THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A FORMANT STUDY OF WHISPERED VOWELS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

WENDELL FRANKLIN SMITH

Oklahoma City, Oklahoma

1973

A FORMANT STUDY OF

WHISPERED VOWELS

APPROVED BY

Floyd W. Emanuel

Donald Counihan

Walter L. Gilliam

Glenda J. Ochsner

Donald C. Parker

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

Sincere appreciation is expressed to Dr. Floyd W. Emanuel, director of this study, for his encouragement and guidance throughout the planning and completion of the investigation. The author is also grateful to the other members of the dissertation committee: Dr. Donald T. Counihan, Dr. Walter L. Cullinan, Dr. Glenda J. Ochsner, and Dr. Donald E. Parker, for encouragement and counsel during this investigation and throughout his training at the University of Oklahoma. Thanks are due also to a fellow doctoral student, Mr. Donald G. Brennan, for his assistance in the collection of the data, and to the subjects and judges who unselfishly participated in this study.

The author wishes to express special appreciation to his parents, Dr. and Mrs. Cloyd V. Smith, Sr., for their moral and financial support during his graduate education. Finally, the author is most profoundly grateful to his wife, Carol, and his son, Jonathan, for their love and patience during this often trying period in their lives.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF ILLUSTRATIONS

A FORMANT STUDY OF

WHISPERED VOWELS

CHAPTER I

INTRODUCTION

Efforts to understand the relationship of acoustic vowel features to the perception of individual vowel phonemes have been traced to the fourteenth century, when investigators attempted to produce vowel sounds artificially with reed pipes (49). During the nineteenth century, investigators including Willis, Wheatstone, and Helmholtz learned that phonemic vowel quality was associated with the natural resonances, or formants, of the supraglottic vocal tract (49). These early scientists did not, however, quantify precisely the resonance characteristics of individual vowels.

In the mid-twentieth century, the development at Bell Telephone Laboratories of a heterodyne-type sound spectrograph, now called the Sonagraph, greatly aided the resolution of vowels (and other speech sounds) into their acoustic energy components and permitted visualization of those components in a graphic plot or sonagram (21). The Sonagraph has since become the primary laboratory instrument for vowel formant studies. A conventional sonagram shows the acoustic components of an analyzed complex signal on a graph which has a vertical frequency scale and a horizontal time scale. The intensity of the acoustic components is indicated by gradations in the darkness of the plot within the frequency-time axes. For sustained, isolated vowels, the acoustic energy at formant frequencies is shown by dark horizontal bars. Vowel formants can also be visualized in a sonagraphic frequency-by-intensity "amplitude section" as peaks in the spectral envelope (20). Fant (12) has indicated that a vowel may be described with respect to its phonemically relevant acoustic features by obtaining measurements, in a Sonagraphic amplitude section, of the frequency, effective bandwidth, and amplitude of the first three or four formants.

Problems are encountered, however, in the delineation of vowel formants by Sonagraphic analysis. In some respects, the filter bandwidths of the Sonagraph (usually 45 or 300 Hz) are too narrow for optimum formant resolution. Analysis of the quasi-periodic complex acoustic waves for phonated vowels which manifest a fundamental frequency equal to or greater than the Sonagraph's filter bandwidth results in the spectral resolution of the fundamental frequency and its higher harmonics. Vocal tract resonances (formants), however, do not necessarily coincide with the harmonic peaks and may occur between the harmonics (17). Thus, harmonic resolution tends to interfere with the accurate delineation of formant frequencies.

It would seem that formants for whispered vowels might be clearly delineated with relative ease by acoustic spectrography. Because whispered vowel acoustic waves lack periodicity, they are characterized by continuous noise spectra which lack the harmonics that obscure formant locations in spectra for phonated vowels. Few investigators have utilized the Sonagraph to study the acoustic features of whispered vowels, however, apparently because of limitations in its power to resolve acoustic components. For example, the acoustic filter bandwidths of the Sonagraph tend to be too broad to resolve individually those whispered vowel formants which occur very close together in frequency, and formant bandwidths and amplitudes for whispered vowels are shown somewhat inaccurately in sonagrams (23, 30, 46). Additionally, some whispered vowels are characterized by relatively low amplitude formants which simply may not be plotted in sonagrams (30, 46). The Sonagraph does not plot low intensity acoustic energy and is not equally sensitive to energy levels across its frequency range.

In the few acoustic studies of whispered vowels which have been reported (23, 26, 30, 39, 46), the Sonagraph was utilized to obtain the vowel spectra; hence, the studies provide only estimates of formant frequencies. Further, those studies were not designed to contribute formant bandwidth and amplitude measures. The findings presently available regarding whispered vowel formant frequencies, bandwidths, and amplitudes are thus incomplete, and studies which provide more complete data regarding the major formant parameters of whispered vowels are needed. The new information provided by such studies would seem important to a comprehensive description of whispered vowels and could aid in understanding vocal tract resonance characterisitics during vowel production.

Advances in acoustic analysis instrumentation since the development of the Sonagraph make possible a more accurate spectrographic representation of vowel components and, thus, enhance the precision and accuracy which may be achieved in obtaining formant frequency, bandwidth, and amplitude measures for whispered vowels. Heterodyne-type spectrographic instruments now available, for example, permit very narrow-band constant-bandwidth acoustic wave analysis and automatic spectral plotting (27). If applied in the study of whispered vowels, such analyzers could overcome major problems associated with Sonagraphic analysis. Specifically, some very narrow-band analyzers can record accurately a range of spectral energy components from very low to high amplitude and can also resolve individual whispered vowel formants which are very close together in frequency. No study has been reported, however, in which whispered vowels were analyzed with such instrumentation.

It was the purpose of this investigation to study the acoustic spectral features of selected whispered vowels produced by adult male and female subjects. A very narrow-band (3-Hz) constant-bandwidth wave analyzer was used to obtain frequency-by-intensity spectra of individual recorded whispered vowel samples. Measurements of the frequency, effective bandwidth, and amplitude of the first three formants of each test vowel production were obtained. In the following chapter, the literature reviewed as background for this study is reported.

CHAPTER II

REVIEW OF THE LITERATURE

Acoustic Theory of Vowel Production

The "mechanism" for vowel production is commonly said to include as major components the lungs and related structures which provide the driving air pressures and flows, the glottic "sound source," the supraglottic resonator, and the oral opening through which sound is transmitted into the external atmosphere. An overview of concepts regarding the function in vowel production of the laryngeal and supralaryngeal "components" of the mechanism is presented in this section.

With regard to vocal sound generation, writers including Fant (10, 11), Curtis (5), and Broad and Peterson (4) have discussed two basic types of human vocalization which may be used to produce vowels. First, vowels may be phonated or "voiced." During phonation, vocal sound results when the exhaled air stream is modulated by the rapid opening-closing movements of the vocal folds. The vibratory action of the folds, powered by subglottic air pressure, causes a quasi-periodic emission of air puffs through the glottis which, in turn, excites the supraglottal air column and produces a complex, audible acoustic wave. That acoustic wave manifests a fundamental frequency which corresponds to the number of glottic opening-closing cycles (or the number of air puffs emitted glottally) per second. According to theory (4, 10, 13, 40), the volume-velocity wave of the air flow through the glottis during phonation may be represented (to a first approximation) by a Fourier line spectrum with components at integral multiples of the fundamental frequency. The amplitude of the harmonic components decreases with increasing frequency at a rate of approximately 12 dB per octave.

Vowels may also be whispered. In whispering, a quasi-random noise is generated when air in the supraglottic spaces is set into vibration by a sustained turbulent air flow driven by subglottic pressure through a narrowly-constricted but partially-open glottis (4, 23, 30, 49). Such sounds have a continuous acoustic spectrum which, in contrast to that for phonated vowels, is comparatively flat across frequencies (4, 5).

The resonator component of the vowel-producing mechanism functions in a manner somewhat analogous to that of an electrical filter circuit to which a complex input wave is applied (10, 41, 44). That is, the supraglottic vocal tract acts as a frequency-selective filter which damps or diminishes, more at some frequencies than at others, the simple (sinusoidal) components of the complex acoustic wave. The filter (or transfer) function of the vocal tract (just described) represents the frequency response of the resonator system. The transfer function is thought to be essentially independent of the previously-described source function and to depend almost exclusively on the shape of the vocal tract (2, 4, 5, 10, 15, 24, 34, 36, 40, 41, 45). The shape of the vocal tract is determined by a number of factors including primarily its length and volume, and lingual posture within the tract (4, 10, 24, 34, 45). The nasal cavity may be coupled slightly to the supraglottic vocal tract during vowel production and, to the extent that it occurs, coupling may affect the resonance characteristics of the tract (4, 5, 10, 11, 14). Because such coupling is normally minimal during production of English vowels, however, it is commonly regarded as being of little practical importance.

It is useful to consider that the vocal tract functions as a continuous acoustic tube, variable in shape, with a number of natural resonances called "formants" which are determined by shape (4, 10, 24, 41). The label "formant" may also be applied to the effects of vocal tract resonance as they are visualized in the acoustic vowel spectrum; that is, to energy peaks within the spectrum (4, 5, 10, 40). The frequency of these formants appears to be the primary acoustic correlate of phonetic vowel quality (4, 5, 10, 40, 43).

The resonator-modulated vocal sound is finally emitted through the opening between the lips into the atmosphere. The radiation of the vocal sound has a further acoustic damping influence which is greater for low than for high frequencies, and radiation is thus associated with a modification in the slope of the output acoustic spectrum (i.e., an increase in the slope from low to high frequencies) of approximately 6 dB per octave (4, 10, 40). The effects of spectral modifications due to radiation appear to have relatively little influence on phonetic vowel quality (4, 11).

To summarize, in human vowel production a sound (either quasi-periodic or noise) is generated by laryngeal action on the expiratory air stream. This sound acquires a phonetically significant quality mainly as the result of frequency selective acoustic damping in the vocal tract and secondarily as the result of damping effects which accompany emission of the sound from the mouth into the atmosphere. The main damping effects are attributable to vocal tract shape.

Vowel Formant Features

It is pertinent to consider in some detail the primary spectral features of vowels, i.e., the resonant peaks or formants. As noted above, the vocal tract manifests several natural resonances which vary in their major parameters with vocal tract shape. Fant (11, 12) suggests that acoustic characteristics essential to vowel phoneme identification may be adequately described by the specification of three formant parameters: the frequency of the formant peak, the half-power (effective) bandwidth, and the amplitude. The results of experiments in vowel synthesis (1, 11, 13, 41, 44) suggest that clearly recognizable vowels may, in most cases, be produced when the frequency, bandwidth, and amplitude of only the two lowest frequency formants are specified. For some vowels, however, the specification of three formants is necessary for optimum vowel representativeness. Thus, in studies of vowel formant properties, investigators usually describe the first three formants (1, 9, 14, 18, 33, 35, 41, 45).

Formant frequency is the formant parameter which has been studied most often. Vowel formant frequencies may be defined operationally by locating the formant peaks on the spectral frequency scale. Major Sonagraphic investigations of formant frequencies for phonated vowels were contributed some time ago by Peterson and Barney (33) and by Fairbanks and Grubb (9). Fant (12), who summarized the findings from those and other previous investigations, reported that the natural range of formant frequency variations for vowels phonated by adult male subjects is approximately as follows: formant one (F1), 150-850 Hz; formant two (F2), 500-2500 Hz; and, formant three (F3), 1700-3500 Hz. Fant noted further that formants for adult female productions are approximately 20% higher in frequency than those for males, and those for children even higher than those for females. The higher formant frequencies for females and children are associated with smaller vocal tract sizes.

Vowel formant frequencies are markedly different for different vowel phonemes, but formants for the same phoneme produced by different subjects vary within narrow limits (9, 14, 33, 35, 45). Slight, statistically nonsignificant formant frequency variations have also been noted among repeated productions of the same vowel by an individual subject due, apparently, to slight, phonemically nonsignificant variations in vocal tract shape across productions (33). Generally, the formant frequency measures given in research reports represent means and ranges over productions by a number of subjects.

The second vowel formant parameter of interest is effective bandwidth. On a frequency-by-intensity plot (linear in dB SPL), bandwidth is defined as the frequency difference between the two points on either side of the formant peak that are 3 dB below the peak level (12). Formant bandwidths, which seem to reflect mainly the selective acoustic damping characteristics of the vocal tract due to resonance, are thought to be affected secondarily by acoustic energy losses due to: sound energy radiation from the mouth; energy absorption by the walls of the vocal tract; energy losses through the glottal opening; and, sound absorption into the nasal cavity (19).

House and Stevens (19) and Fujimura and Lindqvist (14), in separate studies of the response of the vocal tract to externally-applied signals, compared formant bandwidths for open- and closed-glottis conditions. In both of the above-cited studies, the investigators reported wider formant bandwidths for the open-glottis condition which, they concluded, was probably attributable to acoustic energy losses associated with the coupling of the trachea to the supraglottal system. Fujimura and Lindqvist (14) also presented data which suggest that vowel formant bandwidths may vary with vocal tract size. They reported wider formant bandwidths for female than for male productions.

Reports of the magnitude of formant bandwidths characteristic of different vowels have been quite disparate across studies (2, 3, 7, 14, 18, 19, 30). Dunn (7) noted, in a review of several early investigations of formant bandwidths for phonated vowels, that the lack of agreement across studies may relate in part to errors inherent in the methods used to obtain formant bandwidth measures. For example, because measurements of formant bandwidths depend for accuracy on the precise location of the peak formant level, large errors may be made in estimating from acoustic spectra the formant bandwidths for phonated vowels. The actual formant peak for phonated vowels may occur between harmonics and, because the peak is thus not visible in the vowel spectrum, both the formant frequency and the peak formant level may be estimated inaccurately (7, 17).

No study was found in which formant bandwidths were reported for whispered vowels. Fant (12) observed that

40-250 Hz represents probable bandwidth limits for the first three formants of voiced vowels, and that 100 Hz probably represents a typical average formant bandwidth value. He noted, however, that "formant bandwidths are not very critical for the phonetic quality of a sound" (11). House (18), on the other hand, has demonstrated a slight bandwidth influence on the perceived "naturalness" of electrically-synthesized vowels. Generally, synthesized vowels with relatively narrow formant bandwidths were perceived to be more "natural."

The third vowel formant parameter of interest is peak amplitude. The level of the spectral resonant peaks for vowels is measured in decibels relative to a reference level, generally 0.0002 dyne/cm^2 (11, 12). In some investigations, formant amplitudes have been considered relative to the amplitude of a reference formant. Peterson and Barney (33), for example, reported for voiced test vowels the amplitude of each of the first three formants relative to the mean level (over all of their subjects--male and female adults and children) of the first formant of the vowel /ɔ/. The amplitude values they obtained were corrected to compensate for a positive slope (from low to high frequencies) in the frequency response of the Sonagraph.

Stevens and House (40) note that the relative amplitudes of voiced vowel formants vary markedly across vowels, and depend mainly on the frequencies of the vocal tract resonances associated with each vowel. Due to the combined influences of the amplitude decrease in components of the glottal volume-velocity wave with increasing frequency and the effects of vocal tract damping and damping due to acoustic radiation from the mouth, the level of the first formant of voiced vowels is always greater than that of higher-frequency formants (40). As will be shown in a later section, the above observations regarding relative formant amplitudes would not be expected to apply exactly for whispered vowels, but amplitude measures for whispered vowel formants are not presently available.

To summarize, vowel formants may be described acoustically in terms of three major parameters: frequency, bandwidth, and amplitude. The frequency of the formants is the parameter which appears to be influenced most by alterations in the vocal tract "shape" and appears to be the primary acoustic correlate of phonemic vowel quality.

Methods of Vowel Wave Analysis

The present era of investigation into vowel acoustic spectral features began with the development at Bell Telephone Laboratories of the sound spectrograph (21), later marketed as the Kay Electric Sonagraph. This instrument produces a time-frequency-intensity plot of up to 2.4 seconds of an acoustic signal. The Sonagraph typically has two selectable filter bandwidths, 45 and 300 Hz. The spectrogram of an isolated vowel sound is characterized by dark horizontal bands at various levels along the vertical frequency scale. These bands correspond to the vocal tract resonances (formants). The Sonagraph is also capable of producing a spectral frequency-by-intensity plot, or "amplitude section," at selected points along the 2.4-second time scale (20). In spectral sections of vowels, vocal tract resonances are reflected as vertical energy peaks in the spectral envelope (4, 11, 17, 20, 40).

Although the Sonagraph has been the primary instrument for vowel analysis in numerous studies (3, 7, 9, 30, 33, 36, 45), its usefulness for delineating vowel formants is limited by its filter bandwidth characteristics. If the analysis bandwidth is less than or equal to the fundamental frequency of the test vowel phonation, the filter will resolve individual harmonics and the formants may be obscured. Fant (11, 12) concluded that the Sonagraph, used in its 300-Hz mode, is most useful in analyzing the low-frequency adult male vowel productions (with a fundamental frequency well below the analyzer filter bandwidth) and is not equally suited to the analysis of productions by adult females or children which manifest higher fundamentals.

Another method for the study of acoustic vowel features has been termed "analysis-by-synthesis" (1). In such studies, a recorded human vowel production may be compared to a vowel generated within an electronic synthesizer according to rules derived from acoustic vowel theory. The synthesizer generates vowel formant combinations which correspond to computed vocal tract transfer functions, and a measure of error is obtained between the internally-generated signal and the recorded human vowel signal. When a synthesized spectrum that provides a minimum "error" is achieved, the known formant characteristics of the internally-generated spectrum approximate those of the matched human vowel spectrum.

The usefulness in research of the analysis-by-synthesis technique depends in large part on the speed and accuracy of the analysis. The comparison of human and synthesized productions is tedious and time consuming when done manually. Paul, House, and Stevens (28) describe a rapid computer technique for automatic analysis-by-synthesis. The recorded wave of the human vowel sample is automatically digitized and the digital data are stored in computer memory. Spectra of the stored input vowel signal are obtained automatically every 8.3 msec.; frequency is sampled over a 7 KHz range, and the amplitude of components is specified to the nearest decibel. A series of synthesized vowel spectra are generated and compared to the time-averaged input spectrum which is retrieved from temporary storage. The number and frequency of formants in the synthesized spectrum are automatically adjusted until a "best match" between the input and synthesized spectrum is obtained. The formant frequencies thus specified in the synthesized spectrum approximate those for the input spectrum.

Fujimura and Lindqvist (14) describe a method for measuring vocal tract resonance characteristics directly without requiring that the subject actually vocalize a test production. An electrically-generated acoustic wave is introduced into the frequency-selective vocal tract via a moving-coil-type electromagnetic transducer. The subject holds the transducer to his neck in a manner similar to that used in speaking with an electro-larynx. Two types of waves are available as possible inputs to the transducer: a buzz signal from a pulse-train generator and a sinusoidal signal from a beat-frequency-oscillator (BFO). The buzz signal is applied first as the subject sets his articulators for the desired vowel. When the subject has assumed the desired articulatory set, the investigator switches from the buzz to the sinusoidal acoustic input. The subject holds the articulatory set as constant as possible, with his glottis closed, while the BFO sweeps upward in frequency from 100 to 5000 Hz in a time of about 8.5 seconds. The output signal is received at the subject's mouth by a condenser microphone placed one centimeter in front of the lips. The condenser microphone output is then led to a high-speed recorder which plots a frequency-response curve for the vocal tract.

