69-12,445

SANSONE, Jr., Frank Edward, 1939-

SPECTRAL NOISE LEVELS AND ROUGHNESS SEVERITY RATINGS FOR NORMAL AND SIMULATED ROUGH VOWELS PRODUCED BY ADULT MALES.

The University of Oklahoma, Ph.D., 1969
Speech-Pathology

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SPECTRAL NOISE LEVELS AND ROUGHNESS SEVERITY RATINGS FOR NORMAL AND SIMULATED ROUGH VOWELS PRODUCED BY ADULT MALES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

FRANK EDWARD SANSONE, JR.

Oklahoma City, Oklahoma

1969

SPECTRAL NOISE LEVELS AND ROUGHNESS SEVERITY

RATINGS FOR NORMAL AND SIMULATED ROUGH

VOWELS PRODUCED BY ADULT MALES

APPROVED BY

F. W. Emanuel

Donald Cumber

Walter L. Cull

James B. Snow Jr. MD

John F. McCoy

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The writer is deeply indebted to Dr. Floyd W. Emanuel, director of this research, for his guidance, counsel, and criticism during the planning and completion of this study. Appreciation is also expressed to Dr. John F. McCoy of the Department of Biostatistics and Epidemiology, University of Oklahoma Medical Center, for his assistance and counsel concerning the statistical analysis of the data. Thanks are due to Sister Lourdes Leal for her assistance in the collection of the data and to the subjects and judges who unselfishly participated in this study.

The writer also wishes to express his gratitude to his wife Patricia for her continued encouragement and understanding during this period of graduate study.

TABLE OF CONTENTS

TABLE OF CONTENTS--(Continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SPECTRAL NOISE LEVELS AND ROUGHNESS SEVERITY RATINGS FOR NORMAL AND SIMULATED ROUGH VOWELS PRODUCED BY ADULT MALES

CHAPTER I

INTRODUCTION

The voice disorders are usually classified on a perceptual basis as deficiencies in vocal pitch, loudness, and quality. Writers often employ different terms to identify similar voice quality disturbances and efforts to establish a standard terminology for the description of voice disorders have been only partially successful. In conventional terms, however, a rough voice is one perceived to be hoarse or harsh in quality (63).

Studies of the laryngeal structures and their movements (42, 49, 65, 66, 72) suggest that vocal roughness may reflect various anatomical and physiological conditions which interfere with vocal fold vibration during phonation. Rough voice may be a manifestation of structural malrelationships within the larynx or surrounding tissues, deficiencies in phonatory mechanism innervation, tissue masses and ulcerations on the vocal folds or adjacent structures, acute or chronic vocal fold inflammation or edema, or life-threatening malignancies (45, 46, 52, 70, 79). Emotional disorders may also be evidenced in phonatory mechanism dysfunction and vocal roughness (16, 39).

The perceptual assessment of rough voice aids clinicians in detecting, evaluating, and treating conditions underlying this voice abnormality. Individuals vary, however, in their ability to detect vocal roughness and to estimate its severity reliably. The voice sample obtained, the environment in which it is heard, the clinician's prior experience in evaluating voice adequacy, and other inconstancies may influence the perceptual appraisal of roughness. A more objective voice evaluation, than that provided by perceptual assessment alone, is needed to insure early detection of rough voice and to aid in evaluating effects of medical treatment and voice therapy administered for its remediation. In this regard, investigations are of interest which relate acoustic voice features to the perception of normal and rough phonation.

Only recently has there been an extensive research effort to define relationships between perceived vocal roughness and acoustic features of voice. In recent years, the acoustic correlates of vocal roughness have been investigated in studies of synthesized speech and speech-like sounds (9, 10, 73, 74, 75), in studies of the recorded acoustic waves of human phonation (38), and in studies of spectral features of human phonation (30, 36, 44, 64, 69, 77).

