

VOWEL RECOGNITION THRESHOLD
AS A FUNCTION OF TEMPORAL
SEGMENTATION

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY

Richard L. Powell

1967



This is to certify that the

thesis entitled

VOWEL RECOGNITION THRESHOLD
AS A FUNCTION OF TEMPORAL
SEGMENTATION

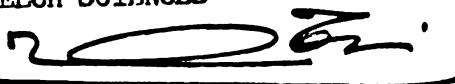
presented by

RICHARD L. POWELL

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in _____

AUDIOLOGY AND SPEECH SCIENCES



Major professor
Oscar I. Tosi, Ph. D.

Date February 16, 1968

JAN 9 1979 1032

~~M 26 1971 247~~

~~E 733~~

ABSTRACT

VOWEL RECOGNITION THRESHOLD AS A FUNCTION OF TEMPORAL SEGMENTATION

by Richard L. Powell

The search for knowledge about how man perceives the world about him has been perpetuated for centuries by man's curiosity. Included in this search are the precise dimensions of the acoustic speech signal. One of the major procedures in this search is modification of the airborne signal that consists of intensity, frequency, and time. The purpose of this study was to determine the vowel recognition threshold as a function of temporal segmentation for the vowels /i /, /ɪ /, /e /, /ɛ /, /æ /, /a /, /ɔ /, /o /, /u /, and /ʊ /. Previous studies had suggested three milliseconds as the vowel recognition threshold as a function of temporal segmentation. Each of the vowels was recorded when phonated by a speaker sitting in a sound treated room. A Lissajous figure produced on an oscilloscope as well as a 125 Hz tone presented through earphones aided the speaker in attaining vowels with 125 Hz fundamentals. These vowels were each presented ten times to eight doctoral students with normal hearing for validation.

A modified psychophysical method of minimal change was selected as the means for attaining the vowel recognition threshold as a function of temporal segmentation. The vowels were then segmented into fifteen different temporal intervals

taken from the middle of the vowel and ranging from four to sixty milliseconds. These stimuli were arranged in random order on magnetic tapes in 220 ascending and descending series and presented to six doctoral students in speech and hearing science. Each listener responded by writing the international phonetic symbols for each of the 3,300 stimuli presented to him through earphones at 70 dB SPL in a sound treated room.

The listener responses were tabulated in an order for each vowel and listeners retaining the temporal segment value, series order and first half second half information. The temporal segment threshold for each series was computed and employed in a four factor analysis of variance ($2 \times 2 \times 6 \times 10$); two series order, two time orders---first half second half--- , six listeners and ten vowels. The vowel /e / was not recognized 50 per cent of the time at the longest temporal segment presented (60 milliseconds) and was excluded from the analysis.

The analysis of variance yielded interactions between and among some of the main effects. In order to visualize these interactions, line graphs of the vowel recognition thresholds for each subject were prepared for ascending, descending, first half, second half, and over the entire experiment. All of these graphs showed the interactions clearly. But a similar pattern held for each graph; the /A /, /a / and /o / had higher recognition thresholds than the other seven vowels, /i /, /ɪ /, /ɛ /, /æ /, /ɔ /, /ʊ /, and /u /. A mean difference test was applied to these data indicating a

significant difference between the / Δ /, / a /, and / $\circ$ / sounds from the other seven vowels. Confusion matrixes were prepared for each subject as well as a composite matrix over all subjects. There was a great deal of confusion among the / Δ /, / a /, and / $\circ$ /. Mis-identification of vowels was most often confused with the vowel sound adjacent to it on the vowel tongue hump diagram. With the interactions and confusion in mind the vowel temporal recognition threshold for each vowel was reported as follows:

Vowel	Milliseconds	Vowel	Milliseconds
i	10.3	a	27.2
ɪ	14.4	ɔ	21.2
ε	13.4	o	11.4
æ	10.6	ʊ	13.8
Λ	18.5	u	9.3

The analysis of variance indicated there was a significant difference between or among the temporal recognition thresholds for: vowels, subjects, ascending and descending series, subject vowel and time order, vowel and series order, vowel and time order. There was no significant difference between the cardinal and non-cardinal vowels.

VOWEL RECOGNITION THRESHOLD
AS A FUNCTION OF TEMPORAL
SEGMENTATION

by
Richard L. ^{LeVri} Powell

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Audiology and
Speech Sciences

1967

G 49257

5-20-68

DEDICATED TO
Richard J. and Muriel,
Ann, Rick, Rob, and Beth

CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	v
Chapter	
I. INTRODUCTION.....	1
Purpose of the Study	
Significance of Project	
Definition of Terms	
Organization of Report	
II. REVIEW OF THE LITERATURE.....	17
Major Studies Probing the Auditory Limits of Sinusoidal Stimuli	
Relation of Pure Tone Acuity to Speech Production	
Speech Sound Discrimination	
Studies Related to Auditory Perception as a Function of Time	
Perceptual Effect of Modifying the Acoustic Vowel Signal	
Summary	
III. EXPERIMENTAL APPARATUS AND PROCEDURES.....	41
Apparatus	
Procedures	
Experimental Procedures	
IV. RESULTS AND DISCUSSION.....	74
Results	
Discussion	
V. SUMMARY AND CONCLUSIONS.....	102
Summary	
Conclusions	
Implications for Further Research	

CONTENTS (Continued)

BIBLIOGRAPHY.....	112
APPENDIX A RAW DATA USED IN THE ANALYSIS OF VARIANCE: THRESHOLDS IN MILLISECONDS TABULATED FOR EACH SERIES.....	119
APPENDIX B SUBJECT VOWEL TEMPORAL SEGMENT RECOGNITION THRESHOLD: SERIES MEANS FOR THE; ASCENDING AND DESCENDING SERIES; FIRST HALF, SECOND HALF, AND OVER ENTIRE EXPERIMENT.....	123
APPENDIX C IDENTIFICATION MATRIXES FOR EACH SUBJECT OVER ENTIRE EXPERIMENT.....	131

LIST OF TABLES

Table	Page
1. Vowel Temporal Segments in Milliseconds.....	64
2. Recognition Threshold in Milliseconds Attained in Pilot Study I.....	65
3. Recognition Thresholds Attained in Pilot Study II.....	67
4. Example of Tabulating Sheet for Ordering the Listeners Responses.....	76
5. Analysis of Variance Table.....	79
6. Vowel Temporal Segment Threshold in Milliseconds Difference Matrix.....	92
7. Identification Matrix Over all Six Subjects.....	96

LIST OF FIGURES

Figure	Page
1. Flow Diagram of Live Vowel Recording Apparatus.....	45
2. Block Diagram of Electronic Vowel Segmenting Apparatus.....	47
3. Diagram of the Vowel Segmenting Apparatus Electronic Function.....	51
4. Vowel Wave Form Taken From Oscilloscope During Live Vowel Recording.....	60
5. Vowel Wave Form Taken From Oscilloscope During Live Vowel Recording.....	61
6. Subject Temporal Segment Recognition Threshold for Each Vowel Presented in Ascending Series.....	81
7. Subject Temporal Segment Recognition Threshold for Each Vowel Presented in Descending Series.....	81
8. Subject Mean Vowel Temporal Segment Recognition Threshold for Ascending Series.....	83
9. Subject Mean Vowel Temporal Segment Recognition Threshold for Descending Series.....	83
10. Subject Temporal Segment Recognition Threshold for Each Vowel During First Half of Experiment.....	84
11. Subject Temporal Segment Recognition Threshold for Each Vowel During Second Half of Experiment.....	84
12. Subject Temporal Segment Recognition Threshold Means and Standard Deviation for Each Vowel for the First Half of Experiment.....	86
13. Subject Temporal Segment Recognition Threshold Means and Standard Deviation for Each Vowel for the Second Half of Experiment.....	86

FIGURES (Continued)

14. Subject Temporal Segment Recognition
Threshold for Vowels Over Entire Experiment..... 87
15. Subject Medians Used to Compute Temporal
Segment Recognition Threshold Rather Than
the Means Over Entire Experiment..... 87
16. Subject Temporal Segment Recognition
Threshold Mean and Standard Deviation for
Each Vowel Over Entire Experiment..... 88
17. Subject Median Temporal Segment Recognition
Threshold Means Over Entire Experiment..... 88
18. The Unsegmented Vowel Wave Forms are
Presented Below the Vowel Segment Presented
to the Subject..... 90
19. The Unsegmented Vowel Wave Forms are
Presented Below the Vowel Segment Presented
to the Subject..... 91

CHAPTER I

INTRODUCTION

Physical parameters of man's environmental percepts have been the focus of many disciplines throughout the centuries. And man's ability to communicate with his environment constitutes one of the major subjects of his investigations since the beginning of time. Even the author of Genesis, one of the older writings, seems to have recognized the importance of sensory perception in attaining information from our environment. In fact, Genesis 3:7 states: "Then the eyes of both were opened, and they knew that they were naked...." More recently, Descartes¹ in 1650, supported the doctrine that man was dominated by two phenomena, mind and body, with the brain being most closely associated with mind. Philosophers, as long ago as Aristotle, advocated that nothing is in man's mind that hasn't entered through the senses. Leibnitz and Hartley supported the parallelism between the experienced world and processes in the body, the nervous system being the pathway into the brain. In 1834, Weber reported his observations that a just noticeable difference (JND) in a stimulus

¹Early history, Descartes, Aristotle, Leibnitz and Hartley, Weber, and Fechner taken from Ira J. Hirsh, The Measurement of Hearing (New York: McGraw-Hill Book Company, Inc., 1952), pp. 8-11.

required an additional amount of stimulus equal to a constant fraction of the original stimulus. He therefore proposed the following general formula $\frac{\Delta I}{I} = K$ (for a JND) where "I" equals the stimulus, " ΔI " equals the change in stimulus, and "K" is a constant that varies with sense modality. With these observations in mind, Fechner, in 1860, proposed that JND's for a particular sense could be calculated for different values on the physical scale, yielding values on a psychological scale. And this theoretical concept can be stated as follows: The responses of human organisms to stimuli vary as the logarithms of those stimuli $S = K \log I$ where "S" equals the magnitude of sensation, "I" is a dimension of the stimulus, and "K" is a constant. This formula is known as Fechner's Law. After much experimentation contemporary knowledge leads us to believe that this function does not hold true for all senses or even for all values within one sense over its entire continuum.

In 1927 Thurston² presented his law of comparative judgment which carried the scaling of psychological behavior ahead one more step. This method is concerned with establishing a basis for scaling psychological events by comparing it with another psychological event rather than with a physical scale. Thurston's law of comparative judgment is stated in the following form

$$R_b - R_a = Z_{ba} \sqrt{\sigma_a^2 + \sigma_b^2 - 2 r_{ab} \sigma_a \sigma_b}$$

²L. L. Thurston, "A Law of Comparative Judgment," Psychology Review, V, No. 34 (1927), pp. 273-286.

where " R_a " and " R_b " = mean response to stimulus "a" and "b", " σ " = standard deviation, " Z_{ab} " = the normal deviation, or standard measure of distance corresponding to " $P_{b > a}$," and " r " = coefficient of correlation between " R_a " and " R_b ." In actual practice we assume all the intercorrelations to be zero and the last term $Z^{rab} \sigma_a \sigma_b$ drops out thus eliminating one of the three unknowns. Methods developed for scaling psychological phenomena are described by Guilford.³

Many scientists have concerned themselves with recognition through various sense modalities. Each sensory organ is receptive to a particular type of physical energy. Many commonly applied diagnostic techniques have been discovered through experimental research of the limits of human sensitivity and perception. Gamble and Cook's study of the olfactory sense in 1898,⁴ Graham's vision studies in 1937,⁵ and the two point tactile study by Boring in 1930⁶ exemplify experiments utilizing psychophysical methods.

³J. P. Guilford, Psychometric Methods (New York: McGraw-Hill Book Company, Inc., 1954), pp. 154-337.

⁴Eleanor Acheson McCulloch Gamble, "The Applicability of Weber's Law to Smell," American Journal of Psychology, X, No. 1 (1898).

⁵C. H. Graham and Carolyn Cook, "Visual Acuity as a Function of Intensity and Exposure---Time," American Journal of Psychology, XLIX, No. 4 (October, 1937), pp. 654-661.

⁶Edwin G. Boring, "The Two-Point Limen and the Error of Localization," American Journal of Psychology, XLII, No. 3 (July, 1930), pp. 446-449.

The three major aspects of auditory stimuli--frequency, intensity, and time--have also been investigated extensively through psychophysical methods. Sivian and White reported finding intensity thresholds attained for 21 frequencies ranging from 100 Hz to 15,000 Hz.⁷ Stevens, Volkman and Newman⁸ developed a pitch scale employing the method of equal sense distance; and this numerical scale was later revised by Stevens and Volkman⁹ and is called the mel scale. Stevens¹⁰ in 1936 devised a numerical scale for loudness and defined one sone as the loudness of a 1,000 Hz pure tone at 40 dB (re 0.0002 dyne/cm²) above absolute threshold. From this reference point he devised a sone scale obtained by using the interval and ratio experimental methods. Equal-loudness contours have been established by Fletcher and

⁷L. J. Sivian and S. D. White, "On Minimum Audible Sound Fields," Journal Acoustical Society of America, IV, No. 4 (April, 1933), pp. 288-321.

⁸S. S. Stevens, J. Volkman and E. B. Newman, "A Scale for the Measurement of the Psychological Magnitude Pitch," Journal Acoustical Society of America, VIII, No. 3 (January, 1937), pp. 185-190.

⁹Stanley Smith Stevens and J. Volkman, "The Relation of Pitch to Frequency, A Revised Scale," American Journal of Psychology, LIII, No. 3 (July, 1940), pp. 329-353.

¹⁰S. S. Stevens, "A Scale for the Measurement of a Psychological Magnitude: Loudness," Psychological Review, XLIII, No. 5 (September, 1936), pp. 405-416.

Munson;¹¹ Stevens;¹² Morgan, Garner and Galambos;¹³ and Thurlow.¹⁴ The phon is the unit of measurement for this intensive scale for loudness. Speech reception thresholds as a function of intensity have also been developed. A composite of these studies plus many more have been presented by Stevens.¹⁵ The results of these studies constitute the bases for the development of audiological diagnostic techniques that are thoroughly described by O'Neill and Oyer.¹⁶

Attempts have been made to find significant differences between individuals with speech defects and those without speech defects on the bases of auditory acuity,^{17,18,19,20,21}

¹¹Harvey Fletcher and W. A. Munson, "Loudness, It's Definition, Measurement, and Calculation," Journal Acoustical Society of America, V, No. 2 (October, 1933), pp. 82-108.

¹²S. S. Stevens, "The Relation of Pitch to Intensity," Journal Acoustical Society of America, VI, No. 3 (January, 1935), pp. 150-154.

¹³C. T. Morgan, W. R. Garner and Robert Galambos, "Pitch and Intensity," Journal Acoustical Society of America, XXIII, No. 6 (November, 1951), pp. 658-663.

¹⁴W. R. Thurlow, "Studies in Auditory Theory: Binaural Interaction and the Perception of Pitch," Journal Experimental Psychology, XXXII, No. 1 (January, 1943), pp. 17-36.

¹⁵Stanley Smith Stevens (ed.), Handbook of Experimental Psychology (New York: John Wiley and Sons, Inc., 1951).

¹⁶John J. O'Neill and Herbert J. Oyer, Applied Audiometry (Dodd Mead and Company, Inc., New York: 1966).

¹⁷Margaret E. Hall, "Auditory Factors in Functional Articulatory Speech Defects," Journal of Experimental Education, VII, No. 2 (December 1938), pp. 110-132.

¹⁸Margaret E. Sullivan, "Auditory Acuity and It's Relation to Defective Speech," Journal of Speech Disorders, IX, No. 2 (June, 1944), pp. 127-130.

speech sound discrimination^{22,23,24,25} and auditory memory span for speech.^{26,27} It is generally recognized that auditory, acuity, sound discrimination, and memory span are essential to learning oral language. These studies and those which follow are more thoroughly discussed in Chapter II. One of the variables of an auditory stimulus is time. What

¹⁹Lee Edward Travis and Mildred G. Davis, "The Relation Between Faulty Speech and the Lack of Certain Musical Talents," Psychological Monographs, XXXVI, No. 2 (March, 1927), pp. 71-81.

²⁰Sara M. Stinchfield, "Some Relationships Between Speech Defects, Musical Ability, Scholastic Attainment, and Maladjustment," Quarterly Journal of Speech Education, XIII, No. 3 (June, 1927), pp. 268-275.

²¹Howard Gilkinson, "The Seashore Measures of Musical Talent and Speech Skill," Journal of Applied Psychology, XXVII, No. 5 (October, 1943), pp. 443-447.

²²Lee E. Travis and Bessie J. Rasmus, "The Speech Sound Discrimination Ability of Cases With Functional Disorders of Articulation," Quarterly Journal, Speech Education, XVII, No. 2 (April, 1931), pp. 217-226.

²³James A. Carrell, "The Etiology of Sound Substitution Defects," Speech Monographs, IV, No. 1 (December, 1937), pp. 17-37.

²⁴Ernest L. Kranvall and Charles F. Diehl, "The Relationship of Auditory Discrimination to Articulatory Defects of Children With No Known Organic Impairment," Journal of Speech and Hearing Disorders, XIX, No. 3 (September, 1954).

²⁵Mildred C. Templin, "A Study of Sound Discrimination Ability of Elementary School Children," Journal of Speech Disorders, VIII, No. 2 (June, 1943), pp. 127-132.

²⁶Virgil A. Anderson, "The Auditory Memory Span for Speech Sounds," Speech Monography, V, No. 1 (December, 1938), pp. 115-129.

²⁷Ruth Watt Metraux, "Auditory Memory Span for Speech Defective Children Compared With Normal Children," Journal of Speech Disorders, VII, No. 1 (March, 1942), pp. 33-36.

effect does the duration of an acoustic stimulus have on auditory perception? Wever and Lawrence²⁸ report Békésy's findings by stating, "Tones well above [intensity] threshold show further that growth of loudness with duration practically attains maximum at 0.2 sec." Munson²⁹ found that a stimulus with constant intensity increases in perceived loudness up to 0.2 sec. or 0.3 sec.

Turnbull³⁰ conducted an experiment requiring six subjects to report the presence of "pitch" or "tonality" when presented three different pure tones at 60 dB above threshold. The time required to perceive pitch was 30 millisecond for 128 Hz, 10 millisecond for 1024 Hz and 20 millisecond for 8192 Hz.

Siegenthaler³¹ investigated the intelligibility of vowel temporal segments presented from 15 to 20 seconds. Vowel /i / was most easily identified and /e / and /u / were most difficult to identify.

²⁸Ernest G. Wever and Merle Lawrence, Physiological Acoustics (New Jersey: Princeton University Press, 1954), p. 6.

²⁹W. A. Munson, "Growth of Auditory Sensation," Journal Acoustical Society of America, XIX, No. 4 (July, 1947), pp. 584-591.

³⁰William W. Turnbull, "Pitch Discrimination as a Function of Tonal Duration," Journal of Experimental Psychology, XXXIV, No. 4 (August, 1944), pp. 302-316.

³¹Bruce M. Siegenthaler, "A Study of the Intelligibility of Sustained Vowels," Quarterly Journal of Speech, XXXVI, No. 2 (April, 1950), pp. 202-208.

Tiffany³² investigated recognition of vowel temporal segments of 0.08, 0.2, 0.5 and 8.0 seconds. Vowel /ɜ / and /i / were most easily identified and /æ / and /u / were most difficult to identify.

Fairbanks and Grubb³³ investigated recognition of 0.3 second vowel temporal segments. The most easily identified were the /i / and, contrary to the above findings, the /u / sound.

Schwartz³⁴ studied thresholds of identification for vowels as a function of their duration, stating, "...results showed that mean sound pressure levels required for threshold of identification varied inversely as a function of the duration of the stimulus." This was true for each of the vowels /i /, /u /, /æ /, /ɔ /, and /a / presented at five temporal segments--32, 48, 64, 80, and 96 millisecond--and phonated by two speakers.

Gray³⁵ reports a significant number of recognitions for

³²William R. Tiffany, "Vowel Recognition Conditions," Journal of Speech and Hearing Disorders, XVIII, No. 3 (September, 1953), pp. 289-301.

³³Grant Fairbanks and Patti Grubb, "A Psychophysical Investigation of Vowel Formats," Journal Speech and Hearing Research, IV, No. 3 (September, 1961), pp. 203-219.

³⁴Martin F. Schwartz, "A Study of Thresholds of Identification for Vowels as a Function of Their Duration," Journal of Auditory Research, III, No. 1 (January, 1963), pp. 47-52.

³⁵Giles W. Gray, "Phonemic Microtomy: The Minimum Duration of Perceptible Speech Sounds," Speech Monographs, IX, No. 1 (December, 1942), pp. 75-90.

vowel temporal segments of 3 millisecond presented at 128 Hz to 13 listeners. Peterson³⁶ reports over 50 per cent recognition of 3 millisecond vowel temporal segments presented at 86 Hz to 15 listeners. Yet Siegenthaler³⁷ found only 52 per cent recognition of vowels presented at 15 to 20 seconds to 14 expert listeners and 12 inexpert listeners. There were no significant differences between expert and inexpert listener scores.

PURPOSE OF THE STUDY

The purpose of this study is to determine the temporal segment threshold required to recognize a vowel correctly. A modified method of minimal change was used to attain the vowel recognition threshold as a function of temporal segmentation. The following questions were formulated to define the problem:

1. What is the vowel temporal segment threshold required to allow 50 per cent vowel recognition?
2. Are the vowel temporal segment recognition thresholds the same for each vowel?
3. Are the vowel temporal segment recognition thresholds the same for each subject?
4. Are there any excessive subject learning or fatigue effects from the beginning to the end of the experiment?

³⁶Gordon E. Peterson, "The Significance of Various Portions of the Wave Length in the Minimum Duration Necessary for Recognition of the Vowel," (Unpublished Ph.D. Dissertation, Department of Speech, Louisiana State University, 1939).

³⁷Siegenthaler, loc. cit.

5. Is there a significant difference in the thresholds of ascending and descending presentation?
6. Are the recognition thresholds for cardinal and non-cardinal vowels significantly different?
7. Are there any interrelationships among or between the subjects, vowel series order, and time presentation?

Question one will be answered by applying the method suggested by Guilford³⁸ to determine a limen when using the method of minimal change.³⁹ The mean of the series will constitute the vowel temporal segment threshold.

The following null hypotheses will be tested to answer questions 2, 3, 4, 5, 6, and 7.

1. There is no significant difference between each of the eleven vowel temporal segment thresholds as obtained in this study.
2. There is no significant difference between the six subjects' temporal segment thresholds for the vowels as obtained in this study.
3. There is no significant differences between the vowel temporal segment thresholds in the first half of the series and second half of the series presented to the subjects.
4. There is no significant differences between the ascending series and descending series vowel temporal segment thresholds.
5. There is no significant differences between the cardinal and non-cardinal vowel temporal segment thresholds.
6. There is no significant interaction between or among the main effects of subject, vowel, series order, and time presentation.

³⁸Guilford, op. cit., pp. 103-106.

³⁹Ibid., pp. 31-33.

SIGNIFICANCE OF PROJECT

Certainly this investigation is akin to what is traditionally categorized as pure research. Determining the minimum duration required to recognize a vowel correctly may lead to defining the distinctive phonemic features as discussed by Jorgensen.⁴⁰ Jorgensen writes, "Only when distinctive features have been phonetically defined for various positions separately can we attempt to find a common denominator." These minimum distinctive features may also contribute to information concerning the differences between allophones, commonly a topic of the linguist in discussions of close phonetic transcriptions. Light may be shed upon the phonemic differences between languages or even within languages such as the possible minute distinctive features between homonyms. Admittedly, it has been hypothesized and demonstrated that syntax is the key to lexical connotations of homonyms, such as "there" and "their;" however, additional identifying information may be revealed by minute differences in the vowel complex periodic wave form.

Peterson⁴¹ states, "It should be observed, however, that techniques of just noticeable difference have not yet been generally applied to the study of the acoustical properties

⁴⁰Sol Saporta, Psycholinguistics: A Book of Reading (New York: Holt, Rinehart and Winston, Inc., 1961), p. 140.

⁴¹Gordon E. Peterson, "Acoustics of Speech, Part II." Handbook of Speech Pathology, ed. Lee E. Travis (New York: Appleton-Century-Crofts, Inc., 1957), p. 151.

of actual speech, and the above method probably does not approach the precision possible with JND techniques." Willfully, the present researcher endeavored to accomplish precisely this, utilizing a modified method of minimal change.

Disregarding the initial and final vowel sound transients, virtually all the wave form variability is produced within one cycle. Hence, the vowel recognition bearing elements may be located within a single cycle. In fact, researchers such as Schwartz⁴² and Gray⁴³ attained 100 per cent vowel recognition from their subjects with vowel stimuli not possessing the initial or final transients. Thus, it seems logical to think these vowel transients are not necessary for vowel recognition. But it should be noted: this does not mean the vowel transients play no part in vowel identification. Also, it is recognized that slight variations in vowel production do not always change the word meaning.⁴⁴ For instance, in the United States a person from the north can be understood by a native of the south even though his vowel production differs, provided he eludes colloquial terms.

Researchers have been striving to gain knowledge about the several aspects of perception. Hopefully the results of

⁴²Schwartz, loc. cit.

⁴³Gray, loc. cit.

⁴⁴John S. Kenyon, American Pronunciation (Ann Arbor, Michigan: George Wahr Publishing Company, 1950), p. 70.

this study will be an additional stepping stone leading to the formulation of diagnostic procedures useful to the speech and hearing clinician. It may also be one of the building blocks in developing a better understanding of the function of phonemic structures in oral language.

DEFINITION OF TERMS

Definition of the major terms employed in this study are as follows:

1. Time: Time is what is measured by a clock. Bindra and Waksberg state, "Elapsed time refers to temporal durations as measured by standard clocks"⁴⁵ A clock is a reliable cyclical phenomenon. For example, the cyclical movement of a planet is a clock. In fact, our time system on earth is defined in terms of the cyclical rotation of this planet. The day is divided into segments called milliseconds which are equal to $\frac{1}{1000 \times 60 \times 60 \times 24}$ = $\frac{1}{86,400,000}$ of an average annual day duration. One millisecond was the minimum time unit employed in this exploration.
2. Frequency: Frequency of a periodic movement is the number of times a cycle occurs in a second. Therefore, if "T" is the period then $f = 1/T$ [cycle/sec.] or cycles per second (CPS), and also

⁴⁵Dalbir Bindra and Helene Waksberg, "Methods and Terminology in Studies of Time Estimation." Psychological Bulletin, LIII, No. 2 (March, 1956), p. 157.

designated by Hz, a more universal symbol.

