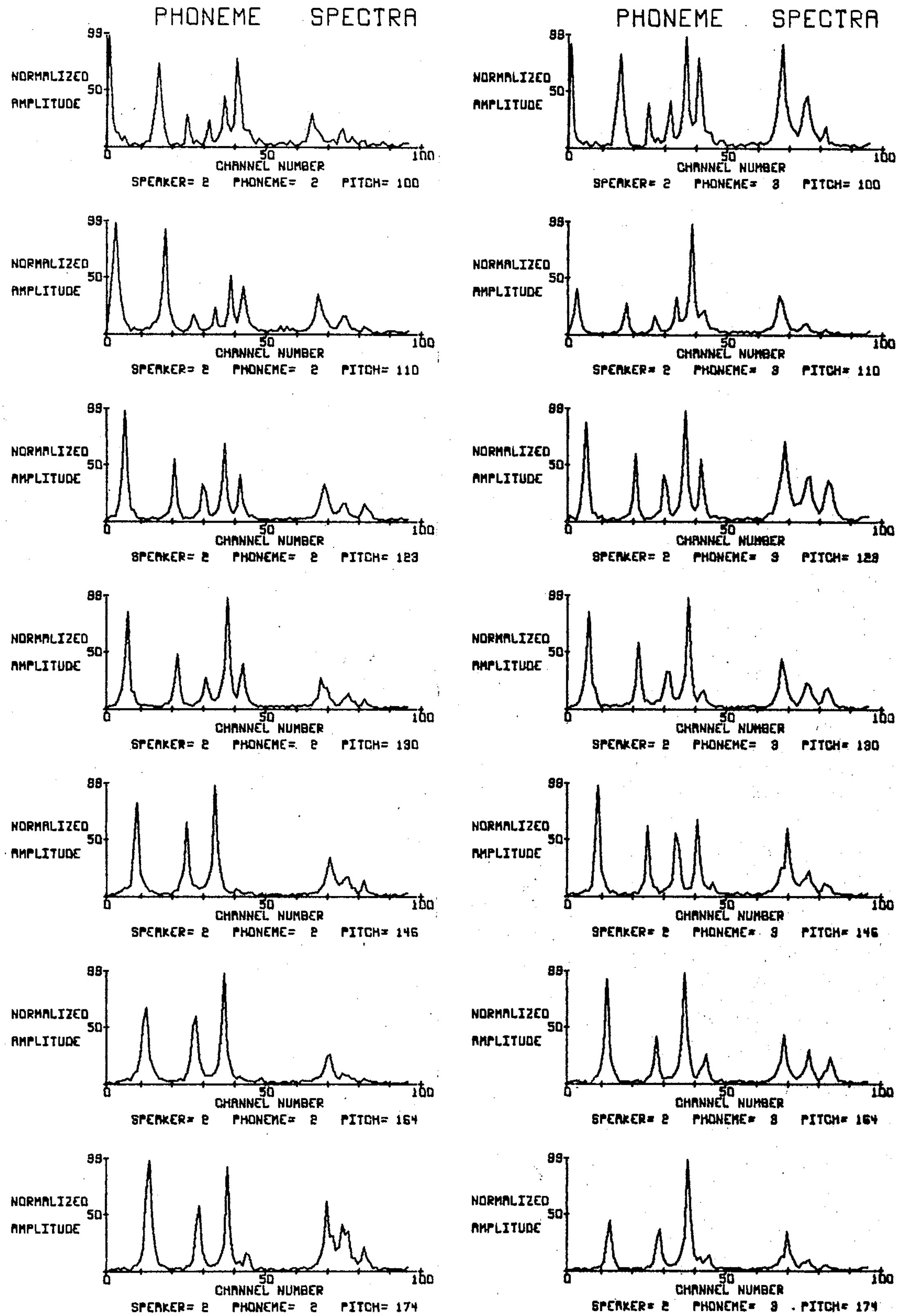


Figure D.7. Spectral Plots of the Phonemes /e/ and /ɛ/ Voiced by Speaker 2 at Several Pitches

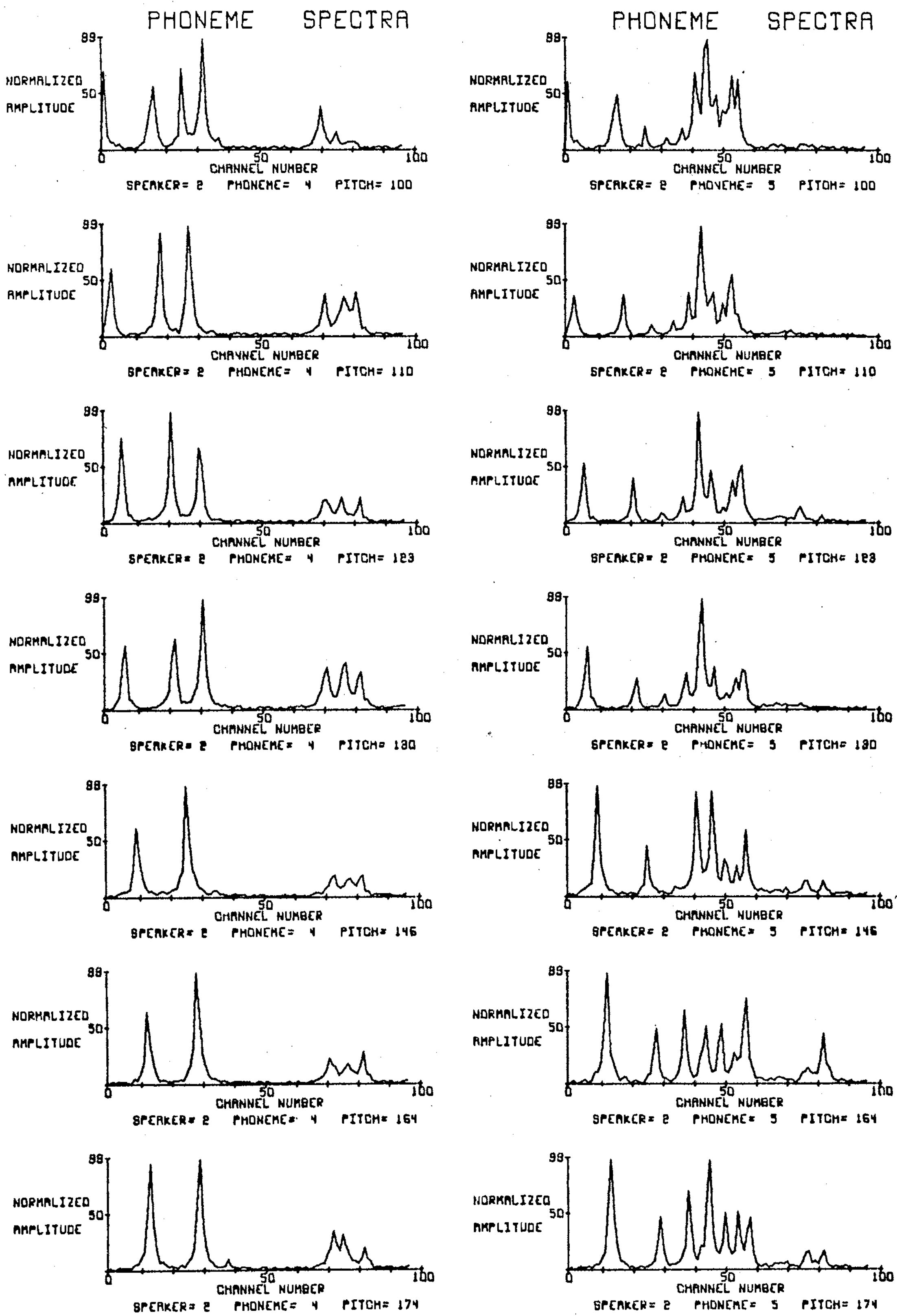


Figure D.8. Spectral Plots of the Phonemes /i/ and /α/ Voiced by Speaker 2 at Several Pitches

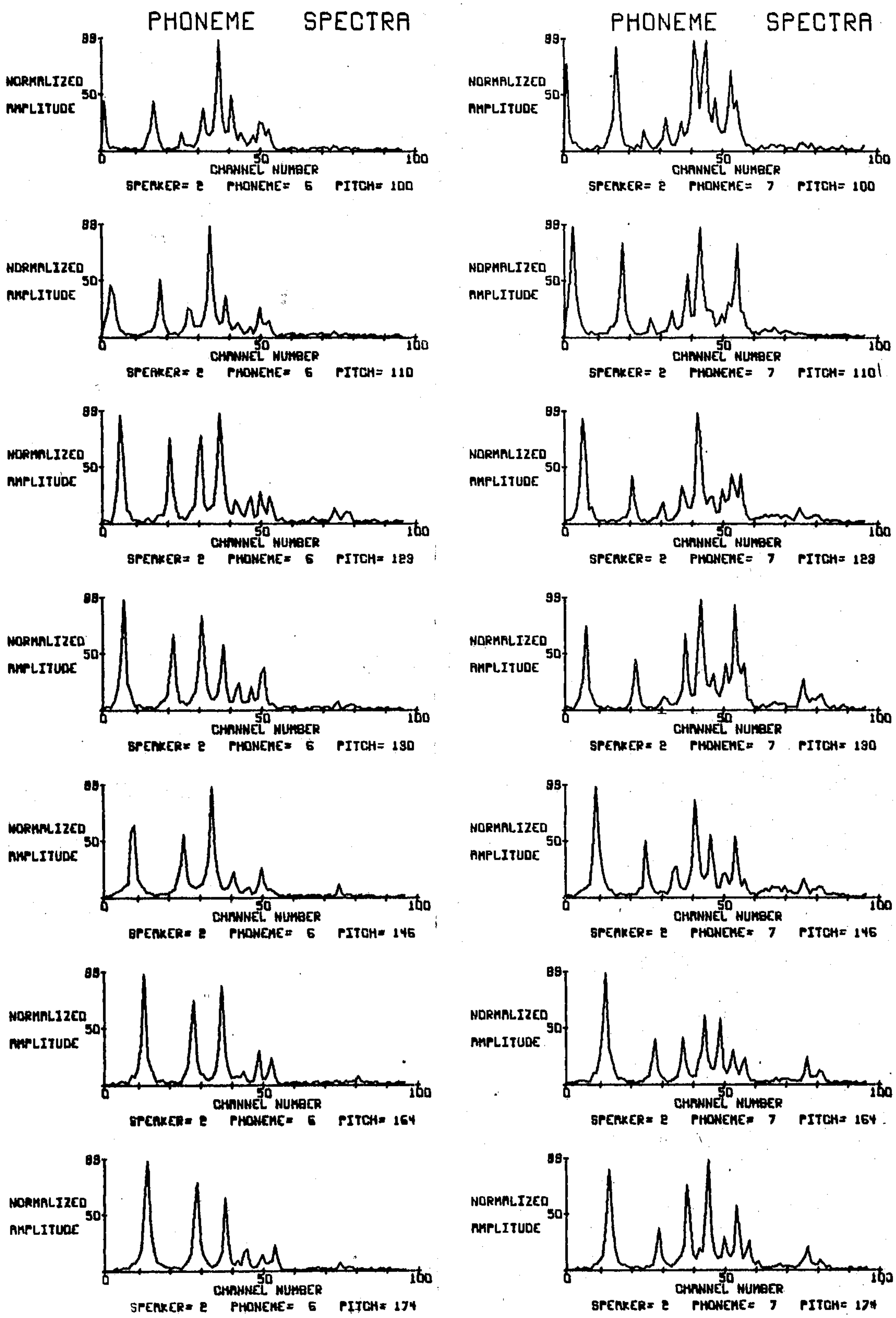


Figure D.9. Spectral Plots of the Phonemes /o/ and /ə/ Voiced by Speaker 2 at Several Pitches