Studies of the acoustic spectra of rough voices have been few, and it appears that further spectrographic investigations may contribute needed information. Recent investigations suggest, for example, that spectrographic analysis of the voice wave may be useful in identifying acoustic features which relate to listeners' perceptions of vocal roughness. Specifically, acoustic spectrography has revealed that the elevation of noise components in vowels is related to a perceived increase in the severity of vocal roughness. On the basis of his observation of elevated noise components in the spectra of rough vowels, Nessel (44) indicated that hoarseness could be defined and differentiated spectrographically. More recently, Isshiki, Yanagihara, and Morimoto (30), and Yanagihara (76, 77) have identified the elevation of vowel noise components as a spectral feature associated with hoarse phonation. Quantitative data delineating precisely the relative magnitude of these noise components for vowels produced normally or with vocal roughness is deficient, however.

The present study sought to investigate quantitatively noise components in narrow-band (3-Hz) spectra of normal and rough productions of selected vowels. It also sought to investigate possible relationships between the spectral noise measures and judgments of vowel roughness. It was thought that the study might contribute information useful in understanding the acoustic features which differentiate normal and rough phonation and, thus, might facilitate the development of improved evaluation techniques and treatment modes for individuals with rough voice.

CHAPTER II

REVIEW OF THE LITERATURE

Vocal roughness is of critical interest to both speech and medical clinicians because of their responsibility for rehabilitation of persons presenting this condition. This interest has engendered a voluminous clinical literature concerning the assessment and treatment of rough voice. Only recently, however, has there been a concerted effort to investigate the acoustics of vocal roughness, a situation largely attributable to instrumentation limitations and attendant difficulties in obtaining refined acoustic measurements. With improved instrumentation, additional data regarding vocal acoustics have been contributed. A deficiency of quantitative data delineating acoustic differences between rough and normal phonation remains, however. Little is known, for example, about possible relationships between spectrographically definable acoustic voice features and the perception of vocal roughness.

The purpose of this investigation was to assess quantitatively the acoustic spectra of vowels phonated normally and with simulated vocal roughness, and to consider possible relationships between vowel spectral features and listener judgments of vowel roughness. The literature reviewed as background for this study is reported under two major headings: (a) physiological features of vocal roughness, and (b) acoustic features of vocal roughness.

Physiological Features of Vocal Roughness

Normal phonation is described by the myoelastic-aerodynamic theory. As presented by van den Berg (67) and others, this theory holds that normal voice production involves laryngeal regulation of the expiratory air stream. Expiratory air emerging through the respiratory passages encounters resistance at the level of the glottis. This resistance is afforded by the vocal folds which, at the initiation of phonation, are completely or nearly completely approximated and held in a state of tension by the laryngeal adductor muscles. As expiratory effort persists, subglottic air pressure increases against the adducted folds until it overcomes their resistance and explosively parts them. The resulting rapid, upward movement of air between the parted folds creates a Bernoulli effect in the glottis. The Bernoulli effect, the elasticity of the displaced folds and surrounding tissues, and the momentary reduction in subglottic pressure due to air escape serve to reapproximate the folds (71). The entire sequence is repeated in rapid succession, releasing through the glottis a series of air puffs which strike the supraglottic air column and set it into vibration at audible frequencies (35).

The number of air puffs released through the glottis per unit time is thought to be dependent upon the effective length, mass, tension, and damping of the folds, the Bernoulli effect between them, and the subglottic pressure (21, 22, 24, 25, 67, 68, 71). The rate at which the puffs strike the supraglottic air column determines the acoustic fundamental vocal frequency. Vocal acoustic intensity is related to the force with which the puffs strike the supraglottic air (28, 31). This force is thought to be determined largely by the interaction of subglottic pressure and glottic resistance (27, 28, 68).

In general, synchronous movements of the vocal folds characterize normal phonation (1, 13, 14). The two folds move simultaneously in opening and in closing and each fold's movement is approximately in phase with the other (43). Differences in the movements of the medial margins of the folds for relatively low and high pitches within an individual's range have, however, been reported (1, 78). At low pitch, the lower margins of the folds are first to part and first to approximate in each vibratory cycle, the upper margins slightly lagging the lower in these movements. At high pitch, the folds are lengthened and thinned (20, 23, 26), and a difference in the parting and approximation times of the upper and lower vocal fold margins is generally not discernable (78).