3. Fundamental Frequency: The fundamental frequency is "...that frequency whose period is equal to that of the whole complex wave."⁴⁶ In the present study all vowels were vocalized with 125 Hz fundamental frequencies. The method for attaining this required fundamental is described in detail later in this report.
4. Vowel Stimulus: A vowel sound is made up of a complex periodic wave form. It refers to any of the pure English vowels, /i /, /ɪ /, /e /, /ɛ /, /æ /, /ʌ /, /ɑ /, /ɔ /, /o /, /u /, or /ʊ / in the present study.
5. Vowel Temporal Segment: The vowel wave form is made up of three temporal divisions: the initial transient, the middle portion and the decay transient. The present study is concerned with short segments taken from the middle portion of the presented vowel. The vowel temporal segments, therefore, refer to a section of the wave form taken from the middle portion of the vowel wave form. These vowel temporal segments range from 2 millisecond to 210 millisecond presented in the pilot studies and from 4 millisecond to 60 millisecond presented in the major experiment.

⁴⁶Hirsh, op. cit., p. 30.

6. Vowel Temporal Segment Threshold or Limen: This refers to the shortest vowel temporal segment at which the vowel can be recognized correctly 50 per cent of the time when presented to the listeners in a manner described within this study.
7. Recognition: For the purpose of this study a vowel stimulus will be considered recognized when the subject responds to the auditory stimulus by writing the correct international phonetic symbol for the auditory vowel stimulus presented to them in a manner described in Chapter III of this report.
8. Psychophysical Methods: They are procedures encountered in the psychophysical sciences; Guilford cites, "From the time of Fechner, psychophysics has been regarded as the science that investigates the quantitative relationship between physical events and corresponding psychological events."⁴⁷ The method of minimal change was modified to accommodate the stimuli employed in this present study. The method of minimal change and its modifications are described in detail within the section devoted to procedures.

ORGANIZATION OF THE REPORT

Chapter I has introduced some of the research which indicated the problems involved in determining the aspects

⁴⁷ Guilford, op. cit., p. 20.

of recognizing acoustic stimuli and in particular speech stimuli. The studies cited which deal with recognition of vowel temporal segments indicate some controversies, and a need to establish recognition thresholds as a function of vowel temporal segmentation is warranted. Terms pertinent to this study are defined.

Chapter II consists of a review of the literature including studies that relate to pure tone acuity, memory span, discrimination, and recognition as well as studies related to recognition of whole vowels and vowel temporal segments.

Chapter III is divided into two major portions, apparatus and procedures, in order to clarify the events. Special technical techniques will be discussed. Selection of listeners and stimuli are discussed in the experimental procedures.

Chapter IV will contain a presentation of the results of the statistical analysis. Appropriate charts and tables will be presented to clarify the results. Also the method of determining the limens will be thoroughly discussed. Discussion of the findings related to previous studies will also be presented.

Chapter V will include a summary of the study and conclusion which can be drawn from the results. Recommendations for further research will also be formulated.

The Bibliography and Appendices containing the raw data utilized in the analysis of variance and in attaining the limens will make up the remainder of the study.

CHAPTER II

REVIEW OF THE LITERATURE

Licklider⁴⁸ in an excellent discussion of speech perception divided the many aspects of the process into three major operations: "... (1) translation of the speech signal into form suitable for the nervous system, (2) identification of discrete speech elements, and (3) comprehension of meaning."

The first operation is carried out by the auditory receptor. The acoustic speech signal is transduced into mechanical energy by the action of the tympanic membrane and middle ear ossicles at the cochlear oval window. This signal is transduced by the cochlear components into a signal suitable for the nervous system to accept. This first stage is likened to a spectrograph, and Licklider cites researchers, such as Bekesy, Ranke, and Fletcher, who support the theory that the cochlea is an analyzer of the acoustic signal. It is probably true that the spectrograph refines the acoustic signal to a higher degree than the auditory receptor. The second operation, a refining of the cochlear output, may be

⁴⁸J. C. R. Licklider, "On the Process of Speech Perception," Journal Acoustical Society of America, XXIV, No. 6 (November, 1952), pp. 590-594.

analogous to a set of matched electronic filters that results in a neural pattern being presented to the higher brain centers. The third step, that of comprehension, appears to involve a neural pattern of cross-correlation similar to the function of the analogous electronic processes. Of course, this last step presupposes the existence of a neural pattern previously learned with which the different neural patterns can be matched. Hebb⁴⁹ has developed a learning theory on the creation of these neural patterns that he terms "cell assemblies." His theory is also related to the knowledge of how the neuron and synapse function, as well as the behavior of persons with brain dysfunction.

Fletcher⁵⁰ has divided the perception of speech into five major groups; interpretation, loudness, pitch, quality, and tempo. The physical parameters of the acoustic speech signal--intensity, frequency and time--are perceived by the individual as loudness, pitch, quality and tempo aspects of the speech signal. It is well recognized that meaning attached to the neural patterns that have been transduced physiologically from the acoustic speech signal is within the beholder and not in the acoustic air born signal.⁵¹ Yet,

⁴⁹Donald Olding Hebb, The Organization of Behavior: A Neuropsychological Theory (New York: John Wiley and Sons, Inc., 1949), pp. 60-78.

⁵⁰Harvey Fletcher, Speech and Hearing in Communication (New York: D. Van Norstrand Company, Inc., 1961), p. ix.

⁵¹John W. Black and Wilbur E. Moore, Speech Code, Meaning and Communication (New York: McGraw-Hill Book Company, Inc., 1955), pp. 126-127.

it is also recognized that the acoustic speech signal must carry variables that can be transduced by the auditory mechanism into neural patterns interpreted by the person. As Stevens⁵² states, "Clearly, the recognizability of the intensity-frequency-time pattern parallels the intelligibility of speech." Peterson⁵³ supports this by stating, "As suggested previously, when the significant dimensions of this phonetic space have been properly identified, the phonetic value of any vowel can be specified quantitatively in these (physical) dimensions." Many researchers have manipulated the acoustic signals attempting to determine the amount of modification required to cause the lack of sensation or misinterpretation of the auditory stimulus. The acoustic signals have been modified by varying the intensities, frequencies and temporal dimensions separately or in various combinations to determine their effect upon human perception.

MAJOR STUDIES PROBING THE AUDITORY LIMITS OF SINUSIODAL STIMULI

Obviously, auditory intensity thresholds are important to recognition of the acoustic speech signal. The intensity thresholds of the human ear were reported by Sivian and White.⁵⁴ The measurements were taken in a specially con-

⁵²S. S. Stevens, Handbook of Experimental Psychology (New York: John Wiley and Sons, Inc., 1963), p. 1059.

⁵³Gordon Peterson, "The Phonetic Value of Vowels," Language, XXVII, No. 4 (1951), p. 543.

⁵⁴Sivian and White, loc. cit.

structed sound absorbant room, called the "sound stage." The free field binaural measurements indicated the sensitivity to be greatest at -10 dB (re: 0.0002 dyne/cm²) for 3000 to 4000 Hz averaged over 22 ears. The sensitivity of the ears progressively decreased as the frequencies varied from these most sensitive frequencies to 20 dB (re: 0.0002 dyne/cm²) for 1800 Hz and 44 dB (re: 0.0002 dyne/cm²) for 60 Hz.

Stevens and Davis⁵⁵ reported the ear to be most sensitive at 2000 Hz and capable of distinguishing 1500 just noticeable differences in pitch and 325 just noticeable differences in loudness.

In 1905, Titchner⁵⁶ determined the threshold of audible pitch to be 14.5 Hz. The method of minimal change was employed, and tones ranging from 25 Hz to 7 Hz were presented in six ascending and six descending series to the observer. The results of this study were confirmed by Wever and Bray⁵⁷ in an extensive experiment utilizing a pistonphone as the sound source. They report that acoustic energy begins to "sound rough" as the frequency falls below 100 Hz. At 30 Hz

⁵⁵S. S. Stevens and Howell Davis, Hearing: Its Psychology and Physiology (New York: John Wiley and Sons, Inc., 1938), p. 59.

⁵⁶Edward Bradford Titchner, Experimental Psychology. Quantitative (New York: Macmillian, 1905), p. 6.

⁵⁷E. G. Wever and C. W. Bray, "The Perception of Low Tones and the Resonance-Volley Theory," Journal of Psychology, III, No. 1 (January, 1937), pp. 101-114.

the stimulus changed to intermittence, and at about 15 Hz the sense of pitch ceased to exist. Wever and Lawrence⁵⁸ cite research demonstrating that tones well above threshold increase in loudness with increase in presentation time up to 0.2 second. Munson⁵⁹ reported a perception of increased loudness as the presentation of a sinusoidal tone increases up to 0.2 or 0.3 second depending upon the frequency. A short exposure to a tone raises the threshold momentarily. Rawnsley and Harris⁶⁰ demonstrated this fatigue phenomenon in 1952. If a listener is presented with a tone well above threshold for 300 milliseconds and then, after a silent period of 80 milliseconds, the threshold will have shifted--the auditory mechanism is less sensitive to the tone presented. The more intense the initial presentation, the greater will be the threshold shift for the second presentation. However, fatigue recovery takes place when the silent period is about 0.5 sec. When the interval between stimuli is at least one second, there is no need to be concerned with this threshold shift.

Relation of Pure Tone Acuity to Speech: A significant relationship between pure tone acuity and ability to articulate

⁵⁸Wever and Lawrence, loc. cit.

⁵⁹Munson, loc. cit.

⁶⁰Anita I. Rawnsley and J. Donald Harris, "Studies in Short Duration Auditory Fatigue: II Recovery Time," Journal of Experimental Psychology, XLIII, No. 2 (February, 1952), pp. 138-142.

speech was reported by Travis and Davis.⁶¹ They investigated three groups of speakers: those with excellent speech, those with fair speech, and those with defective poor speech. All subjects had normal hearing, but those in the poor speech group had significantly higher intensity thresholds when administered "the Seashore Measures of Musical Talent." This same test was used by Gilkinson⁶² and Stinchfield,⁶³ but they failed to support the findings of Travis and Davis.

Sullivan⁶⁴ conducted a hearing survey of Minneapolis school children whose intensity thresholds were 10 dB below normal sensitivity. She found that 18.8 per cent of the children without speech defects and 22.2 per cent of the children with speech defects fell into this category. The difference does not seem to be significant. Hall⁶⁵ and Mase⁶⁶ also found no significant difference in hearing acuity between those with and those without speech defects who had hearing within the normal limits.

Speech Sound Discrimination: Travis and Rasmus⁶⁷ reported significant differences between persons with speech defects

⁶¹Travis and Davis, loc. cit.

⁶²Gilkinson, loc. cit.

⁶³Stinchfield, loc. cit.

⁶⁴Sullivan, loc. cit.

⁶⁵Hall, loc. cit.

⁶⁶Darrel J. Mase, Etiology of Articulatory Speech Defects (New York: Teachers College, Columbia University, 1946).

⁶⁷Travis and Rasmus, loc. cit.

and those without speech defects in their ability to discriminate syllables that sounded alike. Templin⁶⁸ developed a discrimination test that differentiated between children with and without speech defects. Kronvall and Diehl⁶⁹ in 1954 confirmed these findings. Carrell,⁷⁰ Hall,⁷¹ Hansen,⁷² and Mase⁷³ did not find that children with articulation defects differed in their ability to discriminate speech syllables from those children without articulation defects. Spriestersbach and Curtis⁷⁴ directed Anderson in a study in 1949, the results indicated the ability to discriminate "s" sounds was difficult for those having "s" problems. They suggested that general discrimination of sounds may not be as significant as the person's ability to discriminate the particular sound with which he has difficulty.

Studies Related to Auditory Perception as a Function of Time:

Turnbull⁷⁵ first used intensity as the independent variable

⁶⁸Templin, loc. cit.

⁶⁹Kronvall and Diehl, loc. cit.

⁷⁰Carrell, loc. cit.

⁷¹Hall, loc. cit.

⁷²Burrell F. Hansen, "The Application of Sound Discrimination Tests to Functional Articulatory Defectives With Normal Hearing," Journal of Speech Disorders, IX, No. 4 (1944).

⁷³Mase, loc. cit.

⁷⁴D. C. Spriestersbach and James F. Curtis, "Misarticulation and Discrimination of Speech Sounds," Quarterly Journal of Speech, XXXVII, No. 4 (December, 1951), pp. 483-491.

⁷⁵Turnbull, loc. cit.

and adjusted duration in order to determine the minimum number of vibrations needed to produce a sound judged by the subject to have "pitch" or "tonality." Others, Bode, 1907; Buerck, Kotowski, Lichte, and Kuchorski, 1905, are reported by Turnbull to have established the threshold of perceiving "pitch or tonality" at 2 to 15 cycles. This agreed with his finding of from 2.7 to 17.7 cycles as that required to identify pitch. Six subjects were presented three different tones 128, 1024, and 8192 Hz at 60 dB above sensation level. A second experiment was conducted in which the intensity as well as the duration were varied. In this way he demonstrated that perception was a function of the products of intensity and duration; that is, as intensity increases, the duration threshold decreases; and as the duration is decreased, it is necessary to increase the intensity to establish the threshold. Perception of pitch is not affected by durational changes when the signal is presented at 60 dB above normal sensation level. Small, Brandt, and Cox⁷⁶ investigated loudness of noise (a flat spectrum up to 20,000 Hz) as a function of duration. The psychophysical method of adjustment was initiated by asking twelve college students to match the loudness of the variable stimuli to the standard stimulus by increasing or decreasing the sensation level. The standard stimuli were presented at 500 milliseconds and at three

⁷⁶Arnold M. Small Jr., John F. Brandt, and Phillip G. Cox, "Loudness as a Function of Signal Duration," Journal Acoustical Society of America, XXXIV, No. 4 (April, 1962), pp. 513-514.

sensation levels 10, 35 and 60 dB. The durations of the variable ranged from 1 sec. to 4 milliseconds. It was necessary for the subjects to increase the loudness level of a 10 dB signal presented for less than 0.05 sec. to attain equal loudness. It was also necessary for the subjects to increase the loudness level of a 60 dB signal presented for less than 0.15 millisecond to attain equal loudness. "For durations shorter than this critical duration it was necessary for the subjects to increase the level of the short signal in order that it remain equally loud as the standard."

THE PERCEPTUAL EFFECT OF MODIFYING THE ACOUSTIC VOWEL SIGNAL

Fletcher⁷⁷ studied the recognizability of vowels when the upper or lower frequencies were filtered out of the complex vowel signal. The percentage of recognition decreased when the vowel signals were filtered either from the high frequencies down or the low frequencies up. Only the /i / vowel was recognized 98 per cent of the time when the frequencies above or below 1700 Hz were allowed to pass. This can be interpreted to mean that there are elements on either side of 1700 Hz that allow identification of the vowel /i /. Recognition of the other vowels begins to decrease at a point just above the second formant for low pass filtering. High pass filtering causes recognition to decrease at a point just below the first formant for each vowel. Some of the

⁷⁷Fletcher, op. cit., pp. 418-420.

vowels have discriminative characteristics that extended through out the frequency range, and others have characteristics that are localized in a limited frequency range. Miller and Lichlider⁷⁸ filtered on going speech and reported equal amounts of recognition of speech on either side of 1900 Hz.

Stevens and House⁷⁹ found that vowel formant variability differed from vowel to vowel in running speech, but there were certain trends. The variability of formant one is higher for vowels with a high first formant. The variability of second formants is highest for rounded vowels. They suggested that the articulators may undershoot their vowel target in a consonant environment. Lindbloom,⁸⁰ in discussing the undershooting of the articulators in producing a vowel in a consonant environment, relates this to duration. He suggests that with longer vowels there is less tendency for the articulators to miss their target in producing the vowel. Peterson and Lehiste⁸¹ analyzed spectrograms in a study of

⁷⁸G. A. Miller and J. C. R. Lichlider, "The Perception of Speech," Handbook of Experimental Psychology, ed. S. S. Stevens (New York: John Wiley and Sons, 1951), pp. 1040-1074.

⁷⁹Kenneth N. Stevens and Arthur S. House, "Perturbation of Vowel Articulations by Consonantal Context: An Acoustical Study," Journal of Speech and Hearing Research, VI. No. 2 (June, 1963), pp. 111-128.

⁸⁰B. Lindblom, "Spectrographic Study of Vowel Reduction," Journal Acoustical Society of America, XXXV, No. 11 (November, 1963), p. 1885.

⁸¹Kenneth Stevens, "Effect of Duration upon Vowel Identification," Journal Acoustical Society of America, XXXI, No. 1 (January, 1959), p. 109.

the duration of vowels. They concluded that duration of all vowels in English is affected by the nature of the consonants that follow. The vowel is shortest when followed by a voiceless consonant. Therefore, in terms of the last three studies, we may expect the articulators to be more on the vowel target when the vowel precedes a voiced rather than a voiceless consonant.

Stevens⁸² utilized a speech synthesizer to investigate the effect of duration upon vowel identification. He presented four front vowels and four back vowels with durations varying from 20 to 500 milliseconds. The /i / and /ɪ / were only slightly affected by duration when presented in a consonant-vowel-consonant environment. Yet, the /u / and /ʊ / were strongly affected by duration.

Fry⁸³ synthesized these words: object, subject, digest, contract and permit. Each of these words become nouns or verbs depending upon the point of stress. Both noun and verb forms were recorded by adjusting the duration and stress points. These words were then analyzed from spectrograms. The vowel segments showed the major differences in stress with duration being the most effective cue for stress.

⁸²Kenneth Stevens, "Effect of Duration upon Vowel Identification," Journal Acoustical Society of America, XXXI, No. 1 (January, 1959), p. 109.

⁸³D. B. Fry, "Duration and Intensity as Physical Correlates of Linguistic Stress," Journal Acoustical Society of America, XXVII, No. 4 (July, 1955), pp. 765-768.

Broadbent, Ladefoged, and Lawrence⁸⁴ presented a synthesized carrier phrase with three different final words--bet, bit, and bat--to two groups of fifteen listeners. The carrier phrases were made to differ in format frequencies. The first group received the stimulus word with no delay between the carrier phrase and stimulus word, but there was a 10 sec. delay between the carrier phrase and the stimulus words presented to the second group. Fourteen of the first group reported hearing different stimulus words and only seven in the second group reported hearing different stimulus words. They suggest their results indicate that speech patterns immediately preceding the stimulus word influence the perception of the stimulus words.

The intensity relationships between vowels were investigated by Curry.⁸⁵ He found the intensities ranged from lowest to highest as follows: /ɪ /, /i /, /ʊ /, /a /, /u /, /ʌ /, /ɛ /, /ɔ /, and /ə /. The /ə / was recorded at 3.90 dB (SPL) greater than /ɪ /. He suggests that some factor unique to each vowel, in addition to intensity, makes the identification of each vowel possible.

Schwartz⁸⁶ experimented with recognition of the vowel

⁸⁴D. E. Broadbent, Peter Ladefoged, and W. Lawrence, "Vowel Sounds and Perceptual Constancy," Nature, CLXXVIII, No. 4537 (October, 1956), pp. 815-816.

⁸⁵Thayer E. Curry, "An Experimental Study of the Relative Identification Thresholds of Nine American Vowels," Speech Monographs, XVII, No. 1 (December, 1950), pp. 90-94.

⁸⁶Schwartz, loc. cit.

as a function of intensity and time. Five temporal segments: 32, 48, 64, 80 and 96 milliseconds of the vowels: /i /, /u /, /æ /, /ɔ /, and /a / were phonated by two speakers. The mean sound pressure levels required for thresholds of vowel identification varied inversely as function of the duration of the stimulus up to a limit.

Tiffany⁸⁷ presented the /ɜ /, /i /, /æ / and /u / to 18 listeners at four temporal segments: 0.08, 0.2, 0.5 and 8.0 seconds. Four speakers phonated each vowel twice, once with inflection and once without inflection. The inflected vowels were most easily identified. Although he had difficulty recording and standardizing the vowel sounds, he reports the short vowels were more easily identified at shorter durations and the longer vowels were more easily identified at longer durations.

Extensive speed changes in reproducing recorded speech brings about degeneration of vowel intelligibility.⁸⁸ Intelligibility of vowels when the playback speed was decreased was investigated by Tiffany and Bennett.⁸⁹ Knowing that the vowel frequencies would be lowered under this condition, they hoped the individuals with high frequency hearing loss would gain more information from this speech signal. They concluded

⁸⁷Tiffany, loc. cit.

⁸⁸Peterson, Acoustics of Speech, p. 151.

⁸⁹William R. Tiffany and Delmond N. Bennett, "Intelligibility of Slow Played Speech," Journal of Speech and Hearing Research, IV, No. 3 (September, 1961), pp. 248-258.

satisfactory results and recommended further exploration of this procedure.

The duration of homophones was studied by Oyer.⁹⁰ Homophones with the same number of letters tended to be equal in duration. When the spelling of the homophones differed by one or two letters, durations were not as frequently significantly related.

Peterson and Barney⁹¹ recorded 33 men, 28 women, and 15 children phonating the following words: heed, hid, head, had, hod, hawed, hood, who'd, hud, and heard. The total number of recorded stimulus words was 1520. The first formant was plotted as a function of the second formant for men, women and children forming distinct configurations for each group. For each group the configuration was raised proportionately on the graph. This indicated that the relationships between the first and second formants held relatively constant for each group. These stimulus words were presented to seventy adults in a free field manner in an auditorium. Correct recognition of the word was also tabulated as correct recognition of the vowel. The correct vowel recognition scores out of 152 presentations were as follows: /i /, 143, 95 per cent; /ɪ /, 74, 49 per cent; /e /, 52, 34 per cent; /æ /,

⁹⁰Herbert J. Oyer, "Duration of Homophones," Western Speech, XXIII, No. 2 (Spring, 1959), pp. 99-102.

⁹¹Gordon E. Peterson and Harold L. Barney, "Control Methods Used in the Study of Vowels," Journal Acoustical Society of America, XXIV, No. 2 (March, 1952), pp. 175-184.

115, 76 per cent; /ʌ /, 76, 50 per cent; /ɑ /, 9, 6 per cent; /ɔ /, 34, 22 per cent; /ʊ /, 76, 50 per cent, and /u /, 109, 72 per cent. When a vowel was incorrectly recognized, the vowel was usually perceived as the sound adjacent to it on the familiar tongue position graphs. The ease with which the observers classified the various vowels varied greatly.

Fairbanks and Grubb⁹² presented 0.3 sec. temporal segments of nine vowels to eight young graduate students. These vowels were meticulously recorded by seven male professors from the Department of Speech. They were allowed to record each vowel until the speaker was satisfied that his segmented samples were typical of general American production of the vowel. The authors state, "It is plain that the samples were characterized by generally high identifiability." They attained 74 per cent recognition over all vowels, and individual vowels ranged from 53 per cent to 92 per cent correctly recognized. Those most easily identified were /i / and /u /. Those most difficult to identify were /ɪ / and /a /. The vowel identifications tended to be greatest for those vowel temporal segments the speakers preferred. The variability of the formants tended to be less for the speaker's preferred vowel temporal segment.

An interesting experiment with vowel segments sustained for 15 to 20 sec. was conducted by Siegenthaler.⁹³ Recordings

⁹²Fairbanks and Grubb, loc. cit.

⁹³Siegenthaler, loc. cit.

of four male speakers sustaining ten different vowels were presented to fourteen expert and twelve inexperienced listeners. A 128 Hz tone was presented to them, and they were asked to match their phonated pitch to the tone in order to standardize the pitch. The middle section of the tape stimulus was cut from the vowel tape to eliminate the initial and final transients. He found no significant difference between expert and inexperienced listeners who were able to identify correctly only 52 per cent of the sustained vowels. The /i / was most easily identified, and /e / and /u / were most difficult to identify.

Gemelli and Pastori⁹⁴ compared oscillograms of isolated vowels with oscillograms of vowels in context. They discovered, by visual inspection, a minimum isolated vowel configuration of two wave lengths within the vowel in-context oscillograms. Since their research was designed to determine, by visual inspection of the oscillograms, the duration of the isolated vowel sound void of influencing adjacent sound, there is no evidence to support their hypothesis dealing with duration thresholds for aural vowel identification. Although they hypothesized that two wave lengths are required to perceive a vowel aurally, they could have more correctly hypothesized as follows: Two wave lengths is the minimal distance of an isolated vowel configuration identified on the

⁹⁴A. Gemelli and G. Pastori, "La durata minima delle vocali sufficienti alla loro percezione." (The Minimum Duration of Vowels Sufficient to Assure Their Perception." Archivio di Fisiologia, XXXIII, No. 3 (1934), pp. 440-452.

oscillogram of a vocalized vowel produced in context.

A thorough search of the literature revealed only two experimental studies directly concerned with recognition thresholds of vowels as a function of temporal segmentation. Therefore, they will both be discussed in detail. Gray⁹⁵ conducted the only true experiment specifically designed to establish the recognition thresholds of vowels as a function of temporal segmentation in 1939, and it was reported in the literature in 1942. The second study was Peterson's⁹⁶ dissertation that was directed by Gray in 1939. He attempted to determine which segment of a vowel wave form cycle is most easily identified using vowel temporal segments as short as Gray had found to be the minimum temporal segment necessary for vowel recognition.

Gray⁹⁷ utilized three speakers, one male and two female voices, to present eleven vowels to fifteen listeners. The vowel sounds were prolonged by the speaker while a swinging pendulum tripped the on and off switches attaining twelve different duration temporal segments of each vowel ranging from 52 milliseconds to 3 milliseconds. The ingenious switching consisted of a pendulum that swung from a fixed point striking toggles of two switches. The first switch closed the signal circuit when the toggle was struck and the second switch

⁹⁵Gray, loc. cit.

⁹⁶Peterson, "Significance for Recognition of the Vowel."

⁹⁷Gray, loc. cit.

opened the circuit when actuated by the pendulum. The elapsed time between the actuating of the switches determined the length of time the signal was allowed to pass to the loud speaker. The different intervals were attained by lengthening the distance between switches and measuring the signal interval with a timing mechanism. The speakers vocalized the vowel for each stimulus presentation; this may have caused a wide variety in the wave form of the vowels. Also, the listeners were seated in a classroom where the stimulus was reproduced through a loud speaker. But no mention of loudness level or monitoring of the vowel quality was reported in the article. The speakers attempted to match the six different vowel fundamental frequencies to pure tones presented to their ear. This was no verification that the vowel produced actually had the required fundamental frequency. Furthermore, the vowels were not recorded and could not be analyzed at a later date.