This sweep-tone method is claimed to have two advantages. First, the continuous-frequency curves that are obtained can reveal details of the vocal tract transfer function, within the response limits of the recording system, without the obscuring influence of the harmonics of natural voiced vowels. Additionally, the obtained transfer function is unaffected by features of the source function which may vary within and across subjects. One obvious disadvantage of this method, however, is an unspecified alteration in the transducer's frequency response due to transmission of the signal through the body wall. Fujimura and Lindqvist (14) contend that it is not essential to assume a flat transfer characteristic of the body wall to obtain accurate relative response curves for various vowel articulations. They reason that it is only necessary to assume a constant transcutaneous transmission characteristic during the comparatively short period of the experimental session.

Fujimura and Lindqvist indicate that the placement of the vibrator on the neck is critical to the output frequency response. They reported, however, that after several trials each subject was able to locate the transducer placement which produced the most stable vocal tract frequency response. The obtained response curves were matched with comparable curves obtained for electrically-synthesized vowels. Differences between the observed sweep-tone curves and those for synthesized productions were, for the most part, less than ± 1 dB at any test frequency. While this sweep-tone procedure appears to provide a reasonably accurate estimate of vocal tract formants in response to an externally-applied sound source, the results may not describe precisely the formant features of isolated vowels whispered by human subjects.

A relatively recent development in the direct recording of frequency-by-intensity acoustic vowel spectra has been the narrow-band constant-bandwidth wave analyzer which is mechanically synchronized with a graphic level recorder. This instrument has been promoted commercially primarily for its usefulness in the analysis of environmental and industrial noises, but was utilized lately in a series of speech studies by investigators (6, 16, 25, 38, 47, 48) who sought to delineate precisely the levels of inter-harmonic acoustic energy, or spectral noise, associated with vowel productions. The cited investigators did not study the formants in their vowel spectra. It would appear that very narrow-band wave analysis may offer advantages in the study of formant features for whispered vowels. The spectrum of a sustained whispered vowel obtained by such analysis manifests no harmonic components to hinder the spectral delineation of the formants. Narrow-band filtering of whispered productions also permits the resolution of individual formants that are very close together in frequency. Such improved vowel formant delineation enhances not only the precision with which the formant peaks may be specified, but also helps to minimize errors in measuring the half-power (effective) bandwidths of the formants. Additionally, the accurate spectral resolution of low as well as high amplitude energy components makes possible more accurate measurements of formant peak amplitudes. Presently, however, no study of whispered vowels has been reported in which the spectra were obtained by very narrow-band wave analysis. It was the purpose of this investigation to utilize very narrow-band (3-Hz) constant-bandwidth wave analysis to obtain spectra of sustained whispered vowel productions from which measurements of formant frequency, effective bandwidth, and peak amplitude could be obtained. In the following section investigations of the acoustic features of whispered vowels are reviewed.

Studies of Whispered Speech

Zemlin (49) notes that "the essential difference between vocalization and whispering lies in the configuration of the glottis during exhalation, and the resultant acoustic product." Although whispering may be associated with varying degrees of glottal openness, investigators (37) who have studied the physiology of whispering report that the vocal folds tend to be slightly more adducted in a low-volume whisper than in quiet respiration. The air-flow turbulence created when the exhaled air stream is forced through the partially-open glottis produces a "friction" sound which is essentially aperiodic in nature; thus, it possesses no fundamental frequency. Zemlin (49) reports that whispered speech cannot be inflected easily, and only slight modifications in the intensity of whispered sounds are possible.

Meyer-Eppler (26) studied prosodic features of whispered speech. He observed that in languages where vocal pitch is not phonemic (i.e., does not carry linguistic meaning), whispered speech is clearly understood and, thus, must carry the acoustic information of linguistic importance that is also present in voiced speech. He analyzed samples of whispered vowels using the Kay Sonagraph and noted the presence of formants, but he did not measure the major formant features: frequency, bandwidth, and amplitude. He indicated, however, that when his subjects attempted to whisper the vowels at different "pitches" (that is, produce rising and falling inflections), there was an apparent change in the "spectral structure of the vowels within the limits of recognizability" (26).

Peterson (30), on the basis of an analysis of whispered vowels using the Sonagraph, commented on the tendency for some whispered vowel formants to be shifted upward in frequency in comparison to those reported for phonated vowels. Although he did not report measurement data regarding formant bandwidths or amplitudes, he did note that in many instances spectral energy in the region of the first formant of the whispered vowels was relatively weak in intensity with respect to that for the higher-frequency formants. Peterson also noted that the Sonagraph did not produce a clear spectral representation of isolated whispered vowels, and he utilized a mechanical vibrator, held against the throat, to supplement the intensity of the output signal. He suggested that apparent formant shifts in whispered vowels may be attributable, in part, to minor articulatory adjustments made as a result of subjects' conscious efforts to produce whispered vowels which were perceptually equivalent to voiced vowels. A possible additional reason for articulatory adjustments during whispered vowel productions, Peterson suggested, was to compensate for the absence of a fundamental frequency.

Lehiste (23) also studied whispered vowels using the Sonagraph. She reported data regarding formant frequencies but none regarding bandwidths and amplitudes. She noted that F1 was approximately 200-250 Hz higher in frequency, and F2 and F3 were approximately 100-150 Hz higher, in whispered than in voiced vowels. She attributed those formant frequency differences to a difference in the degree of glottal openness during phonation and whispering. That is, she noted that during phonation the glottis may be considered to be effectively closed, while during whispering the glottis is never completely closed. She reasoned that whispered vowels should be expected to manifest formants which are higher in frequency than those for voiced productions, because the resonant frequencies of an acoustic tube open at both ends are higher than those of a tube closed at one end (15, 23).

Schwartz (38) studied the overall intensity range of connected whispered speech samples and noted that the samples showed a reduced intensity range with respect to that for voiced samples. He did not, however, investigate formant features of whispered vowels.

Thomas (46) studied the perceived pitch of whispered vowels produced by one male and one female subject. He also used the Sonagraph to analyze his whispered vowel samples, but he found that the formant resolution was not clear. Hence, the formant frequencies he reported were estimates obtained both from the formant bars on the conventional wide-band sonagram and by multiple amplitude sections of each sample. The formant frequencies obtained by the two analysis methods he employed were different, and the values he reported were means of those obtained by the two methods. Music students who served as listeners matched pure-tones as closely as possible to the "pitch" of the samples. Thomas reported that without exception the perceived "pitch" of each whispered vowel production corresponded very closely to the frequency of its second formant.

Thomas made some general descriptive comments regarding whispered vowel formant bandwidths and amplitudes. He noted with regard to formant bandwidths, for example, that some first formants manifested a relatively broad bandwidth (range undefined) and some third formants manifested a relatively narrow bandwidth (approximately 100-200 Hz). He made no comments regarding F2 bandwidths. With regard to formant amplitudes, Thomas noted that the amplitudes of F1 and F2 were approximately equal for all whispered vowel samples produced by his two subjects. Additionally, he noted that the energy level of F3 for the back vowels was relatively low and that F3 was sometimes not apparent in the amplitude section. Thomas did not report measurement data regarding whispered vowel formant bandwidths and amplitudes.

To summarize, while research studies of whispered vowels have been few, the available data indicate that formant frequencies for whispered vowels tend to be higher than those for comparable voiced productions. Data regarding whispered vowel formant frequencies are presently incomplete due largely, it appears, to limitations in the Sonagraphic presentation of the acoustic components of whispered vowels. Further, no study has been reported which was designed to investigate systematically the bandwidth or amplitude of whispered vowel formants, though such data appear to be necessary to a complete acoustic description of whispered vowels.

In the present investigation, frequency-by-intensity spectra of selected whispered vowels, produced by adult males and females, were obtained by a method (very narrow-band constant-bandwidth wave analysis) selected to avoid limitations associated with the more conventional method of Sonagraphic analysis. In each narrow-band whispered vowel spectrum the frequency, effective bandwidth, and amplitude of each of the first three formants were measured. The following chapter describes in detail the design of the study.

CHAPTER III

DESIGN OF THE INVESTIGATION

It was of primary interest in this study to investigate the formant features of isolated whispered vowels produced by young adult subjects. The research questions considered and the methods employed in the investigation are presented in this chapter.

Research Questions

The following research questions concerning selected whispered vowels were investigated.

1. What is the frequency of each of the first three test vowel formants?
2. What is the effective bandwidth of each of the first three formants?
3. What is the relative amplitude of each of the first three formants?

Subjects

Twenty-four normal-speaking young adults, twelve male and twelve female, served as subjects. The subjects who ranged in age from 21 to 30 years were all graduate students in communication disorders.

Speech Sample

The subjects individually sustained in a whisper, at one intensity and mouth-to-microphone distance, each of four vowels /ɪ/, /æ/, /u/, and /ɔ/. These phonemes were selected for study because they represent relatively stable vowel configurations with respect to lingual posture and degree of lip rounding (9, 15, 24, 33). With the exception of /ɔ/, productions of each of the test vowels selected also tend to be readily distinguished perceptually from productions of other vowels (9, 15, 33). The vowel /ɔ/ may not be easily distinguished from similar vowels by some individuals; specifically, those who have lived mainly in geographic regions where that vowel is infrequently used (9, 15, 33). Hence, it was expected that it might be difficult to obtain perceptually distinct samples of /ɔ/ for the present study. The vowel /ɔ/ was included, nevertheless, to permit a comparison of formant amplitude measures obtained for the present whispered vowel samples to comparable measures obtained by Peterson and Barney (33) for voiced vowel samples. Peterson and Barney reported the amplitude of voiced vowel formants with respect to that of the first formant of /ɔ/.

Instrumentation

Instrumentation used in data collection included an audio recording system, a wave analyzing system, and a playback system. The following is a description of each system.

Audio Recording System

The audio recording system consisted of a sound level meter (General Radio, Type 1551-C) with an attached non-directional PZT ceramic microphone (General Radio, Type 1560-P3); a laboratory-quality magnetic tape recorder (Ampex, Model AG 440); and, a monitoring amplifier (Bruel and Kjaer, Type 2603).

Wave Analyzing System

A recording wave analyzer assembly (General Radio, Model 1910-A) was used to obtain acoustic vowel spectra. Additional instrumentation utilized for frequency and intensity calibration of the wave analyzer included an audio oscillator (Hewlett-Packard, Model ABR200), a universal counter (TSI, Model 361), and the sound level meter listed above. A more detailed description of the wave analyzing system and the procedures employed to insure that the system remained in calibration during data collection are presented in Appendix B.

Playback System

The aforementioned Ampex tape recorder, an impedance matching transformer, and a loudspeaker (Altec, Model 844A) were used as the playback system for vowel judgments.

Procedures

Recording Procedure

All vowel samples were recorded in an acoustically-isolated room with a low ambient noise level. Before recording the test vowel samples, a prepared statement was read to each subject (Appendix A, Instructions to Subjects) which explained the nature of the experimental task and the importance of a careful production of each test sample. The subject was seated in an examination chair and the sound level meter's microphone was placed at a 70° angle of incidence to and three inches in front of the subject's mouth. The Bruel and Kjaer amplifier was utilized to aid in monitoring the intensity of test productions. It was so calibrated that, when a subject's whispered vowel production deflected the amplifier's VU meter to a pre-set mark, the intensity of the production at the microphone was 55 dB (± 1 dB) re 0.0002 dyne/cm^2. Each test vowel production was sustained for five seconds. The duration of test productions was controlled by a system of signal lights which, together with the intensity-monitoring amplifier, was in the subject's field of vision. The signal lights were controlled by a cam timer located outside the test room. All the subjects were able to sustain the whispered vowel productions at 55 dB SPL (mouth-to-microphone distance three inches) for five seconds without "exaggerated" effort.

A microphone wind screen was used to prevent the expiratory air flow associated with the sustained whispered productions from distorting the recorded audio signal at the close mouth-to-microphone distance used. To evaluate effects which the wind screen might have on the acoustic spectrum of an audio signal, the Altec loudspeaker was placed one foot from the microphone of the sound level meter and a white noise signal, produced by a noise generator (Grason-Stadler, Model E5539A), was led to the loudspeaker. With the Ampex tape recorder's VU meter set at -2 dB, recordings of the white noise were made with and without the wind screen in place. Individual tape loops were made of two-second sections of each recorded noise signal, and these were analyzed separately (using the General Radio wave analyzer assembly and procedures discussed later) to obtain an intensity-by-frequency acoustic spectrum of each signal. Noise levels at comparable frequencies in the spectra (spectral frequency range 0-8000 Hz) of the two noise samples were essentially the same. Thus, it appeared that use of the wind screen would not affect the spectra of the whispered test vowels.

Each subject received practice in whispering samples of each test vowel and some additional vowels which were selected. It has been reported previously (9, 33) that isolated test vowel productions are frequently judged by listeners to represent a phoneme other than that intended by the speaker. In most instances, the vowels which listeners commonly confuse occur as adjacent pairs on the traditional vowel triangle. For example, /i/ and /I/, /æ/ and /ɛ/, /ɔ/ and /ɑ/, and /u/ and /U/ are frequently confused phoneme pairs. To increase the probability that representative samples of the four test vowels would be obtained in this study, therefore, each subject practiced producing each test vowel (/i/, /æ/, /ɔ/, /u/) and also the vowel with which each test vowel is most often confused. This procedure appeared useful to emphasize for the subject the care needed for the accurate production of each test vowel.

When the investigator, who was present in the test room with the subject, thought that a subject had received sufficient practice to produce test vowel samples which would be satisfactory with respect to phonetic representativeness and production intensity, two five-second samples of each test vowel were recorded for each subject (vowel order was randomized anew for each subject). Test samples which did not meet the experimental criteria were discarded, and the experimental procedure was repeated until two acceptable samples of each test vowel were recorded for each subject. The second production of each test vowel by each subject was obtained to evaluate intra-subject reliability. In all, 192 whispered vowel test samples were obtained (24 subjects X 4 test vowels X 2 productions = 192).

Wave Analysis

Acoustic spectra of the test vowel productions were required to obtain measurement data relevant to the research questions. To produce the needed spectra, tape loops were constructed from a two-second central portion of the recording of each whispered vowel production. The loops were constructed from the portion of the vowel recording displaying a uniform intensity as monitored on the recorder's VU meter. The loops were then individually played, and the output of the tape recorder was led to the General Radio wave analyzer, to produce frequency-by-intensity spectra showing the first three formants of each production. The analyzer was operated in its 3-Hz bandwidth mode (paper speed 500 Hz/minute; pen speed 20 inches/second). The signal level at each spectral frequency was elevated by approximately 3.7 dB over spectrum level, i.e., the level that would be measured if an analyzer had an ideal response characteristic with a bandwidth of 1 Hz (29). A more complete description of the wave analyzer and its response characteristics will be found in Appendix B. Frequency and intensity calibration of the wave analyzer was performed frequently to insure the accuracy of the obtained spectra (Appendix B). For all the male whispered vowel samples and the majority of the female samples, the first three formants were found to be within a 0-4000 Hz frequency range. Formant three for one female subject's /i/ productions was just above 4000 Hz.

Formant Measures

To obtain frequency, bandwidth, and amplitude measures for the first three formants of the test vowels, a blank spectrogram form was superimposed on each recorded spectrogram and the outline of each formant was traced with a fine-lead drafting pencil. Figure 1 shows the formant outline traced from the spectrum in Figure 11. Measures of the frequency of the formant peaks (in Hz) and the effective formant bandwidth (in Hz) were obtained by the use of a data quantifying device (Data Scaler, Model 400). The Data Scaler is an instrument for manually measuring distances between points on a graph. The points to be measured are located by a fixed and a movable cursor, and the distance between the two cursors is read from a digital dial. For the frequency measures obtained in this study, one unit on the digital dial equaled 1.923 Hz in the vowel spectrum. The frequency measurements were made to the nearest whole scale unit. Amplitude measures for each formant peak were obtained by measurement along the ordinate of the spectrogram, which was scaled at two-decibel intervals by horizontal lines. The formant amplitude measures were recorded to the nearest whole decibel.

The investigator's accuracy in delineating and measuring the vowel formant features of interest was evaluated in the following way. The formant tracing and measurement procedure was repeated independently by another Ph.D.

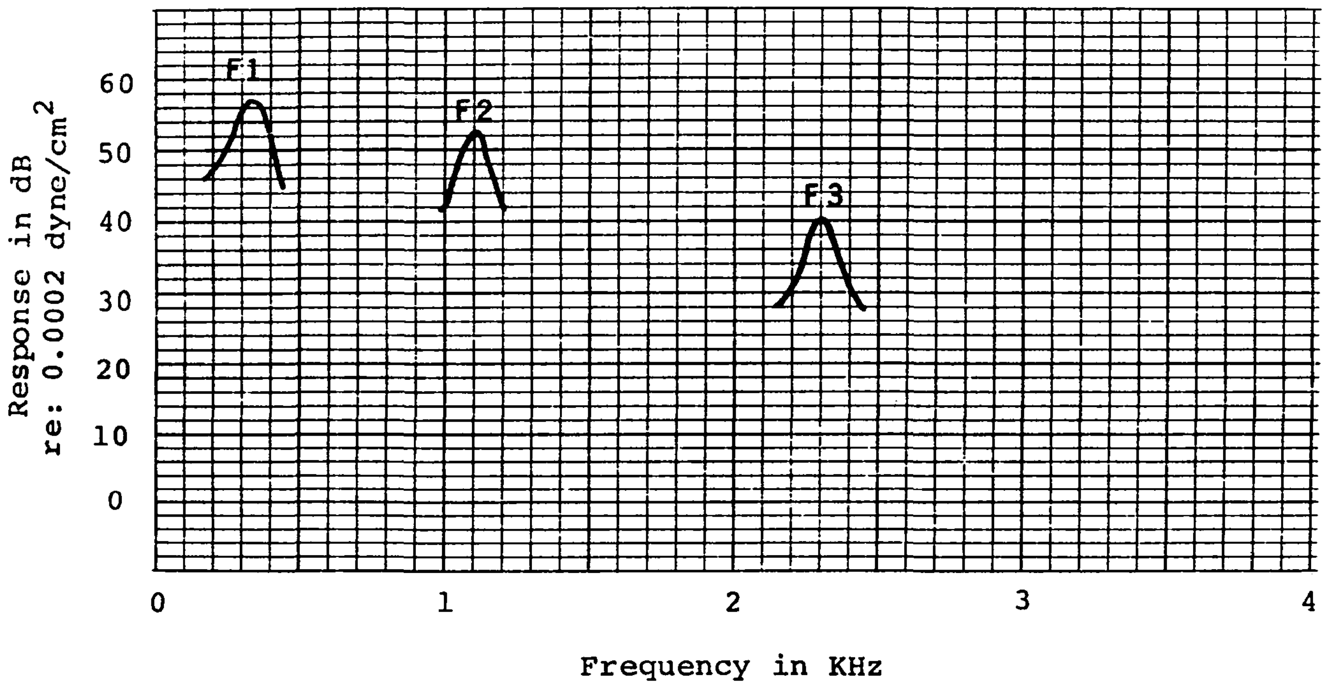

Figure 1.--A tracing (obtained from the spectrum in Figure 11) showing the first three formants for a male whispered /u/ production.

student on one of the eight (randomly selected) spectra for the test productions by each of the 24 subjects. Three formants were measured in each spectrum; thus, a total of 72 measurements each for formant frequency, bandwidth, and amplitude were repeated. The mean formant frequency difference between the "repeat" measures and those obtained earlier by the investigator was 12 Hz; the mean bandwidth difference was 25 Hz; and, the mean amplitude difference was 1.5 dB. The individual measurement differences for each parameter were compared statistically using a paired-"t" test and found to be nonsignificant at the .05 level.