Utilizing high speed photography to slow the apparent motion of the folds in phonation, Timcke, Moore, and von Leden (65) defined three intracycle phases of vocal fold vibration: an opening phase in which the folds move laterally from the approximated position, a closing phase in which the folds move medially toward approximation, and a closed phase in which the folds are approximated. Normal phonation is characterized by the presence of all three phases in each cycle. While the opening phase generally accounts for approximately 25 percent, the closing phase 45 percent, and the closed phase 30 percent of the time of each vibratory cycle, variations in this basic pattern normally obtain. For example, changes in the intensity and frequency of the voice are associated with changes in the relative duration of intra-cycle phases (65).

In general, sequential vocal fold vibratory cycles in sustained normal phonation vary but little in duration, amplitude, and intra-cycle phase durations (43, 49, 65). However, slight variations in the duration and amplitude of vocal fold movements in successive vibratory cycles are associated with normal phonation (43). It is said (14), therefore, that quasi-periodic vibrations, varying slightly in amplitude, are characteristic of the laryngeal sound generator in normal phonation.

In contrast, vocal roughness is associated with deviations from normal vocal fold vibratory patterns. For example, Moore and Thompson (42) found that differences in the periods of adjacent vocal fold vibratory cycles tend to increase as perceived hoarseness increases. They suggested that laryngeal vibratory frequency variations are related to the perceived severity of a hoarse vocal quality. Yanagihara (76) has reported that the amplitude of vocal fold movement is often much less in rough phonation than in normal phonation. Further, investigators (66, 72) have found that in rough phonation the opening and closing phases of vocal fold vibration may account for almost all of the total time of each cycle, and that the closed phase may be unusually short or nonexistant. Von Leden, Moore, and Timcke (72) report that marked asymmetry of vocal fold movements, resulting in different intra-cycle phase durations for each fold, may be associated with vocal roughness. Other intra-cycle irregularities observed in rough phonation include different amplitudes and speeds of movement in each fold and lack of movement in one or both folds (72).

According to Zemlin (78) a consistent opening of the glottis may persist within individual vocal fold vibratory cycles during both rough and normal phonation. During rough phonation, however, the glottal opening may be sufficient to allow excessive air escape, producing turbulence in the expiratory air stream. Isshiki and von Leden (29) suggest that vocal roughness results, in part, from an "...imperfect modulation of the air stream at the glottis and...subsequent turbulence in the flow of air." Yanagihara (76) reported that during hoarse phonation the time patterns of flow rates for orally emitted air are characterized by random variability in shape, intensity, and periodicity. He suggested that the glottal conditions producing such variability favor the creation of turbulent air flows which result in the acoustic noise associated with hoarseness.

Generally, then, variations in the intra-cycle phases of vocal fold vibration, asymmetry of vocal fold movements, variation in the frequency and amplitude of consecutive vocal fold vibratory cycles in a phonatory sequence, and turbulence in the expiratory air stream have been associated with the perception of rough voice quality. Similarly, disturbances in the product of laryngeal vibration, the vocal acoustic wave, are thought to be related to rough voice. Relationships between vocal roughness and certain acoustic voice features are considered in the following section.

Acoustic Features of Vocal Roughness

Acoustic Wave Features

Voicing occurs when puffs of air emitted through the glottis impart kinetic energy to the supraglottic air column causing it to vibrate and, thus, to generate a complex sound wave (31). The wave's shape is determined by the frequencies, amplitudes, and phase relationships of its components (4, 35). Generally, human listeners are not very sensitive to the phase relationships of acoustic wave components, however, and differently shaped waves may be perceived as phonetically equivalent sounds (35).

Cyclic repetitions of equal duration characterize acoustic waves derived from periodic vibrating sources (35). The frequency of glottic puff emissions in phonation, however, reflects the quasi-periodic pattern of vocal fold vibration. It follows, therefore, that successive acoustic cycles in a segment of sustained normal phonation should vary somewhat in period. Lieberman (38) found that differences in the period of successive cycles, or "pitch perturbations", of less than 0.5 ms were typical of isolated vowels phonated normally, while perturbations of less than 1.0 ms were typical of normal vowels in connected speech. Cooper, Peterson, and Fahringer (10) have reported that when period variations are eliminated in synthesized speech, listeners perceive the sample to be mechanical and unnatural. This is consistent with Lieberman's (37) observation that, "Pitch perturbations are apparently essential cues to natural speech quality."