Some of Gray's findings that appear pertinent to the present research are as follows:

1. The minimal perception temporal segments were not the same for all English vowel sounds.
2. Individual differences in ability to recognize the temporal segments exist among subjects.
3. When the cycle length was computed from the data, a significant number of stimuli were identified when less than one complete cycle of the fundamental frequency was present. And 0.24 cycle was correctly identified by two listeners, but recognitions at

0.384 and 0.64 cycle were frequent.

4. Recognition of the vowel at extremely short intervals seems to depend upon the proximity of the pitch to the normal medial pitch of the voice; 128 Hz for the male voice and 256 Hz for the female voice.

Peterson⁹⁸ conducted a follow up study to Gray's experiment by investigating the possible significance of various portions of the wave length presented to listeners at the minimum duration necessary for the recognition of vowel sounds. This dissertation was directed by Gray at the Louisiana State University and conducted in two parts.

The ingenious pendulum switching mechanism reported by Gray⁹⁹ was described in great detail by Peterson. This included a description of the use of a dummy impedance approximately equivalent to the loud speaker impedance. This impedance was switched off when the vowel signal was turned on and switched on when the vowel signal was turned off providing a constant impedance to the amplifier. This prevented "...the initial distortion of the sound which would result if the amplifier were switched to the speaker from a load of a different value."¹⁰⁰ A cathode ray oscillograph was added to the system permitting pictorial recording of the

⁹⁸Peterson, loc. cit.

⁹⁹Gray, loc. cit.

¹⁰⁰Peterson, op. cit., p. 32.

vowel temporal segments.

The experiments were conducted in three rooms; one for the male speaker, one for the listeners and switching mechanism, and a dark room housing the oscillograph. An elaborate system of hand and light signals was devised to coordinate the efforts of various people who were employed to actuate the equipment in the three rooms. The vowel signal from the amplifier was sent to the switching mechanism and the required temporal segment was reproduced by the loud speaker. The cathode ray oscillograph received the signal from a point in the circuit just prior to entering the loud speaker.

Peterson attempted to take the vowel signal from the loud speaker by placing the oscillograph microphone in front of the speaker. But as he states, "Several photographs were made using the short sound from the speaker. The resulting curves showed the sounds at the required short intervals, but extremely small additional vibrations distorted the axis, making accurate calculation of the time interval impossible. The source of this disturbance was never determined."¹⁰¹ This disturbing signal was thought to be due to reverberation or building noise.

Eight vowels /i /, /ɪ /, /e /, /æ /, /a /, /ɔ /, /u /, and /ʊ / were prolonged by one male speaker with a fundamental frequency of 96 Hz in the first study. The method employed

¹⁰¹Ibid., p. 41.

to attain the required vowel fundamental was not reported. Each vowel segment was presented randomly in five series. One of the vowel segments was repeated in each series but not measured in an effort to confound listener guessing possibilities. The temporal segment for the first study was 3.59 milliseconds.

In the second study six vowels were reproduced from a prerecorded disc record. The /ɪ / and /ʊ / were deleted from the stimuli. The vowels were then chopped into 3.1 milliseconds segments for random order presentation to fifteen listeners. The /æ / vowel segments were presented eight times and the /ɔ / vowel segments nine times in a random order along with the other four vowels. The /æ / and /ɔ / segments were recognized by the fifteen subjects 51.7 per cent of the time in the second study. The first study yielded 46 per cent recognition of all vowels. Based on the second study evidence Peterson¹⁰² in his dissertation concludes, "Gray found considerable recognition at .003 of a second, giving .24 of a cycle; in the present investigation, over 50 per cent recognition was obtained at an interval of .0031 of a second, giving approximately .298 of a cycle."

In regard to the significance of various portions of the wave length at the required temporal segments, there was no statistical significance and he reported the following: "It is obviously beyond the scope of this study to determine whether segments coming from one general section of the cycle

¹⁰²Ibid., p. 91.

are generally recognized more easily than others."¹⁰³

SUMMARY

The dimensions of the acoustic signals have been investigated by many researchers in an attempt to determine the acoustic parameters which contribute to the function of human auditory speech perception. It has been demonstrated that reduction of intensity or duration may affect recognition of the speech signal, as well as the recognition of pitch. Studies have also indicated that intensity levels required for vowel recognition or perception of pitch varied inversely as a function of the duration of the stimulus. That is to say, for a vowel to be recognized or a pitch to be perceived, it was necessary to increase the intensity when the duration was decreased; and, conversely, it was necessary to increase the duration when the intensity was decreased. However, the stimulus duration has little effect upon perception of auditory signals above 60 dB (SPL). The relationship of the formant frequencies and especially the relationship between the first and second formants contribute appreciably to the recognition of the speech signal. The vowel formant frequencies differ depending upon the duration of the vowel in running speech. This latter finding has been attributed to possible undershooting of the vowel target by the tongue when the vowel is produced in a consonant environment. The vowel duration is affected by the voicing or

¹⁰³Ibid., p. 83

non-voicing of the consonant which follows it; the vowel duration is longer when followed by a voiced consonant. Vowels that are later judged by the speaker to be most representative of the intended vowel are also recognized most easily by other listeners. The vowels that are preferred by the speaker also show less frequency formant dispersion on graphs representing first formants plotted against second formants. One study indicated that vowel recognition is also affected adversely by prolonging the vowel sound for 15 to 20 sec.

Although the phonetic dimensions of a vowel have not been precisely determined, most of the researchers would agree that there is a phonetic dimension for each vowel; and as Peterson¹⁰⁴ states "...when the significant dimensions of this phonetic space have been properly identified, the phonetic value of any vowel can be specified quantitatively in these dimensions." One of the major variables in the perception of speech is that of duration. The vowel recognition threshold, as a function of temporal segmentation, was investigated by two researchers. Both used the same mechanically actuated switching mechanism to take a section from vowels at the desired length. The control of uniform fundamental frequencies, as well as the intensity, were lacking in these experiments. The first study by Gray, established the vowel recognition threshold at 3 milliseconds

¹⁰⁴Peterson, "Phonetic Value of Vowels," p. 543.

for vowels with a fundamental of 128 Hz. The second experiment failed to attain 46 per cent recognition at 3.1 milliseconds when eight vowels were presented to the listeners. A second part to the experiment employing only two of the vowels did yield a 51 per cent recognition threshold when presented vowel temporal segments of 3 milliseconds. Fairbanks¹⁰⁵ referred to this study by stating "Gordon Peterson, in 1939, did a study on the minimum duration required for vowel recognition and found the value to be as low as 5 milliseconds. From these studies we come out with an average representative value of about 5 milliseconds for duration up to a respectable maximum as high as 200 milliseconds. We realize now that the prior guessing probability of the Peterson study is very high so that the minimum value derived from that study has to be reviewed in that sense."

¹⁰⁵Grant Fairbanks, "Experimental Phonetics: Selected Articles," (Urbana and London: University of Illinois Press, 1966).

CHAPTER III

EXPERIMENTAL APPARATUS AND PROCEDURES

The experimental procedures divide into four major phases: 1) recording and validating the basic vowel sounds, 2) chopping the basic vowel signals into short temporal vowel segments and recording these stimuli, 3) randomizing the stimuli, and 4) presenting the randomized vowel temporal segments to the listeners. To lessen the confusion brought about by discussion of the apparatus and techniques when intermingled with the procedures, this chapter is divided into two sections entitled Apparatus and Procedures.

Apparatus:

The following equipment was employed in the four phases of the experimental procedure:

1. Recorder I (Model AG 350-2, Ampex)
2. Recorder II (Model 1022, Magnecord)
3. Oscilloscope I (Model 0-9 Heath Kit)
4. Oscilloscope II (Model 502A Dual-Beam, Tektronix, Inc.)
5. Oscilloscope III (Model 564 with type 3A72 and 2B67 time and sensitivity plug in units, Tektronic, Inc.)
6. Low frequency oscillator (Model 202C, Hulett Packard)
7. Polaroid Land Camera (Model 12C, Tektronix, Inc.)
8. Electronic switch (Model 829E, Grason-Stadler)
9. Digital Counter (Model 373A, Hulett Packard)
10. Audiometer filter (Model 25, Allison Laboratories, Inc.)
11. Power supply (Model 160A, Tektronix, Inc.)
12. Wave form generator (Model 162, Tektronix, Inc.)

13. Two pulse generators (Model 161, Tektronix, Inc.)
14. Magnetic recording tapes (Low pring type 131, Scotch Brand)
15. Sound Treated Room (Model 10-1304, Industrial Acoustics Company, Inc.)
16. Double Wall Sound Treated Room (Model 10-1052, Industrial Acoustics Company, Inc.)
17. Artificial Ear (Model 4152, Bruel and Kjaer)
18. Sound Pressure Level Meter (Model 2203, Bruel and Kjaer)
19. Earphones (Model TDH-39, Telephonics)

Instrumentation for Recording Basic Vowel Loops:

Attaining the basic vowel tape loops was one of the major tasks undertaken. Undistorted, recognizable vowels with fundamental frequencies of 125 Hz were desired. A technique utilizing an oscilloscope Lissajous figure was the most successful. This common electronic technique enables a sinusoidal wave frequency to be measured very accurately by comparing the unknown frequency sinusoidal wave with a known frequency sinusoidal wave. The known frequency sinusoidal wave is applied to either the horizontal or vertical input of an oscilloscope, and the unknown frequency sinusoidal wave is applied to the opposite input. When the frequency source of the known signal is adjusted to match the unknown frequency, a stable Lissajous circumference is observed on the oscilloscope.

When a vowel signal from an amplifier is filtered through a high cut-off filter, it can be adjusted to allow only the fundamental frequency to pass. This filtered signal can then be fed to the oscilloscope horizontal input and treated as the unknown frequency in the above discussion. By adjusting the audio oscillator output to match the fre-

quency emitted from the filter, a stable Lissajous circumference is observed on the screen. The frequency then read from the oscillator would correspond to the vowel fundamental frequency. Therefore, to attain a vowel with a particular fundamental frequency, the oscillator signal is set at the frequency desired and the vocalization is varied to produce the stable Lissajous circumference. To attain a 125 Hz vowel fundamental in this present study, a 125 Hz oscillator signal was applied to the horizontal input of oscilloscope II. The speaker then varied the tone of the vowel; and when the filtered vowel signal--also the fundamental--attained 125 Hz, a stable Lissajous circumference was observed on the screen. When an audio signal also is presented to the speaker through earphones, as the block diagram in Figure 1 indicates, satisfactory results was obtained.

Pictures of the vowel signal and the 125 Hz pure tone were taken with the polaroid camera attached to oscilloscope I. The radical or screen grid brightness was adjusted to its highest point. The camera settings were $f/5.6$ and shutter speed of $1/25$ of a second. The shutter was then actuated to obtain a picture of the grid. Next, the grid brightness was decreased to its lowest point. With the Schmitt trigger placed in normal position, the time base set at 2 milliseconds per cm., and the sensitivity set at 0.5 volt per cm., the trigger level was set to allow one vowel signal to flash on the screen. The camera shutter speed was then changed to position B and the shutter held open while the trigger was

placed in ready position. When the voltage of the signal reached the level of the trigger, one trace of the pure tone wave and vowel wave form was flashed on the screen and recorded on the film over the radical previously exposed. The shutter was then released and the picture egressed from the cameras. The picture was then inspected for possible distortion and correct frequency.

Instrumentation for Segmenting Basic Vowel Signal:

The equipment for segmenting the vowels was selected to accomplish the following:

1. Ability to select any point from zero time to unity of the total vowel signal as the initial or final chopping point.
2. Provision of an electronic switch with minimum switching transients.
3. Ability to inspect visually and measure the wave form of the short vowel segment before recording it.
4. Control of the intensity to provide uniformity of intensities.

The following equipment was selected to satisfy the above requirements:

1. Dual channel storage oscilloscope III (Model 564 with type 3A72 and 2B67 time and sensitivity plug in units, Tektronix, Inc.)
2. Electronic switch (Model 829E, Grason-Stadler)
3. Wave form generator (Model 162, Tektronix, Inc.)
4. Two pulse generators (Model 161, Tektronix, Inc.)
5. Power supply (Model 160A, Tektronix, Inc.)
6. Recorder I (Model AG 350-2, Ampex)
7. Recorder II (Model 1022, Magnecord)
8. Low frequency oscillator (Model 202C, Hulett Packard)
9. Magnetic recording tape, low print (type 131, Scotch Brand)

The Tektronix, 160A, 161 and 162 series instruments are designed specifically to supply pulse signals having adjustable amplitude and duration. For the present study the

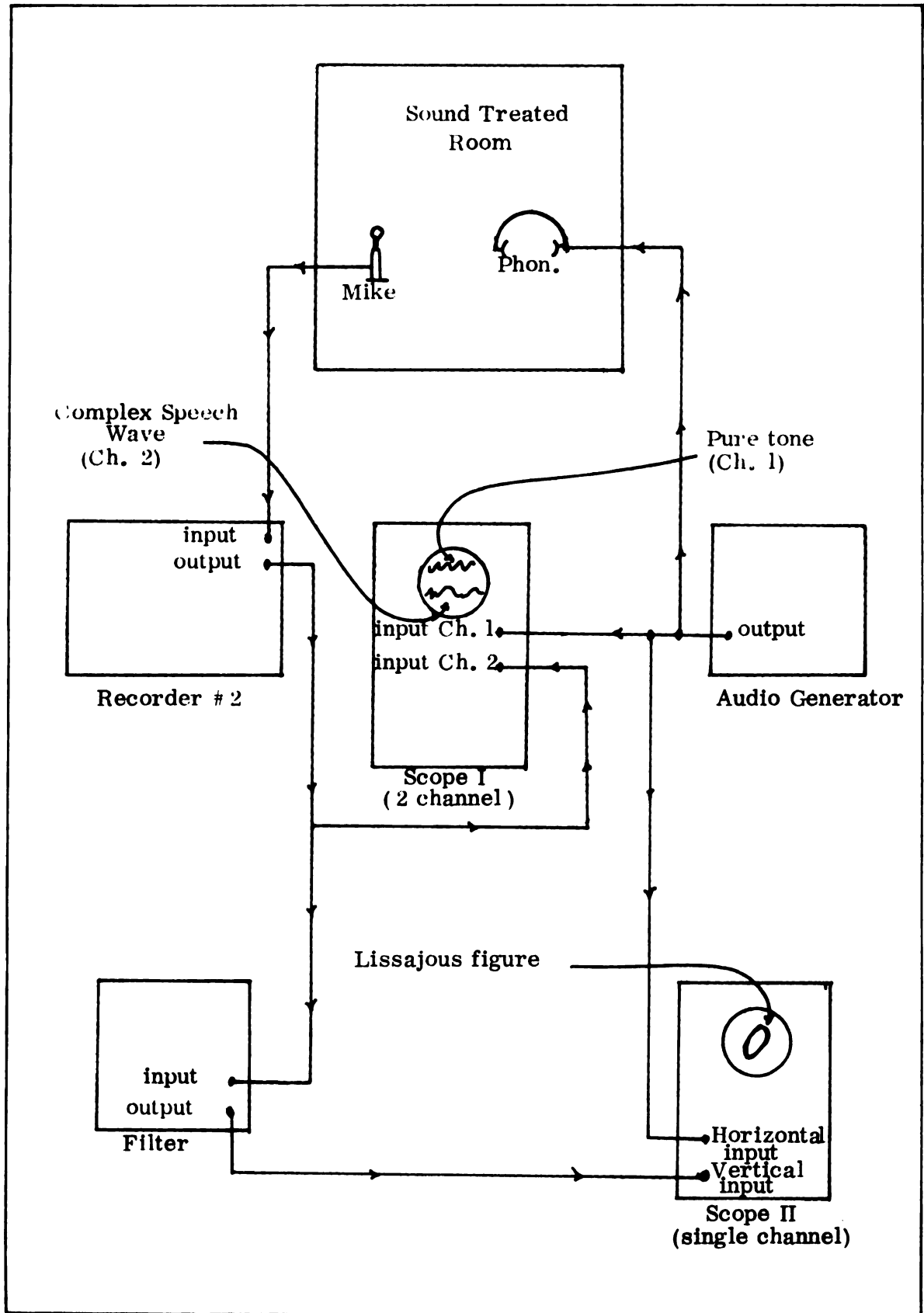


Figure 1. Flow diagram of live vowel recording apparatus.

amplitude and duration of these pulses were not crucial; however, the time elapsed between the pulses from the two pulse generators 161(a) and 161(b) was crucial. It was necessary to adjust the amplitude to 50 volts having a pulse width of 1 millisecond to actuate the electronic switch. The crucial controls for this study were the pulse delay adjustments (Figure 2). These controls actually varied the required trigger voltage needed to produce an output pulse. The wave form generator 162 was capable of producing a sawtooth signal with periods varying from 1 millisecond to 10,000 milliseconds. This negative sawtooth signal had a uniform slope from the beginning to the end of its period. With pulse generator 161(a) delay set at a relatively low trigger level and pulse generator 161(b) delay set at a relatively higher trigger level, the slope of the sawtooth wave could determine the time interval between the output of these two pulse generators. Also, with the period of the sawtooth wave form generator 162 set at 0.5 second, the trigger voltage level of either pulse generator could be adjusted to pulse at any point along the variable voltage of the sawtooth slope. When the trigger level of pulse generator 161(a) was adjusted to actuate its output pulse at the lowest voltage on the sawtooth slope and when the trigger level of pulse generator 161(b) was adjusted to pulse at the highest voltage on the sawtooth slope, the time interval between the two pulses was 0.5 second. By increasing the trigger level of pulse generator 161(b) and decreasing the trigger level

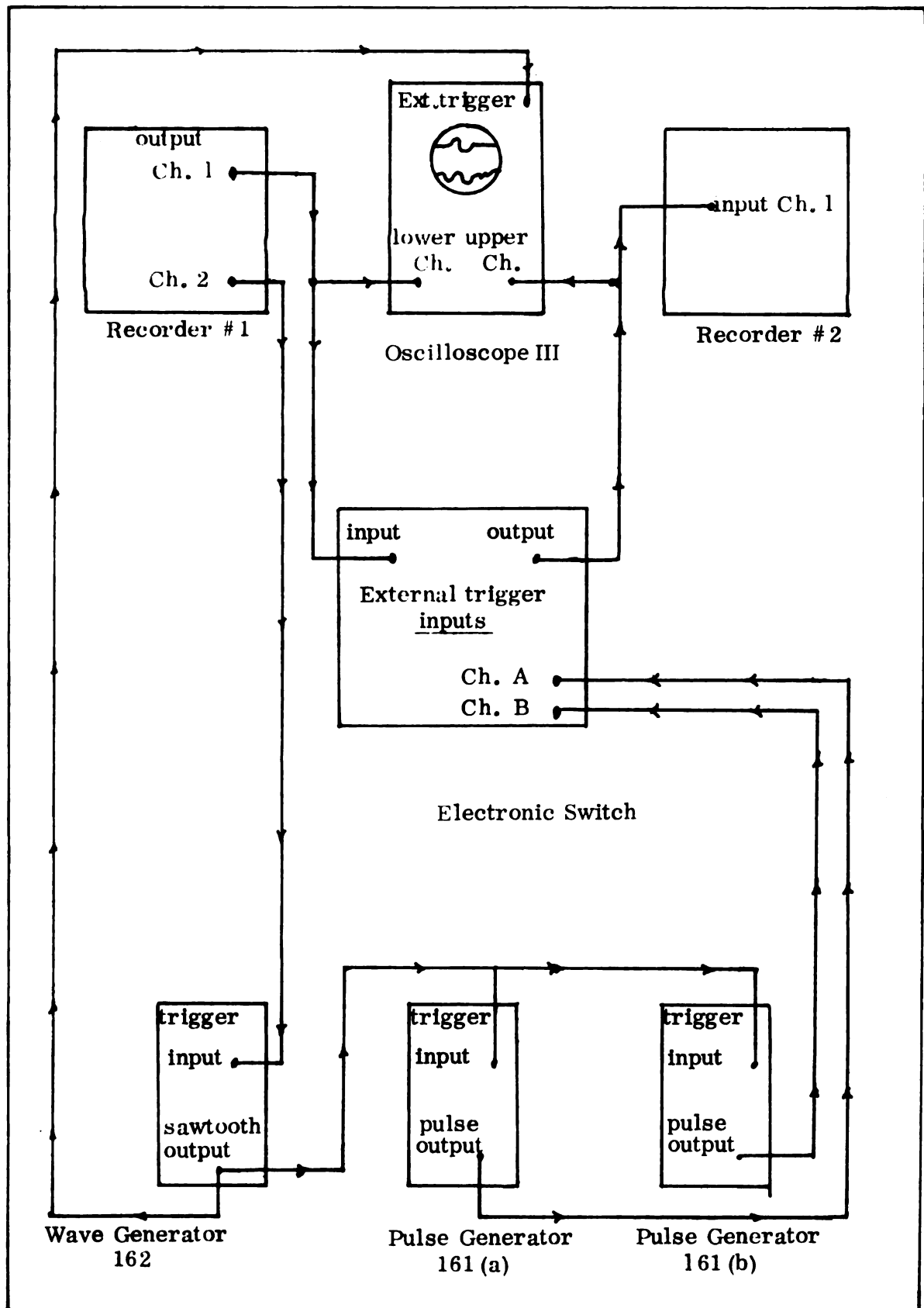


Figure 2. Block diagram of electronic vowel segmenting apparatus. Lower channel is producing the vowel wave form and the upper channel is producing the vowel temporal segment on the scope.

of pulse generator 161(a) the interval could be adjusted from 0.5 second to an infinitely small interval. By adjusting the trigger levels of the pulse generators, this temporal segment could be taken from an infinite number of intervals along the sawtooth slope. The pulse from generator 161(a) closed circuit A of the electronic switch and the pulse from generator 161(b) opened switch circuit A by closing circuit B. The electronic switch was designed so that the two circuits, A and B, were never on or off at the same time. The 50 volts 1 millisecond signal from the pulse generators was sufficient to trigger the switch. The length of time the signal was allowed to pass through the switch was then determined by the temporal interval between the pulses of the two pulse generators. In summary: the wave form generator in turn triggered pulse generators 161(a) triggering switch circuit A on, and pulse generator 161(b) triggering switch circuit A off by closing circuit B, allowing the input signal to the switch to pass through circuit A for the temporal interval determined by the interval between the outputs of the 161(a) and (b) pulse generators.

Trigger Signal Recorded on Basic Vowel Loops:

The wave form generator 162 had a fixed 50 volt trigger voltage that was provided by a signal recorded on channel two of the basic vowel loops. This signal was placed strategically at a point just prior to the initial vowel transient. At first it was thought the vowel wave form could be used to trigger wave generator 162, but the complex vowel

wave form had more than one maximum voltage in the time intervals desired. This created the undesirable multi-triggering within the temporal presentation of the vowel signal. The provision of one signal on channel two of the basic vowel tape loop triggered wave form generator 162 only once per revolution of the tape loop.

This was accomplished in the following manner: a 1000 Hz pure tone from the oscillator was connected to channel one input of the switch. The output of circuit A, channel one, of the switch was connected to channel one of the oscilloscope and also in parallel to the input of recorder I, channel two. The oscillator amplitude was adjusted to 2.8 volts and monitored on the oscilloscope. This voltage was recommended by the switch manufacturer for its most efficient operation. The wave form generator must complete one period before it can be triggered a second time. The period was set to 8 milliseconds allowing sufficient time for the vowel signal to be completed. The vowel signal was sent from recorder I, channel one, to the wave form generator trigger input. The delay controls of the pulse generators were adjusted to allow an 0.5 millisecond pulse to be emitted from circuit A of the switch. The vowel signal from channel one was sent to channel two of the oscilloscope. Therefore, when the tape loop was played, the vowel wave form and the trigger signal could be seen simultaneously. The oscilloscope time base was adjusted so the entire vowel wave form could be seen on the screen.

The first revolution of the tape loop actuated the wave form generator. The pulse of generator 161(a) was delayed causing a pulse during the second revolution of the vowel sound. Pulse generator 161(b) was adjusted to send a pulse to turn circuit A off 0.5 millisecond after 161(a) had turned it on. This 0.5 millisecond signal was recorded on channel two of the tape loop when adjusted to a point just prior to the initial vowel transient.

By utilizing this trigger signal recorded on the second channel of each basic vowel tape loop, it was possible to segment the vowels in the following way: the basic vowel signal on channel one of the tape loop was sent to the switch input and channel two of the oscilloscope. The output of the switch, channel A, was transmitted to channel one of the scope. Therefore, the unsegmented vowel signal from channel one of recorder I and the vowel temporal segment from the switch were monitored on the two channels of the storage scope. The channel two signal from the tape loop triggered the sawtooth generator just before the initial vowel transient. The sawtooth period was adjusted to encompass the desired vowel temporal segment.