The investigator's reliability in obtaining the formant measures was checked by re-tracing and re-measuring one of the eight (randomly selected) spectra for the test productions by each of the 24 subjects. The mean difference between the first and second measurements of formant frequency was 10 Hz, the mean bandwidth difference was 17 Hz, and the mean amplitude difference was 0.5 dB. The individual measurement differences for each parameter were compared statistically using a paired-"t" test and found to be nonsignificant at the .05 level. Thus, it appeared that the original formant measurements were sufficiently accurate and reliable.

Judgment Procedure

To evaluate further the phonetic representativeness of the test vowel samples, listener judgments were obtained. The whispered test vowel samples were randomized and re-recorded onto a continuous tape for presentation to a panel of judges. The judges were eleven graduate students in speech pathology who had not participated as subjects for the study. The listening tape consisted of the 192 two-second whispered test vowels plus the first twenty samples repeated at the end to evaluate intra-listener reliability. These 212 vowel samples were presented individually to the judges for rating.

In the judgment session, the judges reported their perception of the phonetic identity of each whispered vowel production. All of the judges on the panel had recently completed a semester of academic training in phonetics and, thus, had recently received practice in phonetic transcription. Although only four "intended" vowels were represented among the present test samples, the results of previous studies (9, 32) suggested that listeners are likely to identify some vowel productions as allophones of a phoneme other than that intended. Therefore, the judges were not told that there were only four intended vowels but, instead, were given the opportunity to identify a sample as an allophone of any of nine cardinal vowels: /i/, /I/, /ɛ/, /æ/, /ʌ/, /ɑ/, /ɔ/, /U/, and /u/. Key words containing each of the cardinal vowels and the phonetic symbol for each vowel was pre-recorded on the rating-response sheet as a reference for the listeners. The key words were (1) see - /i/; (2) sit - /I/; (3) set - /ɛ/; (4) sat - /æ/; (5) sun - /ʌ/; (6) sob - /ɑ/; (7) saw - /ɔ/; (8) soot - /U/; and (9) soup - /u/.

To increase the probability that listener vowel identifications would be based on the same phonetic criteria across judges, each key word was initially spoken aloud by the investigator and the vowel associated with each was produced both with a sustained voice and a sustained whisper (Appendix A, Instructions to Judges). After the investigator had explained the identification procedure and had demonstrated the characteristic quality of each of the nine vowels on the response sheet, a practice tape of twenty whispered vowel productions drawn at random from the test samples was presented to the listeners. For the first ten practice samples, the recorder was stopped after the presentation of each sample and the judges were allowed to compare their impressions of vowel identity verbally. The second ten practice samples were then played without a pause and the judges recorded their responses without discussion. After completion of the practice session, the test samples were presented in random order for identification. The judges were allowed five seconds between presentations of the two-second samples in which to record their responses.

Reliability Ratings

Subject Reliability in Repeated Productions

It appeared generally that the first and second whispered productions of each test vowel by each subject were closely comparable with respect to the acoustic parameters of formant frequency, bandwidth, and amplitude. To evaluate possible differences between the two productions of each test vowel by each subject, however, the repeated productions were compared with respect to the frequency, bandwidth, and amplitude measures obtained for F1, F2, and F3 (Appendix C). With each of the three formants for each vowel considered separately, the measures of frequency, bandwidth, and amplitude for each of the two productions by subjects of each sex were treated by an analysis of variance to separate the within- and between-subjects variance. The SDs within and between subjects for each formant parameter are presented in Appendix D. For all test vowels, the within-subject SD for each formant parameter was usually considerably smaller than that between subjects.

Judge Reliability

The judges' reliability in performing the judgment task was evaluated by comparing their responses to the first and second presentations of the reliability samples. Table 1 presents the number of identical responses to the two presentations of the same test samples. It may be seen that the number of identical responses to the reliability samples ranged from eight to nineteen, with a median of seventeen identical responses over the eleven judges. The overall judge reliability was considered to be acceptable for this study.

TABLE 1.--Number of identical responses by each of eleven judges to the first and second presentations of the twenty reliability samples.

Judge	Number of Identical Responses
1	19
2	14
3	18
4	8
5	10
6	15
7	18
8	17
9	17
10	19
11	13

CHAPTER IV

RESULTS AND DISCUSSION

This study was designed to investigate the first three formants of selected whispered vowels: /i/, /æ/, /ɔ/, and /u/. Twelve male and twelve female young-adult subjects each produced two samples of each test vowel; each sample was recorded on magnetic tape. Eleven judges independently rated each recorded sample to identify the vowel produced. Each recorded sample was also analyzed to obtain its intensity-by-frequency acoustic spectrum. In each spectrum, the frequency, effective bandwidth, and amplitude of each of the first three formants were measured. The findings are presented in the following sections.

Vowel Identifications

Table 2 shows vowel identifications made by each of the eleven judges for the individual productions of each of the four test vowels. The table reveals that relatively high percentages of "correct" listener identification were obtained for the two "front" vowels /i/ and /æ/ (82% and 78%, respectively), while relatively low percentages were obtained for the two "back" vowels /ɔ/ and /u/ (41% and 57%,

TABLE 2.--Matrix showing the vowel identification results (N = 11 judges X 48 samples = 528 judgments per vowel.)

Intended Vowel	Identified Vowel									Percent Identified As Intended
	i	I	ɛ	æ	ɑ	ʌ	ɔ	U	u	
/i/ (N=528)	431	91	2	1	1			2		82%
/æ/ (N=528)		2	68	412	30	16				78%
/ɔ/ (N=528)			3	3	247	47	218	10		41%
/u/ (N=528)					9	14	7	195	303	57%

respectively). Regarding productions identified as samples of a vowel other than the one intended, /i/ was most frequently identified as /I/, /æ/ as /ɛ/, /ɔ/ as /ɑ/, and /u/ as /U/. Although a substantial number of the test productions were identified as samples of a phoneme other than that intended, the identification results were accepted as satisfactory; they were generally similar to those reported by Peterson and Barney (33) and Fairbanks and Grubb (9) for voiced vowels.

Table 3 presents separately for each test vowel and each sex the number of productions identified as samples of the intended vowel or another vowel by a majority (six or more) of the eleven judges. The table shows that most of the frequent "back" vowel (/ɔ/ and /u/) "misidentifications" (shown in Table 2) were associated with the samples produced by females. It may be seen also that the female productions intended as /ɔ/ and /u/ were predominantly identified as /ɑ/ and /U/, respectively. Most of the female productions of /i/ and /æ/, however, were identified as samples of the intended vowel by a majority of the judges, as were the male productions of all four test vowels.

Acoustic Spectral Measures

Figures 2 through 13 present example spectra for male and female productions of each test vowel. One example spectrum each for /i/ and /æ/ productions by subjects of each sex is presented. Two example spectra of /u/ and /ɔ/

TABLE 3.--Number of samples of each test vowel which were identified as the intended vowel or another vowel by a majority (six or more) of the eleven judges (N = 24 productions of each vowel by subjects of each sex).

Produced by:	Intended Vowel	Identified As Intended Vowel	Identified As Another Vowel
Male	/i/ (N=24)	22	2
	/æ/ (N=24)	19	5
	/ɔ/ (N=24)	18	6
	/u/ (N=24)	20	4
Female	/i/ (N=24)	21	3
	/æ/ (N=24)	21	3
	/ɔ/ (N=24)	2	22[a]
	/u/ (N=24)	6	18[b]

[a]20 female productions intended as /ɔ/ were identified as /ɑ/ by a majority of the judges.

[b]15 female productions intended as /u/ were identified as /U/ by a majority of the judges.

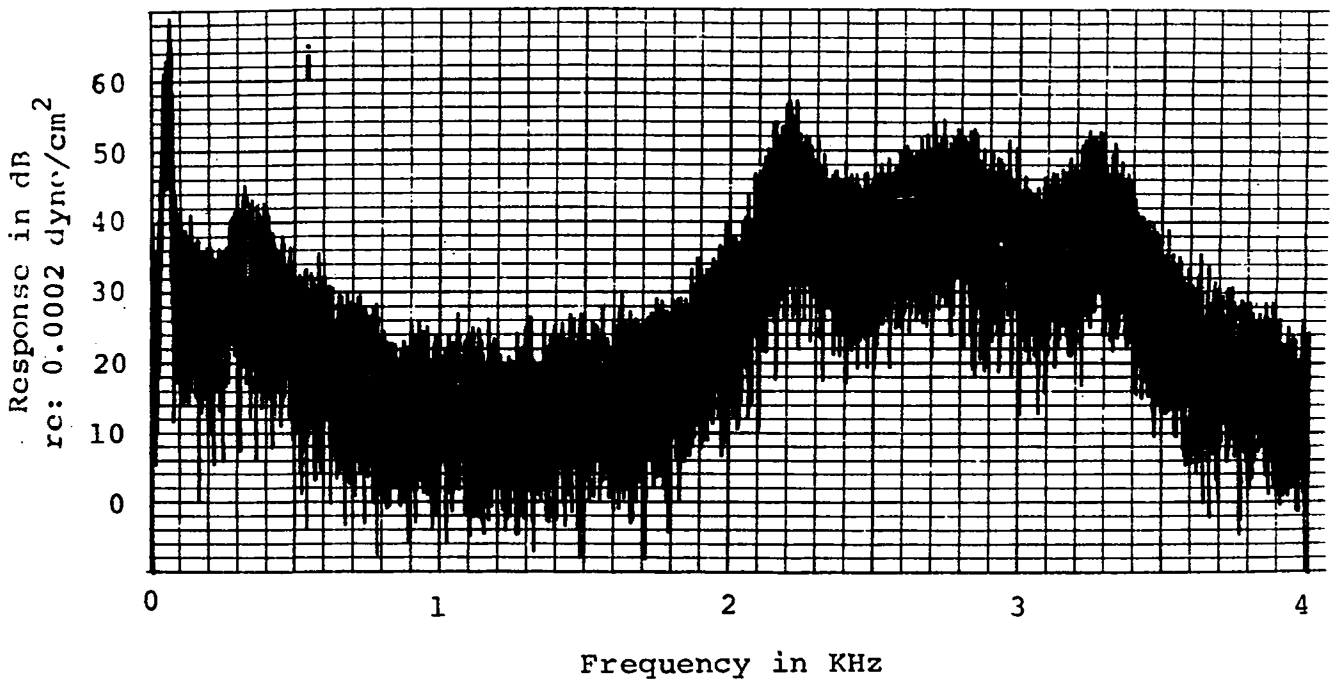

Figure 2.--Narrow-band (3-Hz) spectrum of a male whispered /i/ production.

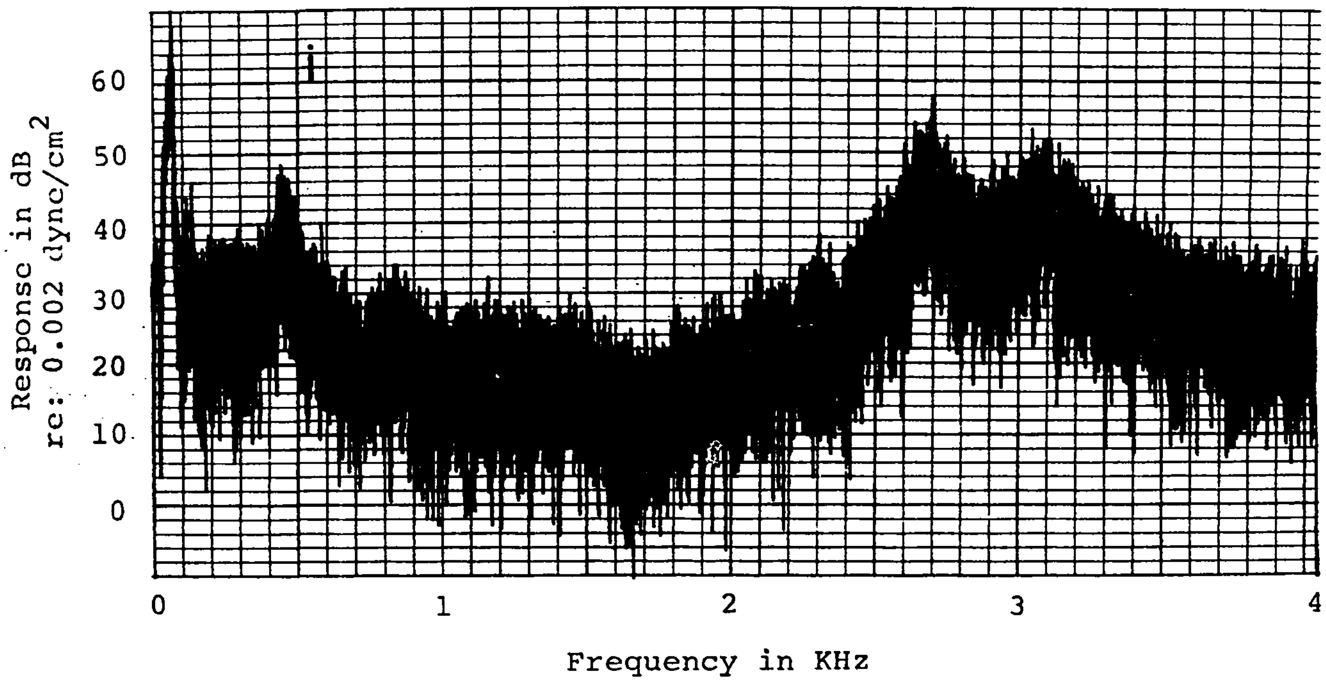

Figure 3.--Narrow-band (3-Hz) spectrum of a female whispered /i/ production.

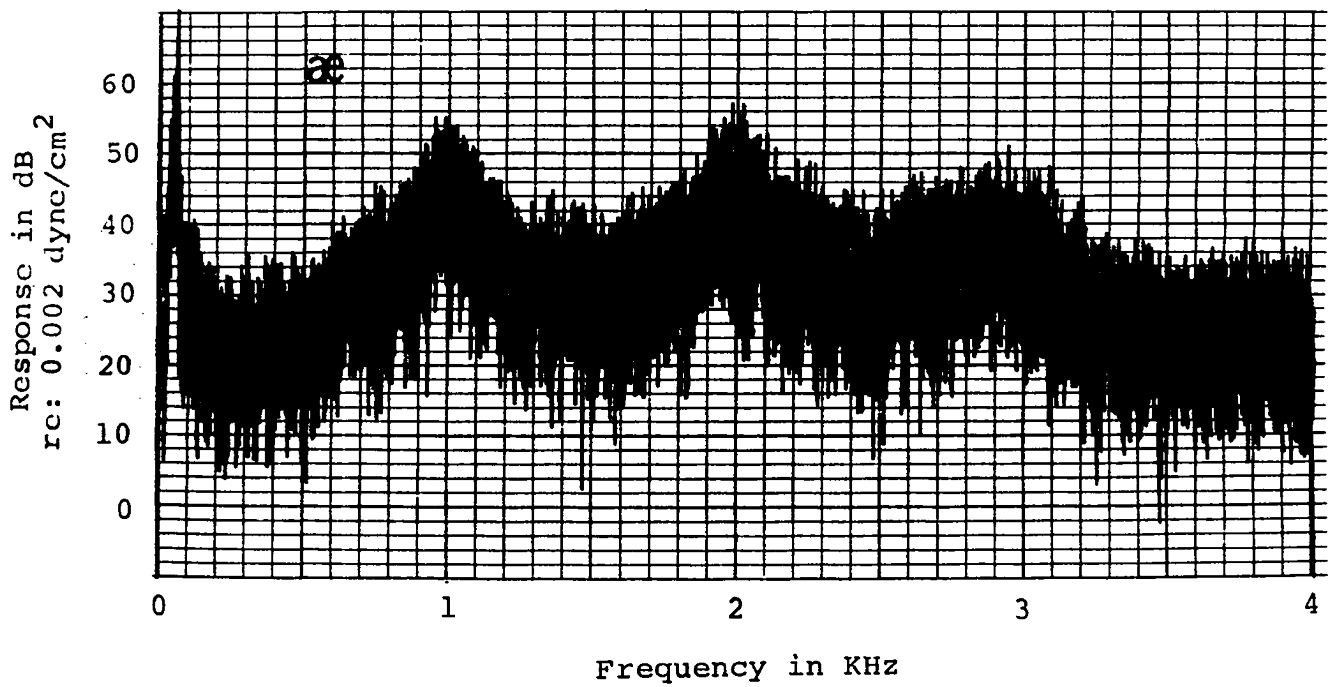

Figure 4.--Narrow-band (3-Hz) spectrum of a male whispered /æ/ production.

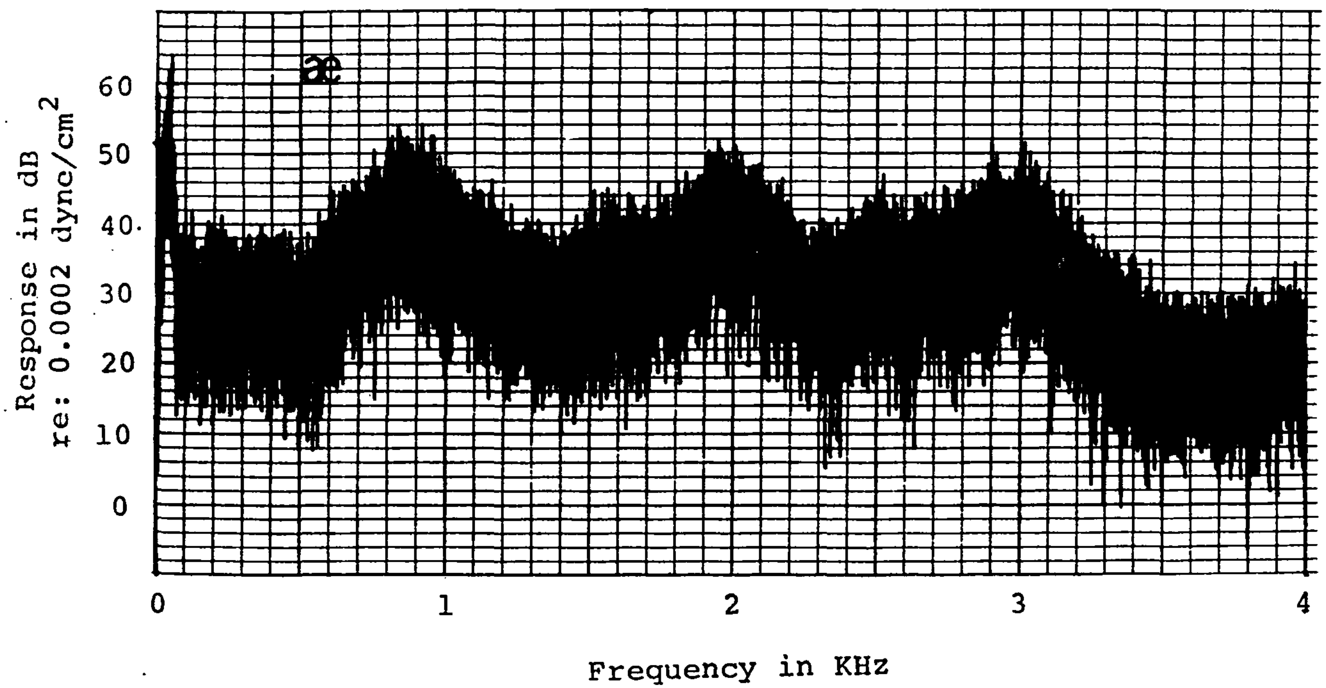

Figure 5.--Narrow-band (3-Hz) spectrum of a female whispered /æ/ production.