Perturbations in normal vowels, however, tend to be small in comparison to those in rough vowels. Lieberman (38) found that pitch perturbations for mildly and moderately rough phonations generally exceed those for normal phonations. He also noted that the acoustic voice wave becomes markedly aperiodic and that individual cycles within it are not discernable when roughness is severe. In general, the wave of rough phonation more nearly approximates the aperiodic wave characteristic of noise than the periodic wave characteristic of tonal sounds. Moore and Thompson (42) also found that differences in the periods of consecutive acoustic cycles are generally greater in severely rough than in mildly rough phonation. Further, Coleman (8) reported that small random changes in fundamental vocal frequency occur less frequently in segments of normal phonation than in hoarse segments of comparable duration. Large frequency breaks, often an octave in extent, were reported by Bowler (5) to characterize harsh vocal quality.

To study the aperiodicity required to cause listeners to judge an acoustic signal rough, Wendahl (73) synthesized complex acoustic stimuli which varied randomly in frequency around a median frequency. When its frequency was varied as little as $\pm$ 1 Hz around the median, the signal was perceived as rough. When the frequency variation was increased, the degree of perceived roughness also increased. While such pitch perturbations in synthetic signals appear to be related to perceived signal roughness (73), only a few relationships between acoustic wave characteristics and roughness in human phonation have been clearly established. It has been demonstrated that human voices may be placed in rough and normal categories on the basis of voice wave frequency variability measures (8, 38). The studies of Coleman (8) and Lieberman (38), however, suggest that acoustic features which relate to the perception of vocal roughness may not be easily or fully identified from inspection of the acoustic wave envelope alone.

Brubaker and Dolpheide (6) suggest that the duration and intensity of the sample may also affect perceived vocal roughness. Sherman and Linke (55) have suggested that listeners may perceive an increase in vocal harshness as the duration of an utterance is lengthened. To determine the effect of stimulus duration on the perception of vocal roughness, Coleman and Wendahl (9) presented jittered synthesized complex acoustic stimuli to listeners for judgment. As the duration of a jittered sample was increased from .16 to .80 seconds, increased roughness was heard by the listeners.

With respect to the effects of varying signal intensity on roughness ratings, Wendahl (74) found that roughness was perceived when the intensities of adjacent cycles of a synthesized complex stimulus were alternately attenuated causing the signal to shimmer. The influence of variations in signal intensity on roughness ratings of human phonation, however, has apparently not been investigated.

Several studies (50, 51, 54, 55) have investigated the relative harshness of common vowels. In general, vowel harshness appears to be related to tongue height in vowel production. In both sustained vowel phonation and in connected discourse, the high vowels /u/, /i/, /ɪ/, and /U/ are perceived as less harsh than the low vowels /ɔ/, /ɑ/, and /æ/, while the mid vowels /ʌ/ and /ɛ/ fall toward the center on this continuum (51, 55).

To generalize, it seems reasonable to expect that specific features of the acoustic voice signal may relate to the perceived presence and relative severity of vocal roughness. There are, however, few clearly defined relationships between vocal roughness and acoustic wave envelope features. This suggests that inspection of the wave envelope alone may not readily reveal the acoustic features of vocal roughness. Research including a more detailed acoustic analysis may be useful.

Spectrographic Features

A detailed analysis of a complex voice wave may include a consideration of the components making up the wave. Periodic complex sound waves consist of a fundamental frequency, the frequency of repetition, and one or more harmonics of the fundamental (31, 55). Conventional Fourier series analysis is useful in determining the relative amplitudes of the fundamental and harmonics in such waves because this analysis assumes the wave to be composed of a series of sinusoidal functions whose frequencies are in harmonic relation (15, 32). In the past, the components of complex sounds were often estimated by the use of graphic analyzers which instrumentally performed a Fourier series analysis of the signal (2, 18, 59, 60). These analyzers provided a spectrum of a complex sound showing harmonic components only (18) regardless of the signal's actual components. Fourier series analysis of waves containing both harmonic and inharmonic components may be misleading, however, because inharmonic components are not represented and the intensity of harmonic components is represented somewhat inaccurately (53).