It was discovered that the shorter sawtooth period made it easier to make fine adjustments of the pulse generator delay controls. This was due to the abruptness of the sawtooth slope. When the slope was steep, there was a greater difference in voltage during a short segment of time than when the slope was less steep as pictured in Figure 3. It

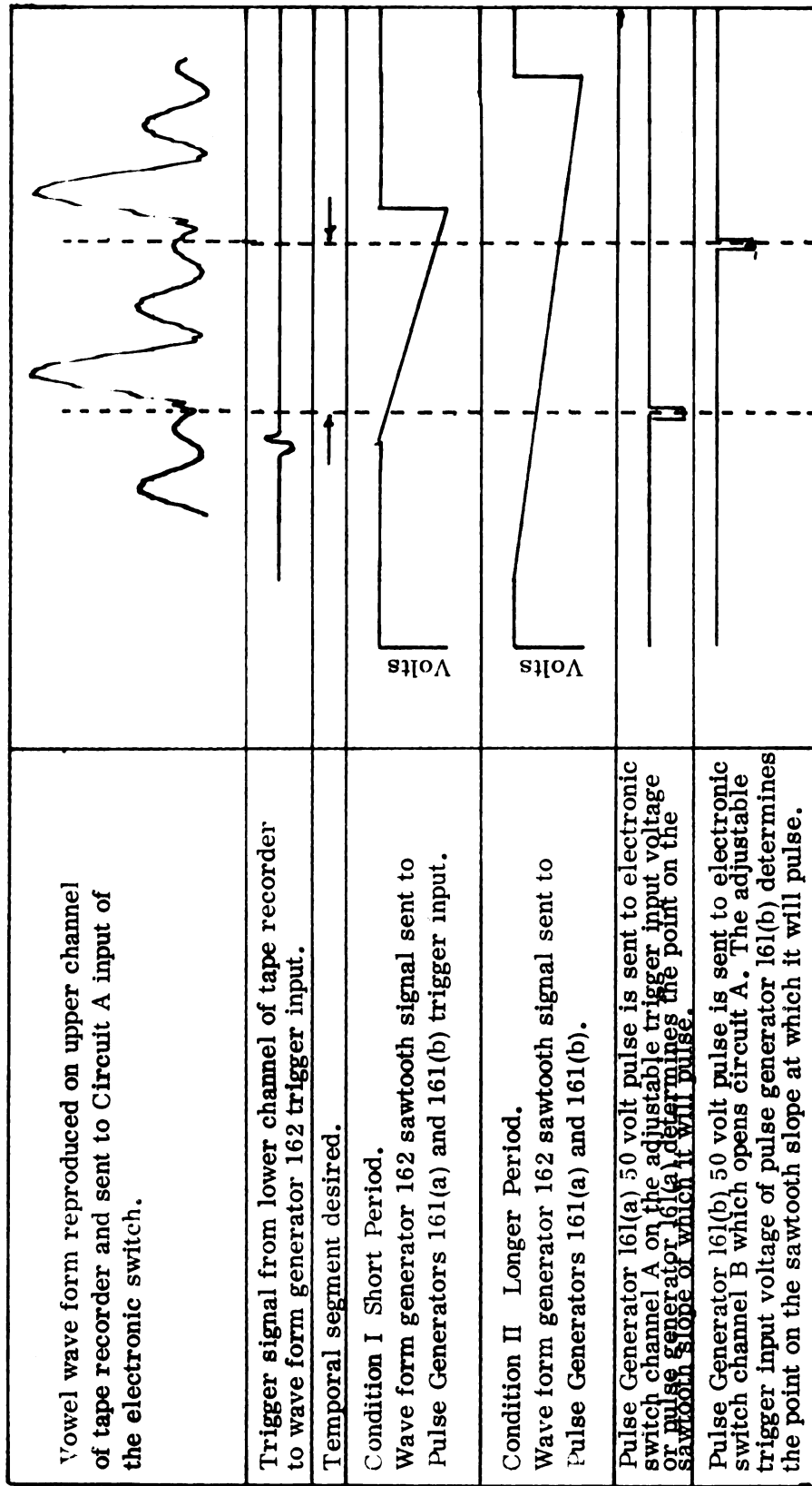


Figure 3. Diagram of the vowel segmenting apparatus electronic function. This pulse generators (PG) 161(a) and 161(b) delay controls adjust the voltage required to actuate their pulse output. In the case illustrated above under condition I the PG 161(a) is adjusted to send a pulse when the sawtooth voltage from wave form generator 162 reaches -10 volts closing the electronic switch. Pulse generator 161(b) is adjusted to pulse when the sawtooth reaches -40 volts, opening the electronic switch. This allows the vowel signal to pass only when the switch is closed. The length of the segment is varied by changing the trigger voltages of the PG 161(a) and PG 161(b). Under condition II the longer period of the sawtooth wave increases the sensitivity of the delay controls.

can be seen that the voltage differences are greater for the short sawtooth period. Thus the delay controls were more sensitive to change as the period became lengthened and less sensitive to change, making it more easily adjusted, as the period was shortened. Yet it was necessary to keep the sawtooth period sufficiently long to include the length of the vowel segment desired. By manipulating the delay controls on the pulse generators, the temporal segment could be varied to the desired temporal length, and the beginning and end of the temporal segments could be adjusted to infinite lengths and points on the vowel wave form.

The storage scope allowed the storage of one sweep across the screen when triggered by the signal emitted from the wave form generator. By placing the scope in storage mode and triggering the scope from the output of the wave form generator, it was possible to inspect visually the output of the switch. This output, of course, was the segmented vowel. The second channel of the scope pictured the output wave form of the basic vowel, and the two signals were compared to insure that distortion had not taken place in the system.

Recording Temporal Segments:

A 0.5 second 1000 Hz tone was recorded on magnetic tape at 5 second intervals. This signal provided a triggering voltage to actuate the digital counter. This second channel signal was also monitored by the experimenter and used to time the interval between the vowel temporal segments.

The low frequency oscillator was set to produce a 1000 Hz tone that entered the electronic switch input channel one. The settings on the switch were as follows: 1) fast decay, 2) period 5000 milliseconds, and 3) per cent A on 10 per cent. This switch is made so that when circuit A is on--circuit B is off. With the above settings the input signal, in this case the 1000 Hz pure tone, was emitted from terminal A for a period of one-half second and then the circuit was opened while circuit B was on for 4.5 seconds. This circuit A signal was recorded on channel two over the entire length of the tape before recording the vowel segments on channel one.

The biased voltage from the record amplifier of recorder II produced a transient clicking sound on the tape when the stop button was actuated. Therefore, it was necessary that the tape be allowed to run and erase the transient click produced by the stop button for the previous recording. Also tape recorder I was stopped immediately after the signal was recorded to prevent it from being recorded a second time. Recorder II was then reversed back to the point beyond the last recorded stop transient click signal and in front of the second channel timing signal. The next temporal segment signal was then recorded repeating the above sequence.

Experimental Procedures:

The psychophysical method of minimal change was selected as the most efficient and accurate means of attaining the vowel recognition threshold as a function of temporal segmentation. The conventional method of minimal change requires

the presented stimuli be arranged in about ten ascending and ten descending series with equal steps between intervals. Guilford¹⁰⁶ recommended that the range should extend from a point where 95 per cent of the responses would be correct at the upper limits and only 5 per cent correct at the lower limits. After considerable deliberations it was decided that approaching the threshold from the upper as well as the lower limits was desirable. "The interdependence of judgments is assured by the serial order of presentation. This helps to stabilize all judgments."¹⁰⁷ However, if the vowel /a / is presented to a subject in a descending series, it is highly probable he might think he recognized the short temporal segment even though, in fact, he could not recognize /a / at this level when presented in isolation. Consequently, a modification of the method of minimal change was needed. Guilford¹⁰⁸ realized the need to modify the method to suit particular circumstances when he stated, "The procedures may have to be modified to some extent to meet certain peculiarities of some sense departments." When establishing a threshold of intensity, the subject is either aware or not aware of the stimulus presented. But in recognition, the task is to identify a particular stimulus as belonging to a definite category. The subject is aware of some presence of

¹⁰⁶ Guilford, op. cit., p. 103.

¹⁰⁷ Ibid., p. 113.

¹⁰⁸ Ibid., p. 111.

the stimulus, but he is required to make a decision as to what particular category it belongs, such as one of the vowels. For example, the subject may be aware of the presence of the sound stimulus but undecided as to whether it is an /i / or an /ɪ / even though he feels positive it is an upper front vowel and even more positive that it is at least a vowel. To alleviate this habituation problem, the vowels within each series could be randomized; but this would not give each stimulus equal opportunity of being presented at each variable temporal segment interval. Furthermore, previous subject knowledge that each vowel would be presented in each series would enhance the chance of guessing. Guilford¹⁰⁹ describes "Haphazard Presentation of Stimuli" as used by Kraepelin in 1891. In this method the stimuli are presented in a haphazard order rather than serial order. The theoretical bases of the method of minimal change seem to be abandoned in this method, and the desire to approach the threshold from either side of the limit is also abandoned. Guilford¹¹⁰ renounced the method by stating, "In both respects the haphazard order of presentation in the method of minimal changes seems inferior to the constant method." Employing this constant method was also considered for the present study. The method of constant stimuli required a limit of four to seven stimuli presented to the subject a large number of times in a pre-

¹⁰⁹Ibid., p. 113.

¹¹⁰Ibid.

arranged order unknown to the subject. The present investigation is concerned with eleven vowels instead of a constant stimulus and would require too many variable stimuli. If each vowel were treated separately to conserve the constant aspect, we would still have the problem as stated earlier: that of the subject possibly thinking he recognized the stimulus at a very short temporal segment because of his pre-knowledge of which vowel was being presented.

To solve these problems, the decision was made to randomize the vowel stimuli within each temporal segment level and present them serially in the method of minimal change because the presentation of each vowel stimulus ten times in descending and ascending series was considered to be important, the eleven vowels were randomized for the ascending series and descending series separately. This yielded a total of 220 stimuli at each temporal segment level. It should be noted that randomizing the 110 stimuli for each temporal segment in this manner also randomized the order of presentation within each series and should minimize the habituation factor in each series as well as curtailing the chances of guessing.

Preparation of the Stimuli:

Initially the entire vowel had to be recorded on tape. Later, segments of these vowels were recorded and randomized without replacement on the master tape to be presented to the listeners. Gray¹¹¹ stated:

¹¹¹Gray, op. cit., p. 87.

At the extremely short interval of .003 second, however, almost fifty percent more recognitions are obtained at 128 c.p.s., which is nearer the normal pitch median of the masculine voice than is either the higher or lower pitch.

Because of this finding, it was decided to record the vowels with fundamentals of 125 Hz.

Equipment:

1. Tape recorder (Model AG 350-2, Ampex)
2. Oscilloscope I
3. Oscilloscope II
4. Low frequency oscillator (Model 202C, Hulett Packard)
5. High and low cut-off filter (Model 25, Allison)
6. Sound treated room X (Model 10, 1304 IAC room)
7. Headphones (Model TDH-39, 300 Ohms telephonics)
8. Microphone (Model 3734A, Electro Vox)
9. Two connection boxes
10. Digital Counter (Model 373A, Hulett Packard)

The following procedure was the most successful in attaining recorded vowels at the required 125 Hz fundamental and consistent wave forms from one cycle to the next. The speaker was a faculty member that has special interests in the area of voice production. He has many years of experience in teaching phonetics and also considerable experience singing in choirs.

The Lissajous technique for matching two pure tones was modified as described earlier under experimental apparatus and utilized to help the speaker attain the desired vowel tone. This technique utilized the formation of a Lissajous circumference produced on oscilloscope II when the desired vowel fundamental was vocalized. By transmitting the speaker's vocalized vowel signal through a low pass filter, the fundamental was allowed to pass. This fundamental was

then transmitted to the vertical input of scope II and a 125 Hz tone from the audio generated was applied to the horizontal input. When the vowel fundamental matched the 125 Hz pure tone, a stable Lissajous circumference was observed on the scope screen. A block diagram of this apparatus is presented in Figure 1. The speaker listened to the 125 Hz pure tone through an earphone and matched his vocalized vowel to this tone. In addition, he monitored the Lissajous pattern on scope number two. The vowel signal was then picked up by the microphone and entered channel one input of tape recorder I. This signal was continually recorded on a tape loop and simultaneously reproduced the vowel signal. This signal entered the filter set to allow the frequencies from 0 to 130 to pass. The filtered signal was itinerated to the horizontal input of scope II. The 125 Hz pure tone being sent to the speaker's headset also traveled to the vertical input of scope II. The speaker then varied the vowel tone until its fundamental was 125 Hz, producing a stable Lissajous pattern on the scope screen. The speaker would say the vowel several times until he felt confident he could produce the required vowel tone. The speaker would then signal he was ready to record and remained silent for one revolution of the tape loop. At that time the experimenter signaled to him to produce the vowel. When the vowel had been vocalized, the recorder was stopped to avoid erasing the signal. This vowel signal was then played back immediately through the same circuit. If the frequency was correct at 125 Hz, a Lissajous

circumference was observed on scope II. If the circle was not evident the vowel was rerecorded. When the circumference was steady the vowel signal was sent to channel one of scope number one along with the pure tone to channel two.

A picture was then taken of the signals on scope I. At this time the frequency was determined by reading the distance from one point on the wave form to the next point where the wave began to repeat itself. This constituted one cycle and needed to be 8 milliseconds to satisfy the requirements of a 125 Hz fundamental frequency vowel. It was necessary to listen to the sound, in order to assure that it sounded like the intended vowel. If the cycle deviated from 8 milliseconds by more than one-tenth of a millisecond or the sound did not sound like a good representation of the intended vowel, it was discarded and the same process repeated. Each of the eleven vowels, /i /, /ɪ /, /e /, /ɛ /, /æ /, /ʌ /, /ɑ /, /ɔ /, /o /, /ʊ /, and /u /, were recorded in the above manner on tape and in picture form as presented in Figures 4 and 5.

Validation of the Recorded Vowels:

At the inception of this study it became apparent that verification of each vowel was essential; that is, each vowel to be segmented must be recognizable by the listeners at least 95 per cent of the time. Each of the eleven vowels was recorded ten times in random order to be presented to eight listeners through earphones. The vowel tape loops were reproduced by recorder I and recorded on recorder II. The output of recorder I has 600 ohm impedance and recorder II

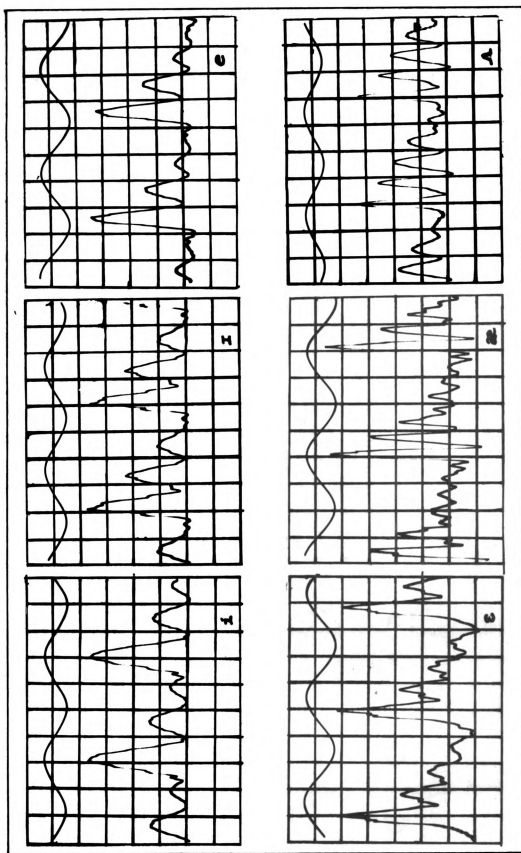


Figure 4. Vowel wave forms taken from oscilloscope during live vowel^o recording. Each division equals two milliseconds.

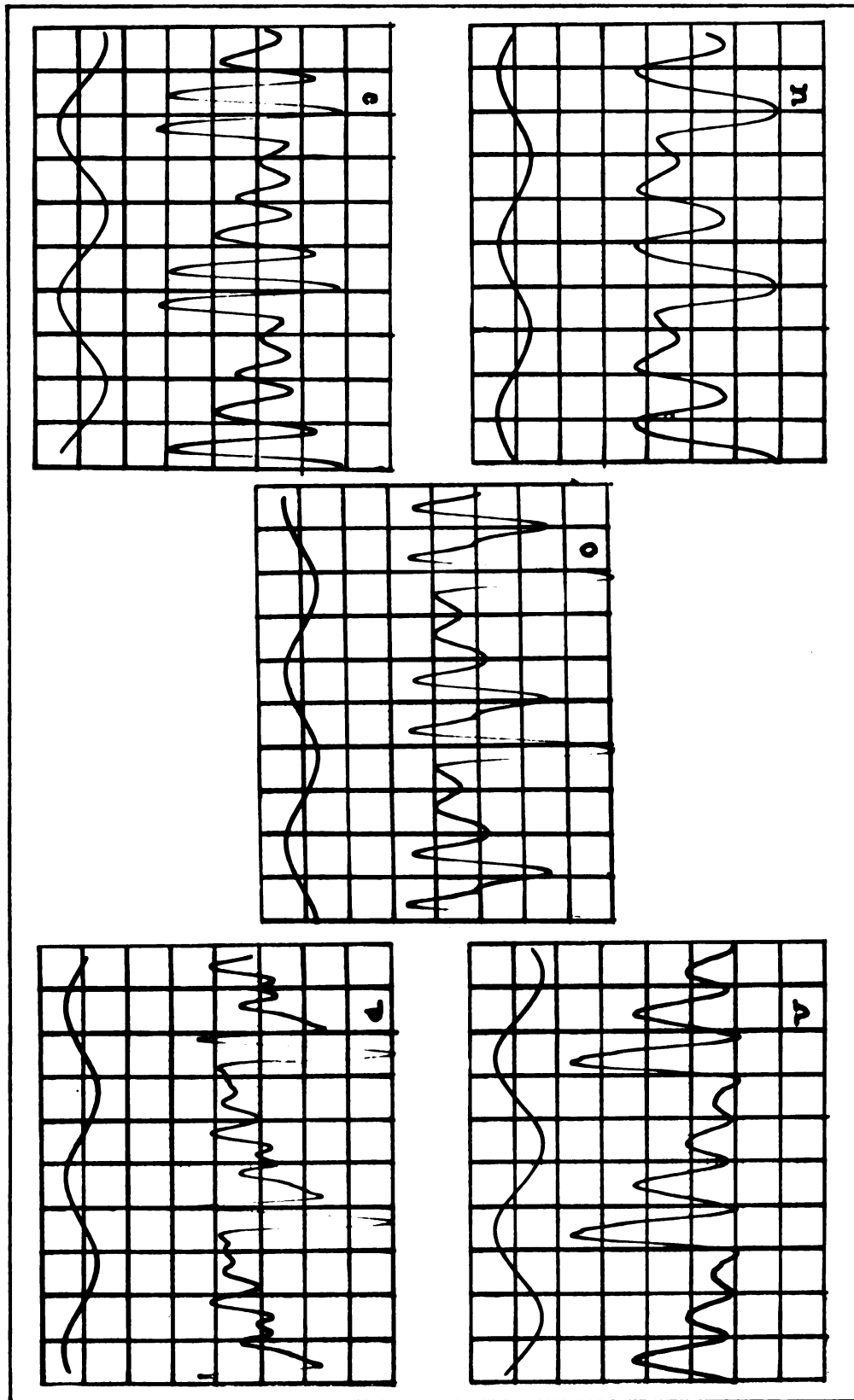


Figure 5. Vowel wave forms taken from oscilloscope during live vowel recording. Each division equals two milliseconds.

has an infinite impedance of 150,000 ohms and therefore no impedance matching devices were required. The vowel tape loop was reproduced on recorder I with the output V U meter peaking at zero. The record gain control of recorder II was adjusted so the V U meter peaked at zero. Channel two of recorder II was monitored and a vowel was recorded on channel one just after the channel two signal passed the playback head.

The eleven vowels were each recorded ten times in random order for presentation to eight listeners. This tape was presented to the listener group through earphones. The listeners were asked to write the phonetic symbol that represented the sound they heard. The subjects were doctoral students in speech and hearing science and all had normal hearing--no hearing loss greater than 25 dB (ISO) on a conventional audiogram. All stimuli were presented at 70 dB (SPL). The pure tone on channel two was used to trigger the digital counter and placed in the visual field of the listener. This counter was modified by placing a diode in the negative line so it was actuated only on a positive pulse. A capacitor was also placed across the terminals to curtail a power build up caused by the diode. The numbers on the counter corresponded to the numbers on the listener response sheets and helped the listener to keep track of stimuli order.

All the vowels except the / Δ / and /a/ reached a 95 per cent recognition level in the first listening session. They were rerecorded in the manner stated above and /a/ was

accepted under the criteria set forth, but /ʌ / was only recognized 90 per cent of the time by the eight listeners. The third attempt was successful and the following recognition scores were attained: /ɪ / 100%, /i / 97%, /e / 100%, /ɛ / 98%, /æ / 100%, /a / 99%, /ʌ / 100%, /ɔ / 100%, /o / 100%, /ʊ / 99% and /u / 100%. It should be pointed out that both /ɪ / and /ɛ / received 100% recognition scores during the first two listening presentations.

Estimating the Probable Recognition Thresholds:

It was necessary to determine the range of the temporal segment values to be presented to the listeners in each series. To estimate the probable recognition thresholds, five individuals were asked to manipulate the delay control on pulse generator 161(b) to produce a vowel temporal segment ranging from 2 milliseconds up to a point where they could recognize it. The eleven basic vowel tape loops were placed on recorder I and reproduced through the apparatus as described earlier. The listeners had no previous knowledge of which vowel they would hear. The cardinal vowels /ɪ /, /e /, /æ /, /o /, and /u / were each presented to the listeners four times. When they had increased the temporal segment by gradually increasing the delay control to a point where they recognized the vowel correctly, the length of time was read from the oscilloscope and tabulated. The mean length of the recognized temporal segments are presented in Table 1 for each of the five listeners.

TABLE 1
VOWEL TEMPORAL SEGMENTS
IN MILLISECONDS

Listener	i	e	a	o	u
1	150	275	100	100	100
2	25	300	75	125	75
3	80	380	225	75	125
4	50	275	75	80	75
5	75	125	75	50	25
	Σ 380	1355	550	330	400
	$\bar{X}$ 76	271	110	66	80

The results of the foregoing attempt to establish the range disagreed a great deal with the findings of Peterson¹¹² and Gray.¹¹³ Because of this, a pilot study was initiated. The five cardinal vowels were chopped into 21 different temporal segments ranging from 10 milliseconds to 210 milliseconds. These temporal segments were recorded with equal intensities in ascending series, and each vowel was presented in ten ascending series. The vowel series were randomized without replacement.

Eight doctoral students listened individually in sound room I to ten ascending series of each vowel presented to them through earphones. Each listener was allowed to listen to one series before the actual study began to reduce the novelty of the experience. The vowel temporal segments were

¹¹²Peterson, "Unpublished Ph.D. Dissertation."

¹¹³Gray, op. cit., pp. 75-90.

presented at 70 dB (SPL) and reproduced by recorder I. The 70 dB (SPL) was determined by adjusting the recorder output applied to the earphones placed on the artificial ear and read from the SPL meter prior to each listening session. The results of this pilot study are presented in Table 2.

Visual inspection of the raw data indicated that once the listener decided upon a particular incorrect response to the vowel stimuli at a short temporal segment, he very seldom changed to the correct response even when presented at longer

TABLE 2
RECOGNITION THRESHOLDS IN MILLISECONDS
ATTAINED IN PILOT STUDY I

Listeners	i	e	a	o	u
1	40	200	180	30	210*
2	40	40	90	20	30
3	40	210*	210*	20	20
4	40	210*	106	30	30
5	40	150	20	30	30
6	40	210*	20	30	210*
7	40	40	30	30	30
8	30	100	50	210*	210*
	Σ 310	1160	776	400	770
Mean	38	145	92	50	91

* was not recognized at longest temporal segment

segments. The listeners expressed the idea that they could not seem to break away from their original response, whether it was right or wrong, even though they had been asked to listen to each stimulus without a previous decision. It

became obvious that there was a habituation effect operating and another pilot study was initiated.

Gray¹¹⁴ presented his subjects with the vowels randomized (without replacement) with long temporal segments in the first series, and he diminished the temporal segment from series to series. He reported 50 per cent recognition for two of his subjects at the 3 milliseconds level.

It was decided to employ this technique in the second pilot study. The range decided upon was from 2 milliseconds to 64 milliseconds; i.e., 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64 milliseconds. All of the basic vowel sounds were segmented into the above temporal segments and recorded in random order (without replacement) for each series. The first series contained vowel temporal segments of 64 milliseconds; each succeeding series was decreased by 8 milliseconds with the final series being 2 milliseconds. There were ten series of eleven vowels each yielding a total of 110 stimuli. The same listeners individually heard the stimuli in the sound treated room I through earphones at 70 dB (SPL) reproduced by recorder I. Again the tape was calibrated with the artificial ear and sound pressure level meter. The results of this study are recorded in Table 3.

¹¹⁴Ibid.

TABLE 3
 RECOGNITION THRESHOLDS IN MILLISECONDS
 ATTAINED IN PILOT STUDY II

Listeners	i	ɪ	e	ɛ	æ	A	a	ɔ	o	ʊ	u
1	16	32	56	16	8	8	40	8	4	8	8
2	16	24	32	4	8	8	16	8	4	4	8
3	16	24	40	40	40	24	56	24	8	16	24
4	24	16	64*	24	32	16	32	40	8	16	32
5	16	8	64*	24	48	16	16	16	16	8	32
6	16	4	56	8	16	16	4	64	16	24	8
7	8	24	64*	24	24	48	2	56	4	16	16
Σ	112	132	376	140	160	136	166	216	52	92	128
$\bar{X}$	16	18	54	20	23	19	24	31	7	12	18

* was not recognized at longest temporal segment

It should be noted that in contrast to the first pilot study only the /e / temporal segment was not recognized by three of the seven listeners at the longest temporal segment of 64 milliseconds. This vowel had been validated as a good representation of the /e / in the validation study cited earlier. Visual inspection of the raw data of both pilot studies indicated the /e / was mistaken for the /ɪ / in almost every instance. It was misunderstood as the /ɛ / only a few times. Therefore, the English speaking Americans, who use this sound rarely, seem to identify it as the /ɪ / vowel. The /ɪ /, of course, makes up a portion of the /eɪ / in the English speaking American pronunciation. The /eɪ / diph-

thong is common in America¹¹⁵ and as Kenyon states,¹¹⁶

"When the sound e loses its accent it is regularly reduced to i, as in daily, deli--- Monday, mandi..." In a foot note Kenyon¹¹⁷ states:

The relative nearness of e to i and of o to u, also confirmed by Parmenter and Trevino (Vowel Positions as Shown by X-rays, Quarterly Journal of Speech, June, 1932) has important bearing on the historical development of these sounds. For example, M.E. i when lengthened often became e:...

Therefore, it does not seem strange for the listeners to perceive the short /e / vowel as an /i /. Peterson,¹¹⁸ alleviated this problem by not presenting the /e /, /o /, /ʌ /, or /ə / which could be mis-identified as /i /, /u /, /a /, or /ə / respectively because of short temporal segment presentations. In the second pilot study, thresholds were established for the /a /, /ʌ /, and /o / stimuli. The /a / and /ʌ / were confused and this did cause the /ʌ / to have a higher estimated recognition threshold than the remaining nine vowels, excluding of course, the /e / stimuli that three subjects out of the seven failed to recognize at 64 milliseconds. On the other hand /o / was mis-identified very few times and, in fact, the estimated recognition threshold of 7 milliseconds was by far the shortest of the eleven vowels presented. Because of the above evidence, it was decided to

¹¹⁵Kenyon, op. cit., pp. 171 174.

¹¹⁶Ibid., p. 173.

¹¹⁷Ibid., p. 60.

¹¹⁸Peterson, loc. cit.

keep all eleven vowels in the final stimuli with the expectation that their recognition threshold may not be attained because the /o/, /e/, /a/ and /A/ may be confused with /v/, /i/, /A/ and /a/ or /e/ respectively. By retaining these vowels, the probability of guessing was deleted and the results might be enlightening. From the information gathered in the pilot studies, it was decided to present the stimuli in series of fifteen steps ranging from 4 to 60 milliseconds in 4 milliseconds steps.