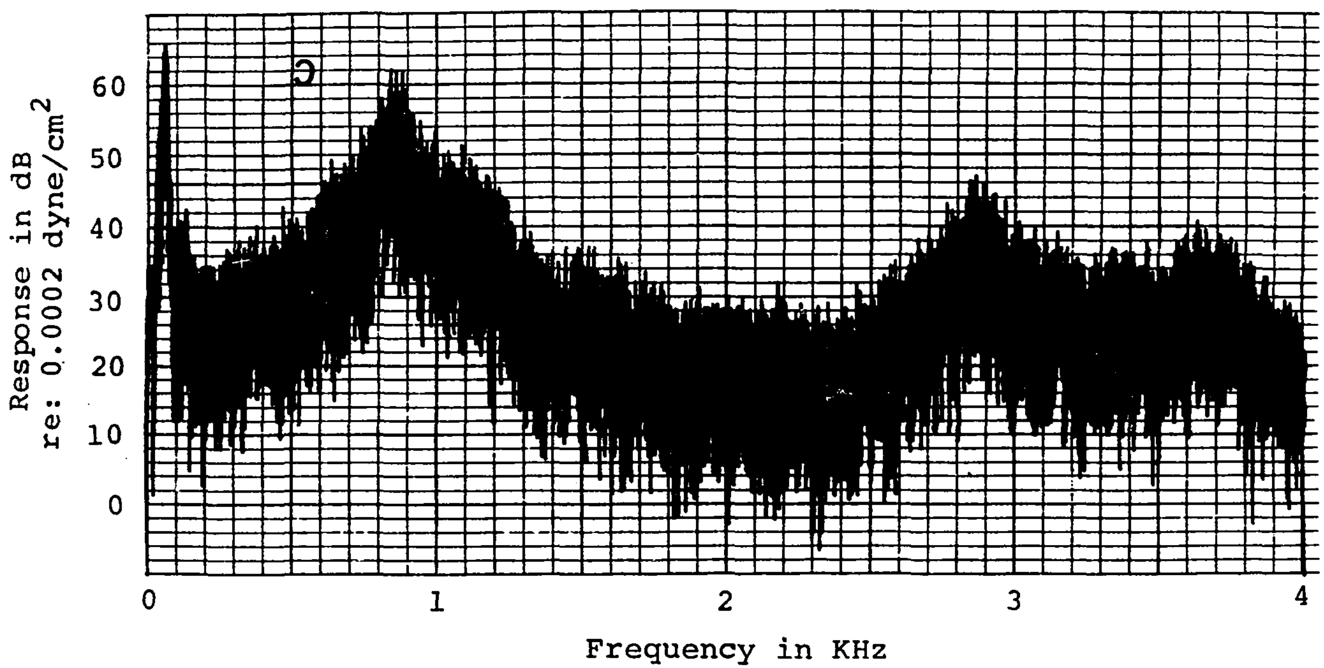

Figure 6.--Narrow-band (3-Hz) spectrum of a male whispered /ɔ/ production showing only one apparent lower formant (i.e., F1 = F2).

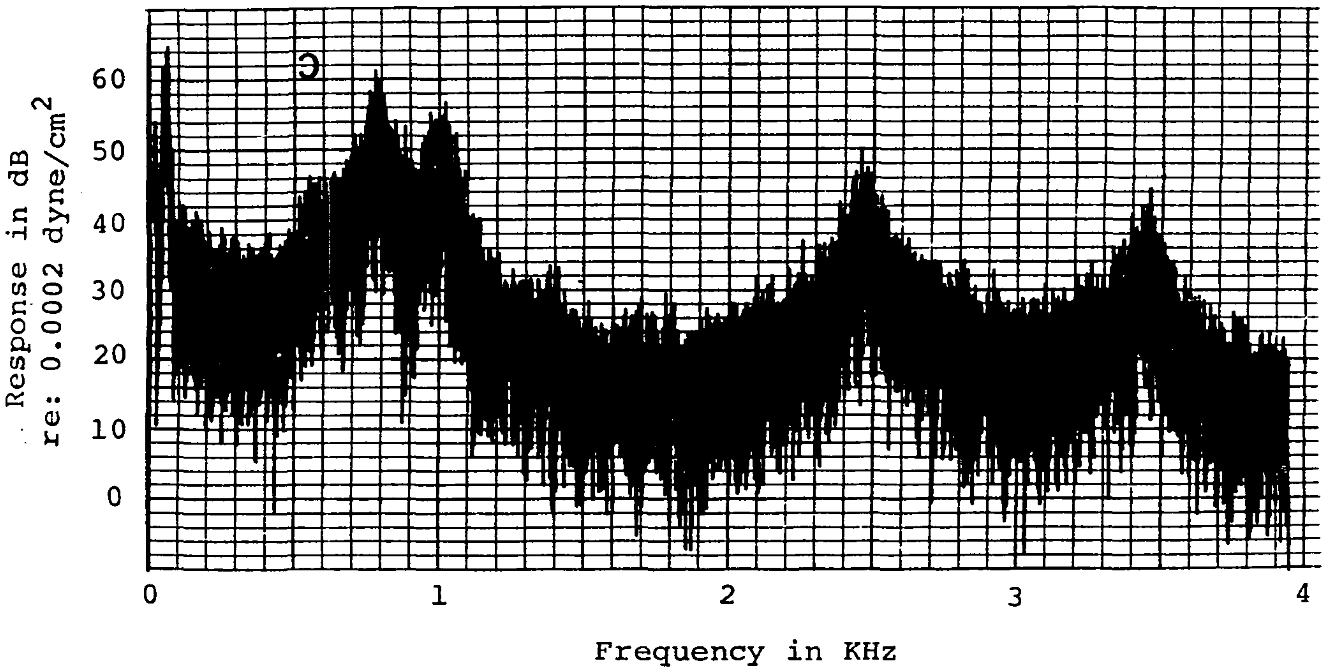

Figure 7.--Narrow-band (3-Hz) spectrum of a male whispered /ɔ/ production showing separate F1 and F2 peaks.

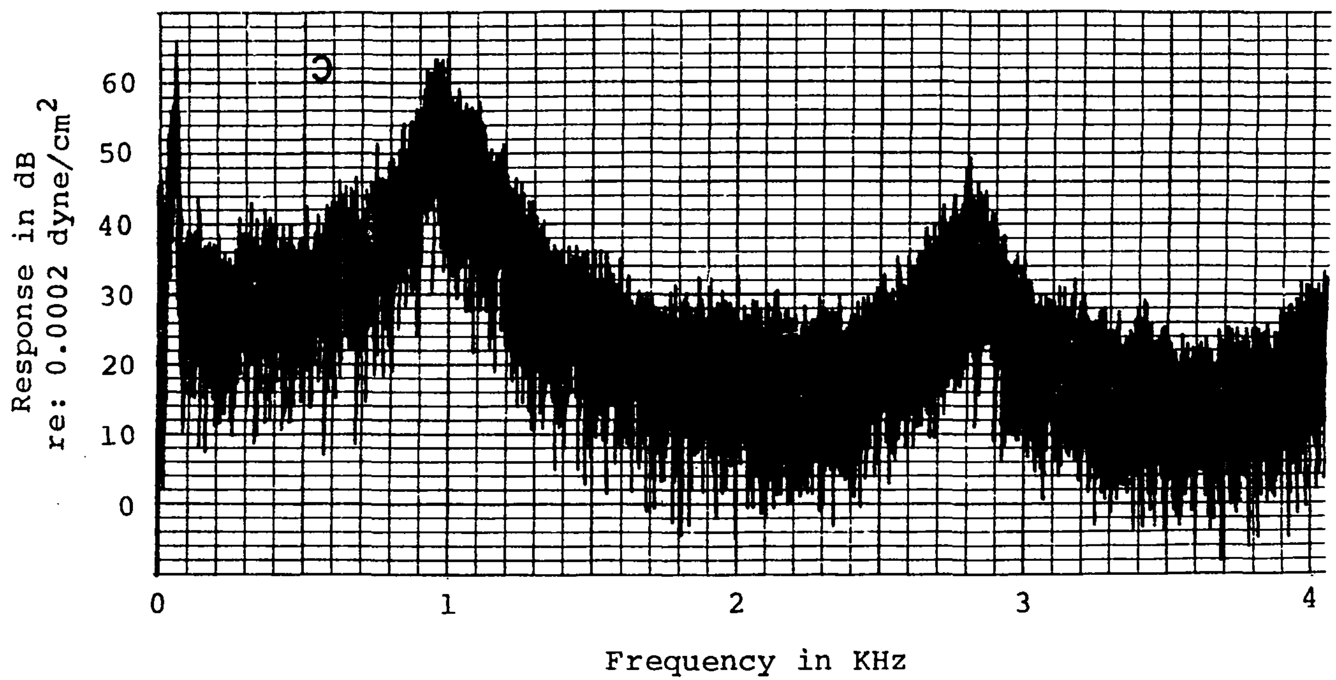

Figure 8.--Narrow-band (3-Hz) spectrum of a female whispered /ɔ/ production showing only one apparent lower formant (i.e., F1 = F2).

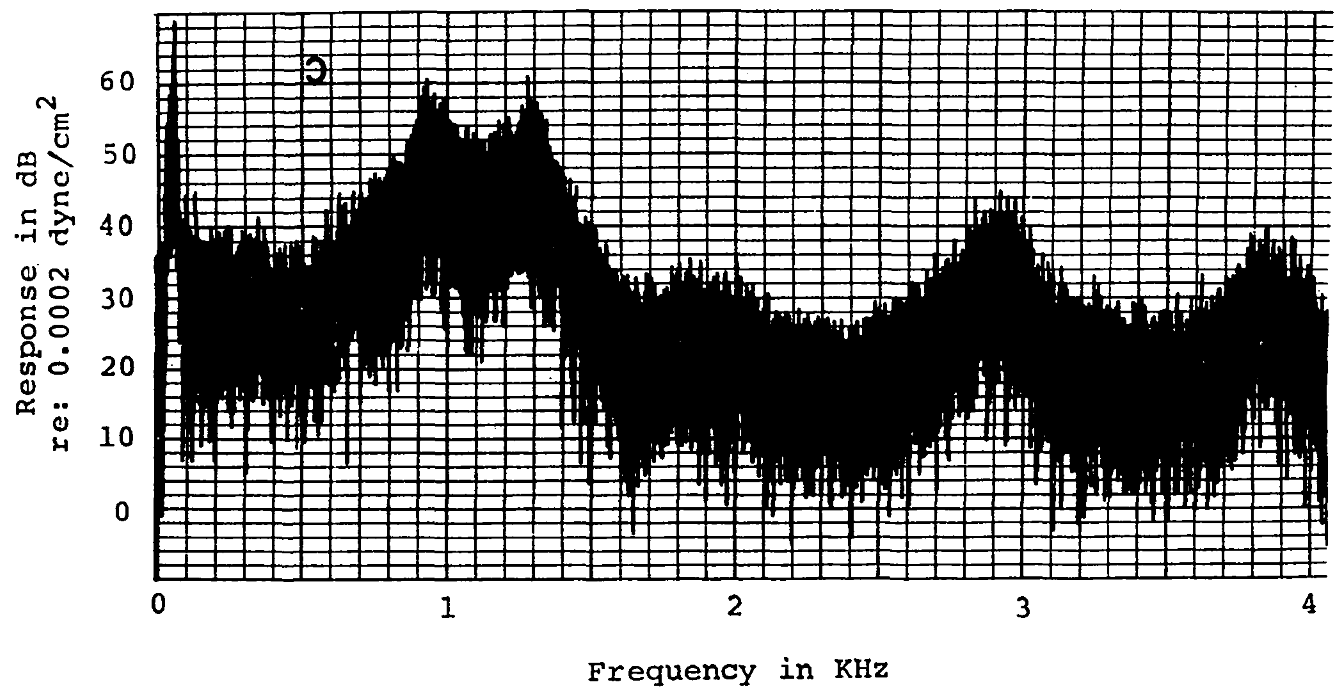

Figure 9.--Narrow-band (3-Hz) spectrum of a female whispered /ɔ/ production showing separate F1 and F2 peaks.

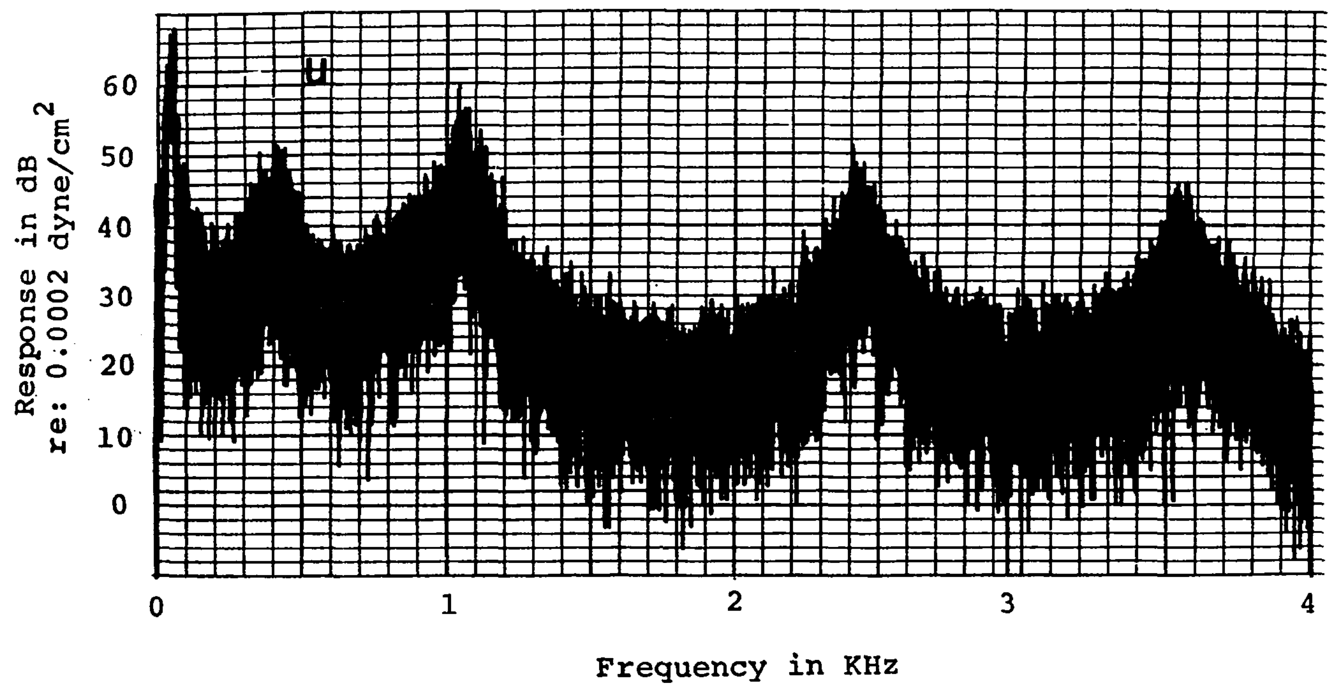

Figure 10.--Narrow-band (3-Hz) spectrum of a male whispered /u/ production with F1 having less amplitude than F2.

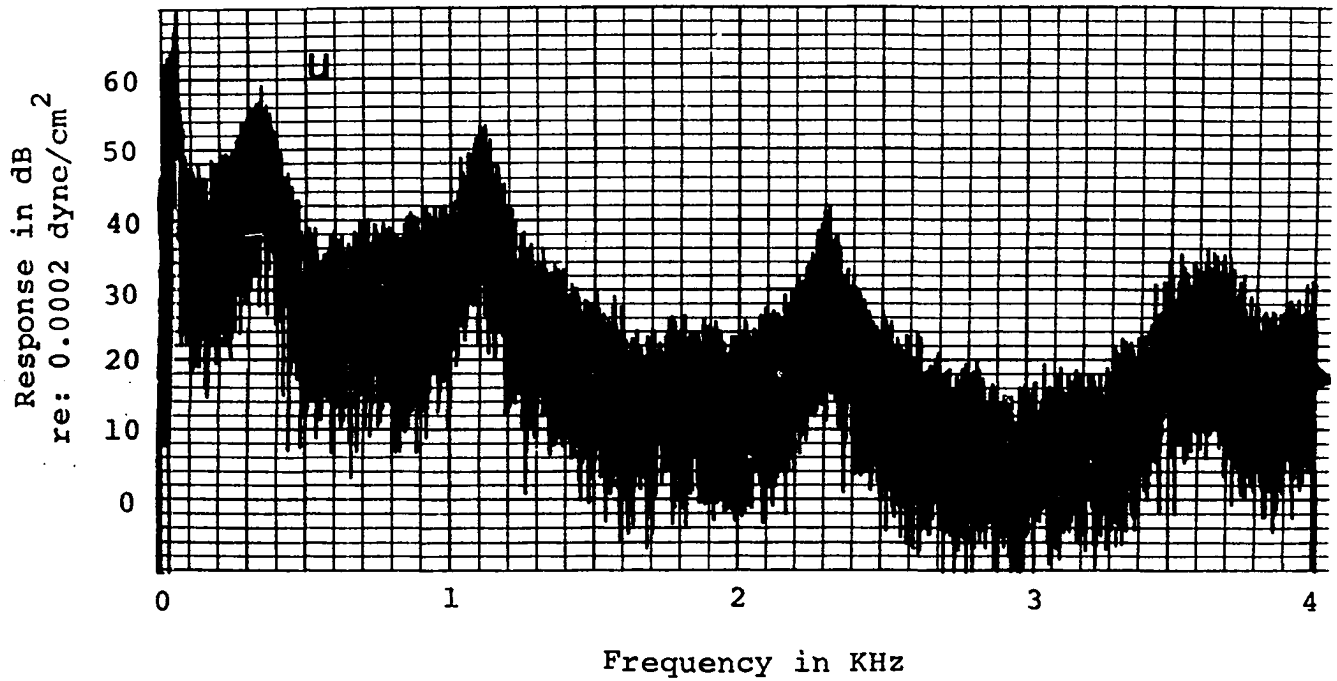

Figure 11.--Narrow-band (3-Hz) spectrum of a male whispered /u/ production with F1 having greater amplitude than F2.