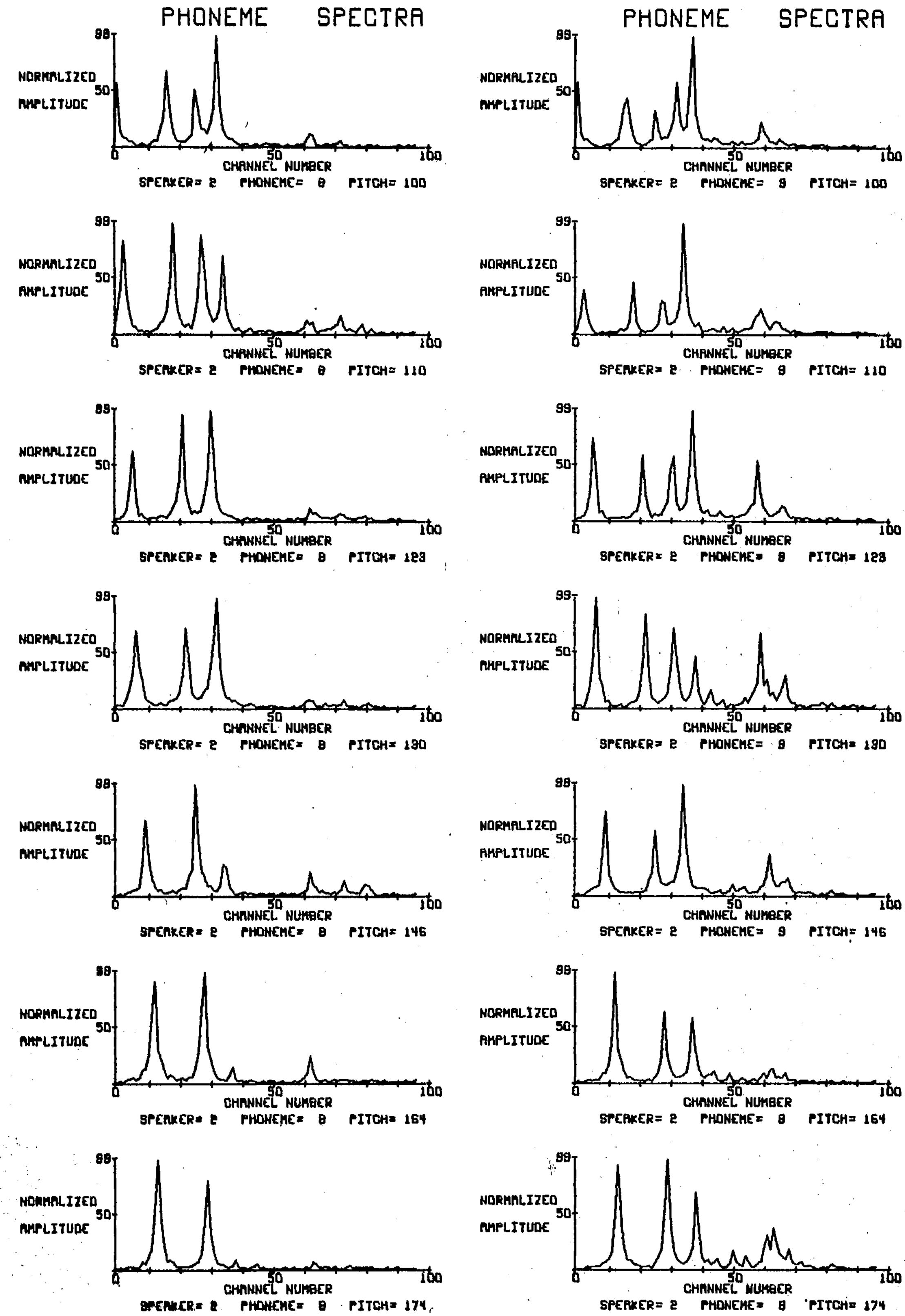


Figure D.10. Spectral Plots of the Phonemes /u/ and /ər/ Voiced by Speaker 2 at Several Pitches

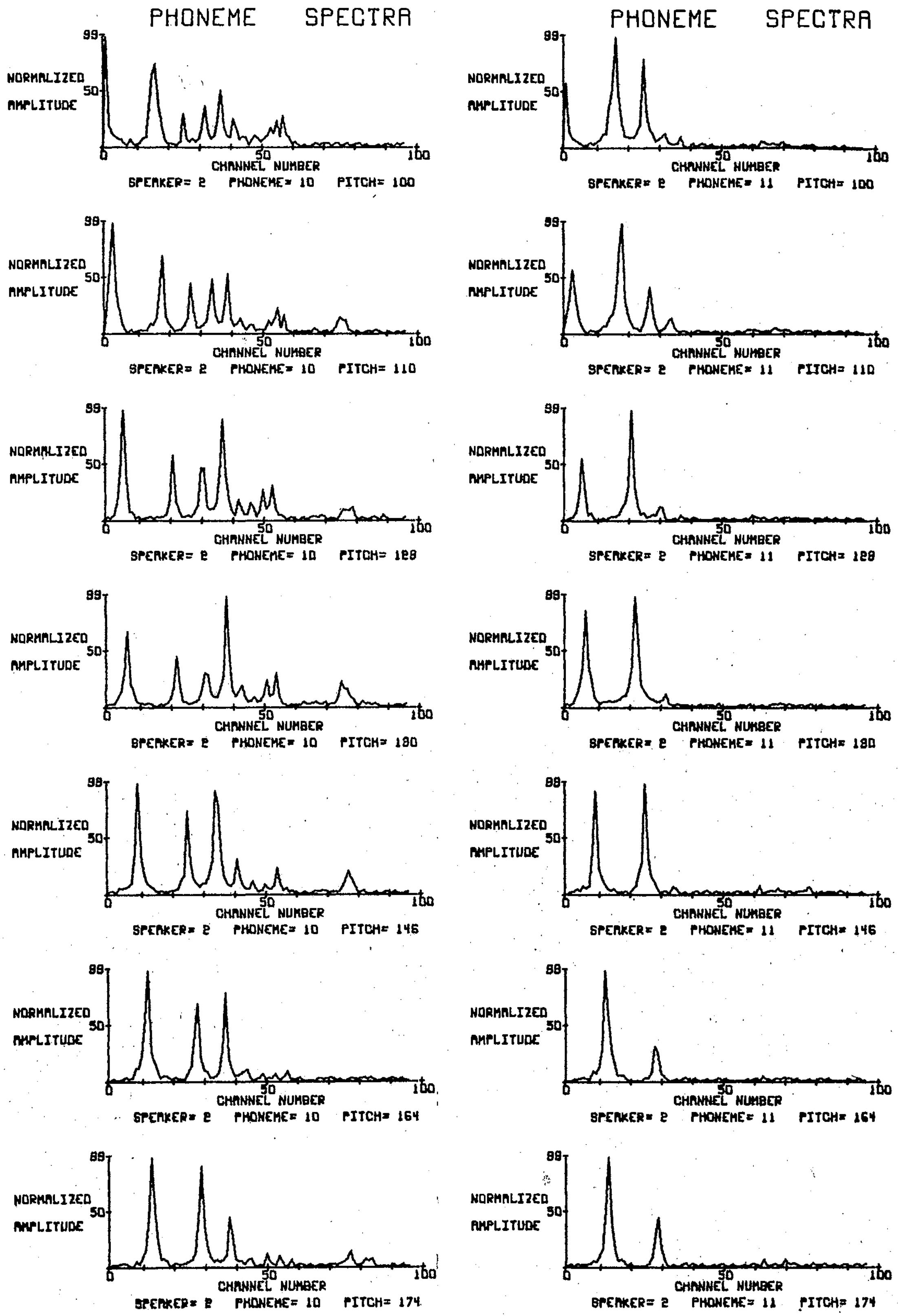


Figure D.11. Spectral Plots of the Phonemes /l/ and /n/ Voiced by Speaker 2 at Several Pitches

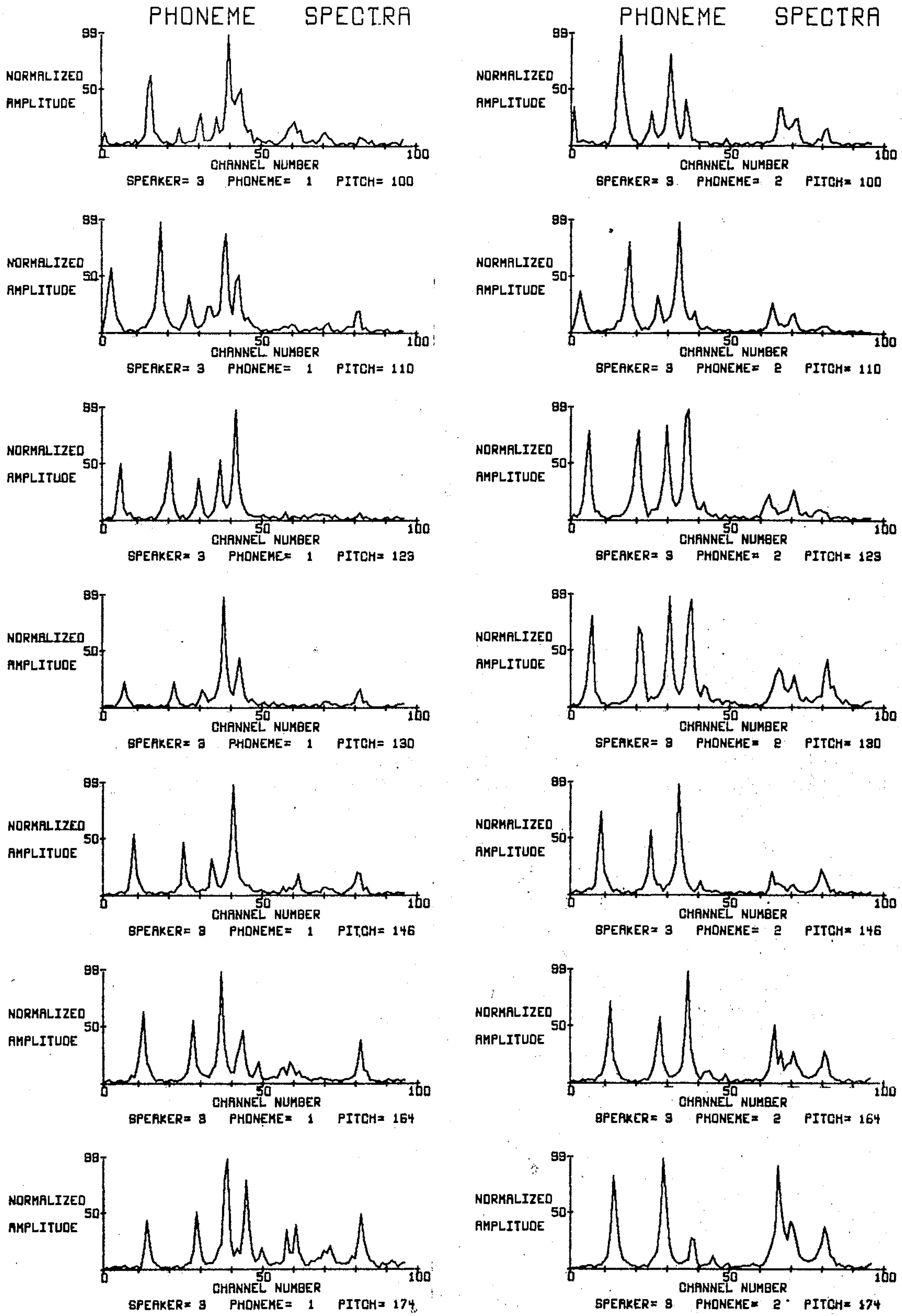


Figure D.12. Spectral Plots of the Phonemes /ae/ and /e/ Voiced by Speaker 3 at Several Pitches