Assuming that harmonic analysis would provide an adequate description of the physical basis of vocal quality, Van Dusen (69) analyzed the phonations of speakers with "metallic" voices. For this analysis, he used a Henrici analyzer which treated the voice waves as though they consisted of harmonic components only. A spread of vowel harmonics into high spectral frequencies was found to be typical of the "metallic" voices. This study by Van Dusen appears to be one of the early attempts to examine abnormal voice spectra. While his analysis may have been somewhat inaccurate, it differentiated normal and deviant voices.

With the advent of electronic acoustic analyzers, instrumentation became available which presented both harmonic and inharmonic components in sound waves (18). As they were developed, bandpass (56, 57), diffraction (41), and heterodyne analyzers (17) were applied in voice investigations. Early versions of these instruments were not without limitations, however. Bandpass analyzers typically had broad filter characteristics and a single amplitude measure represented both harmonic and inharmonic energy over a wide frequency range (18). The diffraction and heterodyne analyzers offered fine spectral resolution (18), but early models analyzed so rapidly that selective representation of individual components in a complex wave was likely to be inaccurate. Verification of vowel formant locations reported in earlier studies was accomplished with these instruments (3, 17, 41, 57, 59, 61), but little information relative to the spectral features of vocal roughness was obtained.

Carhart (7), in 1941, used a heterodyne analyzer to study the spectra of tones from a model larynx vibrating under several different conditions. Manual analyzer tuning permitted recording both the harmonic and inharmonic components of a tone within the analyzer's frequency range. For several model larynx vibratory conditions, predominantly inharmonic spectra were obtained. Carhart reported that these inharmonic tones were perceived as having a rough quality similar to hoarseness. He speculated that, in human phonation, inharmonic components should be expected only in certain voice abnormalities.

An automatic heterodyne analyzer which graphically recorded the time-varying speech spectrum was developed at the Bell Telephone Laboratories in 1946 (34). Its advantages over earlier analyzers included a loop record and playback system which allowed continuous presentation of sound to the analyzer circuit and an analysis time permitting a more detailed portrayal of the components in a sound wave. This spectrograph, now manufactured by the Kay Electric Company and commonly known as the Kay Sonograph, has been employed frequently in the study of speech. Spectra derived with this instrument permit fundamental vocal frequency measurements (48), identification of vowel formant locations (12, 48), observations of changing formant patterns in connected speech (11, 34), and indications of the energy distribution in transient speech signals (19, 62).

Using the Sonograph, Thurman (64) studied the relationships between the acoustic spectra of vowels and listeners' perceptions of voice deviation. He investigated quantitatively formant bandwidth, formant frequency locations, and inharmonic energy levels in vowel spectra. He attempted to examine the relationship between these spectral features and the type of voice quality disturbance perceived. Thurman found, however, that the differentiation of voice quality types, the determination of the degree of voice disturbance, and the measurement of inharmonic energy levels in vowels was essentially impracticable from his Sonograph records. He reported specifically that variability in the Sonograph records across subjects, vowels, and quality types limited inferences from his data. For example, in the vowel spectra for hoarse speakers, formant bandwidth changes and formant frequency shifts occurred, but not consistently. Further, amplitude measures of inharmonic vowel components were of questionable validity because the intensity peaks for these components were not clearly indicated in the spectra.

Recently, however, Yanagihara and others (30, 76, 77) have studied Sonograms of hoarse vowels with more success. Major acoustic features which are related to hoarseness in vowels, as reported by these investigators, include the elevation of spectral noise components and the diminution of harmonic components. Specifically, an increase in hoarseness was accompanied by an increase in spectral noise in the low frequencies, which, when roughness was severe, obscured harmonics. As hoarseness became increasingly severe, noise components spread to the higher frequencies. A .65 correlation between four identifiable spectral patterns and the rated severity of voice disturbance was reported.