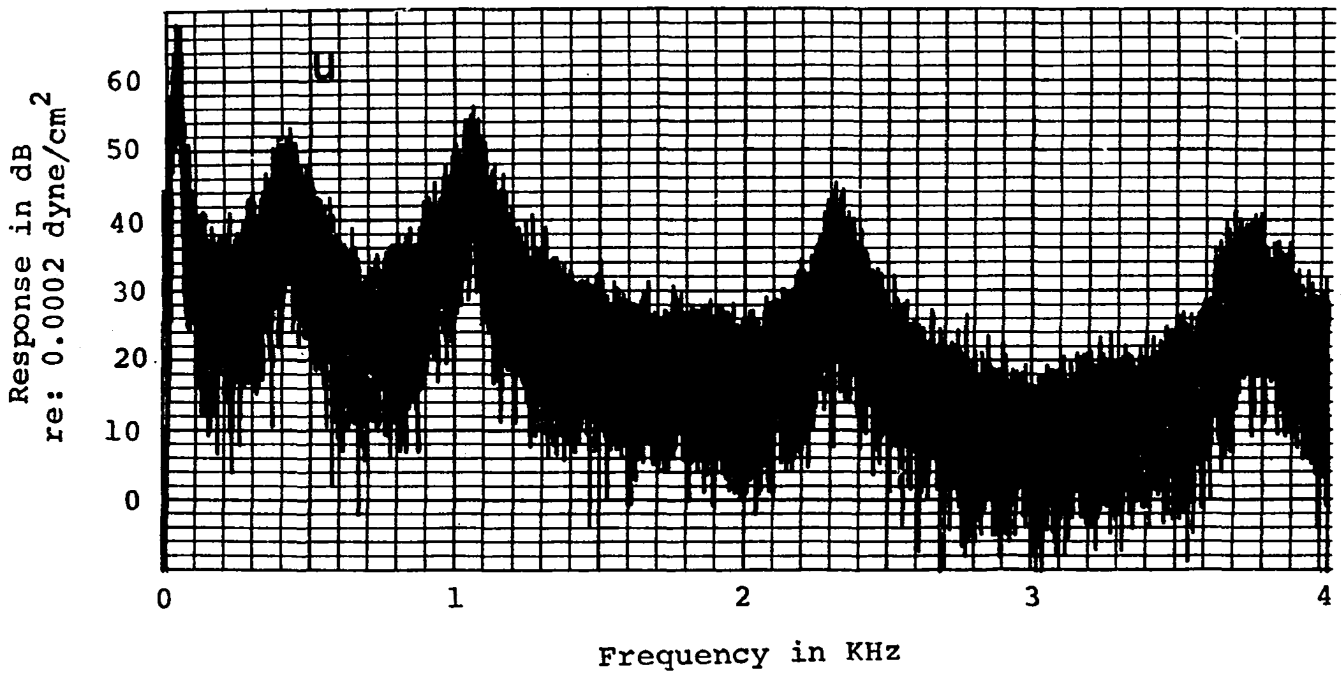

Figure 12.--Narrow-band (3-Hz) spectrum of a female whispered /u/ production with F1 having less amplitude than F2.

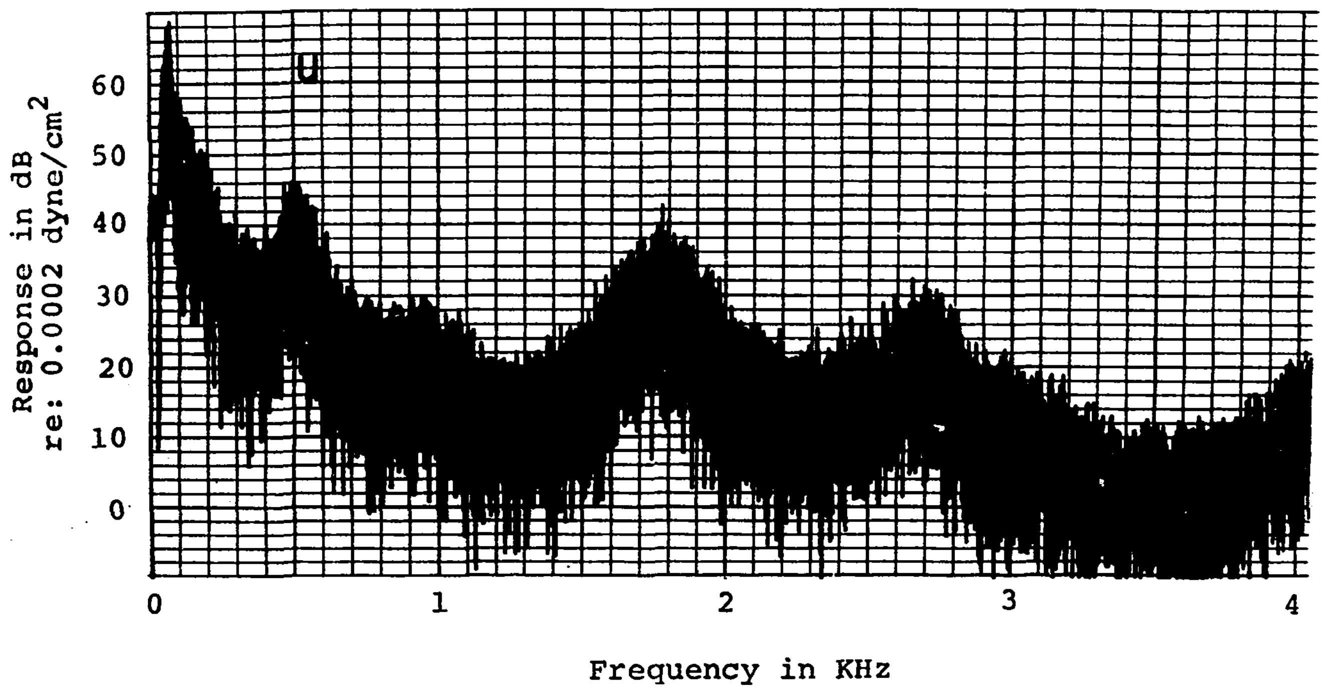

Figure 13.--Narrow-band (3-Hz) spectrum of a female whispered /u/ production with F1 having greater amplitude than F2.

productions by each sex are presented, however. The four spectra of /u/ productions are presented to show that, for both sexes, some /u/ samples (fourteen male and ten female) were characterized by first formant (F1) higher in amplitude than the second formant (F2), while other /u/ samples (ten male and fourteen female) were characterized by an F1 lower in amplitude than F2. The four spectra of /ɔ/ productions are presented to show that, for both sexes, some /ɔ/ samples (ten male and nine female) were characterized by an F1 and F2 at approximately the same frequency; thus, because for those samples F1 and F2 overlapped, only one lower resonance peak was apparent in their spectra. Other /ɔ/ samples (fourteen male and fifteen female) were characterized by an F1 and F2 which, although very close together in frequency, were distinguishable as separate peaks in the narrow-band spectrum. Specific findings regarding formant frequencies, effective bandwidths, and relative amplitudes are presented below (measures of all three formant parameters obtained for individual vowel samples are presented in Appendix C).

Formant Frequencies of the Whispered Vowels

The obtained formant frequency measures were examined in each of three ways according to criteria based on the vowel identification results. The frequency measures were grouped first according to the vowel intended. This treatment seemed reasonable for one presentation because the vowel samples were produced by graduate students in communication disorders who had practiced accurate productions of the samples under the guidance of the investigator; and, at the time it was recorded, each test sample appeared both to the investigator and to the subject who produced it to be a representative whispered production of the intended vowel. Means of those formant frequency measures are reported under the heading Intended. A second set of formant frequency measures was obtained for vowel samples which were identified as the intended vowel by six or more of the eleven judges. Means of those measures are reported under the heading Majority. A third set of formant frequency measures was obtained for samples of each test vowel which were identified as the intended vowel by all eleven judges. Means of those measures are reported under the heading Unanimous.

Tables 4 and 5 show separately by sex of subjects the mean frequency (in Hz) of F1, F2, and F3 for the test productions. The frequency mean (over test productions) and SD for each of the three formants are shown separately for Intended, Majority, and Unanimous vowel samples, as defined above.

Findings regarding relationships between the frequency of F1 and that of F2 for the samples of each test vowel are summarized graphically in Figure 14. The ordinate and abscissa of this figure are logarithmic which, as Pols, Tromp, and Plomp (35) have noted, is "more in line with the

TABLE 4.--The mean frequency and SD (in Hz) of F1, F2, and F3 for male whispered vowels. The means are over all Intended samples, those correctly identified as the intended vowel by a Majority (six or more) of the eleven judges, and those identified as the intended vowel by Unanimous agreement among the eleven judges.

	Vowels							
	/i/		/æ/		/ɔ/		/u/	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formant One								
Intended	385	43	815	95	802	61	406	32
Majority	381	43	818	84	810	57	401	32
Unanimous	361	33	890	104	780	49	400	25
Formant Two								
Intended	2400	145	1903	154	928	56	1180	182
Majority	2410	147	1901	149	915	80	1146	177
Unanimous	2456	128	1968	68	907	87	1302	132
Formant Three								
Intended	2984	139	2601	174	2442	268	2297	116
Majority	2987	145	2601	172	2530	240	2286	114
Unanimous	2998	161	2712	146	2588	295	2300	128

TABLE 5.--The mean frequency and SD (in Hz) of F1, F2, and F3 for female whispered vowels. The means are over all Intended samples, those correctly identified as the intended vowel by a Majority (six or more) of the eleven judges, and those identified as the intended vowel by Unanimous agreement among the eleven judges.

	Vowels							
	/i/		/æ/		/ɔ/		/u/	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formant One								
Intended	435	76	1015	82	952	53	500	71
Majority	429	61	1008	83	915	121	425	23
Unanimous	438	37	1035	66	a		420	21
Formant Two								
Intended	2891	252	2181	252	1119	73	1351	269
Majority	2864	254	2122	208	915	121	1323	228
Unanimous	2900	281	2086	169	a		1167	169
Formant Three								
Intended	3523	346	3164	335	2842	183	2809	223
Majority	3517	353	3094	291	2979	0	2626	254
Unanimous	3536	399	3041	293	a		2464	244

[a]No female /ɔ/ production was identified as /ɔ/ by all eleven judges.

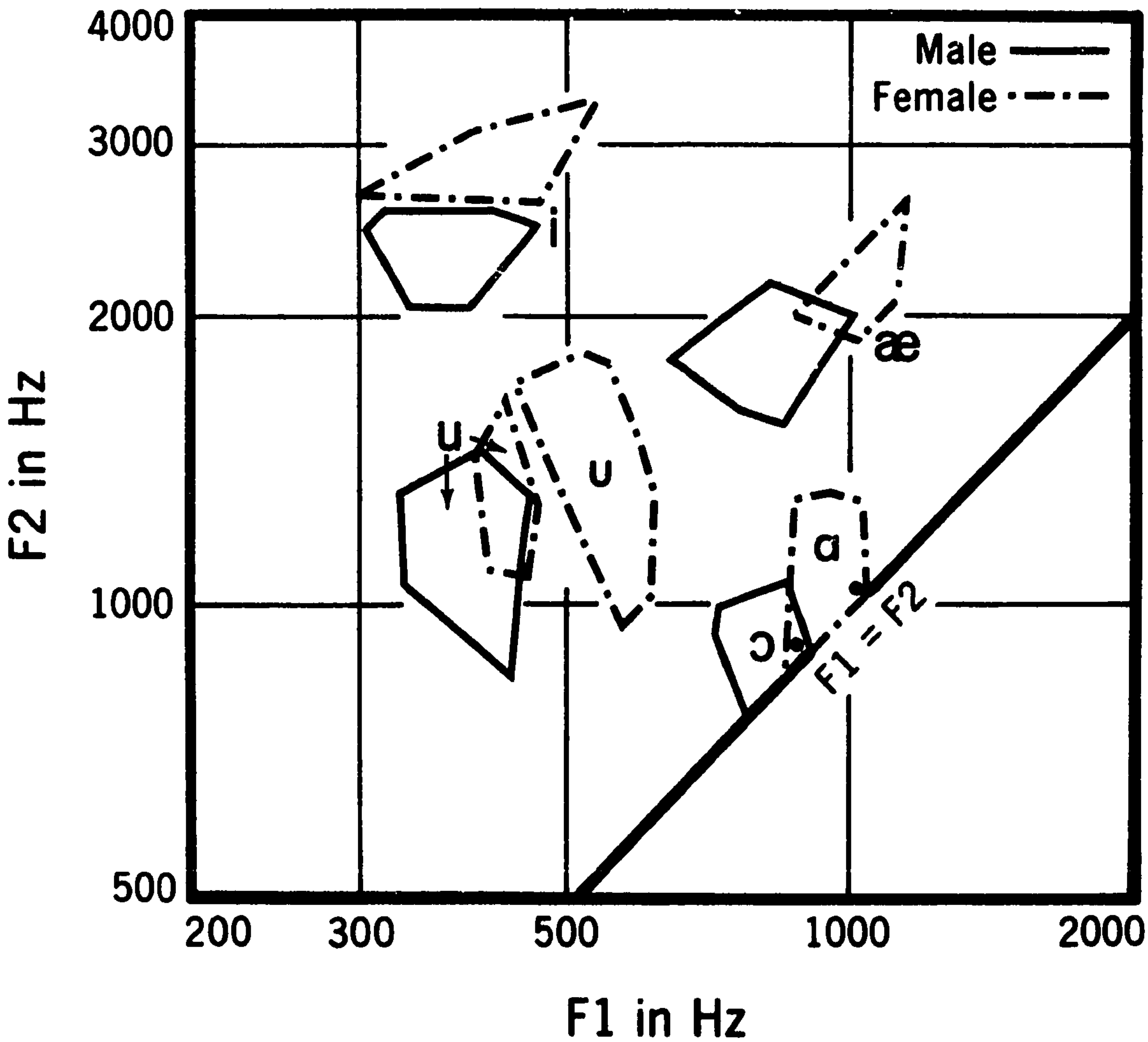

Figure 14.--Frequency areas of formants one and two for Majority whispered vowel samples produced by adult males and females.

hearing process than a linear frequency scale." The figure represents areas enclosing coordinate plots of F1 and F2 for samples of each vowel (male and female data shown separately) which are identified as the intended vowel by a majority of the listeners. Included in the figure are separate plots for /u/ productions (by females) which were identified by a majority of listeners as /U/, and /ɔ/ productions (by females) which were identified by a majority of the listeners as /ɑ/. The two dots on the F1=F2 line represent the only two female productions intended as /ɔ/ which were identified as /ɔ/ by a majority of the listeners. It can be seen in the figure that the F1-F2 area for each test vowel does not overlap that for any other vowel.

Figure 14 shows that the F1-F2 formant plot for female /i/ productions is separate from that for male /i/ productions, with no overlap. The plots are separate primarily because of a sex difference in the frequency range of F2 for /i/. Approximately the same F1 frequency range is associated with the male and the female productions. The F1-F2 plots for the other "front" vowel, /æ/, are also generally separate for the male and female productions. There is, however, a slight overlap in the area for the two sexes in the 900-1000 Hz range for F1 and the 1850-2150 Hz range for F2. The F2 values for female /æ/ productions ranged 500 Hz higher than those for male productions.

The F1-F2 plots for the vowel /u/ show considerable overlap in areas for the male and female productions in the 400-450 Hz range for F1 and the 1050-1450 Hz range for F2. The F1 frequencies for male /u/ productions ranged as high as those for females, but the F2 frequencies for female productions ranged 200 Hz higher than those for males. The separate area representing female productions identified by a majority of listeners as /U/ was characterized by a higher F1 frequency range, but generally the same F2 range, as that for the female productions identified as the intended /u/.

Both male and female productions intended as /ɔ/ included some samples with F1 apparently equal to F2. Among these were two female productions which were identified as /ɔ/ by a majority of the listeners. Other productions (by males) identified by a majority of listeners as /ɔ/ were characterized by clearly separate F1 and F2 frequencies. Most of the female /ɔ/ productions which were identified as /ɑ/ were characterized by a markedly higher frequency for F2 than for F1.

Figure 15 presents separate F2-F3 plots for the male and female Majority samples of each test vowel. Separate plots representing female /u/ productions identified as /U/ and female /ɔ/ productions identified as /ɑ/ by a majority of the listeners are also shown. The two dots in the /ɑ/ area represent the two female productions identified as the intended /ɔ/ by a majority of the listeners. The figure

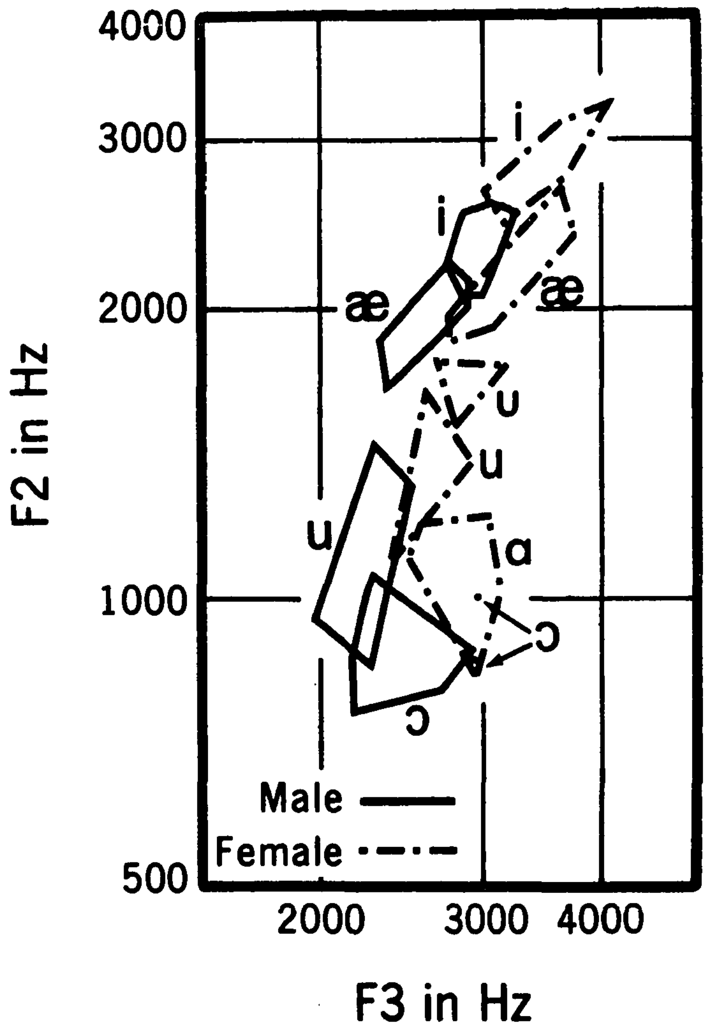

Figure 15.--Frequency areas of formants two and three for Majority whispered vowel samples produced by Adult males and females.

shows a considerable overlap of F3 frequencies for male and female productions in the 2400-3300 Hz range for all test vowels. The separation of the areas for individual vowels is due primarily to the much wider range for F2 than for F3 frequencies. It can be seen for the male productions that third formants lowest in frequency (as low as 2000 Hz) were associated with the high back vowel /u/, whereas the highest third formant frequencies (as high as 3300 Hz) were associated with the high front vowel /i/. The low vowels /æ/ and /ɔ/ show approximately the same F3 frequency range.