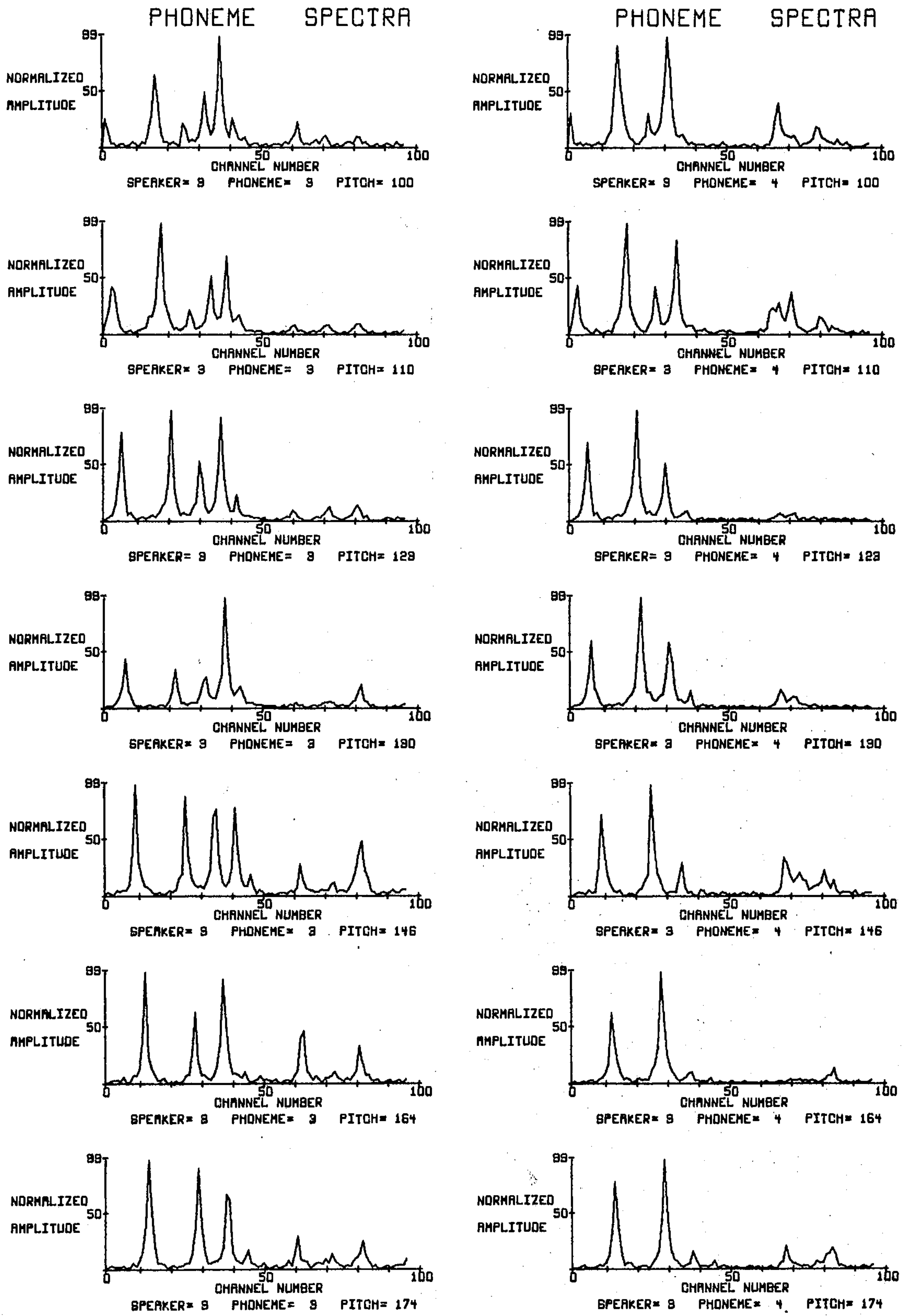


Figure D.13. Spectral Plots of the Phonemes /ɛ/ and /i/ Voiced by Speaker 3 at Several Pitches

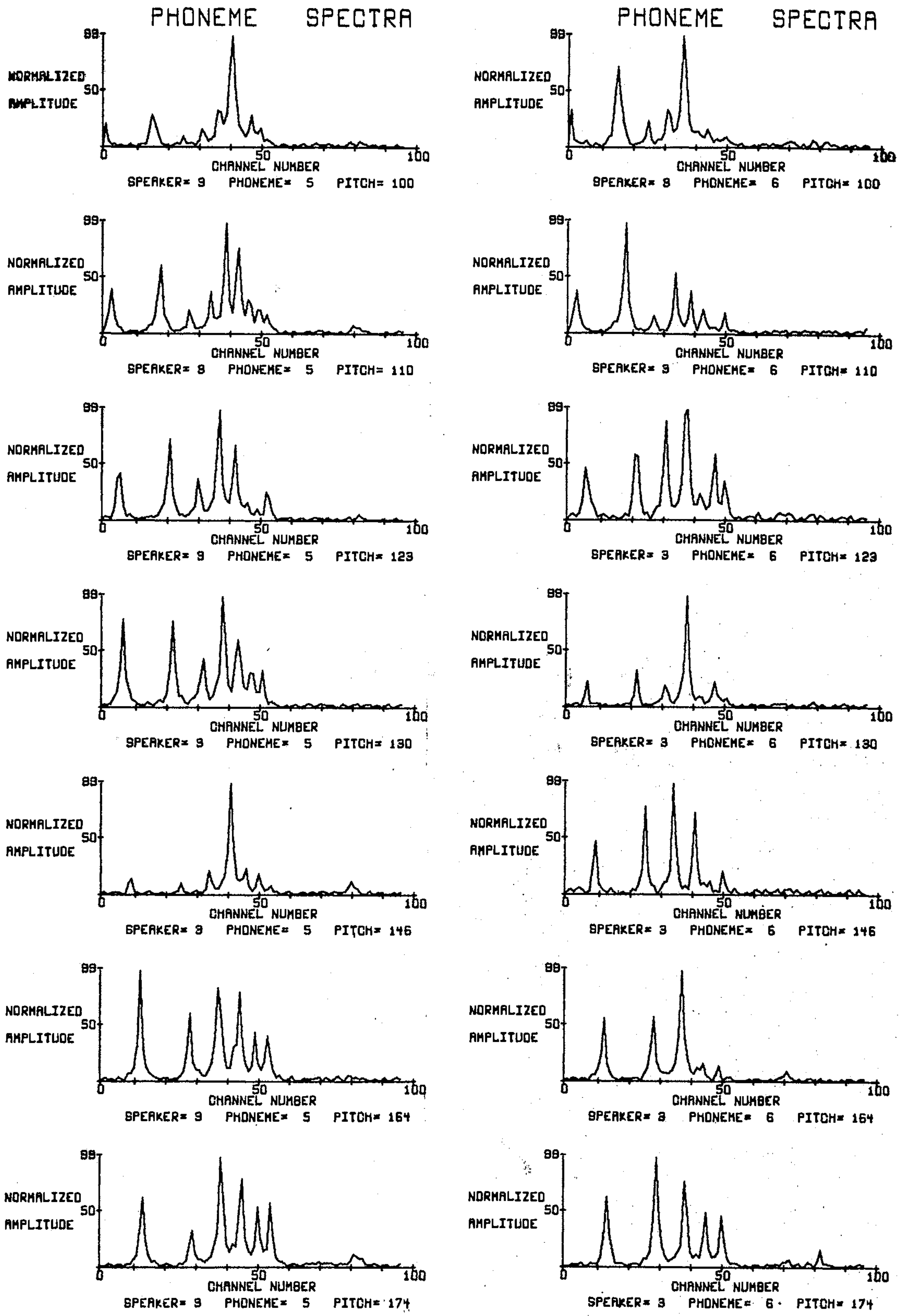


Figure D.14. Spectral Plots of the Phonemes /α/ and /o/ Voiced by Speaker 3 at Several Pitches

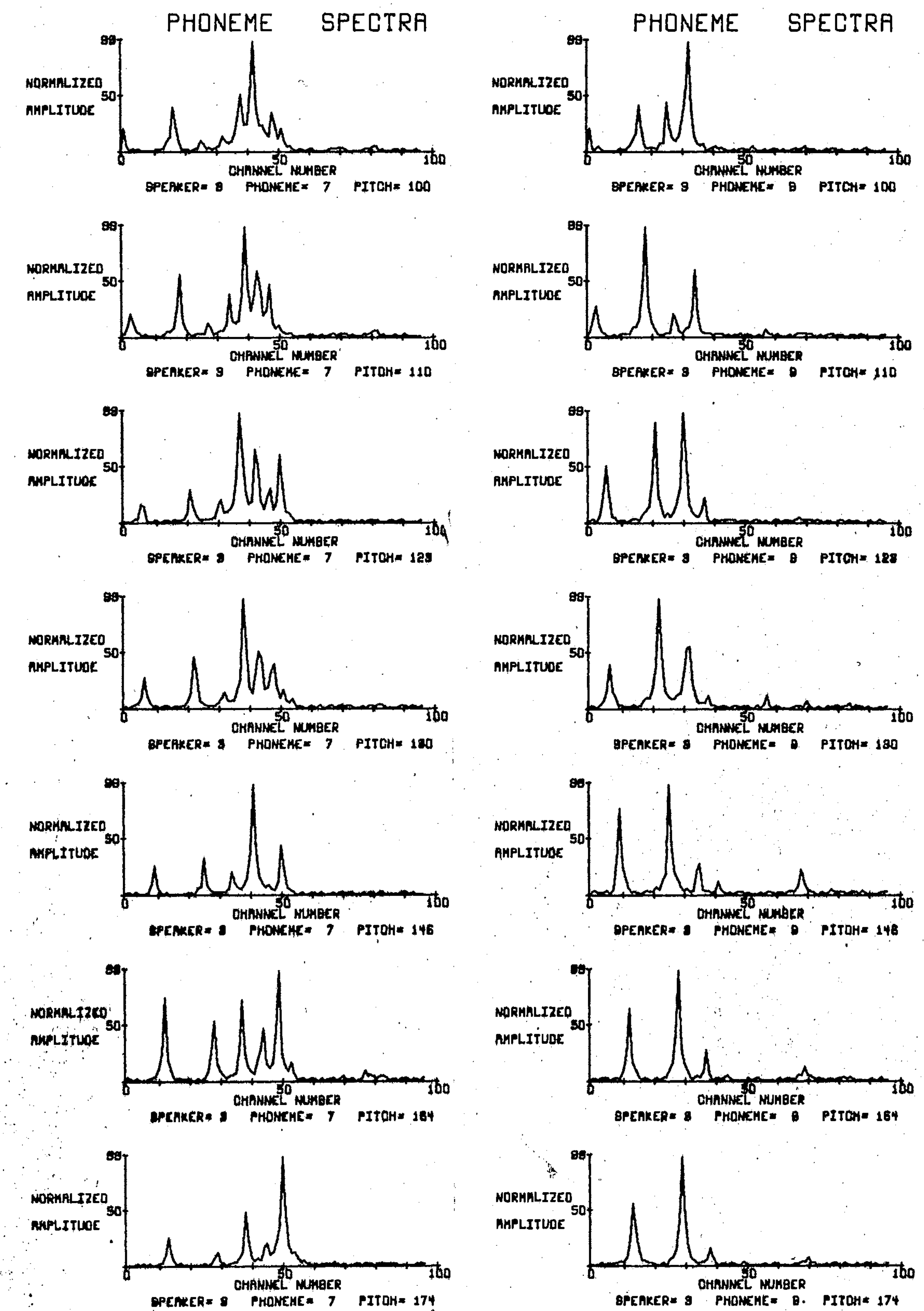


Figure D.15. Spectral Plots of the Phonemes /ə/ and /u/ Voiced by Speaker 3 at Several Pitches