Preparation of Master Stimuli Tape:

A chart was first prepared for the randomization (without replacement) of the stimuli. Fifteen horizontal rows were labeled from 60 milliseconds down to 4 milliseconds in 4 milliseconds steps. Across the top, every other column was labeled descending, and the columns between were labeled ascending. There were 110 descending and 110 ascending columns making a total of 220, and 15 rows yielding a grand total of 3,300 units. Each vowel was then placed on five cards making a total of 55 cards. These cards were shuffled well and then as each vowel card was drawn the vowel was transcribed in the descending square at the 60 milliseconds level. This was done for each of the 15 temporal segment levels down to 4 milliseconds, with cards being shuffled between each temporal segment level. This same procedure was carried out in randomizing (without replacement) the ascending series, and then the whole process was repeated for the second half of the presentations.

Each of the eleven basic vowel signals was pre-recorded on loops reproduced on recorder I. The second channel signal triggered the wave form generator, which in turn triggered pulse generator 161(a) turning circuit A of the switch on and pulse generator 161(b) turning circuit A of the switch off. The segmented vowel wave form emitted from circuit A was sent to channel one of the storage scope and channel two received the whole vowel signal for the purpose of comparing and inspecting segmented sections for possible distortions. The intensity was also adjusted to provide a 2.8 peak to peak voltage for each stimulus recorded. The segmented vowel signal was also sent to the input of channel one of recorder II. Channel two of each tape had been pre-recorded with the signal that triggered the digital counter as described earlier.

Each vowel was recorded at each desired temporal segment in a known order: /i /, /ɪ /, /e /, /ɛ /, /æ /, /a /, /ɔ /, /o /, /ʊ /, /u /. Each temporal segment was inspected very closely for possible distortion both by observing the wave form on the oscilloscope and monitoring through earphones. The temporal segments were next rerecorded on tape loops. These loops were hung on a rack and labeled.

The master tapes to be presented to the subjects were prepared by playing the loops on recorder I and recording them in the prearranged randomized (without replacement) order on eight low print tapes. These tapes were then played on tape recorder II and reproduced through the head sets to

the listeners seated in the double walled sound room II.

The six listeners were doctoral students in speech and hearing science with no hearing loss greater than 25 dB (ISO) on conventional audiograms. The other two listeners used in the pilot studies had left the university and were not available for the final experiment. All listeners had received equal opportunity in listening to vowel temporal segments, because they were members of the listening group utilized in the pilot studies. Four series, two ascending and two descending, were presented to diminish the novelty of the experience. These series were not scored. The electronic digital counter was again used to facilitate the listeners keeping track of the stimuli. Each tape had 32 minutes and 5 seconds of recorded vowel temporal segments, totaling to 4 hours, 8 minutes and 20 seconds of listening time.

The following instructions were read to the listeners just prior to the onset of their listening task:

The purpose of this experiment is to determine the smallest vowel temporal segment required to recognize it correctly. The vowel segments will be presented to you in five second intervals through binaural earphones at a uniform intensity. Each series of fifteen stimuli will become progressively shorter in the descending series and progressively longer in the ascending series. You will be required to listen four half-hour periods in the morning and four half-hour periods in the afternoon with rest periods between each half hour period.

Your task is to listen to the stimuli and write the international phonetic symbol that represents the sound you hear. It is important that you listen very carefully as the sounds become progressively shorter. The vowels will be presented in random order throughout the entire experiment. This means a particular vowel may be duplicated or even triplicated in a given series. It is also possible that a specific vowel may not occur at all in a series. Very probably some

vowels will be recognized at shorter temporal segments than others. Therefore, if you do not recognize a short temporal segment, do not feel that it will be impossible to recognize the next shorter stimulus that may be a vowel with a lower recognition threshold. Therefore, it is essential that each stimulus be scrutinized independently of the preceding stimuli. The electronic digital counter will indicate each stimulus presentation, and the digits will correspond with the numbers on the response sheet.

To familiarize you with the task at hand you will first be presented two ascending and two descending series to delete the novelty of the experience. You may ask questions to clarify the proceedings now or after the first four series.

Plans were made in advance to take rest periods between each tape, but at the end of the first tape the listeners complained about the time. A rest was given after two-thirds of each tape was presented at the listeners request. With break periods ranging from fifteen minutes to one-half hour between tapes and a hour off for lunch between tape 4 and 5. At the end of tape 6, the listeners complained that they were too tired and were not responding as well to the stimuli. It was noticed by the experimenter that during the presentations of tapes 5 and 6, the listeners looked fatigued; but, at the same time, they seemed to be working harder at listening and identifying the stimuli. The decision was made to continue the experiment but allow two rest periods for each of the last two tapes with a break in between each tape. The effect of fatigue is, of course, increased with the lack of sufficient rest periods.¹¹⁹ With the listeners as a group determining the number and length of rest periods,

¹¹⁹Guilford, loc. cit.

the entire 4 hours, 8 minutes and 20 seconds were listened to in one day. Including rest periods, the experiment began at 9:10 a.m. and finished at 4:55 p.m., a total of 7 hours and 45 minutes. Subtracting the listening periods, the subjects rested 3 hours, 36 minutes, and 40 seconds.

CHAPTER IV

RESULTS AND DISCUSSION

The vowel temporal segments were prepared and recorded on a master tape as described in Chapter III. This master stimulus tape was reproduced through earphones to six doctoral students seated in a sound treated room. The vowel temporal segments were presented in 110 ascending and 110 descending series. The subjects knew the vowel temporal segments would become progressively shorter and then progressively longer. The longest vowel temporal segment was 60 milliseconds and the shortest was 4 milliseconds. The vowels were presented to the subjects in random order within each temporal segment value. The ascending and descending series stimuli were independently randomized without replacement. Therefore, the listeners were not aware of the vowel presentation order. The subjects recorded their responses using the International Phonetic Alphabet. Each subject recorded his responses on eight listener response sheets--one sheet for each one half hour tape. A total of 3,300 listener responses to the eleven vowels, each segmented into fifteen different temporal segments, was presented in 110 descending and 110 ascending series. For the six listeners there was a total of 19,800 responses.

The first task was that of recording the listener's re-

sponses on a tabulation sheet for each vowel. An example of one of these tabulation sheets with ten ascending and ten descending series is presented in Table 4. Because the vowels were randomized within each of the fifteen segment temporal levels, it was necessary to rearrange these into an order, being certain that each response retained its ascending or descending order. This procedure for rearranging the listener responses for a particular vowel also preserved the time-order sequence. Therefore, the first response at a particular segment temporal level represented the first presentation of this particular vowel within this level. Continuing in this way, each succeeding appearance of a particular vowel stimulus within that segment temporal level was preserved. A correct response was designated by a plus sign, and the incorrect response was designated by tabulating the incorrect phonetic symbol chosen by the listener. When the listener did not respond to a stimulus, a minus sign was tabulated in this space. Therefore, the responses to the 300 stimuli for each vowel were represented in these tables.

The threshold values for each series were arrived at by the procedure recommended by Guilford.¹²⁰ The sequence of correct listener responses in a particular series was not always continuous as illustrated in Table 4. That is, any given series may have three correct identifications, followed by three misidentifications, then two correct identifications,

¹²⁰ Guilford, op. cit., p. 31.

TABLE 4

EXAMPLE OF TABULATING SHEET FOR ORDERING THE LISTENERS RESPONSES
 MAINTAINING THE SERIES ORDER AND SERIAL ORDER
 FIRST HALF AND SECOND HALF

Milliseconds	VOWEL / s /		LISTENER X											
			Descending						Ascending					
			First Half			Second Half			First Half			Second Half		
60	+	+	+	+	+	+	+	+	+	+	+	+	+	+
56	+	+	+	+	+	+	+	+	+	+	+	+	+	+
52	+	+	+	+	+	+	+	+	+	+	+	+	+	+
48	+	+	+	+	+	+	i*	+	+	+	+	+	+	+
44	+	+	+	+	+	+	+	+	+	+	+	+	+	+
40	+	+	+	+	+	+	+	+	+	+	+	+	+	+
36	+	+	+	+	+	+	+	+	+	+	+	+	+	+
32	+	+	+	o*	+	+	æ*	+	+	+	o*	+	+	æ*
28	+	+	+	+	+	+	+	+	+	+	+	+	+	e*
24	+	+	+	+	+	+	+	e*	+	+	+	+	+	+
20	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16	i*	+	+	+	+	+	+	+	+	+	+	+	+	+
12	+	+	+	+	+	+	i*	+	+	+	+	+	+	+
8	-	-	-	-	-	-	i*	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Series Threshold	14	10	14	10	10	10	18	14	14	18	10	10	18	18

Underline refers to series threshold computed as described in text. Mean descending 12.4,
 Mean ascending 14.4, Mean first half 12.4, Mean second half 14.4, Mean over all 13.4 and
 mis-identification recorded.

e.g. + + + - - - + + - - - - - -. The threshold for a series like this was determined by moving the correct responses toward the highest stimulus value and moving the incorrect responses toward the smallest stimulus values. In this manner, the correct responses were separated from the incorrect responses, and the limen was established at a mid-point intercalating the correct and incorrect responses. When the responses were continuous and changed abruptly from one mode to the next, the limen was established in the conventional manner by selecting the mid-point between the correct and incorrect responses. It was these scores for each series that were used in computing the statistics in this study, and are presented in Appendix 1. The mean of these series' scores was designated the vowel temporal threshold for this individual.

Guilford¹²¹ recommends a study of the homogeneity of series thresholds by analysis of variance with the two series orders and two time blocks--first half and second half--resulting in a 2 x 2 factorial design. These are the series order errors related to habituation and expectation effects provoked by descending and ascending series respectively, and the time error, related to learning and fatigue. In addition to these factors, the present study used six listeners instead of one; and the stimuli consisted of ten vowels, instead of a constant stimulus, such as frequency or

¹²¹Guilford, loc. cit.

intensity. Therefore, the data were subjected to a factorial design ($2 \times 2 \times 6 \times 10$) analysis of variance. A special program was developed for the CDC 3600 Computer by Dr. Lashbrook of the Speech Communication Research Laboratory at Michigan State University. In addition to the error factors between the four main effects, a fifth error factor was injected into the formula to account for errors due to replication. The five ascending series scores and five descending scores in the second half of the presentations were taken as the replication of corresponding scores in the first half of the study. The results of this analysis are found in Table 5.

The results of this analysis showed significant interaction beyond the 0.05 level of significance among subject, vowel and series; and subject, vowel and time order. Interaction between subject and vowel; subject and series; subject and time; vowel and time order; and vowel and series was also indicated beyond the 0.05 level of significance.

The following interactions were not significant at the 0.05 level: subject and time; subject and series; series and time; subject, series and time; vowel, series, and time; and subject, vowel, series, and time.

The main effects of subject, vowel, and series order were each significant beyond the 0.05 level, but the main effect of time--first half and second half--were not significantly different at the 0.05 level. However, the above stated interaction confounded the meaning which can be

TABLE 5
ANALYSIS OF VARIANCE TABLE

SOURCE OF VARIANCE	SUM OF SQUARES	DF	MEAN SQUARE	F STATISTIC	APPROX. SIGN. PROBABILITY OF F STAT.
SUBJECT (S)	4467.94666660	5	893.58933333	27.87385*	0.0005
Z=R+S _{IR} **	769.39999999	24	32.05833333		
VOWEL (V)	34381.946666672	9	3820.21629632	119.52356*	0.000
SxV	46877.853333347	45	1041.73007408	32.59273*	0.0005
Z=SxV _{IR} +V _{IR}	6903.80000007	216	31.96203704		
SERIES ORDER (O)	211.680000000	1	211.680000000	8.37176*	0.008
SxO	285.880000000	5	57.176000000	2.26126	0.081
Z=SxO _{IR} +O _{IR}	606.839999999	24	25.285000000		
TIME ORDER (T)	117.813333333	1	117.813333333	3.42497	0.077
SxT	547.426666666	5	109.485333333	3.18287*	0.024
Z=SxT _{IR} +T _{IR}	825.560000000	24	34.398333333		
SxVxO	1814.053333331	45	40.31229630	1.42285*	0.042
SxVxT	3183.973333330	45	70.75496296	2.49735*	0.0005
SxOxT	134.720000000	5	26.944000000	0.95101	0.449
VxO	3610.386666666	9	401.15407407	14.15904*	0.0005
VxT	2435.986666668	9	270.66518518	9.55333*	0.0005
OxT	5.880000000	1	5.880000000	0.20754	0.653
Z=SxVxO _{IR} +SxVxT _{IR} +SxOxT _{IR} + VxO _{IR} +VxT _{IR} +OxT _{IR}	12919.399999986	456	28.33201754		
SxVxOxT	967.346666666	45	21.49659259	1.02351	0.440
VxOxT	124.453333333	9	13.82814815	0.65840	0.746
Z=SxVxOxT _{IR} +VxOxT _{IR}	4536.599999990	216	21.00277778		
NOT ACCOUNTED FOR	0.000000000	0			
TOTAL	125728.946666290	1199			

*Significant beyond the 0.05 level. **Replication (R).

attached to the significance of the main effects.

Bartlett's¹²² test for homogeneity of variances was applied. Homogeneity of variances is one of the assumptions for the analysis of variance procedures. The test resulted in an $F = 1.11$ with $V_1 = 9$ and $V_2 = 1718.75$ degrees of freedom. An F equal to or greater than 1.88 is needed to reject the hypothesis of homogeneity of variance with $V_1 = 9$ and V_2 greater than 120 at the 0.05 level of significance.¹²³ Since the $F = 1.11$ is smaller, the assumption of homogeneous variance is accepted for the experimental data. Thus, the significant differences and interactions are due more likely to mean differences rather than variance differences.

Next, graphs were plotted to display graphically the interactions among subject, vowel and series order. The temporal segment values are located on the vertical axis and the vowels are located on the horizontal axis. The mean of the series temporal segment recognition thresholds for each of the six subjects was plotted for each vowel arranged on the horizontal axis from high front, low central to high back vowels. Figures 6 and 7 represent ascending and descending temporal segment threshold values respectively. It is readily apparent that Subject Six established a higher temporal segment recognition threshold for /A / on ascending than on

¹²²Wilfrid J. Dixon and Frank J. Massey Jr., Introductions To Statistical Analysis (New York: McGraw-Hill Book Company, 1957), pp. 179-181.

¹²³Ibid., p. 388.

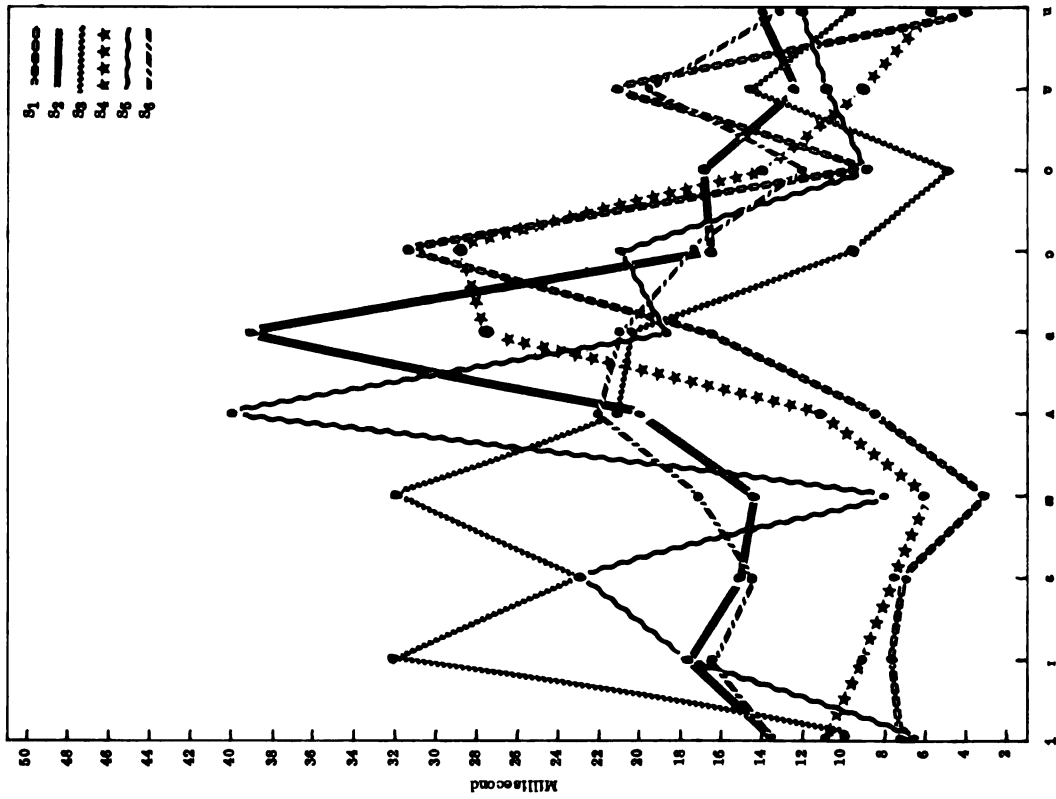


Figure 6. Subject temporal segment recognition threshold for each vowel presented in ascending series.

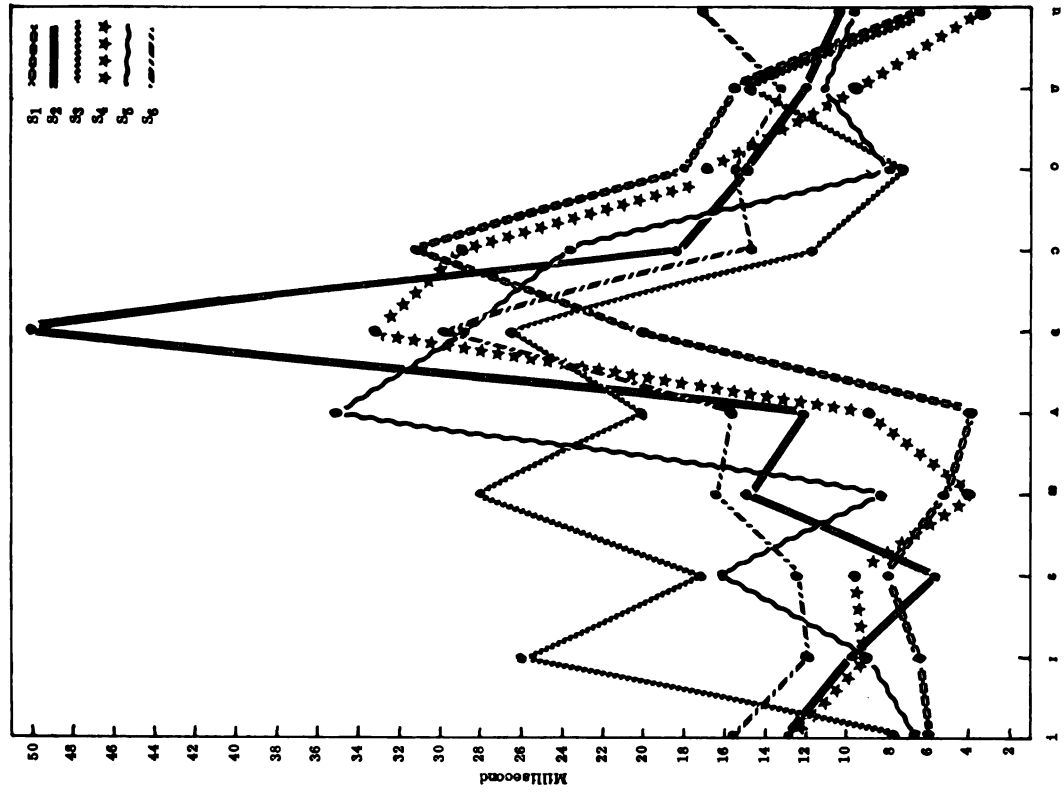


Figure 7. Subject temporal segment recognition threshold for each vowel presented in descending series.

descending series. Yet, Subject Two had the opposite effect by attaining a higher descending temporal segment recognition threshold for the /a / on the descending series than on the ascending series. Close inspection of the graphs indicates some subjects attained high temporal segment thresholds for some vowels when presented in ascending series and other vowels in descending series. One example of this is Subject Six who attained a higher temporal segment recognition threshold in the ascending series for /v / but attained a lower temporal segment recognition threshold for /u / in the ascending series. Realizing that the interaction does exist, a general pattern across the vowels for all subjects is nevertheless apparent. This general pattern indicates lower temporal segment recognition thresholds for the vowels: /i /, /ɪ /, /e /, and /æ /; and vowels /o /, /v / and /u /. The /æ / temporal segment recognition thresholds raises slightly, and the /ʌ /, /a / and /ɔ / have the highest temporal segment recognition thresholds. This aspect will be investigated more thoroughly later in this discussion. Figures 8 and 9 representing the subject mean temporal segment recognition threshold for each vowel in descending and ascending series, also depicts this same pattern. The three way interaction among subject, vowel and time is graphically presented in Figures 10 and 11. The interaction between vowel and subject is evident. The effect of time on subjects and vowels is not consistent. In the second half, Subject Four bettered his temporal segment recognition threshold of the /u /, but

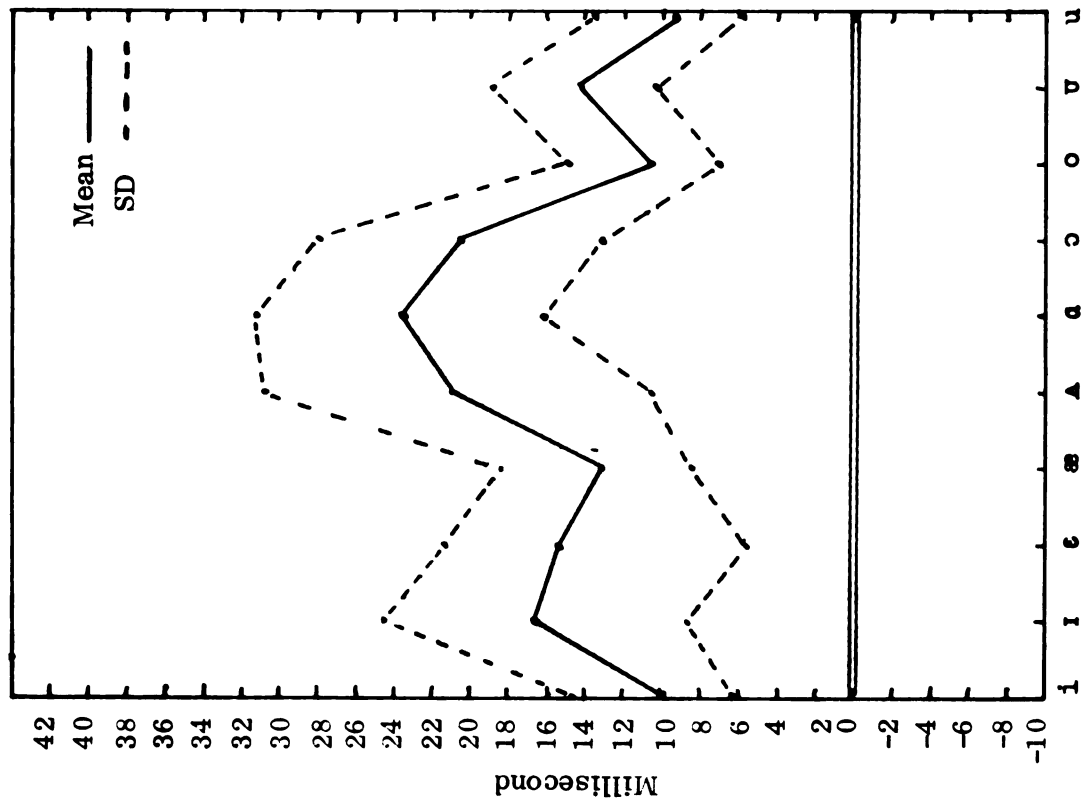


Figure 8. Subject mean vowel temporal segment recognition threshold for ascending series.