The female productions of each test vowel were characterized by generally higher F3 frequencies than those for male productions. The intended /u/ productions by females which were identified as /U/ were generally similar, with respect to F3 frequencies, to the productions that were identified as the intended /u/. The intended /ɔ/ productions by females which were identified as /ɑ/, however, were generally characterized by lower F3 frequencies than those identified as /ɔ/.

Table 6 presents separately for F1, F2, and F3 differences (in Hz) between the formant frequency means obtained for male and female productions of each test vowel. The difference between the frequency means for each sex is also expressed as a percentage of the mean frequency for male productions. The formant means shown in the table are those for the <u>Intended</u> vowel samples. The table shows, for

TABLE 6.--Mean formant frequency differences (in Hz) between male and female whispered vowels and the percentage of the difference relative to the male value (the comparison is between means for the Intended vowel formant frequencies).

		/i/	/æ/	/ɔ/	/u/
F1	Females (N=24)	435	1015	954	500
	Males (N=24)	385	815	802	406
	Difference	50	200	150	94
	% of Male	13.0	24.5	18.7	23.2
F2	Females (N=24)	2891	2181	1119	1351
	Males (N=24)	2400	1903	928	1180
	Difference	491	278	191	171
	% of Male	20.5	14.6	20.6	14.5
F3	Females (N=24)	3532	3164	2842	2809
	Males (N=24)	2984	2601	2442	2297
	Difference	548	563	400	512
	% of Male	18.4	21.6	16.4	22.3

each of the three formants studied and for each of the four test vowels, that the frequencies for female productions were higher than those for corresponding formants for male productions. The table also shows that formant frequencies for female productions ranged from 13% (/i/, F1) to 24.5% (/æ/, F1) higher than those for male productions. The mean sex difference over all formants and test vowels was 19%. This relationship of male to female whispered vowel formant frequencies is generally comparable to that reported by Fant (11) for voiced vowels. Fant indicated that the relatively small vocal tract of adult females (with respect to that of adult males) is associated with approximately 20% higher formant frequencies for females than for males.

The relationship of vowel formant frequencies to presumed vocal tract configuration during vowel production is of interest. Stevens and House (41) and Ladefoged (22), for example, have reported with respect to voiced vowels that relatively high first formant frequencies are associated with lingual constrictions of the vocal tract relatively near the glottis (i.e., the low tongue position usually associated with /æ/ and /ɔ/), while relatively low first formant frequencies are associated with constrictions farther away from the glottis (i.e., the high tongue position usually associated with /i/ and /u/). Variations in second formant frequencies, on the other hand, are associated primarily with the anterior-posterior location of the lingual constriction within the vocal tract and with the degree of lip rounding. Relatively high second formant frequencies accompany an anterior lingual placement and diminished lip rounding (e.g., as for /i/ and /æ/) while relatively low second formant frequencies are associated with a more posterior lingual placement and greater lip rounding (e.g., as for /ɔ/ and /u/). These relationships held generally for the whispered vowel productions in this study. The third formant frequencies obtained for whispered samples were relatively constricted in range over the four test vowels but tended, generally, to increase slightly as the lingual constriction moved anteriorly and lip rounding was diminished.

Whispered vs Voiced Vowel Formants

Table 7 compares the mean frequencies (in Hz) of F1, F2, and F3 for whispered and voiced vowel productions by adult males. The formant frequencies for voiced vowels were reported by Fairbanks and Grubb (9) and by Peterson and Barney (33). Fairbanks and Grubb reported formant frequencies for voiced vowels grouped into three categories: Self-Approved, Identified, and Preferred. The Self-Approved samples were those accepted by the speakers, after considerable practice, as representative productions of the intended vowels. The Identified samples were those correctly identified as the intended vowel by 75% or more of a group of eight listeners. The Preferred samples were the four Identified productions of each vowel which were evaluated by the

TABLE 7.--Mean frequencies (in Hz) of F1, F2, and F3 for whispered and voiced vowels produced by adult males. The voiced vowel formant frequencies shown were reported by Fairbanks and Grubb (F-G) and by Peterson and Barney (P-B).

		Whispered				Voiced (F-G)			Voiced (P-B)		
		F1	F2	F3		F1	F2	F3	F1	F2	F3
/i/	Intended	385	2400	2984	Self-Approved	267	2551	2974	270	2290	3010
	Majority	381	2410	2987	Identified	264	2284	2991			
	Unanimous	361	2456	2998	Preferred	263	2378	3099			
/æ/	Intended	815	1903	2601	Self-Approved	660	1569	2464	660	1720	2410
	Majority	818	1901	2601	Identified	700	1606	2468			
	Unanimous	890	1968	2712	Preferred	773	1654	2510			
/ɔ/	Intended	802	928	2442	Self-Approved	612	778	2664	570	840	2410
	Majority	810	915	2526	Identified	592	690	2615			
	Unanimous	780	907	2588	Preferred	600	846	2636			
/u/	Intended	406	1180	2297	Self-Approved	276	840	2517	300	870	2240
	Majority	401	1146	2286	Identified	272	806	2518			
	Unanimous	400	1302	2300	Preferred	279	825	2496			

listeners as the most representative samples of the intended vowel.

Table 7 shows that, with few exceptions, the formant frequencies obtained in the present study for whispered male productions tended to be higher than those for the corresponding formants of voiced vowels reported for males by the cited authors (females were not studied by Fairbanks and Grubb). With regard to the _Unanimous_ whispered samples and the _Preferred_ voiced samples, which were apparently the samples most representative of the intended vowels in this study and in the Fairbanks and Grubb study, respectively, the table shows that mean F1 and F2 frequencies for whispered productions were consistently higher than those for voiced productions of the test vowels.

Figure 16 shows for male subjects the F1-F2 plots for _Majority_ whispered vowels based on data from the present study and _Identified_ voiced vowels based on the data reported by Fairbanks and Grubb. The figure shows for each test vowel, except /æ/, that the whispered and voiced vowel plots occupied separate areas. The vowel /æ/ manifested a small overlap of formant frequencies common to both whispered and voiced productions (the 700-850 Hz range for F1 and the 1550-1700 Hz range for F2). For /i/, the difference in the plots for whispered and voiced productions is attributable to differences in F1; the range of F2 frequencies is essentially the same for productions of both types. The other three test

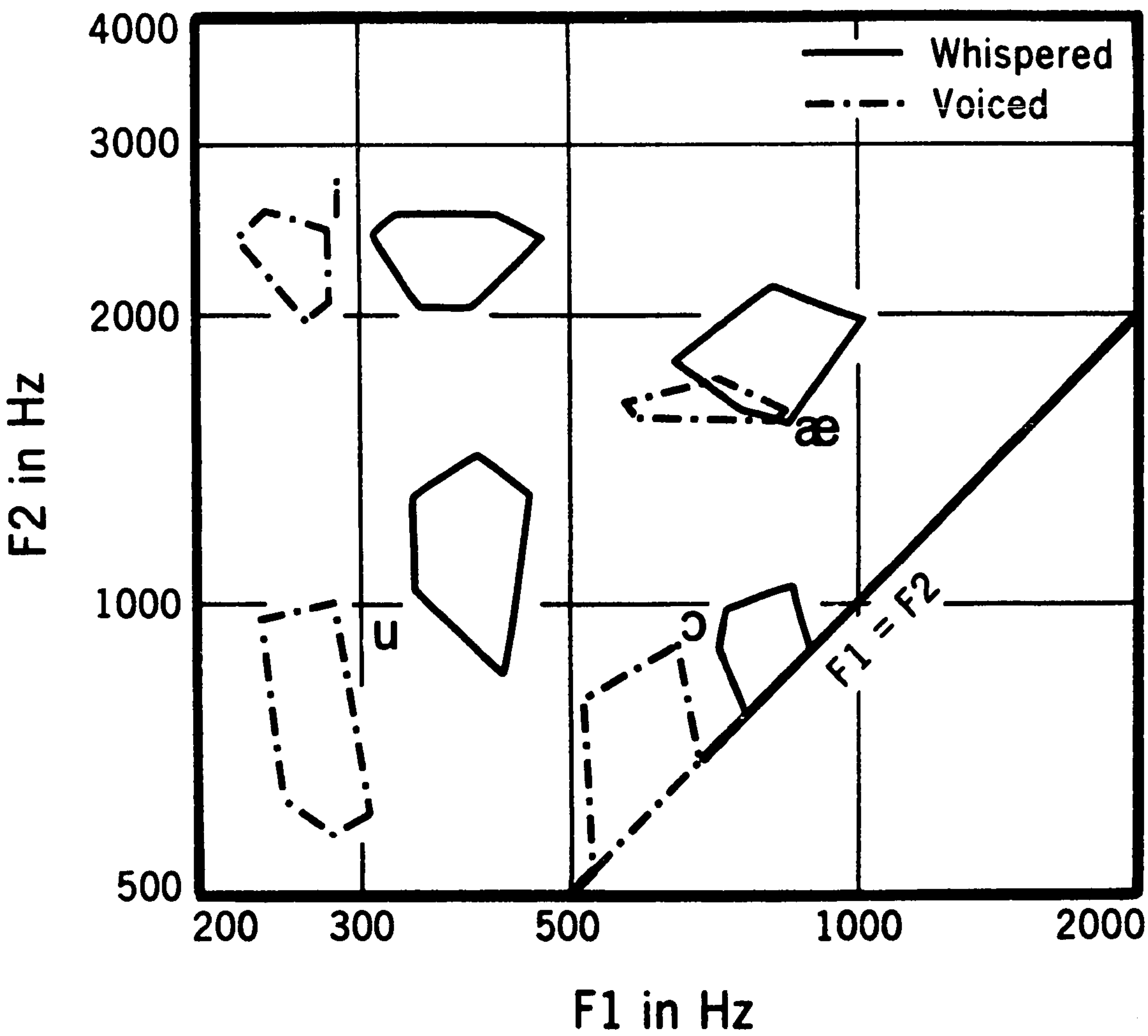

Figure 16.--Frequency areas of formants one and two for Majority whispered vowel samples (this study) and Identified voiced vowel samples (Fairbanks and Grubb). The data are for male subjects only.

vowels are characterized by higher F1 and F2 frequencies for whispered than for voiced productions.

Table 8 presents mean frequencies (in Hz) of F1, F2, and F3 for the female whispered vowel productions and comparable data for female voiced productions reported by Peterson and Barney (33). The authors in the cited study reported mean formant frequencies for female productions grouped according to the vowel intended; hence, the Intended vowel formant frequency means for this study are shown in the table for comparison with those from the Peterson and Barney study. It may be seen, for the female productions, that without exception the whispered vowel formant frequencies were higher than those for corresponding formants for voiced vowels.

Figure 17 shows the F2-F3 plots for Majority whispered vowels (present study) and Preferred voiced vowels (Fairbanks and Grubb study). Data for male productions only are presented. This figure shows, with respect to F3, that although the whispered vowels tended to have slightly lower F3 frequencies, there was considerable overlap between F3 frequencies for whispered and voiced vowels.

The relationship of whispered to voiced vowel formant frequencies reported here is generally consistent with that suggested by Lehiste (23) and Peterson (30). Lehiste noted that the vocal tract is characterized by higher resonant frequencies when open at both ends (as for sustained whispering) than when closed at one end (as for sustained

TABLE 8.--Mean frequencies (in Hz) of F1, F2, and F3 for whispered and voiced vowels produced by adult females. The voiced vowel formant frequencies shown were reported by Peterson and Barney (P-B).

	Vowels			
	/i/	/æ/	/ɔ/	/u/
Formant One				
Whispered	435	1015	952	500
Voiced (P-B)	310	860	590	370
Formant Two				
Whispered	2891	2181	1119	1351
Voiced (P-B)	2790	2050	920	950
Formant Three				
Whispered	3532	3164	2842	2809
Voiced (P-B)	3310	2850	2710	2670

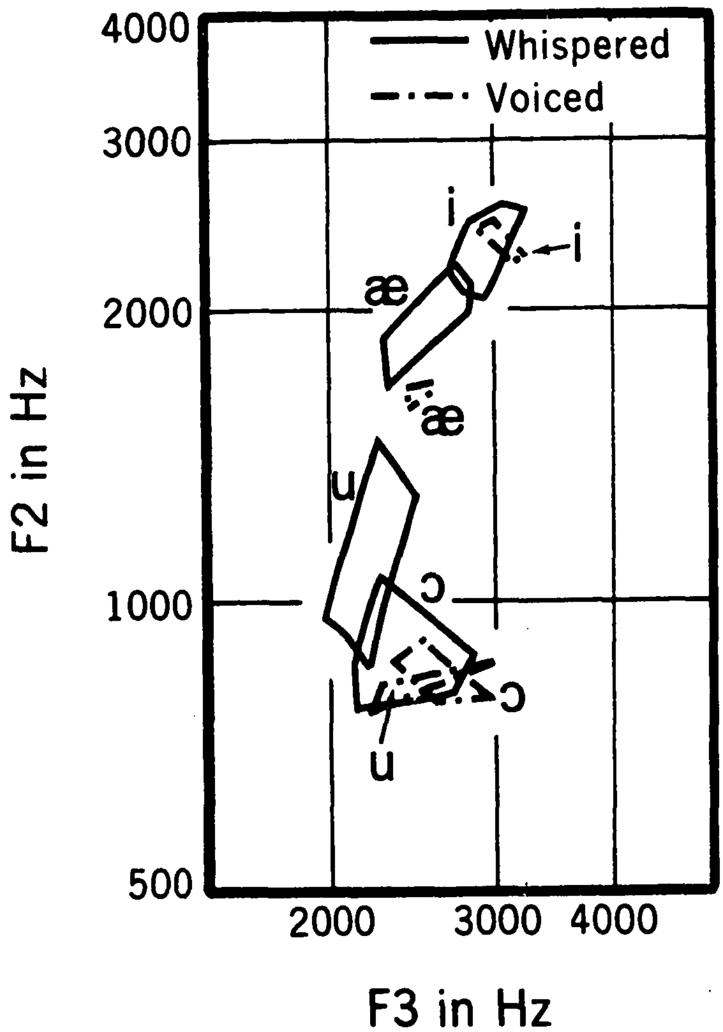

Figure 17.--Frequency areas of formants two and three for Majority whispered vowel samples (this study) and Preferred voiced vowel samples (Fairbanks and Grubb). The data are for male subjects only.

vocalization). Whether or not articulatory postures are different for whispered and voiced vowels, as Peterson (30) suggested, was not evaluated in this investigation, but might be usefully explored in further research.

Whispered Vowel Formant Bandwidths

Tables 9 and 10 present separately for each test vowel and for male and female productions the mean bandwidth (in Hz) and the SD for each of the first three formants of the test vowels. The bandwidth values were grouped according to the Intended, Majority, and Unanimous categories discussed earlier, to permit an examination of possible relationships between formant bandwidth measures and the degree of agreement among listeners regarding vowel phonetic identity. The tables show, however, that neither narrow nor wide formant bandwidths appeared to be consistently associated with correct listener identification of the intended vowel. That is, for some formants of some vowels, the mean bandwidth of the Intended samples was larger than that for the Unanimous samples, but for others the mean bandwidth for the Unanimous samples was larger.

Tables 9 and 10 show that, for productions by subjects of both sexes, the front vowels tested (/i/ and /æ/) showed the widest formant bandwidths and the back vowels tested (/ɔ/ and /u/), the narrowest. The third formant bandwidth for both front vowels was markedly wider than that for the first and second formants of those vowels. In contrast,

TABLE 9.--The mean bandwidth and SD (in Hz) of F1, F2, and F3 for male whispered vowels. The means are over all Intended samples, those correctly identified as the intended vowel by a Majority (six or more) of the eleven judges, and those identified as the intended vowel by Unanimous agreement among the eleven judges.

	Vowels							
	/i/		/æ/		/ɔ/		/u/	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formant One								
Intended	133	41	139	60	118	40	106	28
Majority	135	41	139	62	115	42	104	25
Unanimous	140	44	123	28	113	50	87	21
Formant Two								
Intended	113	46	160	84	112	32	103	24
Majority	115	47	149	73	111	35	104	26
Unanimous	124	60	117	25	87	13	104	26
Formant Three								
Intended	214	61	240	80	138	43	97	22
Majority	220	60	246	72	136	48	96	22
Unanimous	224	61	261	110	156	57	94	18

TABLE 10.--The mean bandwidth and SD (in Hz) of F1, F2, and F3 for female whispered vowels. The means are over all Intended samples, those correctly identified as the intended vowel by a Majority (six or more) of the eleven judges, and those identified as the intended vowel by Unanimous agreement among the eleven judges.

	Vowels							
	/i/		/æ/		/ɔ/		/u/	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formant One								
Intended	145	54	175	69	146	69	133	36
Majority	153	52	182	70	319	136	116	13
Unanimous	178	57	177	63	a		118	17
Formant Two								
Intended	161	61	192	59	147	69	149	67
Majority	163	66	186	60	319	136	109	55
Unanimous	118	41	190	77	a		113	44
Formant Three								
Intended	235	49	256	67	142	52	127	50
Majority	242	50	248	54	90	35	108	25
Unanimous	267	31	246	57	a		95	14

[a]No female /ɔ/ productions were identified as /ɔ/ by all eleven judges.

F3 was generally similar in bandwidth to F1 and F2 for the back vowels. The formant bandwidths ranged from 44 to 460 Hz over all samples of all test vowels, with an overall mean for male productions of 139 Hz (SD = 45 Hz) and for female productions of 167 Hz (SD = 41 Hz).

The magnitude of formant bandwidths obtained in this study was generally greater than that reported previously for phonated vowels (3, 7, 30), and there are no previously reported whispered vowel bandwidth data with which to compare those obtained in this study. With few exceptions, however, the formant bandwidths for the present whispered vowel samples were within the 40-250 Hz frequency range suggested by Fant (12) as probable bandwidth limits for the first three formants of voiced vowels. The partial glottal opening characteristic of whispering may account, in part, for the slightly wider formant bandwidths for whispered than for voiced vowels. As was noted in the review of literature, both House and Stevens (19) and Fujimura and Lindqvist (14) reported wider formant bandwidths for the open- than for the closed-glottis conditions and suggested that such a difference probably was attributable to additional damping at formant frequencies associated with coupling the trachea to the supraglottic spaces.

Relative Amplitudes of Whispered Vowel Formants

The mean formant amplitude measures were grouped first according to the previously-defined Intended, Majority, and Unanimous categories (see Table 23, Appendix E). There was little difference, however, among these mean amplitude values for any vowel formant and, in the following discussion, only the Intended vowel means are considered. The Intended means for whispered vowels are compared with similar means reported by Peterson and Barney (33) for voiced vowels.

As Peterson and Barney (33) found in their earlier study of phonated vowels, the mean amplitudes of individual formants for the whispered vowels in this study were generally similar across the two sexes (Appendix C). The greatest mean formant amplitude (over all vowel productions by subjects of both sexes) was associated with F1 for /ɔ/ productions, and this value served as the reference for obtaining the relative amplitudes of all other vowel formants. This procedure was essentially the same as that used by Peterson and Barney, and it facilitated the comparison of whispered with voiced vowel formant amplitudes.