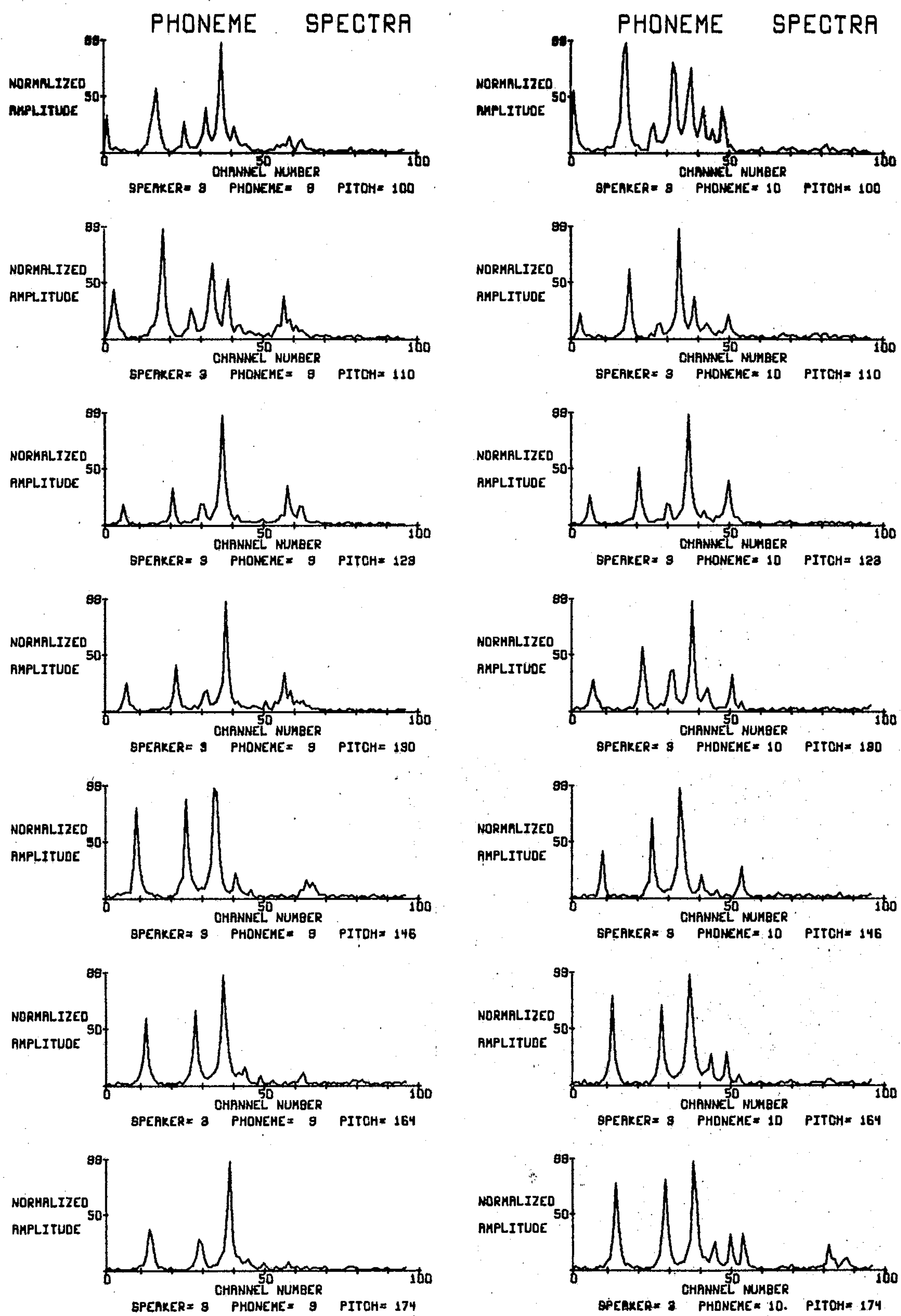


Figure D.16. Spectral Plots of the Phonemes /ɜr/ and /l/ Voiced by Speaker 3 at Several Pitches

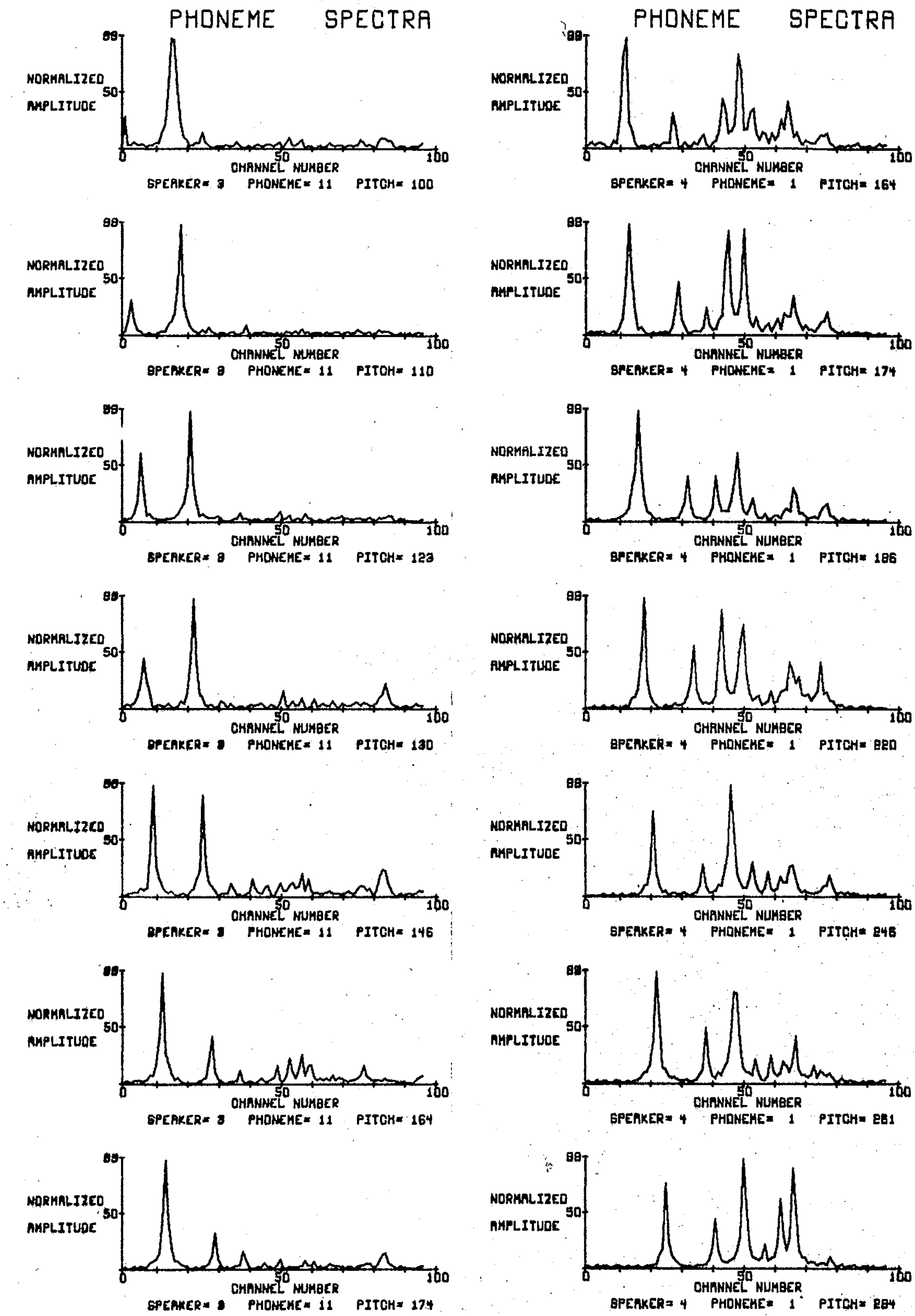


Figure D.17. Spectral Plots of the Phonemes /n/ Voiced by Speaker 3 and /ae/ Voiced by Speaker 4 at Several Pitches

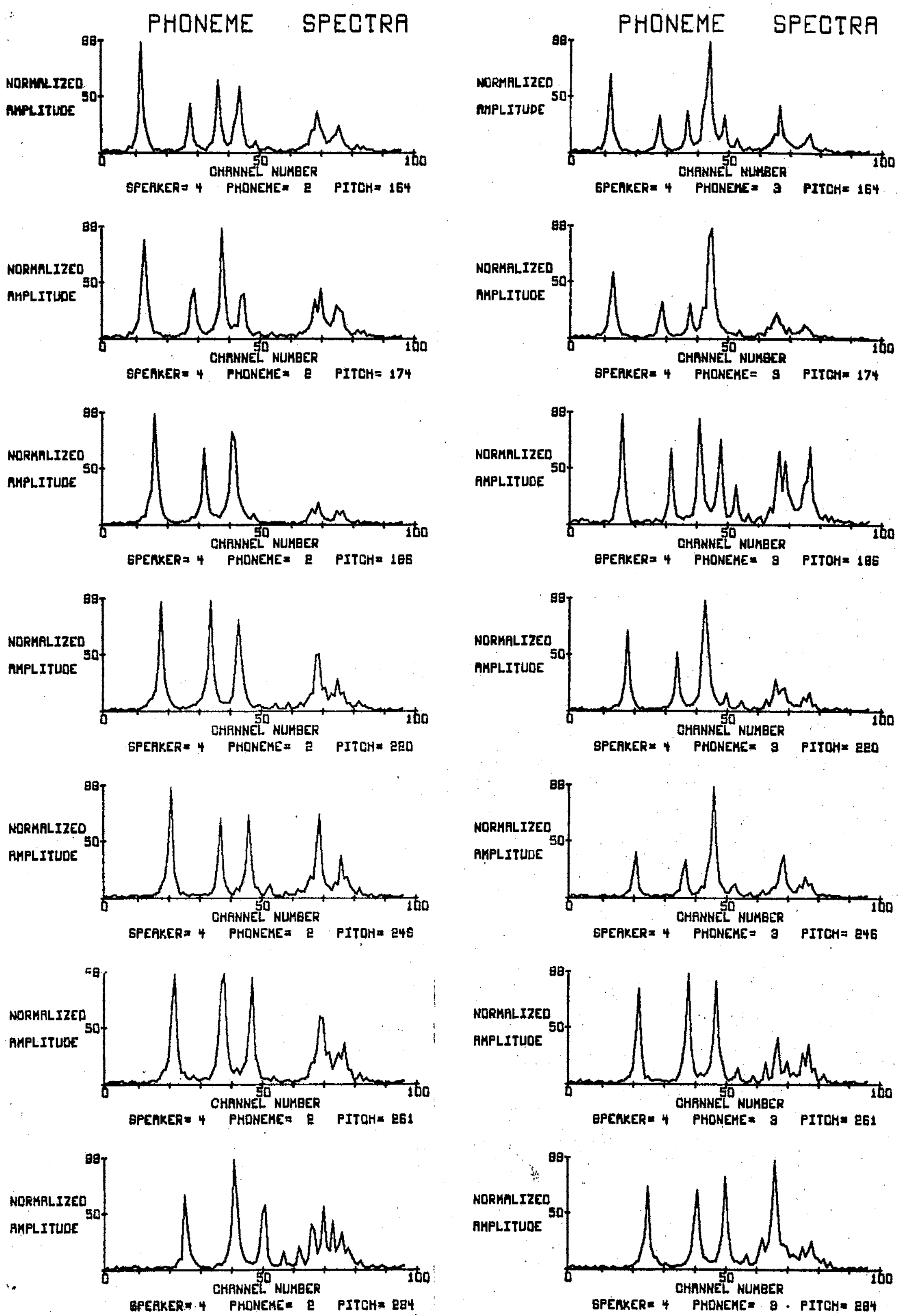


Figure D.18. Spectral Plots of the Phonemes /e/ and /ε/ Voiced by Speaker 4 at Several Pitches