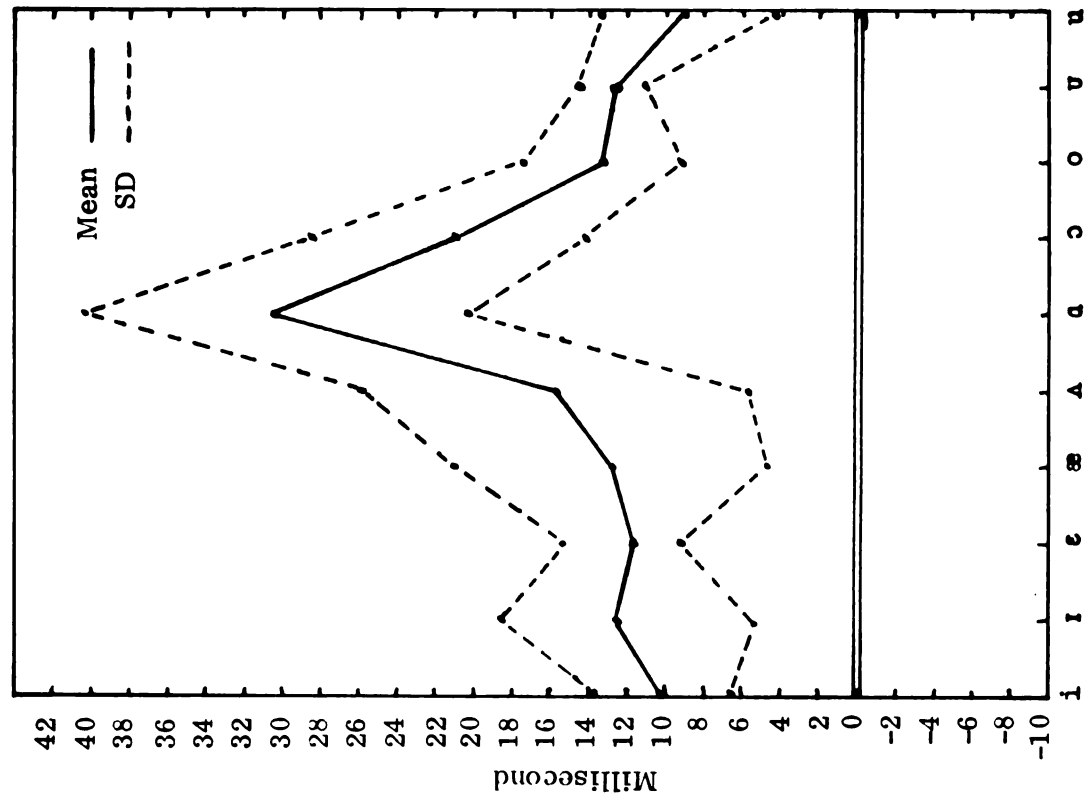
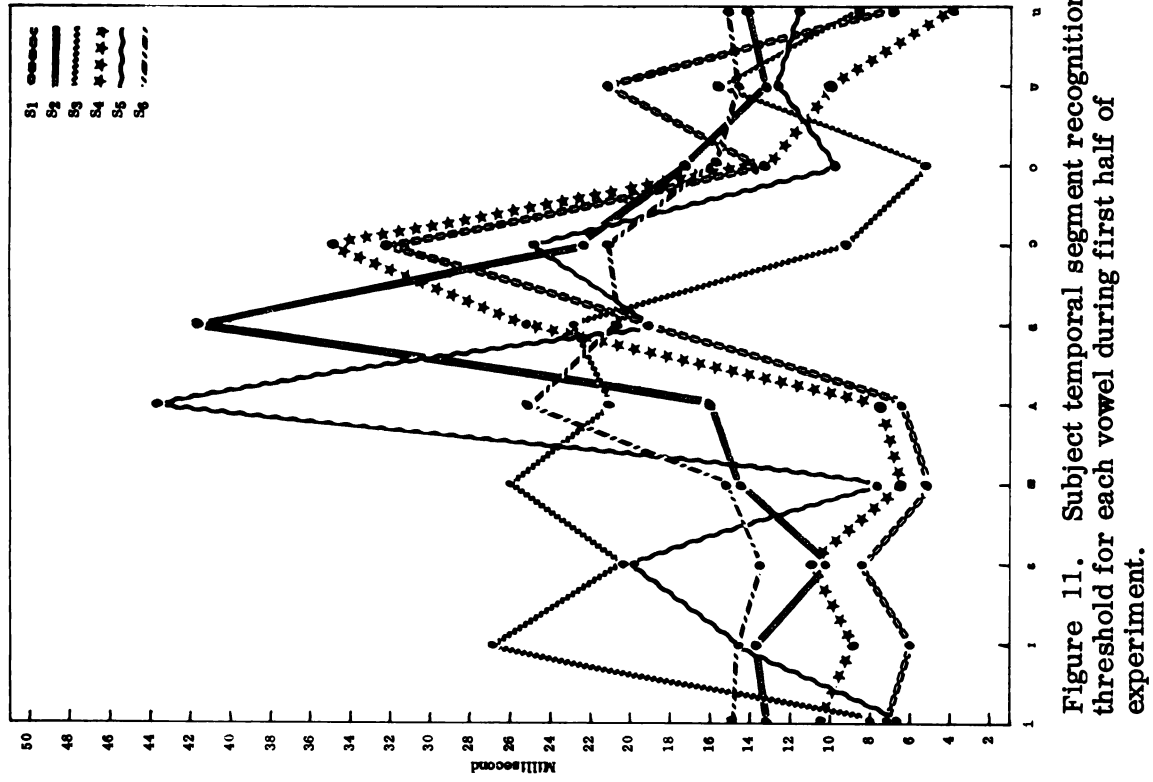
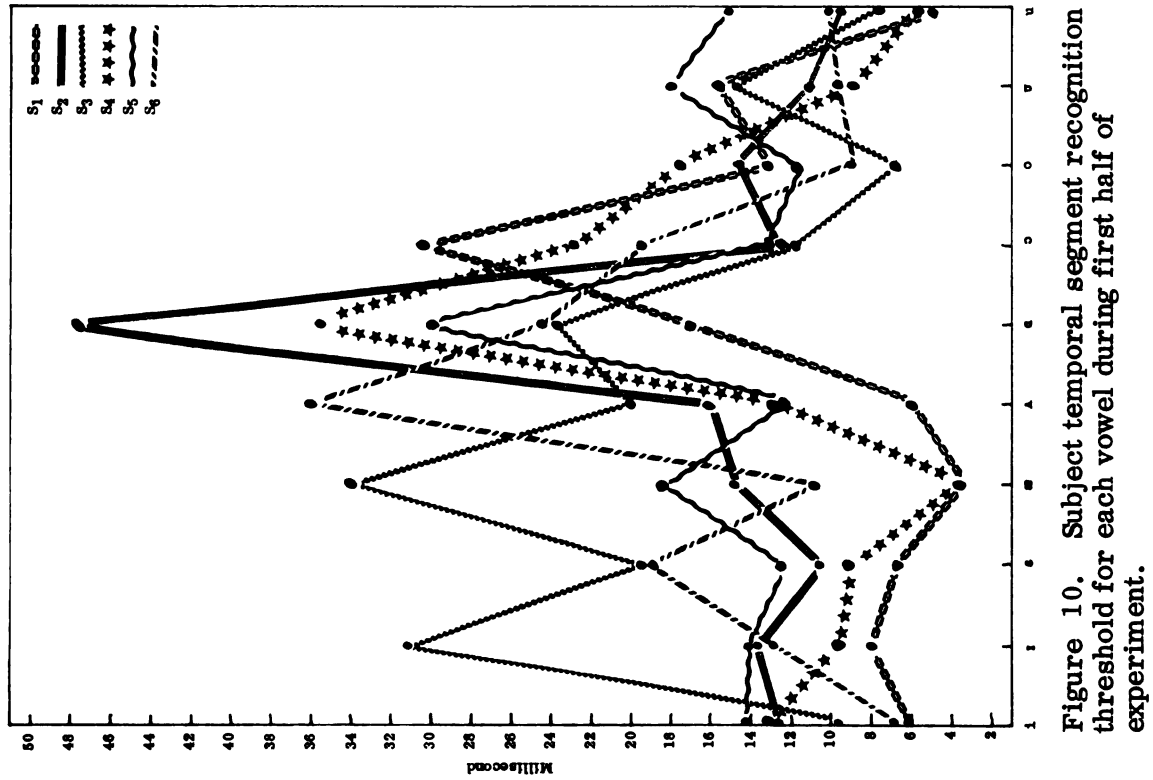


Figure 9. Subject mean vowel temporal segment recognition threshold for descending series.



increased the temporal segment recognition threshold of the /ɔ / sound. The /a / and /ɔ / temporal segment recognition threshold for Subject Four completely reversed between first and second half of the listening experience. The subject mean temporal segment recognition threshold and standard deviations for each vowel are plotted in Figures 12 and 13. The temporal segment recognition thresholds were slightly lower for the following vowels: /i /, /ɪ /, /e / and /o / but were slightly higher for /ɛ /, /æ /, /ʌ /, /ɔ /, /ʊ / and /u /. These differences do not look significant, and the results of the analysis of variance did not show significant difference in the main effect of time--first half and second half. Therefore, the interaction of subject, vowel, and time is most likely due to individual differences between subjects and vowels.

The interaction between subject and vowel is also evident in Figure 14. This graph represents the subject temporal segment recognition thresholds for each vowel over the entire experiment. The mean and standard deviation for each vowel is plotted in Figure 16. The general pattern is again evident, with the temporal segment recognition thresholds being higher in the middle of the graph. The medians were plotted in Figure 15 and indicate the same pattern as that in Figure 14. This provided evidence that the use of medians in place of means in the analysis would not appreciably change the results.

Cognizant of the above discussed interactions the sub-

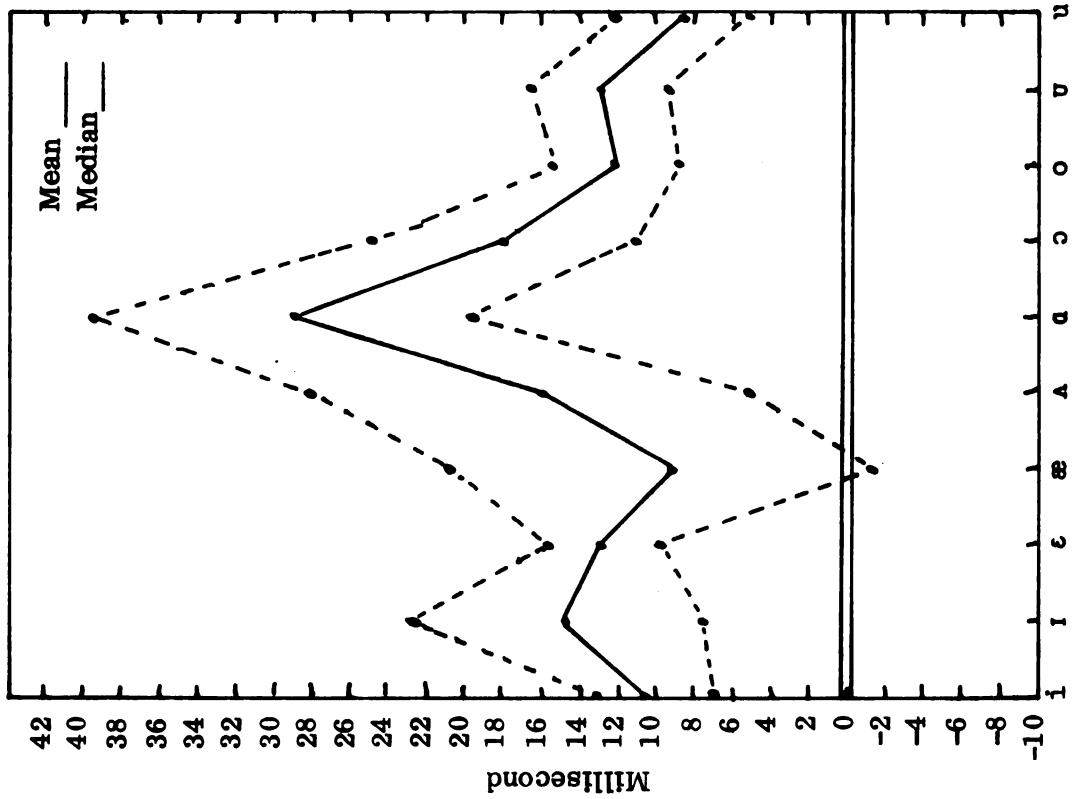


Figure 12. Subject temporal segment recognition threshold means and standard deviation for each vowel for the first half of experiment.

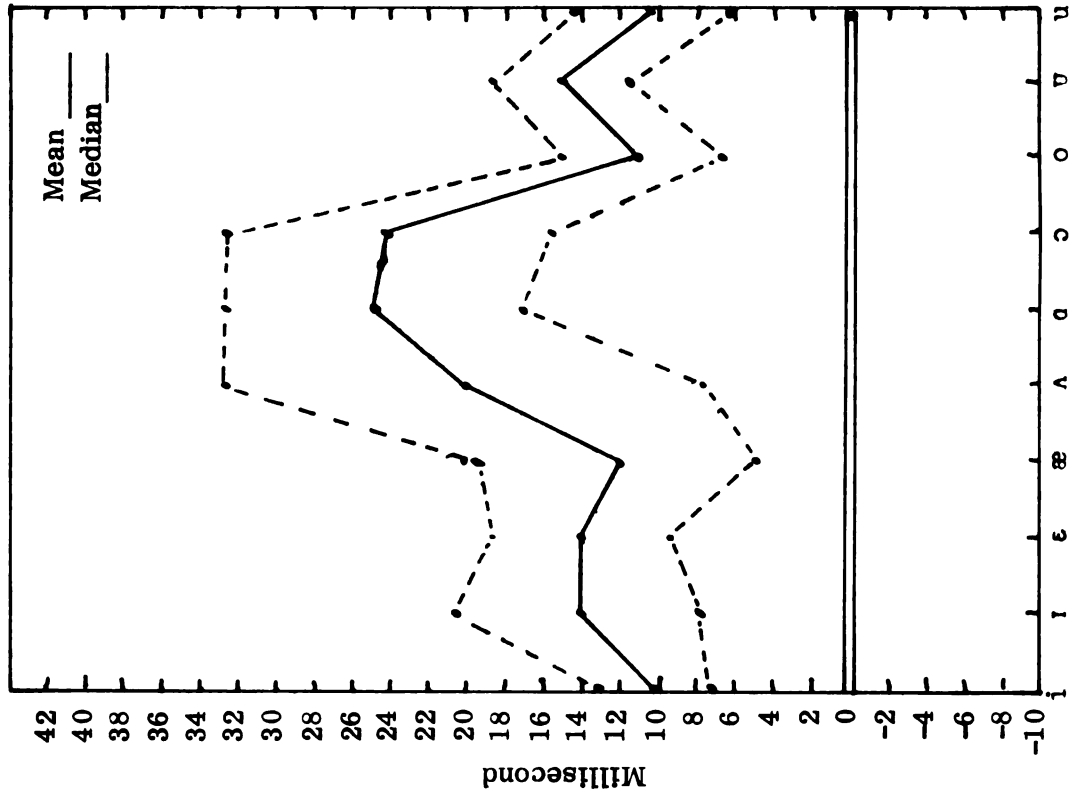


Figure 13. Subject temporal segment recognition threshold means and standard deviation for each vowel for the second half of experiment.

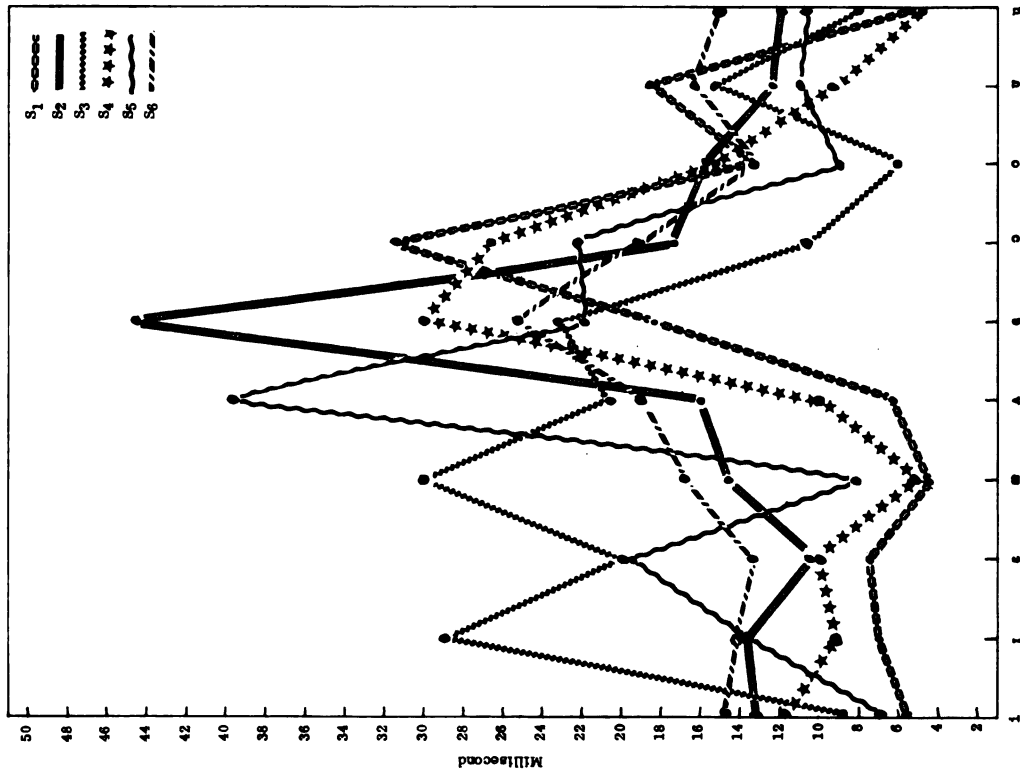


Figure 14. Subject temporal segment recognition threshold for each vowel over entire experiment.

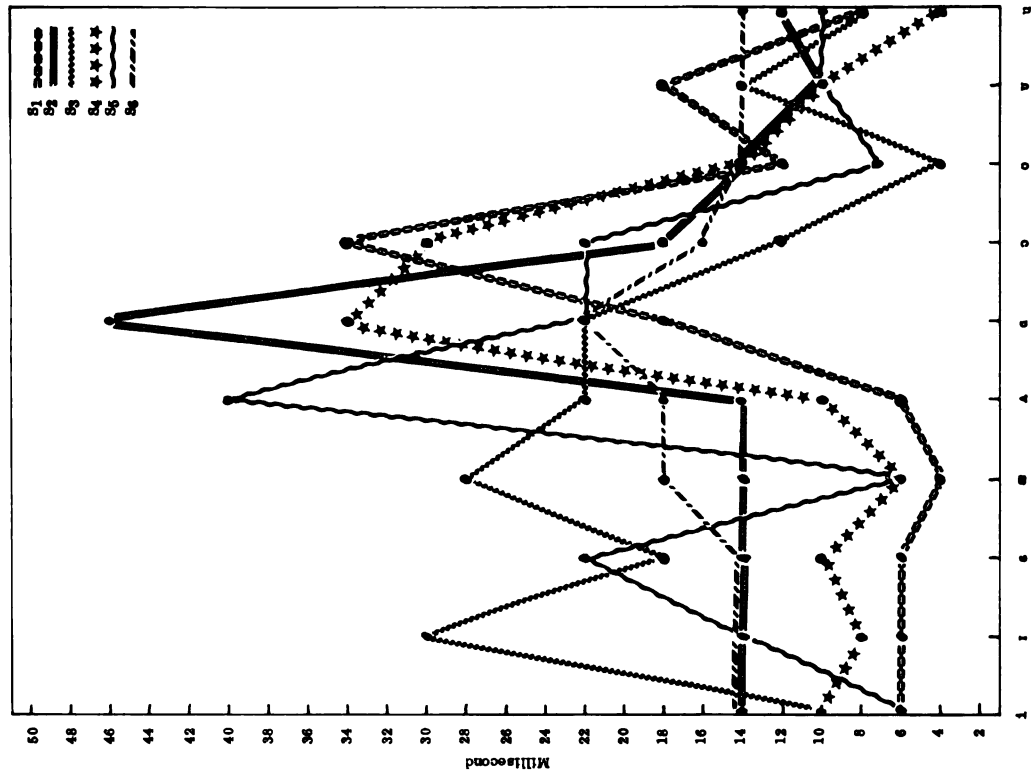


Figure 15. Subject medians used to compute temporal segment recognition threshold rather than means over entire experiment.

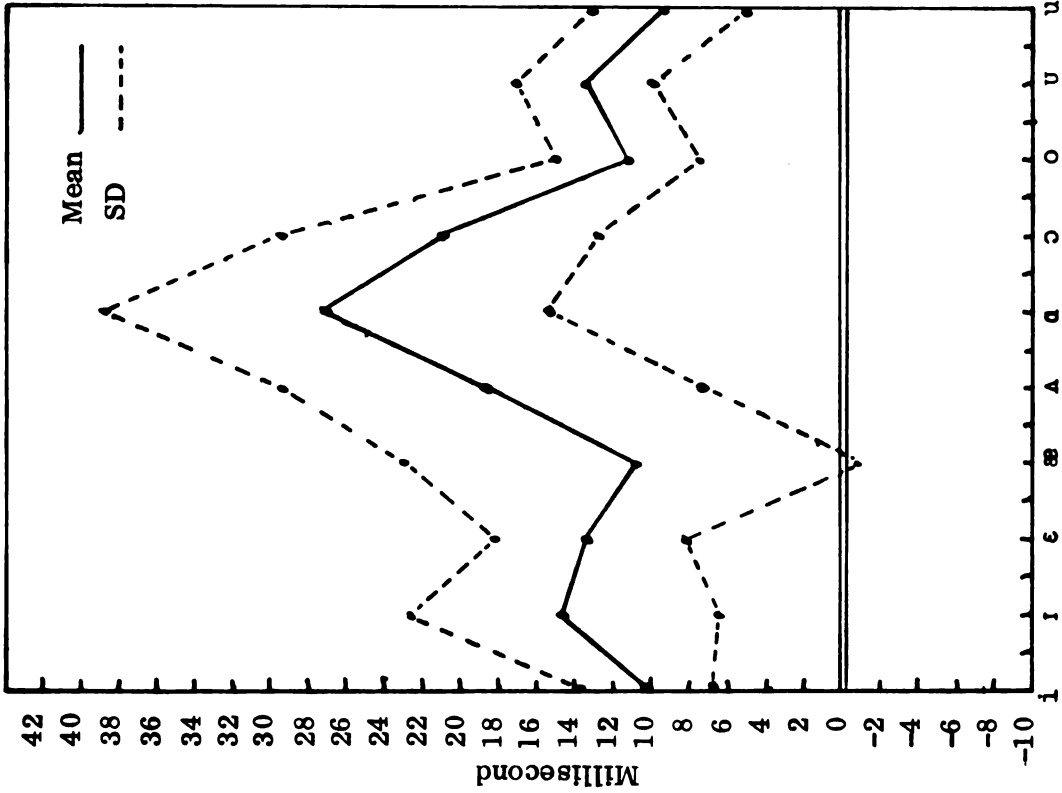


Figure 16. Subject temporal segment recognition Threshold mean and standard deviations for each vowel over entire experiment.

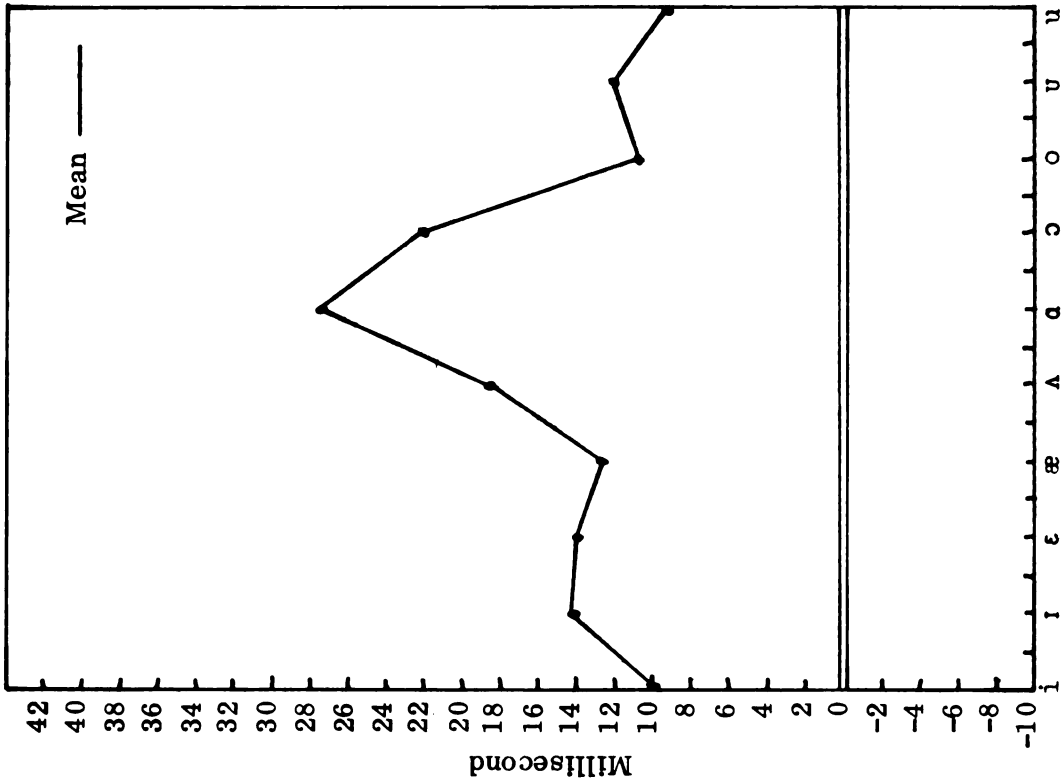


Figure 17. Subject median temporal segment recognition threshold means over entire experiment.

ject's temporal segment recognition threshold mean was computed to attain the temporal segment recognition threshold for each vowel. These thresholds arranged from lowest to highest threshold as established under the conditions of this study, were: /u /, 9.3 milliseconds: /i /, 10.3 milliseconds: /æ /, 10.6 milliseconds: /o /, 11.4 milliseconds: /e /, 13.5 milliseconds: /v /, 13.8 milliseconds: /r /, 14.4 milliseconds: /ʌ /, 18.5 milliseconds: /ɔ /, 21.2 milliseconds: and /a /, 27.2 milliseconds. A graph depicting these vowel temporal segment recognition thresholds is presented in Figures 18 and 19 and pictures of the wave form and segments of the wave forms closest to these values are presented with arrows indicating the computed temporal segment recognition threshold for the particular vowel. A critical difference of the means test as described by Lindquist¹²⁴ was applied to the data. The results of this test indicated that a 6.3 difference between means was significant at the 0.05 level. The temporal segment recognition threshold for each vowel is plotted on the vertical and horizontal axes in Table 6. The numbers in the cells represent the differences between the temporal segment recognition threshold means in milliseconds for the corresponding vowels.

The pattern of significant differences in Table 6 show that the /a / and /ɔ / are significantly different at the

¹²⁴E. F. Lindquist, "Design and Analysis of Experiment in Psychology and Education" (Boston: Houghton Mifflin Company, 1953), pp. 90-96.