The narrowest filter bandwidth generally available in the Sonograph is a relatively wide 45 Hz. It appears possible that filter bandwidths this wide obscure spectral information which may be related to the perception of vocal roughness. Studies employing very narrow-bandwidth spectrographs seem to support this view. For example, Nessel (44), using a sound-frequency spectrograph of narrow frequency selectivity, compared the spectra of hoarse vowels to those for normal vowels. In the spectra for rough phonations, he found that harmonic energy below 5000 Hz was replaced by noise and that noise components were elevated in the frequencies above 5000 Hz. Nessel indicated that his spectrographic results "...prove that the phenomenon of 'hoarseness' can be defined and differentiated when using a suitable method of frequency analysis." Since Nessel's work was published several years ago, there does not appear to have been a follow-up investigation employing a similar approach.

The present study sought to quantify the intensity of noise components in normal and in rough vowel spectra and to relate these measures to listener judgments of vowel roughness. The study included a very narrow-band (3-Hz) spectrographic analysis of normal and rough vowels produced by the same speakers, a quantitative analysis of vowel spectral noise, a comparison of the noise levels associated with rough and normal vowels, and a study of the relation of vowel noise levels to listener judgments of vowel roughness. A description of the experimental apparatus and procedures used in the study is presented in the following chapter.

CHAPTER III

DESIGN OF THE INVESTIGATION

This study was designed to investigate noise components in the spectra of vowels produced normally and with simulated vocal roughness, and possible relationships between the vowel spectral noise and listener judgments of vowel roughness. Normal-speaking adult males produced selected vowels both normally and with simulated vocal roughness at one intensity. A magnetic tape recording was made of each vowel production. These recordings were played in random order to a panel of eleven judges who rated the vowels for vocal roughness. A tape loop was also constructed from the recording of each vowel and these were individually analyzed to produce very narrow-band (3-Hz) vowel spectra. As an index of vowel spectral noise, the lowest peak of energy in each of seventy-nine successive 100-Hz spectral sections from 100 Hz to 8000 Hz was measured in each vowel spectrum.

Research Questions

The following research questions regarding the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/ were investigated:

1. What is the relative roughness of the vowels produced normally and with simulated roughness?

2. What are the spectral noise features of normal and of rough productions of the vowels?

3. What are the relationships between measures of spectral noise and roughness ratings for each of the vowels?

The selection of subjects, the experimental apparatus, and the data collection procedures are discussed in the following sections.

Subjects

Twenty normal-speaking white male adults, selected primarily on the basis of their ability to perform the experimental task, served as subjects in this study. The investigation was limited to adult males to provide homogeneity of the subject sample with regard to vocal pitch. Each subject produced selected vowels under two experimental phonatory conditions, normal and rough. Thus, each subject could serve as his own control.

Subjects ranged in age from twenty-three to thirty-three years. The subjects were, therefore, persons who had undergone pubescent voice change, but who had not undergone significant physiological changes in laryngeal structures due to advanced age. Only normal-speaking subjects apparently representative of the normal speaking population were studied and, to insure this, the voice quality and speech of each potential subject were evaluated by a trained speech pathologist.

Speech Sample

Subjects individually sustained each of five vowels for seven seconds, first normally and then with simulated vocal roughness. The vowels /u/, /i/,

/ʌ/, /ɑ/, and /æ/, articulated with tongue positions ranging from high to low (33), were used because previous studies suggest that perceived roughness in vowels may be related to tongue height in their articulation (51, 55). Isolated sustained vowels offered a further advantage; they were suitable for 3-Hz bandwidth spectrographic analysis. A single intensity level, 75 dB re .0002 dyne/cm^2 (SPL), was employed in recording the vowel samples.

Instrumentation

Instrumentation used in data collection included: a signal system, an audio recording system, a wave analyzing system, a playback system, and a calibration system.

Description

Signal system. A simple electro-mechanical cam timer, activated by the experimenter, controlled the illumination of two panel lights used to signal subjects to begin and terminate test vowel phonation.

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

The PZT microphone had a flat frequency response ($\pm$1 dB) from 20 Hz to 8000 Hz when at a 70° angle of incidence to the sound source. Its sensitivity was -60.3 dB re lv/microbar. The sound-level-meter indicated the sound-pressure level at its PZT microphone with an average signal-to-noise ratio in octave bands from 20 Hz to 10000 Hz of at least 66 dB. The tape recorder had a flat frequency response (±2 dB) from 40 Hz to 12000 Hz with a signal-to-noise ratio of at least 65 dB at a tape speed of 15 ips.