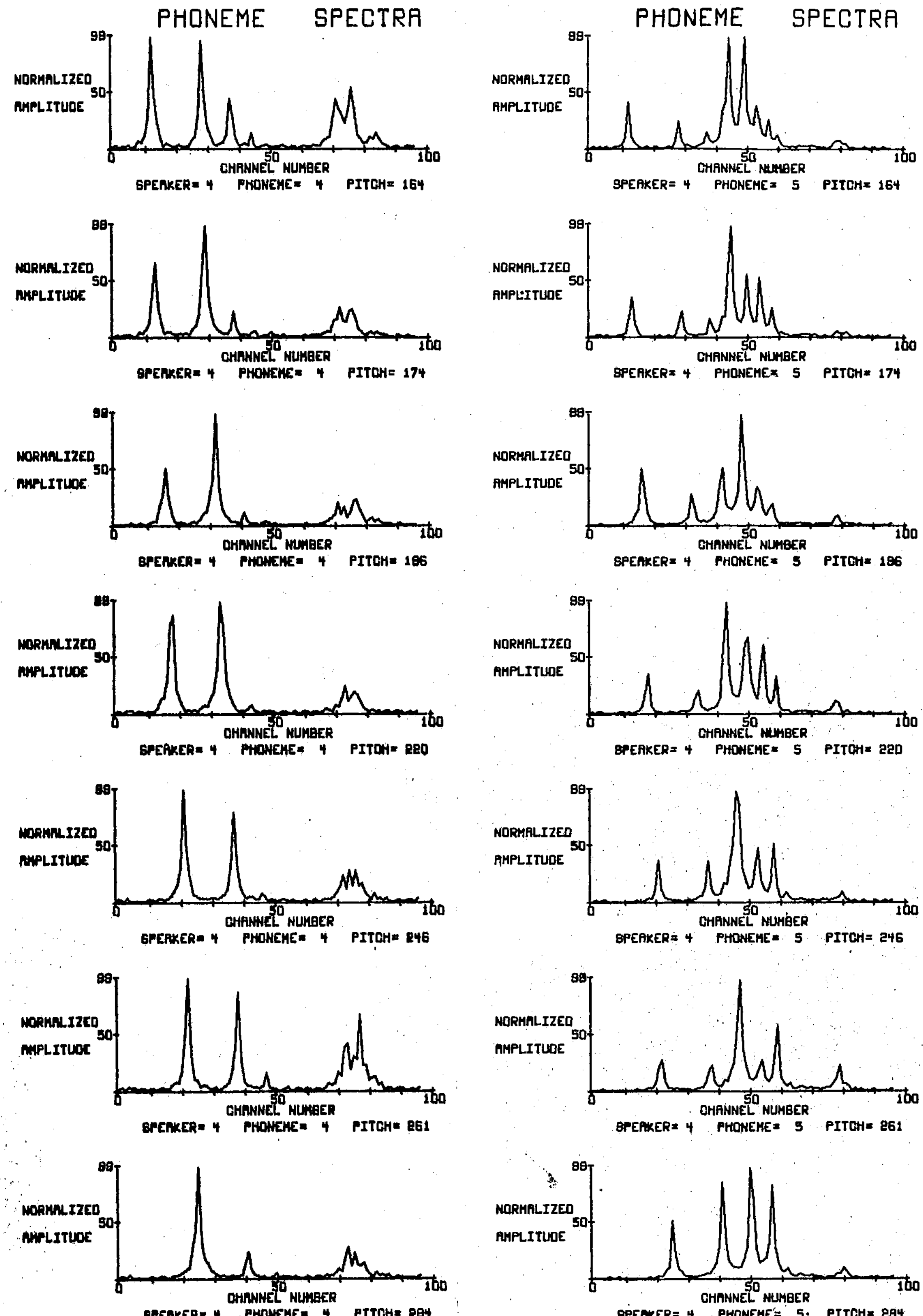


Figure D.19. Spectral Plots of the Phonemes /i/ and /ɑ/ Voiced by Speaker 4 at Several Pitches

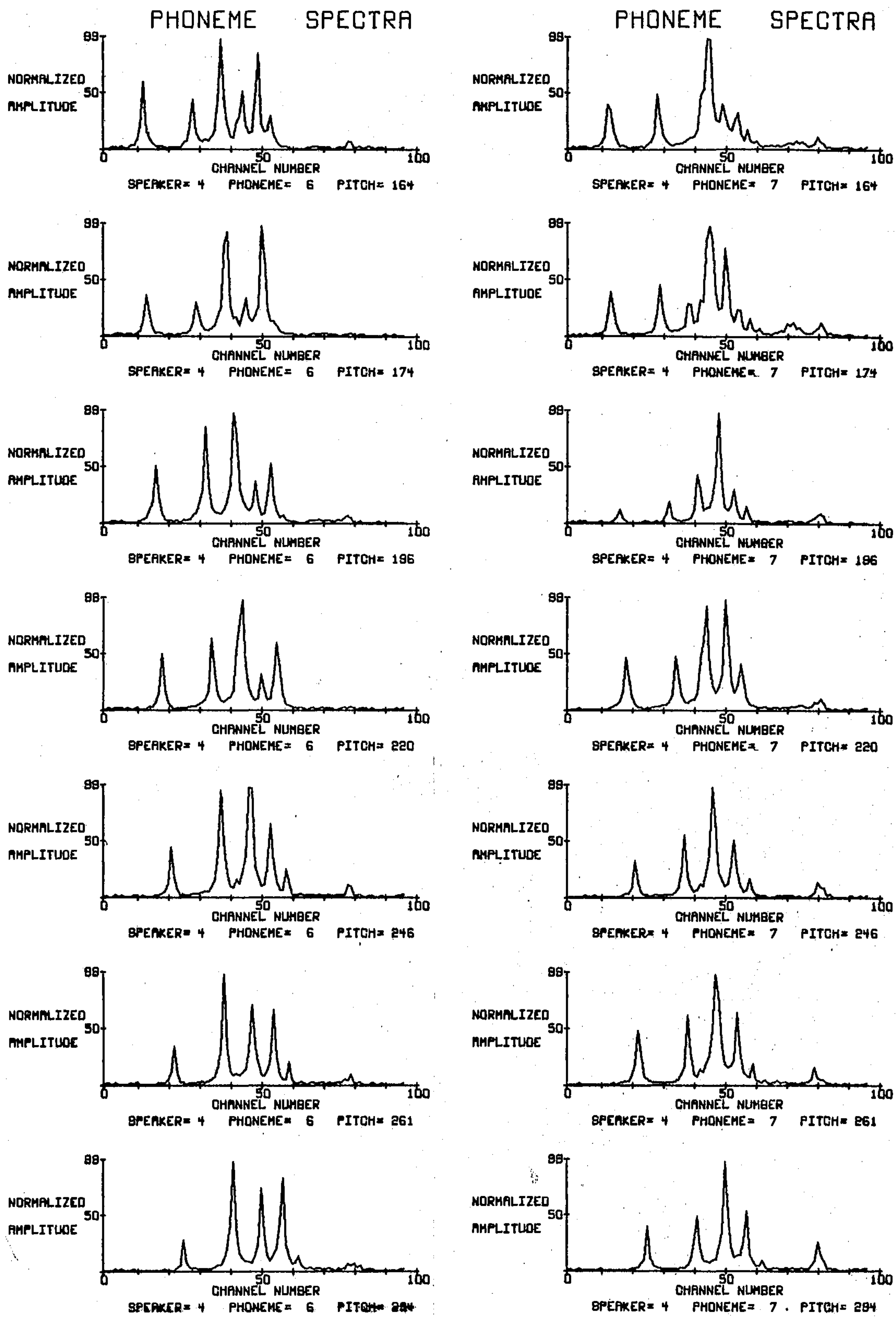


Figure D.20. Spectral Plots of the Phonemes /o/ and /ə/ Voiced by Speaker 4 at Several Pitches

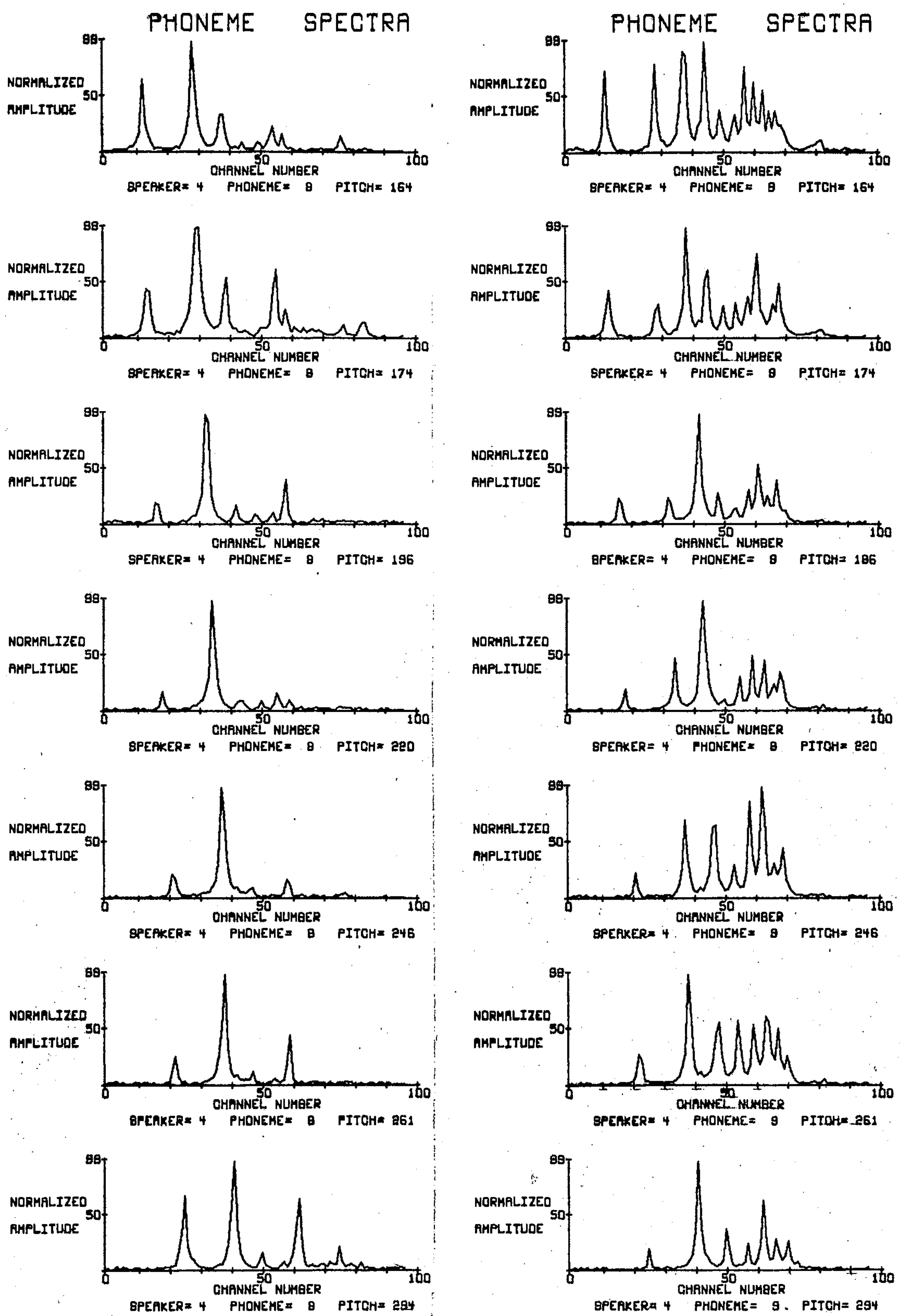


Figure D.21. Spectral Plots of the Phonemes /u/ and /ər/ Voiced by Speaker 4 at Several Pitches