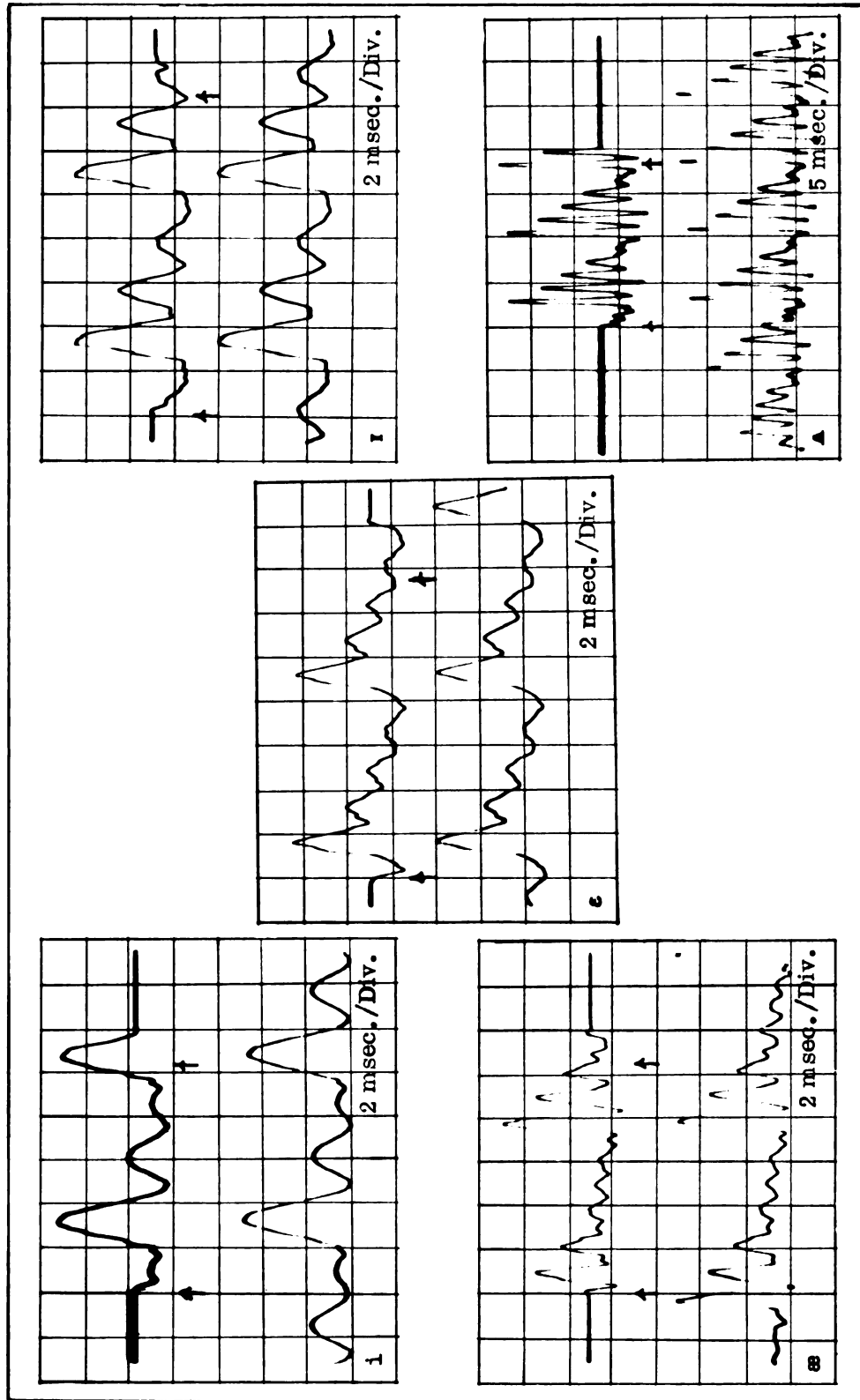


Figure 18. The unsegmented vowel wave forms are presented below the vowel segment presented to the subjects. The phonetic symbol and the horizontal scale are presented below each set of wave forms. The arrows indicate the computed temporal segment threshold for each vowel.

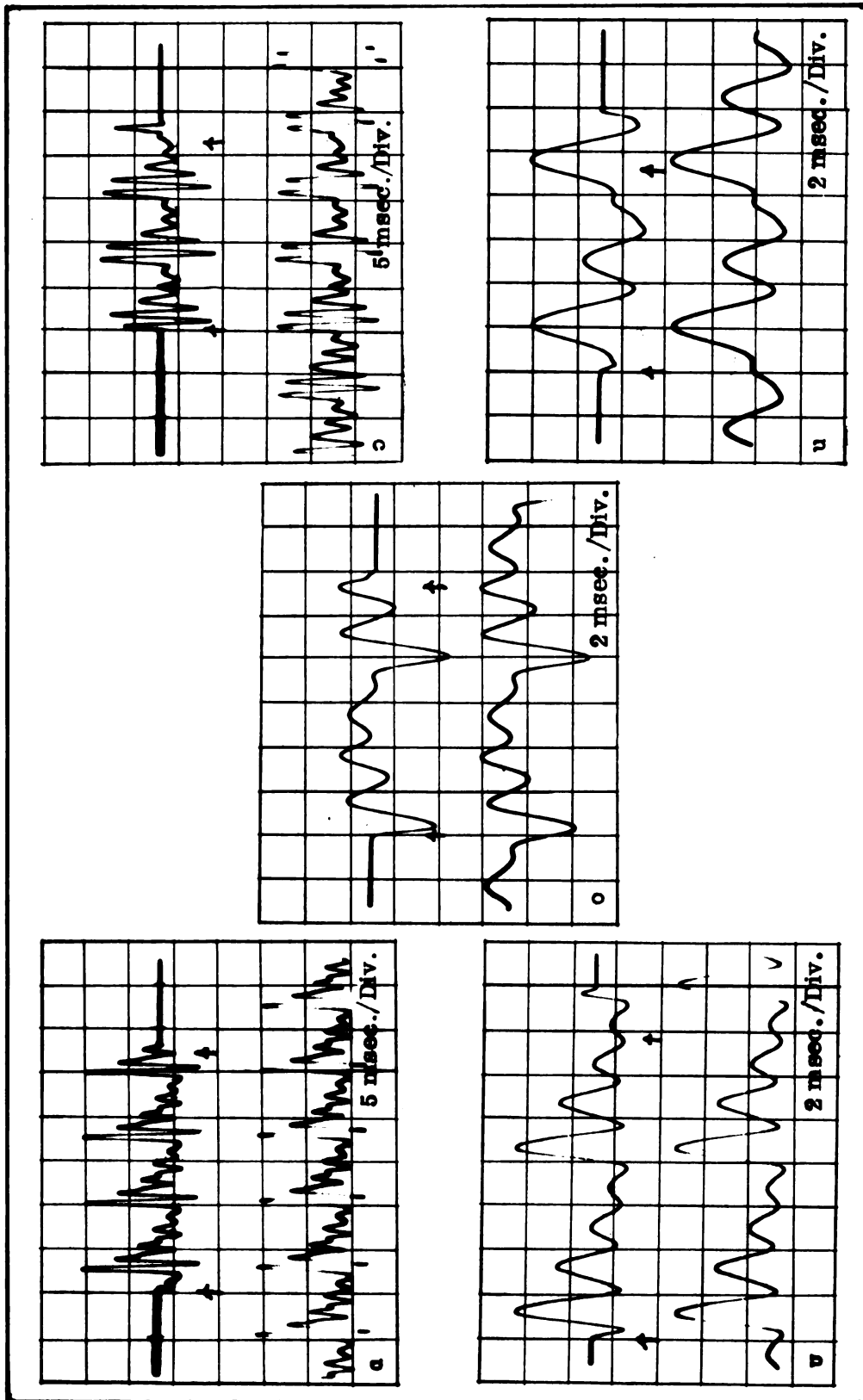


Figure 19. The unsegmented vowel wave forms are presented below the vowel segment presented to the subjects. The phonetic symbol and the horizontal scale are presented below each set of wave forms. The arrows indicate the computed temporal segment threshold for each vowel.

TABLE 6
VOWEL TEMPORAL SEGMENT THRESHOLD IN
MILLISECONDS DIFFERENCE MATRIX

	14.4	13.4	10.6	18.5	27.2	21.2	11.4	13.8	9.3
	i	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
10.3	i	4.1	3.4	8.2*	16.9*	10.9*	1.1	3.5	1.0
14.4	i	x	1.0	4.1	12.8*	6.8*	3.0	0.6	5.1
13.4	e		x	5.1	13.8*	7.8*	2.0	0.4	4.1
10.6	æ		x	7.9*	16.6*	10.5*	0.8	3.2	1.3
18.5	ʌ			x	8.7*	2.6	7.1*	4.7	9.2*
27.2	ɑ				x	6.0	15.8*	13.4*	17.9*
21.2	ɔ					x	7.5*	7.4*	11.8*
11.4	o						x	2.2	2.4
13.8	ʊ							x	4.5
9.3	u								x

*Temporal segment recognition thresholds difference greater than 6.3 and significantly different at the 0.05 level as determined by the critical difference of means test.¹²⁵

¹²⁵Ibid.

0.5 level from all the other vowels, but are not significantly different from each other. The vowel / Λ / temporal segment recognition threshold is significantly different from / $\circ$ /, / u /, / æ /, / a /, and / i /, but not significantly different from / ɔ /, / ʊ /, / i /, and / ɛ /. There is a strong tendency for the temporal segment recognition threshold of / Λ /, / a / and / ɔ / to be significantly different from the other vowels, and the / a / significantly different from both / Λ / and / ɔ / at the critical level previously stated.

After close examination of the graphs and tables presented above, as well as the results of the analysis of variance, the following null hypotheses are rejected at the 0.05 level of significance under the conditions set forth in this study:

1. There is no significant difference among each of the ten vowel temporal segment recognition thresholds as obtained in this study.
2. There is no significant difference among the six subjects' temporal segment recognition thresholds for vowels as obtained in this study.
3. There is no significant differences between the ascending series and descending series vowel temporal segment recognition thresholds.
4. There is no significant difference among subjects, vowel, and series order.
5. There is no significant difference among subject, vowel and time order.
6. There is no significant difference between subject and vowel.
7. There is no significant difference between subject and time order.
8. There is no significant difference between vowel and series order.
9. There is no significant difference between vowel and time order.

The following null hypotheses were not rejected at the 0.05 level of significance after due consideration:

1. There is no significant difference between the first half of the series and the second half of the series, and the subjects' temporal segment recognition thresholds.
2. There is no significant difference between the first half and second half of the temporal segment recognition thresholds.
3. There is no significant interaction among the subject, series order, and time order.
4. There is no significant interaction between the series order and time order.
5. There is no significant interaction among the subject, vowel, series order and time order.
6. There is no significant interaction among vowel, series order and time order.

The null hypothesis, there is no significant difference between the temporal segment recognition thresholds for the cardinal and non-cardinal vowels, was tested by employing a t-test.¹²⁶ The cardinal vowel temporal segment recognition thresholds were /u / 9.3 milliseconds, /i / 10.3 milliseconds, /o / 11.4 milliseconds, and /a / 27.2 milliseconds. And the non-cardinal vowel temporal segment recognition thresholds were /m / 10.6 milliseconds, /s / 13.5 milliseconds, /v / 13.8 milliseconds, /r / 14.4 milliseconds, /l / 18.5 milliseconds and /e / 21.2 milliseconds. The mean cardinal vowel temporal segment recognition threshold equals 19.5 milliseconds and the mean non-cardinal vowel temporal segment recognition threshold is equal to 15.3 milliseconds. The computed t-test equaled 0.81 which was not significant at the 0.05 level for eight degrees of freedom.¹²⁷ Thus, the null hypothesis was not rejected.

¹²⁶Dixon and Massey, op. cit., pp. 121-122.

¹²⁷Ibid., p. 384.

DISCUSSION

An identification matrix was constructed from the data for each subject to visualize which vowels were mis-identified most frequently. These tables are found in Appendix C. A composite of these identification matrices is presented in Table 7. The /a / and /ɔ / vowels were most often mis-identified as /ʌ /, and /ʌ / was most often mis-identified as /a /. This illustrates the high degree of confusion among these vowels and may account for the significantly higher temporal segment recognition thresholds for these vowels as cited earlier. It should also be noted that when vowels were mis-identified, they were most often mis-identified as the vowel adjacent to the intended vowel located on the vowel tongue hump position diagram. This same phenomenon was reported by Gray,¹²⁸ Peterson and Barney,¹²⁹ and Fairbanks and Grubb.¹³⁰ Peterson's¹³¹ data did not concur with this finding.

It is interesting to note in Table 7 that the /u / was mis-identified as /i / 37 times and only mis-identified more often as /ʊ / 39 times. On the other hand /i / was mis-identified as /u / 51 times and only mis-identified as /ɪ / 74 times. Inspection of the vowel identification matrix

¹²⁸Gray, loc. cit.

¹²⁹Peterson and Barney, loc. cit.

¹³⁰Fairbanks and Grubb, loc. cit.

¹³¹Peterson, loc. cit.

TABLE 7

IDENTIFICATION MATRIX OVER
ALL SIX SUBJECTS

Intended Vowel	Identified as										
	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u	
i	(1556)	74	11	1	5	0	0	5	2	51	
ɪ	63	(1467)	23	8	21	3	4	34	5	17	
e	4	50	(1454)	15	24	1	11	20	12	5	
æ	2	9	85	(1469)	16	49	68	4	0	0	
ʌ	0	7	15	10	(1410)	317	15	2	13	3	
ɑ	1	11	36	5	582	(1038)	14	10	9	0	
ɔ	13	7	127	21	189	56	(1296)	62	6	34	
o	1	14	17	1	36	7	10	(1504)	10	68	
ʊ	7	25	24	2	47	1	3	122	(1447)	26	
u	37	23	1	0	12	1	1	10	39	(1578)	
ΣX-C*	129	220	339	156	933	434	126	269	155	142	

ΣX-C* Sum of the column minus the number of correct responses.

presented by Fairbanks and Grubb¹³² reveals this same slightly confused pattern between these vowels. The magnitude of this confusion is very small considering the number of times the vowels were presented. Yet, even this slight confusion may seem unusual until we consider that as Cooper et. al.¹³³ state, "...the ear can and sometimes does perform an averaging operation on two formants which lie close together; thus the first and second formants of back vowels may at times be replaced by a single formant...." If the first two formants of the /u / sound were averaged in this manner, they could be heard as the first formant of the /ɪ / sound and the third formant heard as the second formant of the /ɪ / sound accounting for this confusion.

Statistically significant interaction between subject and vowel has not been reported by other researchers. Yet, Peterson and Barney¹³⁴ stated, "The ease with which the observers classified the various vowels varied greatly," after they had asked subjects to identify ten words in which the vowel was varied in a /h-d/ consonant environment. One of Gray's¹³⁵ conclusions states, "Individual differences

¹³²Fairbanks and Grubb, op. cit., p. 207.

¹³³Franklin S. Cooper, Pierre C. Delattre, Alvin M. Liberman, John M. Borst and Louis J. Gerstman, "Some Experiments on Perception of Synthetic Speech Sounds." Journal Acoustical Society of America, XXIV, No. 6 (November, 1952), p. 603.

¹³⁴Peterson and Barney, loc. cit.

¹³⁵Gray, op. cit., p. 89.

exist among the subjects, some of them being able to identify a significant number of vowels at shorter periods than others." He also reports that vowels are not equally recognizable by individuals. Interactions similar to those found in the present study were probably also present in these studies.

The /ʌ /, /a / and /ɔ / were the most difficult vowels to recognize and attained the highest temporal segment recognition thresholds. The /i /, /u / and /ə / vowels were the most easily identified and attained the lowest temporal segment recognition thresholds in the present study. This agrees with the findings of Peterson and Barney who report the /ʌ /, /a / and /ɔ / as the most difficult to recognize and the /i /, /ə / and /u / as the most easily identified. This supports a suggestion made by Stevens and House¹³⁶ after they compared the data from Peterson and Barney's study of the vowel in /h-d/ context with studies of vowels in isolation. They suggested, "...that the /h-d/ context has a negligible effect upon the articulation during the central portion of the vowel, that is, the vowel in context /h-d/ is generated with essentially the same articulatory configuration as the vowel in isolation." A study cited earlier initiated by Fairbanks and Grubb¹³⁷ also report the /i / and /u / as the most easily recognized vowels when presented to listeners at

¹³⁶Stevens and House, op. cit., p. 16.

¹³⁷Fairbanks and Grubb, op. cit., pp. 203-219.

0.3 second temporal segments. They reported the /r / and /a / as being most difficult to recognize.

As reported earlier, the study by Gray¹³⁸ was the only previous study attempting to establish vowel recognition thresholds as a function of temporal segmentation. The temporal segment recognition thresholds for each vowel were not reported by Gray. Yet, he did conclude that, "Some subjects are able to identify a significant number of speech sounds when presented with a duration of as little as .003 second. Duration minima of $1/200$ to $1/333$ second permit recognition of some of the vowels by some of the subjects." Two of the six subjects in the present study attained temporal segment recognition thresholds near 5 milliseconds ($1/200$ second) for particular vowels: S_1 /a / 4.4 milliseconds and /u / 5.2 milliseconds; S_2 /a / 5.0 milliseconds and /u / 4.8 milliseconds. The mean subject vowel temporal segment recognition thresholds, as previously reported in this report, ranged from 9.3 milliseconds for the /u / to 27.25 milliseconds for the /a /. Thus the temporal segment recognition thresholds for the present study do not agree with Gray's study or with Peterson's¹³⁹ report of 50 per cent recognition of the /o / and /a / vowels presented at 3.1 milliseconds temporal segments.

The present study agrees with Gray's findings on the

¹³⁸Gray, op. cit., pp. 75-90.

¹³⁹Peterson, loc. cit.

following points:

1. The temporal segment recognition thresholds are not the same for all vowels, some being recognized at shorter intervals than others.
2. Subjects differ in their ability to recognize vowels presented at short temporal segments, some being able to identify vowels at shorter periods than others.
3. One of the most unstable vowels in American English is the /e/.
4. When a vowel was mis-identified, it was most often mis-identified as one adjacent to it as indicated on the tongue hump position graph.

The disagreement between the temporal segment recognition thresholds as reported in this study and the "duration minima" for the perception of the vowels reported by Peterson in 1939 and Gray 1942 may be due to the differences in methodology. As cited earlier Gray presented the vowel stimuli consecutively at one temporal segment value and the subject realized that each vowel would be presented during the particular temporal segment presentation. This limited the field of choice to the eleven vowels and after each presentation the field of choice was decreased by one. In the present study the vowels were randomized throughout the entire experiment and the subjects had no knowledge of which vowel would be presented next. Yet, they did know the sounds would be presented in ascending and descending order. Peterson attained a 50 per cent recognition of the vowels presented at 3.1 milliseconds only after decreasing the vowel field of choice from eight vowels presented in the first portion of the study to six vowels in the second portion of his study. Even in the second study he attained the 50 per cent on only two vowels: /a/ and /o/ when presented to

fifteen rather than the eighteen initial subjects at 3.1 milliseconds segments. It was this methodology that Fairbanks¹⁴⁰ was referring to when he suggested that the results of Peterson's study may have been influenced by a prior guessing.

¹⁴⁰Fairbanks, loc. cit.

CHAPTER V

SUMMARY AND CONCLUSIONS

The study of how man interprets the world around him has been the quest of researchers through the ages. Psychophysical methods, as described by Fechner and Weber, have been utilized in defining subject response scales in relation to physical scales in attempts to measure and define the dimensions of human perception. In addition to these psychophysical scaling devices, methods have been developed to scale psychological phenomena that are not directly related to physical scales. The limits of various sense modalities, such as, olfactory, visual, tactile, and auditory, have been established by researchers employing psychophysical methods. Many diagnostic procedures, commonly administered by professionals in the various disciplines, are based upon the earlier research that established absolute and difference thresholds for the senses employing these methods.

Measuring and defining the dimensions of the acoustic speech signal have been the goal of researchers for many years. Attempts have been made to show that persons with functional speech disorders have deficits in auditory acuity or general discrimination. Although the findings conflict, most authorities would agree the research is not conclusive. Research has indicated that persons with functional speech disorders

do have a hard time discriminating the sounds they have difficulty with, even though their general acoustic discrimination is not different from those without speech defects.

One of the prime avenues utilized by researchers to investigate the dimensions of the speech signal is that of the airborne acoustic signal which is accessible to modification by various research devices. This acoustic signal varies in terms of intensity, frequency and time. The intensity threshold varies with the frequency; the lower frequency extending to 15 Hz and the upper frequency extending to 20,000 Hz require the greatest intensity. The frequencies of 1000 Hz to 3000 Hz require the least intensity to be perceived. There are 1500 difference thresholds for pitch and 325 difference thresholds for loudness. Researchers have provided evidence indicating that perception of these acoustic properties is contingent upon time or length of the stimulus presentation, as well as the frequency and intensity. That is, the perception of an acoustic stimulus is dependent upon the length of time the frequency is presented at intensities near threshold. Researchers have concluded that this holds true for pure tones as well as vowel acoustic signals.

Only two former studies have been focused on establishing temporal recognition thresholds for vowels. The first of these studies was conducted by Gray¹⁴¹ in 1937. This study incorporated an ingenious mechanical switch actuating device

¹⁴¹Gray, loc. cit.

which segmented live productions of vowels phonated at six different fundamental frequencies. These temporal segments were simultaneously presented through a loud speaker to his subjects. Although he did not publish the vowel temporal segment recognition thresholds, he did report that many of his subjects recognized the vowel temporal segments at 3 milliseconds.

The second study was directed by Gray and conducted by Peterson.¹⁴² This study failed to support the hypothesis that various portions of a vowel wave length were more easily recognized than other portions. The first portion of his study did not support Gray's findings that the vowel temporal segment threshold for vowels was 3 milliseconds. However, the second portion of the study did result in over 50 per cent recognition of two out of six vowel temporal segments presented at 3.1 milliseconds.

The purpose of the present study was to establish temporal segment recognition thresholds for the following vowels: /i/, /ɪ/, /e/, /ɛ/, /æ/, /ʌ/, /ɑ/, /ɔ/, /o/, /ʊ/, and /u/. The vowels were phonated with 125 Hz fundamentals and recorded. The speaker monitored a Lissajous figure on an oscilloscope and a 125 Hz pure tone through earphones to help him attain the 125 Hz fundamental. Each vowel was then presented ten times in random order (without replacement) to eight listeners for the purpose of validation.

¹⁴²Peterson, loc. cit.

Each complete vowel was recognized at least 95 per cent of the time before it was accepted as a valid stimulus to be segmented for presentation to the listeners in this study. The pilot studies provided information that was used to determine the number and length of stimuli to be presented. These vowel stimuli were then chopped into fifteen different temporal segments ranging from 4 to 60 milliseconds in 0.4 millisecond steps based upon the information attained from the pilot studies. Next the segmented vowels were randomized at each temporal segment level in 110 ascending and 110 descending series and presented to six listeners in a modified method of minimal change. These eleven vowels presented at fifteen different temporal segments in 220 series yielded a total of 3,300 stimuli present to the six listeners. The listeners were all doctoral graduate students, with normal hearing acuity, in speech and hearing science. The stimuli were presented to the listeners through earphones in a sound treated room. They responded to the stimuli by writing the international phonetic symbol that represented their perception of the vowel temporal segment presented. They had no knowledge of the vowel presentation order, but they did know the stimuli would be presented in temporal descending and ascending series. Together all six subjects listened to the stimuli for four hours over an eight hour period with appropriate rest periods.

The subject responses were then tabulated for each vowel maintaining the descending and ascending order as well as the

time order--first half and second half--of the experiment. The threshold for each series were then computed. The /e/ sound was eliminated from the analysis because 50 per cent recognition of this vowel was not attained at the longest temporal segment of 60 milliseconds. The Bartlett's test for homogeneity of variance supported the assumption of homogeneity of variance for the data. These series temporal segment threshold values, were subjected to a four way analysis of variance ($2 \times 2 \times 6 \times 10$) corresponding to the two series orders, two time orders, six listeners and ten vowels. The results of this analysis pointed to interactions significant at the 0.05 level among and between some of the main effects. To envision these interactions line graphs were prepared employing the temporal segment recognition thresholds for each subject plotted for each vowel. Although the interactions were apparent, a general pattern emerged when these graphs were plotted for descending and ascending, first half and second half, and over the entire experiment. A critical difference test of the means indicated the essential differences to be higher temporal segment recognition thresholds for the / Δ /, /a / and /o / vowels. A vowel identification matrix was also prepared for each subject. These tables indicated a great deal of confusion among these three vowels.

CONCLUSIONS

The following conclusions are presented within the limitations set forth in this study.

1. Vowel temporal segment recognition thresholds were affected by the particular vowel presented to the listeners in this study.
2. Vowel temporal segment recognition thresholds are significantly different for individual listeners.
3. The vowel temporal segment recognition thresholds were higher when presented in ascending series than in descending series.
4. Vowel temporal segment thresholds were not affected by serial position. That is, the vowel temporal segment thresholds were essentially the same during the first half and the second half of the experiment.
5. There was no significant difference for the temporal segment thresholds of recognition for cardinal and non-cardinal vowels.
6. Differences among temporal segment recognition thresholds for subjects, vowel and series order seems to be the effect of individual listener differences.
7. A low or high vowel temporal segment recognition threshold for a particular vowel does not have a direct relationship to the performance of the same subject when recognizing other vowels.
8. The vowels /**ʌ** /, /**a** / and /**ɔ** / have the longest

temporal segment recognition thresholds; these thresholds were essentially different than the thresholds of other vowels. This may have been due to the mis-identification among these three vowels.

9. Vowels mis-identified were most frequently mis-identified as adjacent vowels on the tongue hump diagram. This phenomenon has also been reported by previous researchers.
10. The American English speaking subjects in this study mis-identified the /e / vowel as /ɪ / over 50 per cent of the time even when presented at long durations (60 milliseconds).

SUGGESTIONS FOR FURTHER STUDY

Further experimentation utilizing essentially the same design and instrumentation as the present study could be carried on by using ongoing speech rather than vowels in isolation. The fundamental frequencies of these vowels would present more natural variations than the isolated presentations in the present study.

The intensity is another variable that could be increased from the 70 dB (SPL) utilized in the present study to ascertain the possibility that this would decrease the vowel temporal segment recognition thresholds.

Another variation of the present design could be that of training the subjects to a criterion. The training could consist of presenting each vowel at 24 milliseconds segments to each subject, informing them as to whether they are right

or wrong until each subject is able to identify each vowel 75 per cent of the time. It is important to realize, however, that this procedure may and probably would lead to vowel recognition based upon different variation in the acoustic wave form than ordinarily utilized to differentiate the vowels. That is, the subjects may learn to identify the vowel based upon auditory differences not attributed to discriminating the vowel in natural speech. The telegraph operator learns to "hear" /i / when auditorily stimulated with a particular sequence of dots and dashes, but this auditory stimulus is not the same as the oral /i /.

Psychophysical methods other than minimal differences could be employed. The method of adjustment could be initiated with modifications of the present apparatus. In particular, the delay control knob on the pulse generator could be enlarged so it would not be as sensitive to change. This would make it easier for the subject to adjust the temporal segments in smaller temporal amounts. This method could only utilize the ascending series because of the nature of the stimuli and the recognition task. The recognition task is essentially a cortical function, whereas the intensity detection is essentially a neural function.

Vowel temporal difference limens for the various vowels could be established through the use of the constant stimulus methods. This, of course, would yield the difference limens in terms of the subject's ability to distinguish differences in temporal segment presentations but would not yield infor-

mation concerning the ability to recognize a particular vowel presented at a particular temporal segment.

In the present study the initial and final transients of the vowel were eliminated from the stimuli by taking the vowel temporal segment from the middle of the vowel. To establish the function of these transients in the process of recognition, it would be possible to chop off varying portions of the vowel beginning at the foremost portion of the wave form including the initial transients and present these stimuli to the subjects with increasing amounts chopped off. The same procedure could be followed by chopping the wave form from the final toward the initial portion of the vowel.