The output of the sound level meter was connected directly to the input of the tape recorder. The recorder's output served as input to the monitoring amplifier which functioned as a vocal-intensity indicator. The calibrated scale on the amplifier's voltmeter indicated when subjects were phonating at the required vocal intensity. A simplified diagram of the audio recording system is presented in Figure 1.

Wave analyzing system. A constant bandwidth wave analyzer (General Radio, Type 1910-A) was used in vowel spectrum analysis. The analyzer's frequency range was from 0 Hz to 54000 Hz. Its frequency accuracy to 50000 Hz was ±1/2% of frequency dial reading plus 5 Hz. In its 3-Hz bandwidth mode, 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 center frequency. 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

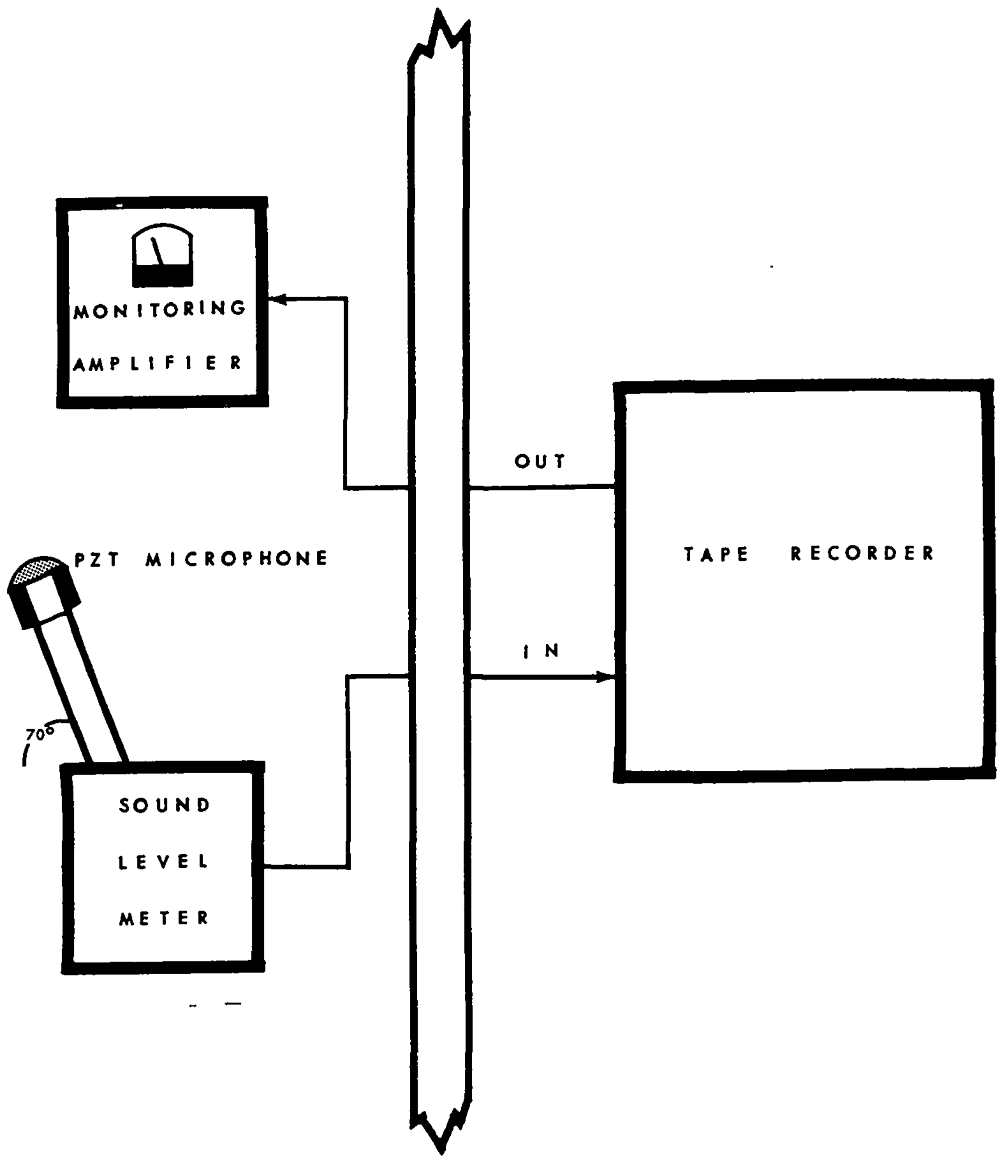

Figure 1.--Simplified diagram of the audio recording system.

3-Hz band of the complex signal under analysis, served as input to the graphic level recorder. The graphic 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. A simplified diagram of the wave analyzing system is presented in Figure 2.