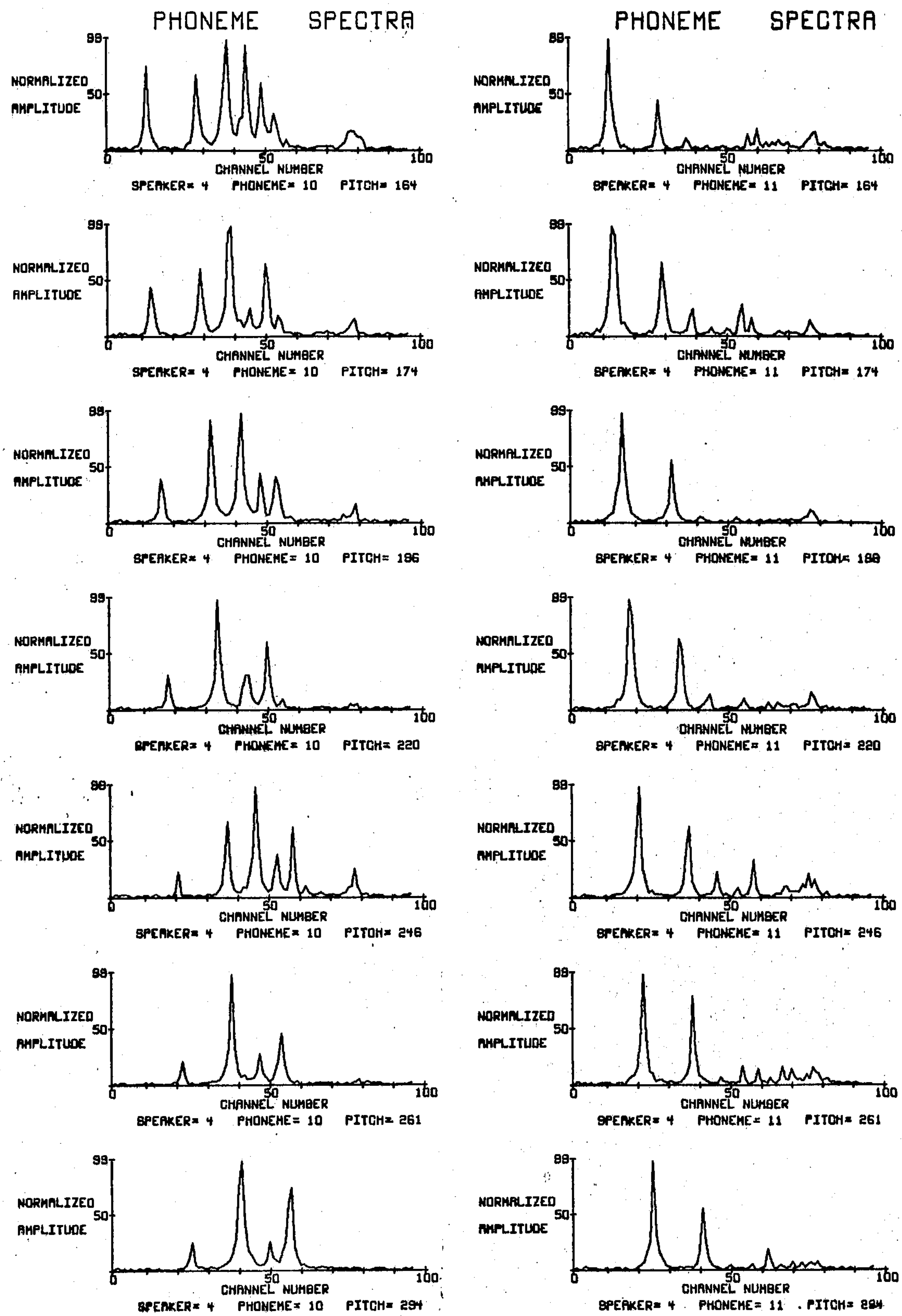


Figure D.22. Spectral Plots of the Phonemes /l/ and /n/ Voiced by Speaker 4 at Several Pitches

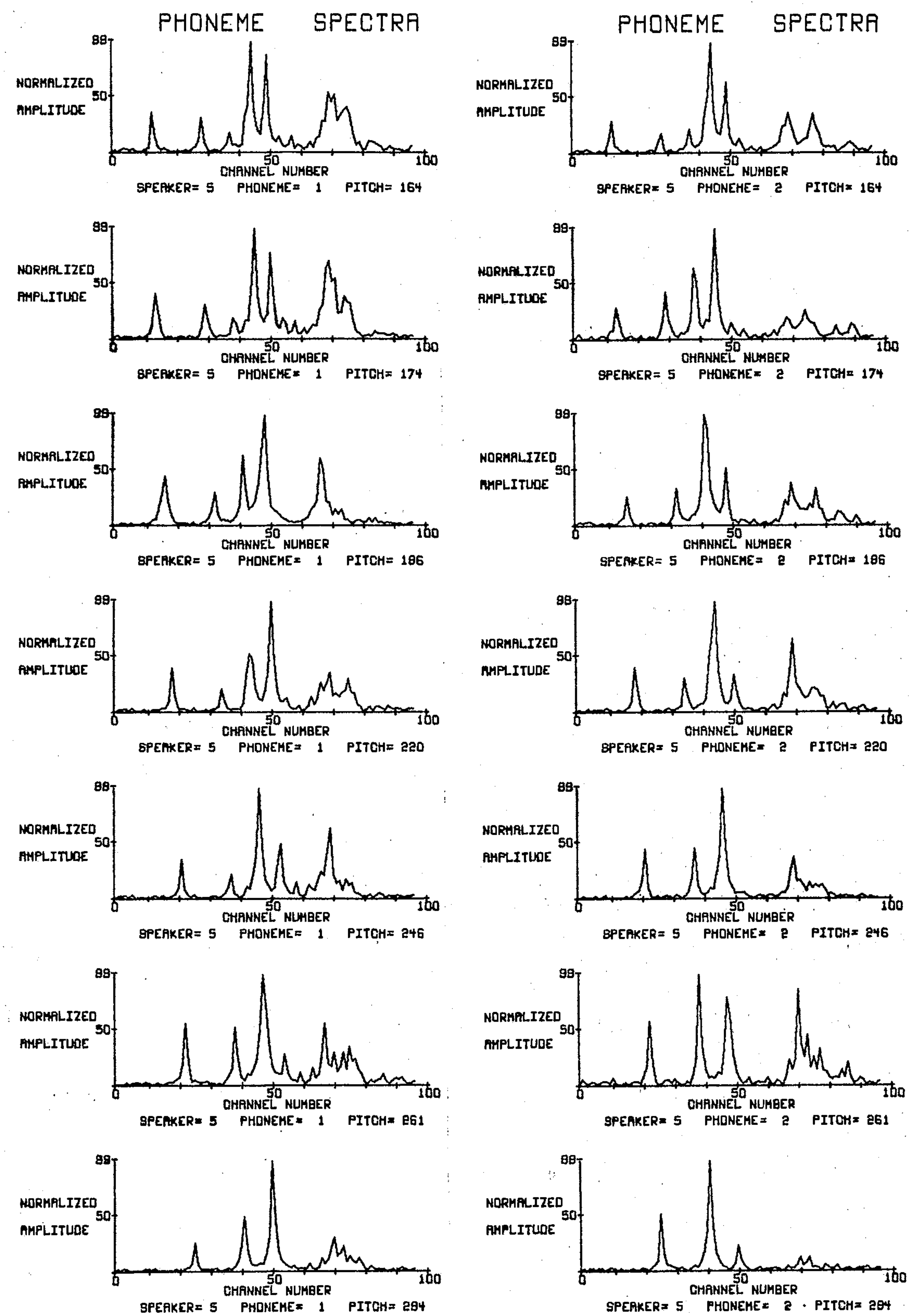


Figure D.23. Spectral Plots of the Phonemes /ae/ and /e/ Voiced by Speaker 5 at Several Pitches

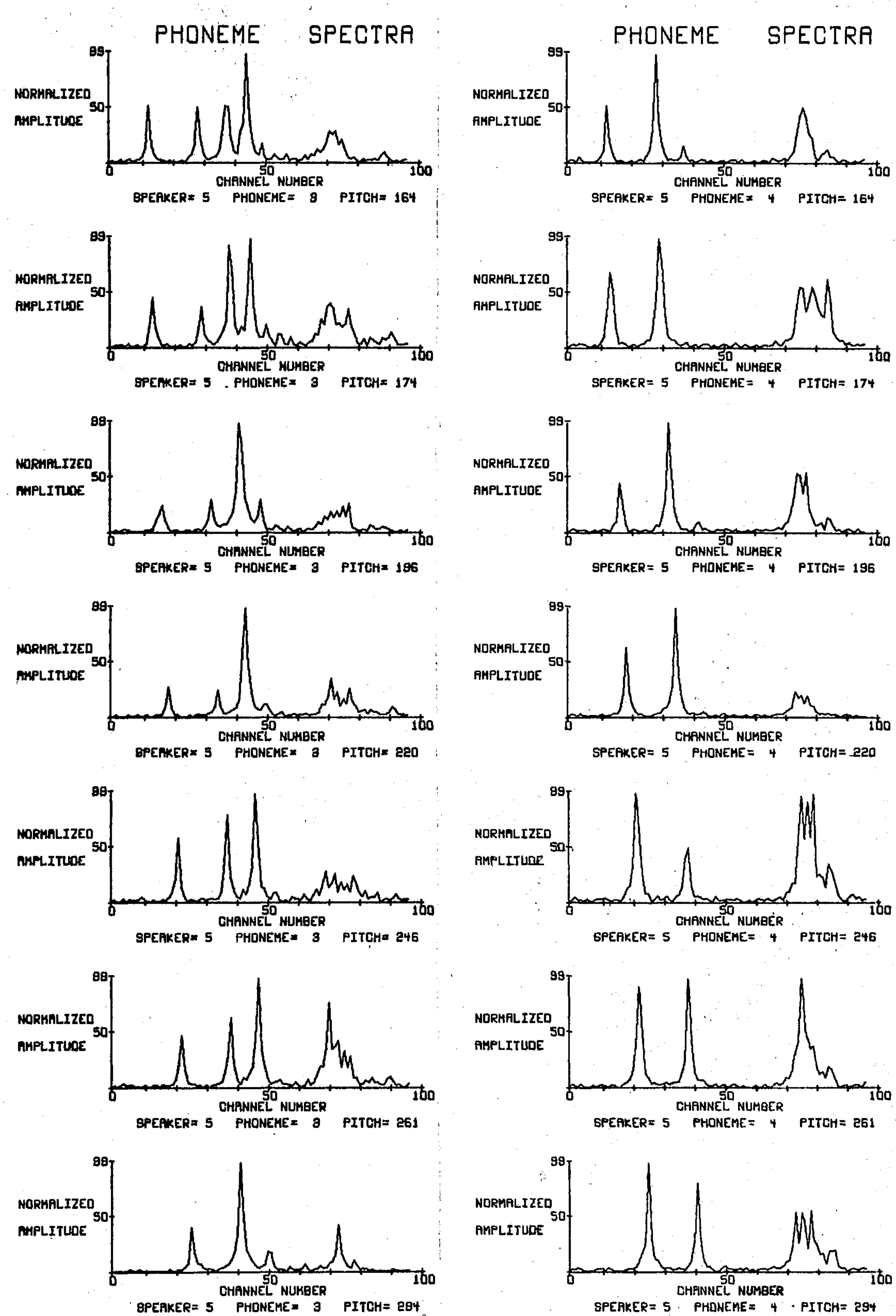


Figure D.24. Spectral Plots of the Phonemes /[illegible]/ and /i/ Voiced by Speaker 5 at Several Pitches