Since the present study revealed a strong confusion factor between the / Δ /, / α / and / $\circ$ /, it would be interesting to investigate possible existence of this phenomenon in ongoing speech. Sentences could be prepared containing these vowels. These vowels could then be interchanged to determine their effects upon intelligibility. If intelligibility of the sentences was not affected by this change, it would indicate that for some words, the vowels may be interchanged without changing the word recognition. It may be that the distinct characteristics of these vowels are not essential for the recognition of some spoken words but are essential for the recognition of other spoken words.

Various vowel temporal segments could be placed in (h-d) consonant environments and presented to listeners. The listener's recognition score at particular vowel temporal

segments would yield the vowel temporal segment recognition threshold.

A comparison of temporal segment recognition thresholds or temporal segment recognition threshold subject ranges between expert and non-expert or young listeners must be accomplished before any standardization of vowel temporal segment recognition threshold norms would be feasible. This type of study may indicate the degree to which past experience influences the ability to recognize the vowels presented at short temporal segments.

BIBLIOGRAPHY

BOOKS

- Black, John W. and Moore, Wilbur E. Speech Code, Meaning and Communication. New York: McGraw-Hill Book Company, Inc., 1955.
- Dixon, Wilfrid J. and Massey, Frank J., Jr. Introductions to Statistical Analysis. New York: McGraw-Hill Book Company, Inc., 1957.
- Fairbanks, Grant. Experimental Phonetics: Selected Articles. Urbana and London: University of Illinois Press, 1966.
- Fletcher, Harvey. Speech and Hearing in Communication. Princeton, New Jersey: D. Van Nostrand Company, Inc., 1961.
- Guilford, J. P. Psychometric Methods. New York: McGraw-Hill Book Company, Inc., 1954.
- Hebb, Donald Olding. The Organization of Behavior: A Neuropsychological Theory. New York: John Wiley and Sons, Inc., 1949.
- Hirsh, Ira J. The Measurement of Hearing. New York: McGraw-Hill Book Company, Inc., 1952.
- Kenyon, John S. American Pronunciation. Ann Arbor, Michigan: George Wahr Publishing Company, 1950.
- Lindquist, E. F. Design and Analysis of Experiment in Psychology and Education. Boston: Houghton Mifflin Company, 1953.
- Mase, Darrel J. Etiology of Articulatory Speech Defects. New York: Teachers College, Columbia University, 1946.
- Miller, G. A. and Lichlider, J. C. R. "The Perception of Speech," Handbook of Experimental Psychology. S. S. Stevens, editor. New York: John Wiley and Sons, Inc., 1951.
- O'Neill, John J. and Oyer, Herbert J. Applied Audiometry. New York: Dodd Mead and Company, Inc., 1966.
- Peterson, Gordon E. "Acoustic of Speech, Part II." Handbook of Speech Pathology. Lee Edward Travis, editor. New York: Appleton-Century Crofts, Inc., 1957.
- Saporta, Sol. psycholinguistics: A Book of Readings. New York: Holt, Rinehart and Winston, Inc., 1961.
- Stevens, Stanley Smith, (ed.). Handbook of Experimental Psychology. New York: John Wiley and Sons, Inc., 1951

_____. Handbook of Experimental Psychology. New York: John Wiley and Sons, Inc., 1963.

_____. and Davis, Howell. Hearing; Its Psychology and Physiology. New York: John Wiley and Sons, Inc., 1938.

Titchener, Edward Bradford. Experimental Psychology. Quantitative. New York: Macmillian, 1905.

Wever, Ernest G. and Lawrence, Merle. Physiological Acoustics. Princeton, New Jersey: Princeton University Press, 1954.

ARTICLES AND PERIODICALS

Anderson, Virgil A. "The Auditory Memory Span for Speech Sounds," Speech Monographs, V, No. 1 (December, 1938), pp. 115-129.

Bindra, Dalbir and Waksberg, Helene. "Methods and Terminology in Studies of Time Estimation," Psychological Bulletin, LIII, No. 2 (March, 1956), pp. 155-159.

Boring, Edwin G. "The Two-Point Limen and the Error of Localization," American Journal of Psychology, XLII, No. 3 (July, 1930), pp. 446-449.

Broadbent, D. E., Ladefoged, Peter and Lawrence, W. "Vowel Sounds and Perceptual Constancy," Nature, CLXXVIII, No. 4537 (October, 1956), pp. 815-816.

Carrell, James A. "The Etiology of Sound Substitution Defects," Speech Monographs, IV, No. 1 (December, 1937), pp. 17-37.

Cooper, Franklin S., Delattre, Pierre C., Liberman, Alvin M., Borst, John M. and Gerstman, Louis J. "Some Experiments on Perception of Synthetic Speech Sounds," Journal Acoustical Society of America, XXIV, No. 6 (November, 1952), p. 603.

Curry, Thayer E. "An Experimental Study of the Relative Identification Thresholds of Nine American Vowels," Speech Monographs, XVII, No. 1 (December, 1950), pp. 90-94.

Fairbanks, Grant and Grubb, Patti. "A Psychophysical Investigation of Vowel Formats," Journal Speech and Hearing Research, IV, No. 3 (September, 1961), pp. 203-219.

Fletcher, Harvey and Munson, W. A. "Loudness, It's Definition, Measurement, and Calculation," Journal Acoustical Society of America, V, No. 2 (October, 1933), pp. 82-108.

- Fry, D. B. "Duration and Intensity as Physical Correlates of Linguistic Stress," Journal Acoustical Society of America, XXVII, No. 4 (July, 1955), pp. 765-768.
- Gamble, Eleanor Acheson McCulloch. "The Applicability of Weber's Law to Smell," American Journal of Psychology, X, No. 1 (1898), pp. 82-142.
- Gemelli, A. and Pastori, G. "La durata minima delle vocali sufficienti alla loro percezione." (The Minimum Duration of Vowels Sufficient to Assure Their Perception), Archivio di Fisiologia, XXXIII, No. 3 (1934), pp. 440-452.
- Gilkinson, Howard. "The Seashore Measures of Musical Talent and Speech Skill," Journal of Applied Psychology, XXVII, No. 5 (October, 1943), pp. 443-447.
- Graham, C. H. and Cook, Carolyn. "Visual Acuity as a Function of Intensity and Exposure-Time," American Journal of Psychology, XLIX, No. 4 (October, 1937), pp. 654-661.
- Gray, Giles W. "Phonemic Microtomy: The Minimum Duration of Perceptible Speech Sounds," Speech Monographs, IX, No. 1 (December, 1942), pp. 75-90.
- Hall, Margaret E. "Auditory Factors in Functional Articulatory Speech Defects," Journal of Experimental Education, VII, No. 2 (December, 1938), pp. 110-132.
- Hansen, Burrell F. "The Application of Sound Discrimination Tests to Functional Articulatory Defectives With Normal Hearing," Journal of Speech Disorders, IX, No. 4 (December, 1944), pp. 347-355.
- Kranvall, Ernest L. and Diehl, Charles F. "The Relationship of Auditory Discrimination to Articulatory Defects of Children With No Known Organic Impairment," Journal of Speech and Hearing Disorders, XIX, No. 3 (September, 1954), pp. 335-338.
- Licklider, J. C. R. "On the Process of Speech Perception," Journal Acoustical Society of America, XXIV, No. 6 (November, 1952), pp. 590-594.
- Lindbloom, B. "Spectrographic Study of Vowel Reduction," Journal Acoustical Society of America, XXXV, No. 11 (November, 1963), pp. 1773-1778.
- Metraux, Ruth Watt. "Auditory Memory Span for Speech Defective Children Compared With Normal Children," Journal of Speech Disorders, VII, No. 1 (March, 1942), pp. 33-36.

- Morgan, C. T., Garner, W. R., and Galambos, Robert. "Pitch and Intensity," Journal Acoustical Society of America, XXIII, No. 6 (November, 1951), pp. 658-663.
- Munson W. A. "Growth of Auditory Sensation," Journal Acoustical Society of America, XIX, No. 4 (July, 1947), pp. 584-591.
- Oyer, Herbert J. "Duration of Homophones," Western Speech, XXIII, No. 2 (Spring, 1959), pp. 99-102.
- Peterson, Gordon E. "The Phonetic Value of Vowels," Language, XXVII, No. 4 (1951), pp. 541-553.
- _____, and Barney, Harold L. "Control Methods Used in the Study of Vowels," Journal Acoustical Society of America, XXIV, No. 2 (March, 1952), pp. 175-184.
- _____, and Lehiste, Ilse. "Duration of Syllable Nuclei in English," Journal Acoustical Society of America, XXXII, No. 6 (June, 1960), pp. 693-703.
- Rawnsley, Anita I. and Harris, J. Donald. "Studies in Short Duration Auditory Fatigue: II Recovery Time," Journal of Experimental Psychology, XLIII, No. 2 (February, 1952), pp. 138-142.
- Schwartz, Martin F. "A Study of Thresholds of Identification for Vowels as a Function of Their Duration," Journal of Auditory Research, III, No. 1 (January, 1963), pp. 47-52.
- Siegenthaler, Bruce M. "A Study of the Intelligibility of Sustained Vowels," Quarterly Journal of Speech, XXXVI, No. 2 (April, 1950), pp. 202-208.
- Sivian, L. J. and White, S. D. "On Minimum Audible Sound Fields," Journal Acoustical Society of America, IV, No. 4 (April, 1933), pp. 288-321.
- Small, Arnold M. Jr., Brandt, John F. and Cox, Phillip G. "Loudness as a Function of Signal Duration," Journal Acoustical Society of America, XXXIV, No. 4 (April, 1962), pp. 513-514.
- Spiestersbach, D. C. and Curtis, James F. "Misarticulation and Discrimination of Speech Sounds," Quarterly Journal of Speech, XXXVII, No. 4 (December, 1951), pp. 483-491.
- Stevens, Kenneth. "Effect of Duration upon Vowel Identification," Journal Acoustical Society of America, XXXI, No. 1 (January, 1959), p. 109.

- _____, and House, Arthur S. "Perturbation of Vowel Articulations by Consonantal Context: An Acoustical Study," Journal of Speech and Hearing Research, VI, No. 2 (June, 1963), pp. 111-128.
- Stevens, Stanley Smith. "The Relation of Pitch to Intensity," Journal Acoustical Society of America, VI, No. 3 (January, 1935), pp. 150-154.
- _____. "A Scale for the Measurement of a Psychological Magnitude: Loudness," Psychological Review, XLIII, No. 5 (September, 1936), pp. 405-416.
- _____, and Volkman, J. "The Relation of Pitch to Frequency, A Revised Scale," American Journal of Psychology, LIII, No. 3 (July, 1940), pp. 329-353.
- _____, Volkman J. and Newman, E. B. "A Scale for the Measurement of the Psychological Magnitude Pitch," Journal Acoustical Society of America, VIII, No. 3 (January, 1937), pp. 185-190.
- Stinchfield, Sara M. "Some Relationships Between Speech Defects, Musical Ability, Scholastic Attainment, and Maladjustment," Quarterly Journal of Speech Education, XIII, No. 3 (June, 1927), pp. 268-275.
- Sullivan, Margaret E. "Auditory Acuity and It's Relation to Defective Speech," Journal of Speech Disorders, IX, No. 2 (June, 1944), pp. 127-130.
- Templin, Mildred C. "A Study of Sound Discrimination Ability of Elementary School Children," Journal of Speech Disorders, VIII, No. 2 (June, 1943), pp. 127-132.
- Thurlow, W. R. "Studies in Auditory Theory: Binaural Interaction and the Perception of Pitch," Journal Experimental Psychology, XXXII, No. 1 (January, 1943), pp. 17-36.
- Thurston, L. L. "A Law of Comparative Judgment," Psychology Review, XXXIV, (1927), pp. 273-286.
- Tiffany, William R. "Vowel Recognition Conditions," Journal Speech and Hearing Disorders, XVIII, No. 3 (September, 1953), pp. 289-301.
- _____, and Bennett, Delmond N. "Intelligibility of Slow Played Speech," Journal of Speech and Hearing Research, IV, No. 3 (September, 1961), pp. 248-258.
- Travis, Lee Edward and Davis, Mildred G. "The Relation Between Faulty Speech and the Lack of Certain Musical Talents," Psychological Monographs, XXXVI, No. 2 (March, 1927), pp. 71-81.

- _____, and Rasmus, Bessie J. "The Speech Sound Discrimination Ability of Cases With Functional Disorders of Articulation," Quarterly Journal, Speech Education, XVII, No. 2 (April, 1931), pp. 217-226.
- Turnbull, William W. "Pitch Discrimination as a Function of Tonal Duration," Journal of Experimental Psychology, XXXIV, No. 4 (August, 1944), pp. 302-316.
- Wever, E. C. and Bray, C. W. "The Perception of Low Tones and the Resonance--Volley Theory," Journal of Psychology, III, No. 1 (January, 1937), pp. 101-114

UNPUBLISHED MATERIALS

- Deal, Leo V. "The Effects of Duration, Frequency, and Loudness Upon the Reproduction of Temporal Intervals," Unpublished Ph.D. Dissertation, Department of Speech, Michigan State University, 1965.
- Hartbauer, Roy Elden. "The Effects of Duration, Frequency, and Loudness Upon the Reproduction of Temporal Intervals by Sensori Neural Hearing Impaired Subjects," Unpublished Ph.D. Dissertation, Department of Speech, Michigan State University, 1967.
- Peterson, Gordon E. "The Significance of Various Portions of the Wave Length in the Minimum Duration Necessary for Recognition of the Vowel," Unpublished Ph.D. Dissertation, Department of Speech, Louisiana State University, 1939.

APPENDIX A

**RAW DATA USED IN THE ANALYSIS OF VARIANCE:
THRESHOLDS IN MILLISECONDS TABULATED
FOR EACH SERIES**

		DESCENDING SERIES						ASCENDING SERIES					
		First Half			Second Half			First Half			Second Half		
S ₃	i	10	10	10	6	6	10	14	14	10	6	14	10
	r	30	30	26	26	26	18	42	34	38	30	30	34
	e	18	14	18	14	18	18	26	22	26	18	22	18
	a	34	34	26	38	26	22	46	22	38	38	38	26
	u	34	14	22	10	26	30	26	22	6	22	18	26
	o	22	18	22	38	28	18	18	18	18	30	26	22
	u	14	18	6	18	14	2	14	18	2	14	14	6
	o	14	6	6	10	6	6	2	10	2	10	2	6
	v	6	14	14	18	18	22	14	14	22	14	14	18
	u	6	6	6	2	6	6	6	14	10	10	10	6
S ₄	i	22	18	10	14	14	10	10	10	14	14	6	10
	r	14	22	6	6	6	14	10	10	6	6	10	6
	e	6	6	6	14	14	14	10	6	10	10	10	6
	a	2	6	2	2	6	6	2	6	6	2	6	10
	u	14	14	6	10	10	10	22	14	10	18	6	6
	a	42	42	38	42	34	26	22	30	38	38	30	18
	u	14	22	30	22	22	30	26	26	22	14	34	34
	o	18	18	10	26	22	14	22	22	10	14	14	14
	v	10	10	10	10	10	6	14	6	6	6	14	6
	u	10	10	2	2	2	2	10	6	10	2	6	6

DESCENDING SERIES				ASCENDING SERIES			
	First Half	Second Half		First Half	Second Half		
S ₅	i	6 6 2 10 6	6 6 6 10 10	10 10 6 6 6	2 6 10 6 6		
	r	14 6 6 10 2	10 10 10 14 10	14 14 18 22 22	14 22 22 14 18		
	e	6 6 18 22 30	42 10 6 6 14	6 6 38 26 34	30 26 22 18 26		
	æ	18 18 10 2 6	2 6 6 6 10	22 10 10 6 6	6 6 6 6 2		
	ʌ	34 26 26 38 30	46 46 22 46 38	34 42 50 42 38	50 54 50 46 38		
	ɑ	30 22 18 22 30	22 26 34 26 18	34 30 18 18 22	18 10 10 18 10		
	ɔ	34 26 26 18 10	26 22 22 22 30	18 18 22 14 10	18 34 22 26 26		
	o	10 14 6 6 10	6 2 6 14 14	6 10 6 6 10	2 2 10 26 14		
	u	6 10 10 6 10	18 10 10 18 14	10 10 10 10 14	6 18 10 10 10		
	u	14 14 6 2 6	10 14 10 6 14	10 14 14 10 10	10 14 10 18 10		
S ₆	i	18 14 18 6 22	18 10 22 14 14	14 14 10 18 10	22 14 18 6 10		
	r	10 14 10 18 10	6 10 10 14 18	14 18 14 18 14	14 18 14 22 18		
	e	14 10 14 10 10	10 18 14 14 10	14 14 18 10 10	18 18 10 14 18		
	æ	14 14 18 18 22	14 14 18 14 18	18 14 18 22 26	18 18 10 18 10		
	ʌ	18 10 10 6 10	30 14 22 18 18	22 14 14 10 10	34 38 30 30 18		
	ɑ	46 36 38 34 22	22 18 22 22 34	30 26 22 26 18	18 22 14 22 10		
	ɔ	14 14 10 10 14	26 18 22 14 26	18 14 10 14 10	18 18 22 22 26		
	o	10 14 10 14 10	26 22 14 14 18	14 14 10 10 10	14 10 10 14 14		
	u	10 22 14 18 10	10 10 10 14 14	18 22 26 26 14	18 22 14 22 14		
	u	22 18 14 10 18	22 18 14 18 18	22 14 14 10 22	14 22 10 10 6		

APPENDIX B

**SUBJECT VOWEL TEMPORAL SEGMENT RECOGNITION THRESHOLD:
SERIES MEANS FOR THE; ASCENDING AND DESCENDING
SERIES; FIRST HALF, SECOND HALF,
AND OVER ENTIRE EXPERIMENT**

TEMPORAL SEGMENT RECOGNITION THRESHOLD
FOR ASCENDING SERIES IN MILLISECONDS

	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
S1	7.2	7.6	7.2	3.0	8.4	16.4	31.4	8.8	21.2	4.0
S2	13.6	17.6	15.2	14.4	20.2	39.2	16.4	16.8	12.4	14.0
S3	10.0	32.0	22.8	32.0	21.2	20.4	9.6	4.8	14.6	9.6
S4	10.8	9.2	7.6	6.0	11.2	27.6	28.8	14.0	9.2	5.6
S5	6.8	17.6	23.0	8.0	40.6	18.6	20.8	9.2	10.8	12.0
S6	13.6	16.4	14.4	17.2	22.0	20.8	17.2	12.0	19.6	13.2
Total	62.0	100.4	90.2	80.6	123.4	143.0	124.2	65.6	87.8	58.4
$\bar{X}$	10.33	16.73	15.03	13.43	20.57	23.83	20.70	10.93	14.63	9.73
SD	4.2	7.9	6.3	4.9	10.3	7.7	7.5	3.9	4.4	3.8

TEMPORAL SEGMENT RECOGNITION THRESHOLD
FOR DESCENDING SERIES IN MILLISECONDS

	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
S1	6.0	6.4	8.0	5.2	4.0	20.0	31.2	18.0	15.6	6.4
S2	12.8	9.6	5.6	14.8	12.0	50.0	18.4	14.8	12.0	10.4
S3	7.6	26.0	17.2	28.0	20.0	26.4	11.6	7.2	14.8	6.4
S4	12.8	9.2	9.6	4.0	8.8	33.2	28.8	16.8	9.6	3.3
S5	6.8	9.2	16.0	8.4	35.2	24.8	23.6	7.8	11.2	9.6
S6	15.6	12.0	12.4	16.4	15.6	29.6	14.6	15.2	13.2	17.2
Total	61.6	72.4	68.8	76.8	95.6	184.0	128.2	79.8	76.0	53.3
$\bar{X}$	10.27	12.07	11.47	12.80	15.93	30.67	21.37	13.30	12.73	8.88
SD	3.6	6.4	4.2	8.2	10.0	9.6	7.2	4.2	2.1	4.8

FIRST HALF MEANS

	i	l	s	ss	A	a	o	o	u	u
S1	6.0	8.0	6.7	3.6	6.0	17.2	30.4	13.2	15.6	3.6
S2	12.8	13.6	10.4	14.8	16.0	47.6	12.4	14.4	11.2	9.6
S3	9.6	31.2	19.6	34.0	20.0	23.8	12.0	6.8	14.8	7.6
S4	13.2	9.6	9.2	3.6	12.4	35.6	22.8	17.6	8.9	5.6
S5	6.8	12.8	19.2	10.8	36.0	24.4	19.5	8.4	9.6	10.0
S6	14.1	14.0	12.4	18.4	12.4	30.0	12.8	11.6	18.0	15.2
Total	62.5	89.2	77.5	54.6	102.4	178.6	108.9	72.0	78.0	51.6
$\bar{X}$	10.4	14.9	12.9	9.1	17.0	29.8	18.1	12.0	13.0	8.6
SD	3.2	7.6	3.0	11.6	11.0	9.8	6.9	3.6	3.5	3.6

SECOND HALF MEANS

	i	l	s	ss	A	a	o	o	u	u
S1	7.2	6.0	8.4	5.2	6.4	19.2	32.3	13.6	21.2	6.8
S2	13.6	13.6	10.3	14.4	16.0	41.6	22.4	17.2	13.2	14.4
S3	8.0	26.8	20.4	26.0	21.2	22.8	9.2	5.2	15.6	8.4
S4	10.4	8.8	10.8	6.4	7.6	25.2	34.8	13.2	10.0	4.0
S5	6.8	14.4	20.0	5.6	43.6	19.2	24.8	9.6	12.4	11.6
S6	14.8	14.4	14.4	15.2	25.2	20.4	21.2	15.6	14.8	15.2
Total	60.8	89.0	84.3	72.8	120.0	148.4	144.7	64.1	87.2	60.4
$\bar{X}$	10.1	14.0	14.0	12.1	20.0	24.7	24.1	10.7	14.5	10.1
SD	3.1	6.5	4.7	7.4	12.7	7.8	8.3	4.3	3.5	4.0

MEANS FOR COMPLETE EXPERIMENT

	i	l	e	e	A	a	c	o	u	u
S1	6.6	7.0	7.5	4.4	6.2	18.2	31.4	13.4	18.4	5.2
S2	13.2	13.6	10.4	14.6	16.0	44.6	17.4	15.8	12.2	12.0
S3	8.8	29.0	20.0	30.0	20.6	23.3	10.6	6.0	15.2	8.0
S4	11.8	9.2	10.0	5.0	10.0	30.4	26.6	15.4	9.4	4.8
S5	6.8	13.6	19.6	8.2	39.8	21.8	22.2	9.0	11.0	10.8
S6	14.6	14.2	13.4	16.8	18.8	25.2	17.0	13.6	16.4	15.2
Total	123.3	173.2	161.8	127.4	222.4	327.0	253.6	136.4	163.8	112.0
$\bar{X}$	10.27	14.43	13.48	10.64	18.53	27.25	21.22	11.36	13.8	9.33
SD	3.65	8.1	4.8	11.2	11.2	11.6	8.2	3.7	3.6	3.9

APPENDIX C

IDENTIFICATION MATRIXES FOR EACH SUBJECT OVER ENTIRE EXPERIMENT

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 1

Intended	Identified as									
Vowel	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(277)	10	1					1	1	7
ɪ	9	(276)	1		2			1	2	6
e		15	(274)		3		2	1	2	
æ		1	4	(288)	1	3	1	1		
ʌ			7	1	(281)	6	3		1	1
ɑ		4	11	2	62	(218)	1	1		
ɔ	2	1	11	1	116	8	(151)	9		1
o		6	5		19	1		(255)	3	
ʊ		10	8		22			38	(218)	3
u	4	9	1		2					(284)
ΣX-C*	17	56	49	4	328	18	7	52	9	18

*C equals number correct

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 2

Intended	Identified as									
Vowel	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(244)	13		1						1
ɪ	9	(243)	2							
e	1	6	(258)	1	1					
æ	1		5	(237)	1	5	2			
ʌ					(231)	20	2			
ɑ		1			167	(85)	3			
ɔ	2	1	9	4	9	1	(225)	5		
o		1		1	1			(231)	8	2
ʊ					2			10	(248)	3
u	1	2						2	1	(249)
ΣX-C*	14	24	16	7	181	26	7	17	9	6

*C equals number correct

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 3

Intended Vowel	Identified as									
	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(269)		7		4			3	1	15
ɪ	4	(178)		1	17	3	4	30		4
e	1	2	(213)	8	16		8	16	8	2
æ		1	14	(164)	13	38	63	3		
ʌ		2	1	2	(209)	77	4		2	2
ɑ	1	2	4	1	87	(194)	3	5	2	
ɔ	2		2		15	9	(260)	12	1	
o					6	2	2	(280)	7	
ʊ	2	3	3	2	10	1	1	35	(236)	2
u	9				8	1	1	5	5	(269)
ΣX-C*	19	10	31	14	176	130	86	109	26	25

*C equals number correct

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 4

Intended Vowel	Identified as									
	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(252)	15	1		1					24
ɪ	5	(264)	13		2			2	3	3
e	2	16	(260)		3			1	1	3
æ		5	6	(285)						4
ʌ		5	4	5	(262)	7	3		9	
ɑ		2	2		120	(160)	2		7	
ɔ	2	5	72	9	22	4	(169)	9	5	
o		2			7	3	5	(234)	39	4
ʊ	1	4	2		4		1	5	(264)	12
u									4	(286)
ΣX-C*	10	54	100	14	159	14	11	17	68	46

*C equals number correct

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 5

Intended	Identified as									
Vowel	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(277)	16	2							
ɪ	29	(265)	5					1		
e		5	(207)	4	1					
æ			28	(267)	1	1	1			
ʌ			3	2	(111)	154	2			
ɑ		2	19	2	71	(198)	1	4		
ɔ	2		23	5	19	30	(261)	17		
o		5	11		1			(264)	1	
ʊ	3	8	11					9	(255)	4
u	22	11			2			2	7	(256)
ΣX-C*	56	47	102	113	95	185	4	33	8	4

*C equals number correct

NUMBER OF TIMES EACH VOWEL WAS IDENTIFIED
AS EACH OTHER VOWEL BY SUBJECT 6

Intended	Identified as									
Vowel	i	ɪ	e	æ	ʌ	ɑ	ɔ	o	ʊ	u
i	(237)	20						1		4
ɪ	7	(241)	2						1	4
e		6	(242)	2		1	1	2	1	
æ		2	28	(228)		2	1			
ʌ					(216)	53	1	2	1	
ɑ					75	(183)	4			
ɔ	3		10	2	8	4	(230)	10		33
o	1		1		2	1	3	(240)	10	
ʊ	1				9		1	25	(226)	2
u	1	1						1	22	(234)
ΣX-C*	13	29	41	4	94	61	11	41	35	43

*C equals number correct

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03175 5600