Table 11 presents the relative amplitudes of the first three formants of the whispered vowels for this study (obtained according to the above criteria) and comparable data for voiced vowels reported by Peterson and Barney. The table shows that the whispered vowel spectra were generally flatter (i.e., showed less difference between the relative

TABLE 11.--Relative formant amplitudes (in dB re:amplitude of F1 for /ɔ/) of whispered (W) and voiced (V) vowels (values for voiced vowels were obtained from Peterson and Barney).

	/i/		/æ/		/ɔ/		/u/	
	W	V	W	V	W	V	W	V
F1	-9	-4	-5	-1	0	0	-7	-3
F2	-4	-24	-5	-12	-1	-7	-7	-19
F3	-8	-28	-13	-22	-17	-34	-18	-43

amplitudes of the three formants) than those for the voiced vowels. This finding was not unexpected in view of the energy decrease (with increasing frequency) in the voiced sound source spectrum and the relatively flat continuous noise spectrum associated with whispering. It was of interest, however, that the highest peak formant amplitude for the whispered vowels studied was associated with the same vowel as that reported by Peterson and Barney for phonated vowels; that is, F1 for /ɔ/.

Figure 18 presents "mean" spectra of each of the four test whispered vowels, representing the mean frequency, bandwidth, and relative amplitude values presented in Tables 4, 9, and 11, respectively. The inter-formant levels are based on estimates, but preserve the "natural curvature" of the formants. The figure shows that for /i/, the amplitude of F1 is lower than that for F2. Both /æ/ and /u/ are characterized by an F1 level which approximates that for F2, while F1 for /ɔ/ is slightly higher in amplitude than F2. For all test vowels, F3 presents a lower amplitude than F2, with the lowest F3 amplitude being associated with the back vowels /ɔ/ and /u/.

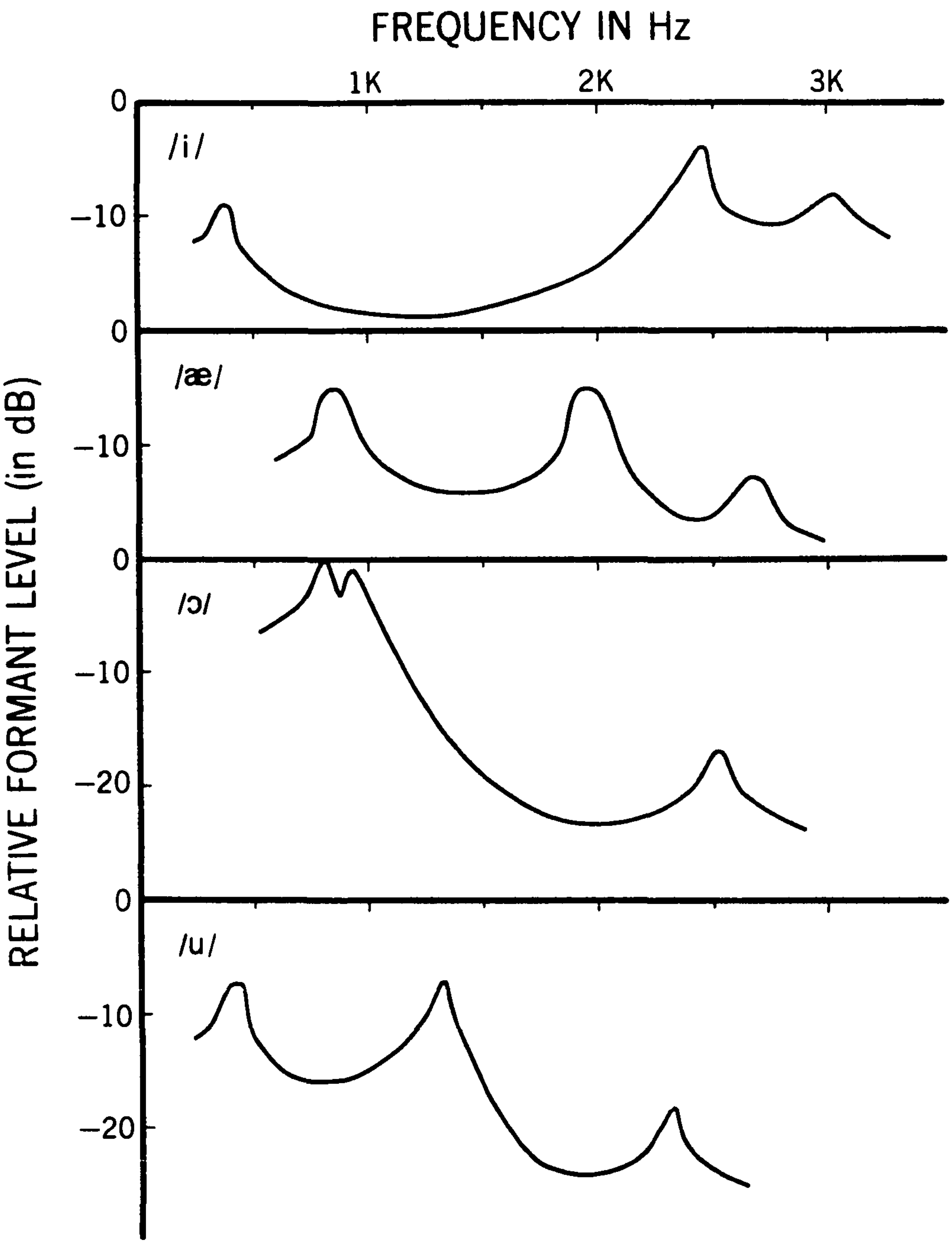

Figure 18.--"Mean" spectra for whispered /i/, /æ/, /ɔ/, and /u/ productions.

CHAPTER V

SUMMARY

The purpose of this study was to investigate the formant frequency, bandwidth, and amplitude characteristics of selected whispered vowels (/i/, /æ/, /ɔ/, /u/) produced by adult male and female subjects. The whispered vowel samples, each produced at the same controlled intensity and mouth-to-microphone distance, were individually recorded on magnetic tape. A two-second central portion of each five-second recorded vowel sample was obtained for perceptual judgments and acoustic wave analysis. Eleven judges independently but simultaneously rated each two-second vowel sample with regard to its phonetic identity. Tape loops of the two-second sustained vowel samples were also analyzed spectrographically by very narrow-band (3-Hz) constant-bandwidth wave analysis. Measures of formant frequency, effective bandwidth, and amplitude were obtained from each whispered vowel spectrum.

The narrow-band (3-Hz) acoustic wave analysis was found to produce a clearer spectral representation of the vocal tract resonance features of whispered vowels than has been reported previously. The analysis method used appeared to overcome certain difficulties associated with Sonagraphic analysis of whispered vowels (26, 30, 46). Specifically, the narrow-band analysis made possible the resolution of individual formants which were very close together in frequency (e.g., F1 and F2 for some /ɔ/ productions). Additionally, the extended (80 dB) dynamic range of the graphic level recorder component of the analyzer assembly aided in showing the level of acoustically weak formants (e.g., F3 for some /u/ and /ɔ/ productions). It was thus possible to make more precise measurements of the whispered vowel formant frequencies than had been possible in previous studies employing Sonagraphic analysis (23, 30, 46). The narrow-band analysis also facilitated fairly precise measurements of formant bandwidths and amplitudes. Such measures have not been reported previously for whispered vowels. A slight elevation of approximately 3.7 dB in the overall level of spectral components (relative to that which would be obtained using a 1-Hz bandwidth filter) was attributable to the 3-Hz bandwidth filter (29). It appeared reasonable to assume, however, that this difference would not affect the relative amplitude level obtained for individual formants.

The present study revealed that formants for whispered vowels tend to be higher in frequency than those for the same vowels produced with voicing. This difference appeared to relate, at least in part, to differences in the degree of glottal openness associated with whispering and voicing. It appeared, however, that the general relationship of formant frequencies to the lingual posture and degree of lip rounding associated with each test whispered vowel was similar to that reported previously for voiced vowels (22, 41, 45). In comparing the whispered vowel formant frequencies for male and female productions, it was found that mean formant frequencies for the female productions were approximately 19% higher than those for the male productions.

The findings regarding formant bandwidths indicated that bandwidths were slightly wider for whispered vowels than had been previously reported for voiced vowels. This difference may relate to additional vocal tract damping associated with the coupling of the trachea to the supraglottal spaces. Formant bandwidths for whispered productions of the front vowels /i/ and /æ/ were generally wider than those for the back vowels /ɔ/ and /u/. Additionally, for the front vowels, the bandwidth of F3 was markedly and consistently wider than that for F1 and F2, whereas for the back vowels, the bandwidths for all three formants were of similar magnitude. No consistent relationship was apparent between formant bandwidths and the degree of agreement among judges regarding phonetic identity of the test vowel productions.

Regarding relative formant amplitudes, the present findings revealed that the whispered test vowels were generally characterized by relatively flat spectra in comparison to those for voiced vowels. That is, there was relatively little difference between the peak amplitudes of the three formants of all four test vowels. The greatest peak amplitude difference between the three formants was found for the back vowels tested. The low back vowel /ɔ/ was characterized by the highest first formant amplitude, and both back vowels /ɔ/ and /u/, presented very low amplitude third formants. In contrast, there was comparatively little difference in the relative amplitudes of the three formants of the front vowels, /i/ and /æ/.

The present investigation demonstrated the capability of a modern acoustic analyzer to show relatively clearly the acoustic spectral features of whispered vowel productions. The findings suggest several possibilities for further research using similar methods. A comparable wave analysis method might be applied, for example, in studies of the effects of nasalization on the acoustic spectra of whispered vowels. Additionally, very narrow-band acoustic analysis may aid in delineating more precisely (than was previously possible) the acoustic spectral features of voiceless continuant (fricative or sibilant) consonants. Such consonants are similar to whispered vowels in that they are produced by a sustained, aperiodic noise which is generated when the expiratory air flow is forced through a narrow aperture.

Concepts regarding the acoustic source, transfer, and radiation function characteristics during vowel production set forth in current acoustic vowel theory (10, 40) have been predicated mainly on data for phonated vowels. The data presented in this study suggest, however, that the formant parameters of whispered vowel productions are different from those for voiced vowel productions. The present data might be useful, therefore, in a systematic extension of acoustic vowel theory to include a consideration of whispered vowels. Finally, the experience gained in this study suggested also that whispered vowels might be studied more readily by using real-time analysis instruments which permit vowel spectra and spectral measures to be obtained more rapidly.

BIBLIOGRAPHY

1. Bell, C. G., Fujisaki, H.; Heinz, J. M.; Stevens, K. N.; and House, A. S. Reduction of speech spectra by analysis-by-synthesis techniques. J. Acous. Soc. Amer., 33, 1725-1736 (1961).

2. Berg, J. van den. Transmission of the vocal cavities. J. Acous. Soc. Amer., 27, 161-168 (1955).

3. Bogert, B. P. On the bandwidth of vowel formants. J. Acous. Soc. Amer., 25, 791-792 (1953).

4. Broad, D. J., and Peterson, G. E. The acoustics of speech. In L. Travis, ed., Handbook of Speech Pathology and Audiology. New York: Appleton-Century-Crofts (1971).

5. Curtis, J. F. Acoustics of speech productions and nasalization. In D. C. Spriestersbach and D. Sherman, eds., Cleft Palate and Communication. New York: Academic Press (1968).

6. Deal, R. E. Some waveform and spectral features of vowel roughness. Ph.D. dissertation, University of Oklahoma (1971).

7. Dunn, H. K. Methods of measuring vowel formant bandwidths. J. Acous. Soc. Amer., 33, 1737-1746 (1961).

8. Fairbanks, G. A physiological correlative of vowel intensity. Speech Monographs, 17, 390-395 (1950).

9. Fairbanks, G., and Grubb, P. A psycho-physical investigation of vowel formants. J. Speech Hearing Res., 4, 203-219 (1961).

10. Fant, G. Acoustic Theory of Speech Production. The Hague: Mouton (1960).

11. ________. Analysis and synthesis of speech processes. In B. Malmberg, ed., Manual of Phonetics. Amsterdam: North-Holland Publishing Company (1968).

12. ________. On the predictability of formant levels and spectrum envelopes from formant frequencies. In M. Halle, H. Lunt, and H. MacLean, eds., For Roman Jakobson. The Hague: Mouton (1956).

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

14. Fujimura, O., and Lindqvist, J. Sweep-tone measurements of vocal-tract characteristics. J. Acous. Soc. Amer., 44, 541-558 (1971).

15. Gray, G. W., and Wise, C. M. The Bases of Speech. New York: Harper and Brothers (1959).

16. Hanson, W. Vowel spectral noise levels and roughness severity ratings for vowels and sentences produced by adult males presenting abnormally rough voices. Ph.D. dissertation, University of Oklahoma (1969).

17. House, A. S. A note on optimal vocal frequency. J. Speech Hearing Res., 2, 55-60 (1959).

18. ________. Formant bandwidths and vowel preference. J. Speech Hearing Res., 3, 3-8 (1960).

19. House, A. S., and Stevens, K. N. Estimation of formant bandwidths from measurements of transient response to the vocal tract. J. Speech Hearing Res., 1, 309-315 (1958).

20. Kersta, L. G. Amplitude cross-section representation with the sound spectrograph. J. Acous. Soc. Amer., 20, 796-801 (1948).

21. Koenig, W.; Dunn, H. K.; and Lacy, L. Y. Sound spectrograph. J. Acous. Soc. Amer., 18, 19-49 (1946).

22. Ladefoged, P. Elements of Acoustic Phonetics. Chicago: University of Chicago Press (1962).

23. Lehiste, I. Suprasegmentals. Cambridge: The MIT Press (1970).

24. Lindblom, B., and Sundberg, J. Acoustical consequences of lips, tongue, jaw, and larynx movement. J. Acous. Soc. Amer., 50, 1166-1179 (1971).

25. Lively, M. A., and Emanuel, F. W. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult females. J. Speech Hearing Res., 13, 503-517 (1970).

26. Meyer-Eppler, W. Realization of prosodic features in whispered speech. J. Acous. Soc. Amer., 29, 104-106 (1957).

27. Operating Instructions: Type 1900-A Wave Analyzer and Type 1910-A Recording Wave Analyzer. West Concord, Mass.: General Radio Company (1964).

28. Paul, A. P.; House, A. S.; and Stevens, K. N. Automatic-reduction of vowel spectra. J. Acous. Soc. Amer., 36, 303-308 (1964).

29. Peterson, A. P. G., and Gross, E. E. Handbook of Noise Measurement. West Concord, Mass.: General Radio Company (1967).

30. Peterson, G. E. Parameters of vowel quality. J. Speech Hearing Res., 4, 10-29 (1961).

31. ________. Vowel formant measurements. J. Speech Hearing Res., 2, 173-183 (1959).

32. ________. The information-bearing elements of speech. J. Acous. Soc. Amer., 24, 629-637 (1952).

33. Peterson, G. E., and Barney, H. L. Control methods used in a study of the vowels. J. Acous. Soc. Amer., 24, 175-184 (1952).

34. Peterson, G. E., and Shoup, J. The elements of an acoustic phonetic theory. J. Speech Hearing Res., 9, 68-99 (1966).

35. Pols, L. C. W.; Tromp, H. R. C.; and Plomp, R. Frequency analysis of Dutch vowels from 50 male speakers. J. Acous. Soc. Amer., 53, 1093-1101 (1973).

36. Potter, R. K., and Peterson, G. E. The representation of vowels and their movements. J. Acous. Soc. Amer., 20, 528-535 (1948).

37. Pressman, J., and Kelemen, G. Physiology of the larynx. Physiological Reviews, 35, 506-554 (1955).

38. Sansone, F. E., and Emanuel, F. W. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult males. J. Speech Hearing Res., 13, 489-502 (1970).

39. Schwartz, M. Power spectral density measurements of oral and whispered speech. J. Speech Hearing Res., 13, 445-446 (1970).

40. Stevens, K. N., and House, A. A. An acoustical theory of vowel production and some of its implications. J. Speech Hearing Res., 4, 303-320 (1961).

41. ________. Development of a quantitative description of vowel articulation. J. Acous. Soc. Amer., 27, 484-493 (1955).

42. ________. Perturbation of vowel articulations by consonantal context: An acoustical study. J. Speech Hearing Res., 6, 111-128 (1963).

43. Stevens, K. N.; House, A. S.; and Paul, A. P. Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation. J. Acous. Soc. Amer., 40, 123-132 (1966).

44. Stevens, K. N.; Kasowski, S.; and Fant, G. An electrical analog of the vocal tract. J. Acous. Soc. Amer., 25, 734-742 (1953).

45. Sundberg, J. Formant structure and articulation of spoken and sung vowels. Folia phoniat., 22, 28-48 (1970).

46. Thomas, I. B. Perceived pitch of whispered vowels. J. Acous. Soc. Amer., 46, 468-470 (1969).

47. Wayte, R. F. Some relationships between spectral harmonic levels and vocal roughness for vowels produced by adult females. Ph.D. dissertation, University of Oklahoma (1971).

48. Whitehead, R. L. Some spectrographic and perceptual features of normal, vocal fry, and simulated abnormally rough vowel phonations. Ph.D. dissertation, University of Oklahoma (1970).

49. Zemlin, W. R. Speech and Hearing Science: Anatomy and Physiology. Englewood Cliffs, N. J.: Prentice-Hall, Inc. (1968).

APPENDIX A

Instructions to Subjects and Instructions to Judges

Instructions to Subjects

You are to produce some isolated vowel sounds in a sustained whisper. To insure correct pronunciations, you will practice producing each test vowel and other vowels which are very similar to the test vowels with respect to tongue position. It is expected that by practicing and listening carefully, you will be able to produce samples which are as representative of each test vowel as possible. The intensity of each whispered vowel production will be maintained throughout at the level indicated by the black mark on the VU meter before you. After each test vowel has been practiced sufficiently, two five-second productions of that vowel will be recorded. The amber light on the signal box will be your signal to take a breath and prepare for the vowel production. The red light, two seconds later, will signal you to initiate the test vowel production. Peak the needle on the VU meter at the black mark and hold it as steady as possible for the duration of the red light (five seconds). If you make a mistake during the recording, we will repeat the test procedure. Do you have any questions?

Instructions to Judges

You are asked to listen to sustained vowel samples and to identify the vowel being produced. Each vowel has been produced in an isolated sustained whisper. As a guide to your vowel identifications, you will find on your answer sheet nine key words with the phonetic symbol for the vowel in each word. Your task is to listen to each whispered vowel sample and then to write on the appropriately-numbered line the phonetic symbol which identifies the vowel as you perceive it. You may hear individual samples a second time if the intended vowel is not clear to you on the first presentation. Do not leave any test vowel unidentified. It is important that you rely on your own individual judgment for each vowel identification. Do not be influenced by what you may see of the responses of your neighbors. Do you have any questions?