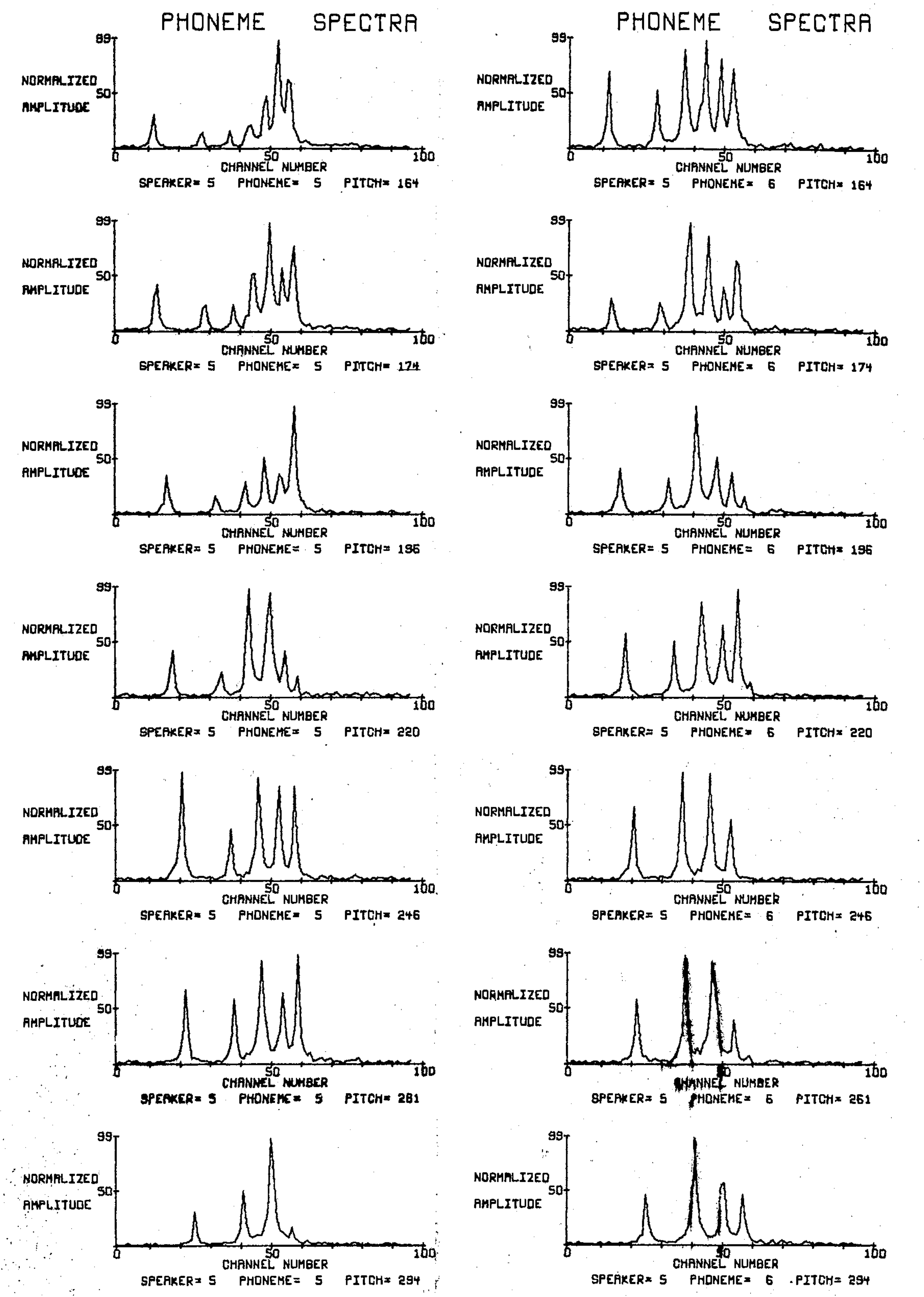


Figure D.25. Spectral Plots of the Phonemes /ɑ/ and /o/ Voiced by Speaker 5 at Several Pitches

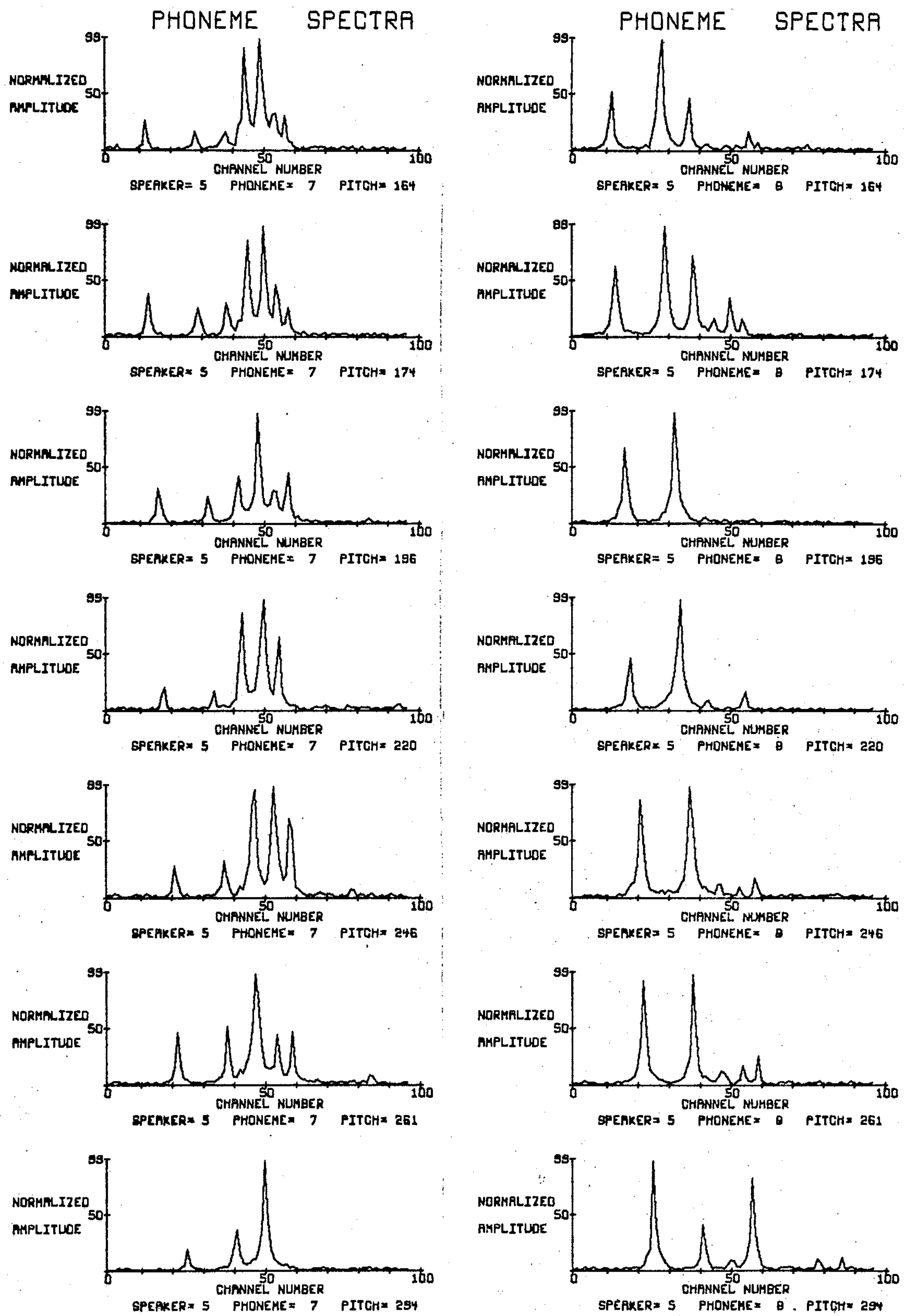


Figure D.26. Spectral Plots of the Phonemes /ə/ and /u/ Voiced by Speaker 5 at Several Pitches

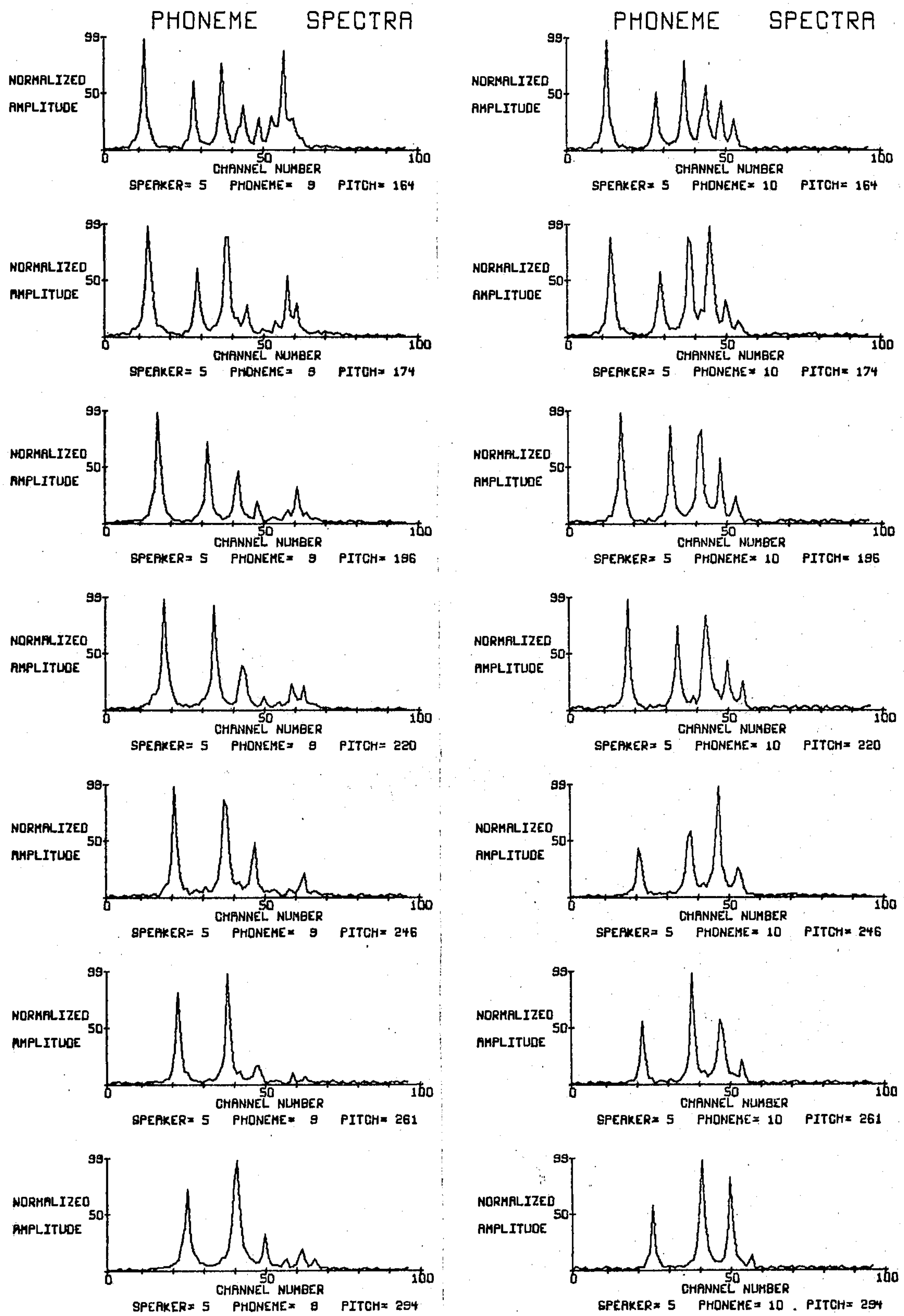


Figure D.27. Spectral Plots of the Phonemes /ər/ and /l/ Voiced by Speaker 5 at Several Pitches

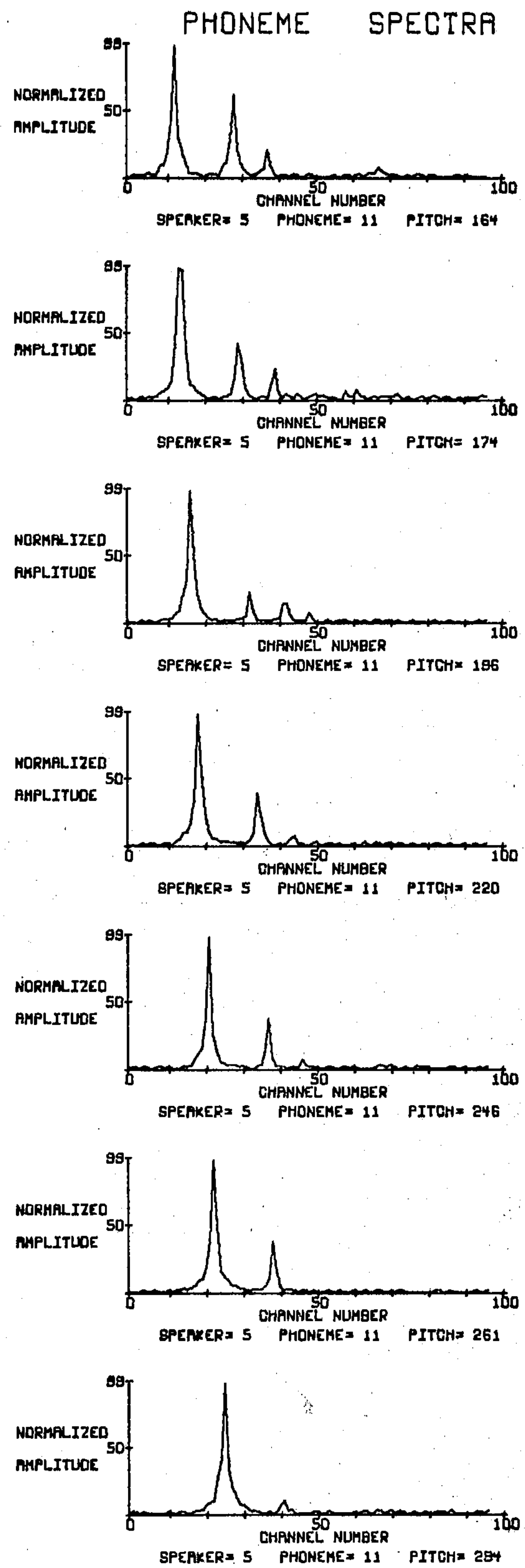


Figure D.28. Spectral Plots of the Phoneme /n/ Voiced by Speaker 5 at Several Pitches

APPENDIX E

COMPUTER PROGRAMS

This appendix contains listings of the computer programs which were used in reading the phoneme samples into the computer, plotting the spectral data and recognition of phonemes by the computer.

TABLE XIII

PROGRAM FOR READING PHONEME SPECTRA

```
C      PROGRAM FOR READING PHONEME SPECTRA
       DIMENSION PHN(100),KPHN(100)
       CALL SETCL(2)
 10    PAUSE
       CALL SETCL(1)
       CALL RSETCL(1)
       CALL SETCL(1)
       CALL RSETCL(1)
 30    DO 60 J=1,98
       CALL SETCL(1)
       CALL RSETCL(1)
       CALL GETADF(1,X)
       PHN(J)=X
 60    CONTINUE
       CALL RSETCL(2)
 70    FORMAT(I1,1X,I2,1X,I3)
 71    READ 70,SPK,PHO,PIT
 75    TM=PHN(1)
       DO 150 I=2,97
       IF (TM-PHN(I))140,150,150
 140   TM=PHN(I)
 150   CONTINUE
       DO 171 K=1,97
       PHN(K)=(PHN(K)/TM)*99.0
       KPHN(K)=PHN(K)
 171   CONTINUE
 174   FORMAT(13X,8HSPEAKER=,I1,3X,8HPHONEME=,I2,3X,6HPITCH=,I3,3X,3HTM=,
      1F6.2,3X,3HST=,F6.2)
 175   FORMAT(I1,1X,I2,1X,I3,1X,I1,2X,20(I3))
 176   FORMAT(I1,1X,I2,1X,I3,1X,I1,2X,17(I3))
       PUNCH 174,SPK,PHO,PIT,TM,PHN(98)
       N=1
       PUNCH 175,SPK,PHO,PIT,N,(KPHN(M),M=1,20)
       N=2
       PUNCH 175,SPK,PHO,PIT,N,(KPHN(M),M=21,40)
       N=3
       PUNCH 175,SPK,PHO,PIT,N,(KPHN(M),M=41,60)
       N=4
       PUNCH 175,SPK,PHO,PIT,N,(KPHN(M),M=61,80)
       N=5
       PUNCH 176,SPK,PHO,PIT,N,(KPHN(M),M=81,97)
       CALL SETCL(2)
       GO TO 10
       END
```

TABLE XIV

PROGRAM FOR PLOTTING PHONEME SPECTRA

```
C     PROGRAM FOR PLOTTING PHONEME SPECTRAL DATA
  50  READ 40,Z
  40  FORMAT(80I1)
      READ 9,SPK,PHO,PIT,TM,ST
   9  FORMAT(21X,I1,11X,I2,9X,I3,6X,F6.2,6X,F6.2)
      DIMENSION XX(97)
      READ 20,(XX(J),J=1,97)
  20  FORMAT(12X,20I3)
      CALL PLAT(201,0.0,100.0,1.428,1.0,0.0,100.0,4.0,1.0)
      IF(PIT-164.0)31,32,31
  32  DO 36 J=1,2
      PUNCH TAPE 18
  36  CONTINUE
      CALL PLAT(99,-11.0,15.0)
      CALL CHAR(0,.20,1)
  35  FORMAT(17HPHONEME    SPECTRA)
      CALL PLAT(89)
  31  CALL PLAT(90,0,0)
      CALL PLAT(90,100.0,0.0)
      CALL PLAT(90,100.0,100.0)
      CALL PLAT(89)
      CALL PLAT(99,40.0,-31.0)
      CALL CHAR(0,.1,1)
   1  FORMAT(10HNORMALIZED)
      CALL PLAT(89)
      CALL PLAT(99,65.0,-31.0)
      CALL CHAR(0,.1,1)
   2  FORMAT(9HAMPLITUDE)
      CALL PLAT(89)
      CALL PLAT(99,52.0,-6.5)
      CALL CHAR(0,.1,1)
   3  FORMAT(3H50-)
      CALL PLAT(89)
      CALL PLAT(99,2.0,-6.5)
      CALL CHAR(0,.1,1)
   4  FORMAT(3H99-)
      CALL PLAT(89)
      Y=-1.5
      DO 10 J=10,110,10
      CALL PLAT(99,103.0,Y)
      CALL CHAR(0,.1,1)
   5  FORMAT(1H+)
      Y=J
  10  CONTINUE
      CALL PLAT(89)
      CALL PLAT(99,110.0,-1.25)
      CALL CHAR(0,.1,1)
   6  FORMAT(1H0)
      CALL PLAT(89)
      CALL PLAT(99,110.0,49.0)
      CALL CHAR(0,.1,1)
   7  FORMAT(2H50)
      CALL PLAT(89)
      CALL PLAT(99,110.0,98.0)
      CALL CHAR(0,.1,1)
   8  FORMAT(3H100)
      CALL PLAT(89)
      CALL PLAT(99,120.0,35.0)
      CALL CHAR(0,.1,1)
  25  FORMAT(14HCHANNEL NUMBER)
      CALL PLAT(89)
      CALL PLAT(99,135.0,8.0)
      CALL CHAR(3,.1,1,SPK,PHO,PIT)
  11  FORMAT(8HSPEAKER=,I2,3X,8HPHONEME=,I3,3X,6HPITCH=,I4)
      CALL PLAT(89)
      CALL PLAT(90,100.0,0.0)
      Y=1.0
      DO 15 J=1,97
      X=99.0-XX(J)
      CALL PLAT(90,X,Y)
      Y=Y+1.0
  15  CONTINUE
      CALL PLAT(89)
      CALL PLAT(99,0.0,0.0)
      M=5
  51  M=M-1
      PUNCH TAPE 18
  18  FORMAT(48H333333333333333333333333333333333333333333333333)
      IF(M)51,60,51
  60  CONTINUE
      GO TO 50
      END
```

TABLE XV

PROGRAM FOR MACHINE RECOGNITION OF SUSTAINED PHONEMES

```
      PROGRAM FOR MACHINE RECOGNITION OF SUSTAINED PHONEMES
      DIMENSION A(3,40),T(3,40),R(3,40),AMAX(3),AMAX1(40),D(3),S(40),W(3
     1),AMIN(40),C(40)
  175 FORMAT(F12.6)
  180 FORMAT(I2,E9.2)
      M=0
      I=1
  115 FORMAT(////,33X,13I3)
  125 FORMAT(12X,7(I3))
  135 FORMAT(12X,20I3)
  100 M=M+1
      READ 115,(A(M,J),J=1,13)
      READ 135,(A(M,J),J=14,33)
      READ 125,(A(M,J),J=34,40)
      IF(M-2)100,100,59
   59 DO 61 J=1,40
   61 R(I,J)=(A(1,J)+A(2,J)+A(3,J))/3.0
      DO 80 J=1,40
      AMAX1(J)=0.0
      DO 80 M=1,3
      IF(AMAX1(J)-A(M,J))70,80,80
   70 AMAX1(J)=A(M,J)
   80 CONTINUE
      DO 110 J=1,40
      AMIN(J)=50.0
      DO 110 M=1,3
      IF(AMIN(J)-A(M,J))110,110,90
   90 AMIN(J)=A(M,J)
  110 CONTINUE
      DO 120 J=1,40
  120 T(I,J)=(AMAX1(J)-AMIN(J))/2.0
      I=I+1
      M=0
      IF(I-3)100,100,130
  130 CONTINUE
      READ 115,(S(J),J=1,13)
      READ 135,(S(J),J=14,33)
   99 READ 125,(S(J),J=34,40)
      E=1.0
      DO 171 M=1,3
      D(M)=0
      DO 170 J=1,40
      C(J)=(S(J)-R(M,J))/(E+T(M,J))
      C(J)=C(J)*C(J)
      D(M)=D(M)+C(J)
  170 CONTINUE
  171 CONTINUE
      DO 190 M=1,3
      D(M)=(D(M)/40.0)
      W(M)=99.0/(1.0+D(M))
      TYPE 180,M,W(M)
      CALL RETYPE
  190 CONTINUE
      GO TO 130
      END
```

VITA

James Edgar Baker

Candidate for the Degree of

Doctor of Philosophy

Thesis: EXPERIMENTAL ADVANCEMENTS IN SPEECH SIGNAL REDUCTION TECHNIQUES

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Columbus, Ohio, April 16, 1931, the son of Joseph E. and Sally P. Baker.

Education: Graduated from Chester High School, Chester, Vermont in June, 1950; received the Bachelor of Science degree from the United States Naval Academy, Annapolis, Maryland in 1954. Commissioned as a regular officer in the United States Air Force in June, 1954; continuous active duty to the present. Current rank is Lieutenant Colonel. Graduated from the Air Force Ground Radar Maintenance School at Keesler Air Force Base, Mississippi in August, 1955; received the Bachelor of Science in Electrical Engineering in May, 1962 and Master of Science degree in August, 1962, from Oklahoma State University. Completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1971.

Professional Experience: Served as Radar Maintenance Officer in Korea, Japan and Florida from August, 1955 to July, 1960. Performed duties of Ballistic Missile Guidance Engineer from September, 1962 to April, 1969.

Professional Organizations: Eta Kappa Nu and Institute of Electrical and Electronic Engineers.