APPENDIX B

The Wave Analyzer and Its Calibration

The Wave Analyzer and Its Calibration

The General Radio, Type 1910-A wave analyzer assembly (27) used in this investigation included a constant-bandwidth heterodyne-type wave analyzer with a filter bandwidth tuneable to 3-, 10-, and 50-Hz. The analyzer's frequency range was from 0 Hz to 54,000 Hz, with frequency accuracy to 50,000 Hz of ± ½% of the frequency dial reading plus 5 Hz. In the 3-Hz bandwidth mode, which was used in this investigation, the intensity of frequency components in a complex signal was at least 30 dB down at ±6 Hz, at least 60 dB down at ±15 Hz, and at least 80 dB down at ±25 Hz from the center frequency. The spectrum level was elevated approximately 3.7 dB (relative to that which would be obtained using a 1-Hz filter) due to the 3-Hz filter bandwidth (29). Because of the quasi-random nature of the whispered (noise) signal, with no prominent pure-tone components, it seemed reasonable to assume that the correction factor was approximately constant across all test frequencies and, thus, would have no major influence on the relative formant amplitude levels. The analyzer's signal-to-noise ratio was at least 75 dB.

An electric motor drive system mechanically tuned the wave analyzer through its frequency range. This drive system also moved the chart paper in a component graphic level recorder, thus synchronizing movements of the chart paper and the wave analyzer's frequency-tuning dial. The wave analyzer's output voltage, which was proportional to the intensity of the frequency components in a 3-Hz band of the complex signal under analysis, served as input to the graphic level recorder component of the analyzer assembly. The level recorder was equipped with an 80 dB input potentiometer which was accurate within ± 1% of full scale decibel value. The recorder output was linear in decibels and was plotted as a function of frequency on the chart paper. The chart paper was ruled in 2-dB intervals vertically and 100-Hz sections horizontally.

The frequency and intensity calibration of the wave analyzer was accomplished by the following procedure. First, with the wave analyzer set in its 3-Hz bandwidth mode, a 1000-Hz tone, checked for accuracy with a universal counter (TSI, Model 361), was introduced from an oscillator (Hewlett-Packard, Model ABR200) into the wave analyzer, with the wave analyzer's frequency dial set at 1000 Hz. The F-Zero dial was then used to tune the wave analyzer for maximum volt meter deflection, thus completing the frequency calibration. Next, to accomplish the intensity calibration, a previously-recorded 75 dB SPL 1000-Hz reference tone, recorded as -2 dB VU deflection, was introduced from the tape recorder (Ampex, Model AG 440) into the wave analyzer. The gain of the analyzer and the pen excursion of the coupled graphic level recorder were then adjusted to 75 dB. With this adjustment completed, the wave analyzer's full-scale attenuator could be altered without changing the relationship of the tape recorder output (VU meter set at -2 dB) to the wave analyzer intensity calibration.

APPENDIX C

Formant Frequencies, Bandwidths, and Amplitudes for Individual Whispered Vowel Productions

TABLE 12.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /i/ by each of twelve male subjects.

Subject		Frequency (Hz)			Bandwidth (Hz)			Amplitude (dB)		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
1	a	387	2079	2946	115	133	215	55	43	39
	a	342	2104	2847	188	104	215	51	40	39
2	a	448	2437	2921	121	129	248	55	55	52
	a	462	2479	2917	81	92	117	60	58	53
3	b	385	2537	3056	242	131	167	42	55	52
	b	373	2554	3019	198	256	192	43	56	52
4	b	367	2544	3219	92	73	217	55	59	48
	b	394	2523	3235	121	115	285	54	56	48
5	a	431	2277	2752	110	123	181	56	55	53
	a	417	2385	2971	85	125	160	59	54	50
6	b	400	2500	2913	165	206	181	49	55	52
	b	315	2519	2860	146	160	175	51	56	52
7	b	329	2365	3117	123	71	196	46	58	55
	a	367	2331	3000	154	104	321	44	55	48
8	b	400	2548	3052	100	154	167	53	55	52
	b	365	2508	3050	154	79	194	47	56	50
9	a	433	2500	3200	106	90	163	53	59	51
	a	419	2475	3112	150	65	285	52	58	48
10		427	2300	2962	79	98	162	59	59	54
		425	2269	2929	154	88	144	53	58	52
11	b	308	2465	2967	115	75	354	55	57	46
	a	354	2488	3073	177	75	231	44	58	50
12	b	329	2200	2767	100	88	298	44	57	52
	b	363	2212	2721	125	77	267	46	58	52
Mean		385	2400	2984	133	113	214	51	55	50

[a]Majority Sample

[b]Unanimous Sample

TABLE 13.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /æ/ by each of twelve male subjects.

Subject		Frequency (Hz)			Bandwidth (Hz)			Amplitude (dB)		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
1		773	1600	2475	358	98	210	48	54	52
		858	1577	2337	233	88	217	50	55	51
2	a	912	1988	2633	65	227	185	62	48	44
	b	887	2038	2679	77	117	249	61	51	45
3	a	650	2000	2567	115	390	138	49	52	56
	a	640	1885	2565	102	96	177	44	58	54
4		800	1940	2558	146	160	142	57	53	51
	b	777	1912	2554	121	77	127	54	58	52
5	a	833	2100	2862	138	217	275	56	53	47
		800	2038	2900	135	160	408	55	56	45
6	a	800	1854	2327	125	163	298	56	50	48
	a	800	1900	2400	100	73	300	61	57	49
7	b	977	1979	2873	154	140	377	55	57	48
	b	1012	2027	2856	129	138	365	59	56	41
8	a	835	2100	2700	98	235	329	58	48	44
	a	827	2165	2775	140	385	238	60	46	42
9	a	767	1885	2681	187	129	202	54	54	40
	a	800	1912	2700	175	135	287	60	54	38
10	b	800	1885	2600	133	113	185	58	56	45
	a	769	1900	2600	146	121	271	60	55	47
11	a	887	1827	2525	119	217	142	58	52	51
	a	858	1800	2529	58	142	177	61	53	50
12	a	746	1662	2363	127	104	281	52	58	45
	a	748	1687	2354	150	125	188	51	56	45
Mean		815	1903	2601	138	160	240	56	54	47

[a]Majority Sample

[b]Unanimous Sample

TABLE 14.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /ɔ/ by each of twelve male subjects.

Subject		Frequency (Hz)			Bandwidth (Hz)			Amplitude (dB)		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
1		723	1000	2148	165	121	117	57	55	50
		713	987	2000	156	131	150	53	51	44
2	a	777	979	2515	125	94	62	60	57	44
	a	788	980	2375	108	125	185	60	57	39
3	a	881	881	2590	71	71	141	62	62	48
	a	800	800	2725	112	112	125	61	61	50
4		735	1071	2167	98	135	175	57	54	41
		800	1050	2129	148	104	102	55	58	46
5	a	838	838	2200	169	169	140	60	60	36
	a	888	888	2187	185	185	165	59	59	34
6	b	769	769	2160	71	71	163	61	61	40
		800	800	2158	115	115	158	60	60	37
7	a	829	829	2800	106	106	165	61	61	43
	a	863	863	2867	106	106	154	60	60	48
8	a	869	1077	2292	88	110	181	60	54	40
		900	900	2448	73	73	146	61	61	38
9	a	900	900	2500	177	177	100	60	60	40
	b	856	975	2548	100	102	117	62	61	33
10	b	787	923	2775	92	100	250	61	56	36
	b	769	885	2940	100	83	108	60	58	37
11	a	737	969	2465	60	79	65	62	56	44
	a	785	1017	2458	73	94	94	60	55	48
12	a	729	917	2640	129	135	100	59	58	36
	b	719	983	2517	200	79	142	56	58	38
Mean		802	928	2442	118	112	138	59	56	41

[a]Majority Sample

[b]Unanimous Sample

TABLE 15.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /u/ by each of twelve male subjects.

Subject		Frequency (Hz)			Bandwidth (Hz)			Amplitude (dB)		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
1	a	383	1233	2250	87	142	62	47	37	34
	b	383	1163	2179	65	144	115	50	36	29
2		430	1208	2217	125	100	110	60	54	46
	b	413	1192	2262	96	77	102	58	54	43
3	a	421	1000	2287	127	100	90	56	58	46
	a	400	985	2277	142	135	71	58	58	42
4	a	360	1079	2340	110	85	112	57	52	34
	a	346	1100	2300	92	83	87	57	53	40
5	a	448	1323	2362	73	142	117	57	49	52
		431	1462	2313	75	102	79	55	44	47
6		408	1367	2367	87	83	96	58	54	46
	a	425	1338	2354	106	69	100	57	58	47
7	a	383	1027	2267	115	96	158	50	57	37
	a	417	1058	2431	110	100	100	51	56	48
8	a	435	848	2279	79	108	112	54	55	38
	a	435	883	2300	108	106	112	52	55	42
9	a	337	1300	2354	100	133	73	63	46	43
	a	375	1275	2360	94	129	94	60	44	31
10	b	435	1292	2515	81	87	100	57	56	46
		454	1350	2519	171	110	117	55	55	45
11	b	371	1388	2244	75	96	81	52	54	52
	b	400	1477	2300	119	115	71	47	50	50
12	a	413	958	2000	169	65	63	50	60	47
	a	438	1010	2060	129	69	100	50	59	44
Mean		406	1180	2297	106	103	97	55	52	43

[a]Majority Sample

[b]Unanimous Sample

TABLE 16.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /i/ by each of twelve female subjects.

Subject		Frequency (Hz) F1	F2	F3	Bandwidth (Hz) F1	F2	F3	Amplitude (dB) F1	F2	F3
1	a	517	3300	4169	169	192	194	52	52	50
	b	513	3300	4121	225	208	300	56	50	49
2	b	385	3046	3758	188	112	256	50	58	52
	b	421	3058	3762	144	85	254	53	58	50
3	a	446	2677	3087	62	137	165	48	55	52
	b	429	2623	3038	83	138	233	47	55	53
4	a	448	2485	3119	102	135	158	43	55	51
	b	458	2500	3113	96	248	198	46	54	53
5	b	462	3063	3773	215	85	304	43	59	51
	b	413	3112	3854	262	92	252	42	56	52
6	a	487	2877	3677	125	304	212	54	43	55
	a	485	2725	3662	144	240	213	51	44	54
7		529	3156	3740	140	112	227	54	54	51
		588	3187	3919	87	152	204	50	54	53
8	a	329	2687	3231	112	185	181	43	56	52
	a	310	2733	3319	92	140	269	44	58	49
9	a	463	3165	3900	140	131	204	52	56	54
	a	500	3121	3742	210	215	260	52	54	54
10	b	442	2565	3087	169	96	229	40	58	52
	a	346	2637	3427	173	127	171	41	54	50
11		313	2900	3265	54	204	215	51	52	49
	a	321	2800	3273	131	271	310	46	52	49
12	a	417	2837	3425	121	127	333	45	57	50
	b	419	2831	3321	135	129	304	44	57	50
Mean		435	2891	3533	141	161	235	48	54	51

[a]Majority Sample

[b]Unanimous Sample

TABLE 17.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /æ/ by each of twelve female subjects.

Subject		Frequency (Hz) F1	F2	F3	Bandwidth (Hz) F1	F2	F3	Amplitude (dB) F1	F2	F3
1		1056	2575	3463	102	227	462	62	48	36
		988	2654	3800	102	215	248	63	47	37
2	b	1000	2352	3700	125	292	190	53	56	30
	a	1000	2367	3737	113	238	298	52	56	31
3	a	923	2000	2823	175	173	223	53	52	44
	b	962	2212	3058	221	156	285	53	53	43
4	a	929	1900	2867	373	96	213	47	58	47
	b	973	1885	2900	275	92	279	49	58	46
5	b	1100	2235	3162	83	317	223	60	51	45
	a	1083	2354	3300	87	238	279	60	52	42
6	a	977	2185	3112	238	173	212	51	53	50
	a	954	2200	3077	208	173	267	48	52	49
7	a	1142	2685	3617	140	204	217	59	51	43
		1135	2550	3712	171	267	221	57	47	44
8	b	1046	1987	2760	113	92	319	55	58	50
	b	1065	2044	2900	158	190	315	53	56	51
9	a	1100	2240	3100	106	202	380	62	53	50
	b	1135	2173	2975	185	210	242	56	53	49
10	b	950	1846	2735	192	185	192	46	60	54
	a	950	1900	2762	244	100	219	43	61	50
11	a	1069	1969	3148	173	160	277	54	56	45
	b	1083	2042	3183	237	173	165	54	52	46
12	a	862	2013	3050	192	221	240	56	52	49
	a	877	1967	3000	185	213	169	54	51	51
Mean		1015	2181	3164	175	192	256	54	54	45

[a]Majority Sample

[b]Unanimous Sample

TABLE 18.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /ɔ/ by each of twelve female subjects.

Subject		Frequency (Hz)			Bandwidth (Hz)			Amplitude (dB)		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
1		925	1319	2800	110	221	192	61	53	46
		871	1265	2823	85	179	119	60	51	46
2	a	1000	1000	2979	223	223	65	59	59	37
	a	829	829	2979	415	415	115	56	56	40
3		912	1177	2754	160	129	83	55	56	45
		925	1169	2713	158	121	129	56	55	42
4		1015	1015	3000	140	140	119	61	61	42
		985	985	3012	156	156	96	59	59	47
5		937	1142	2977	137	102	160	57	61	42
		900	1148	2910	210	165	140	56	55	47
6		942	1248	2779	148	115	248	57	55	42
		948	1269	2787	92	138	231	58	55	44
7		1000	1000	2923	115	115	109	59	59	23
		985	985	3065	100	100	158	61	61	27
8		967	967	2817	102	102	119	63	63	47
		981	981	2771	200	200	113	59	59	52
9		942	1283	2927	104	133	169	58	57	44
		888	1273	2923	154	117	221	58	58	42
10		892	1065	2542	100	100	119	60	53	51
		942	1100	2473	142	131	92	56	56	58
11		1015	1015	3125	87	87	250	62	62	41
		1017	1242	3063	113	121	148	60	54	42
12		1019	1173	2515	123	146	100	57	56	45
		1013	1200	2554	121	81	106	58	56	44
Mean		952	1119	2842	146	147	142	59	55	43

[a]Majority Sample

TABLE 19.--Frequencies (in Hz), bandwidths (in Hz), and amplitudes (in dB) of the first three formants of each of two productions of the vowel /u/ by each of twelve female subjects.

Subject		Frequency (Hz) F1	F2	F3	Bandwidth (Hz) F1	F2	F3	Amplitude (dB) F1	F2	F3
1		537	1583	2937	129	140	106	60	50	46
		592	1746	3146	104	62	75	60	55	40
2	a	400	1377	2785	106	188	112	52	45	46
	b	400	1362	2746	137	163	111	52	46	45
3		427	1685	2785	106	104	81	49	52	46
	a	431	1671	2637	112	85	154	48	52	48
4		410	1100	2815	183	112	77	45	56	42
		433	1081	2837	181	129	63	46	56	43
5		456	1427	2900	160	87	123	48	58	44
	a	456	1390	2942	127	44	94	42	60	42
6		477	1758	2662	104	273	246	49	45	36
		513	1779	2700	113	198	198	45	39	31
7		583	1317	2669	94	92	110	48	50	28
		600	1369	2525	138	269	173	45	44	26
8		537	1523	2848	108	160	100	49	51	50
		552	1631	2900	98	87	135	45	52	46
9		533	1046	3085	210	210	125	54	53	34
		608	1029	3210	106	267	138	54	54	34
10	b	419	1063	2321	104	79	87	52	55	44
	b	442	1075	2325	112	96	87	52	56	45
11		535	1181	3000	148	185	188	49	47	32
		510	1200	3056	196	196	240	47	45	31
12		575	1048	2773	196	177	140	48	55	36
		583	983	2825	110	175	96	47	50	33
Mean		500	1351	2810	133	149	127	49	51	40

[a]Majority Sample

[b]Unanimous Sample

APPENDIX D

Within- and Between-Subjects SDs for Measures of Formant Frequency, Bandwidth, and Amplitude

TABLE 20.--Frequency standard deviations (in Hz) within (W) and between (B) subjects for each of three formants of each test vowel, sexes treated separately.

Vowel	Sex	F1 W	F1 B	F2 W	F2 B	F3 W	F3 B
/i/	Male	26.34	55.32	28.98	207.96	63.93	189.88
	Female	27.84	101.81	45.20	362.01	91.92	491.29
/æ/	Male	23.94	120.07	36.41	219.24	37.82	248.44
	Female	21.35	116.34	65.30	357.36	100.04	473.67
/ɔ/	Male	29.75	82.01	48.70	114.61	74.97	373.43
	Female	41.07	63.00	60.42	179.52	40.47	260.91
/u/	Male	15.72	43.67	41.18	257.71	43.52	161.36
	Female	22.56	99.29	44.98	385.72	69.44	314.14

TABLE 21.--Bandwidth standard deviations (in Hz) within (W) and between (B) subjects for each of three formants of each test vowel, sexes treated separately.

Vowel	Sex	F1		F2		F3	
		W	B	W	B	W	B
/i/	Male	33.29	47.39	34.91	54.93	54.21	66.77
	Female	30.53	69.56	37.66	79.62	43.21	55.41
/æ/	Male	31.72	80.34	77.87	89.28	48.52	103.29
	Female	34.67	92.77	35.76	77.01	67.45	67.26
/ɔ/	Male	27.51	50.29	26.25	37.09	44.18	38.95
	Female	50.40	84.20	49.94	84.75	33.82	66.87
/u/	Male	24.80	31.56	13.67	31.95	21.28	22.05
	Female	32.53	39.19	47.94	82.16	28.60	65.85

TABLE 22.--Amplitude standard deviations (in dB) within (W) and between (B) subjects for each of three formants of each test vowel, sexes treated separately.

Vowel	Sex	F1		F2		F3	
		W	B	W	B	W	B
/i/	Male	3.28	7.07	1.35	6.48	1.96	5.48
	Female	1.93	6.49	1.27	5.64	1.09	2.36
/æ/	Male	2.37	6.44	2.41	4.09	1.75	6.29
	Female	1.73	7.38	1.32	5.29	1.26	8.91
/ɔ/	Male	1.34	2.99	1.95	3.58	2.90	6.79
	Female	1.56	2.61	2.48	3.69	2.65	9.88
/u/	Male	1.65	5.95	1.66	9.27	4.12	7.95
	Female	1.87	6.15	2.36	7.05	2.06	9.86

APPENDIX E

Mean Amplitudes (in dB) for _Intended_, _Majority_, and _Unanimous_ Vowel Productions

TABLE 23.--The mean amplitude and SD (in dB) of F1, F2, and F3 for the whispered test vowels. The means are over all Intended samples, those correctly identified as the intended vowel by a Majority (six or more) of the eleven judges, and those identified as the intended vowel by Unanimous agreement among the eleven judges.

	Vowels							
	/i/		/æ/		/ɔ/		/u/	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Formant One								
Intended	50	5.3	54	8.8	59	2.2	52	5.1
Majority	49	5.3	55	5.0	60	1.6	53	4.7
Unanimous	48	5.0	55	4.0	60	2.3	53	3.5
Formant Two								
Intended	55	4.3	54	3.5	58	3.0	52	5.9
Majority	55	4.4	54	3.4	58	2.3	52	6.5
Unanimous	56	2.0	55	2.9	59	2.2	51	6.9
Formant Three								
Intended	51	3.2	46	5.5	42	6.0	41	6.8
Majority	51	3.3	46	5.3	41	5.0	43	6.0
Unanimous	51	2.0	46	5.9	37	2.6	44	6.9