Playback system. A dual-track tape recorder (Ampex, Model 354) with a flat frequency response (±2 dB) from 40 Hz to 12000 Hz at 15 ips was used in conjunction with an amplifier (Sherwood, Model 59900a) and a loud-speaker (Altec, Model 844A) as the playback system for vowel judgments.

Calibration system. A pure tone oscillator (Hewlett Packard, Model ABR200) which drove a loud-speaker (Altec, Model 844A), a sound level meter (General Radio, Type 1551-C), a pulse generator assembly (Tektronix, 160 Series), and a frequency-calibrated condenser microphone assembly (Bruel and Kjaer, Type 2603), were used in instrument calibration. A simplified diagram of the calibration system is presented in Figure 3.

Calibration

Audio recording system. The vocal-intensity-monitoring section of the audio recording system was calibrated to indicate when the subject's vocal intensity had reached 75 dB SPL. The monitoring amplifier's voltmeter was used as the intensity indicator. To calibrate this meter, a 1000-Hz reference tone produced by the pure tone oscillator was led to the loud-speaker. The sound-

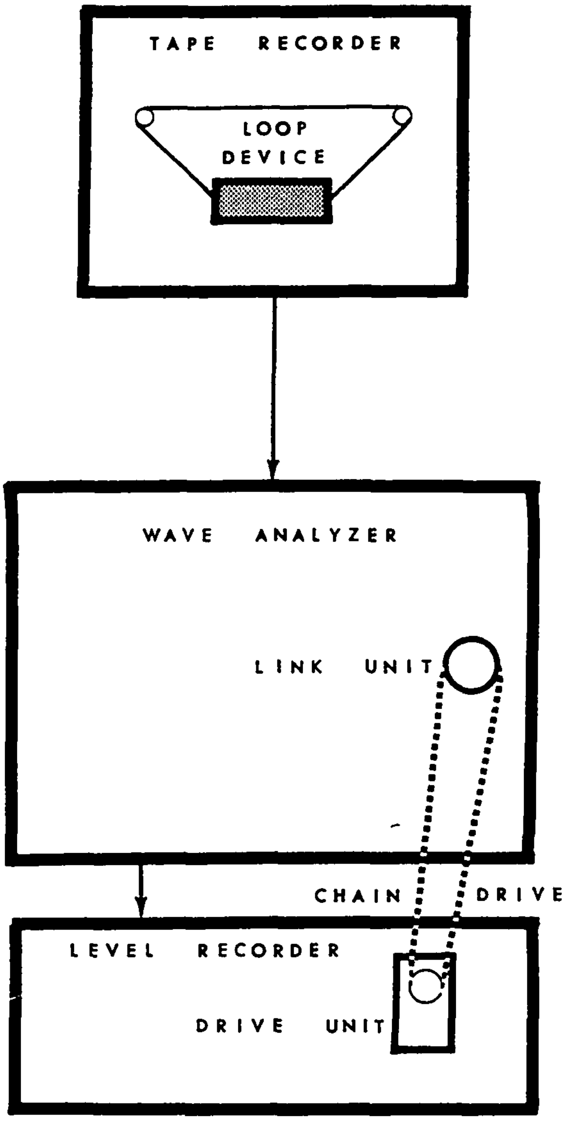

Figure 2.--Simplified diagram of the wave analyzing system.

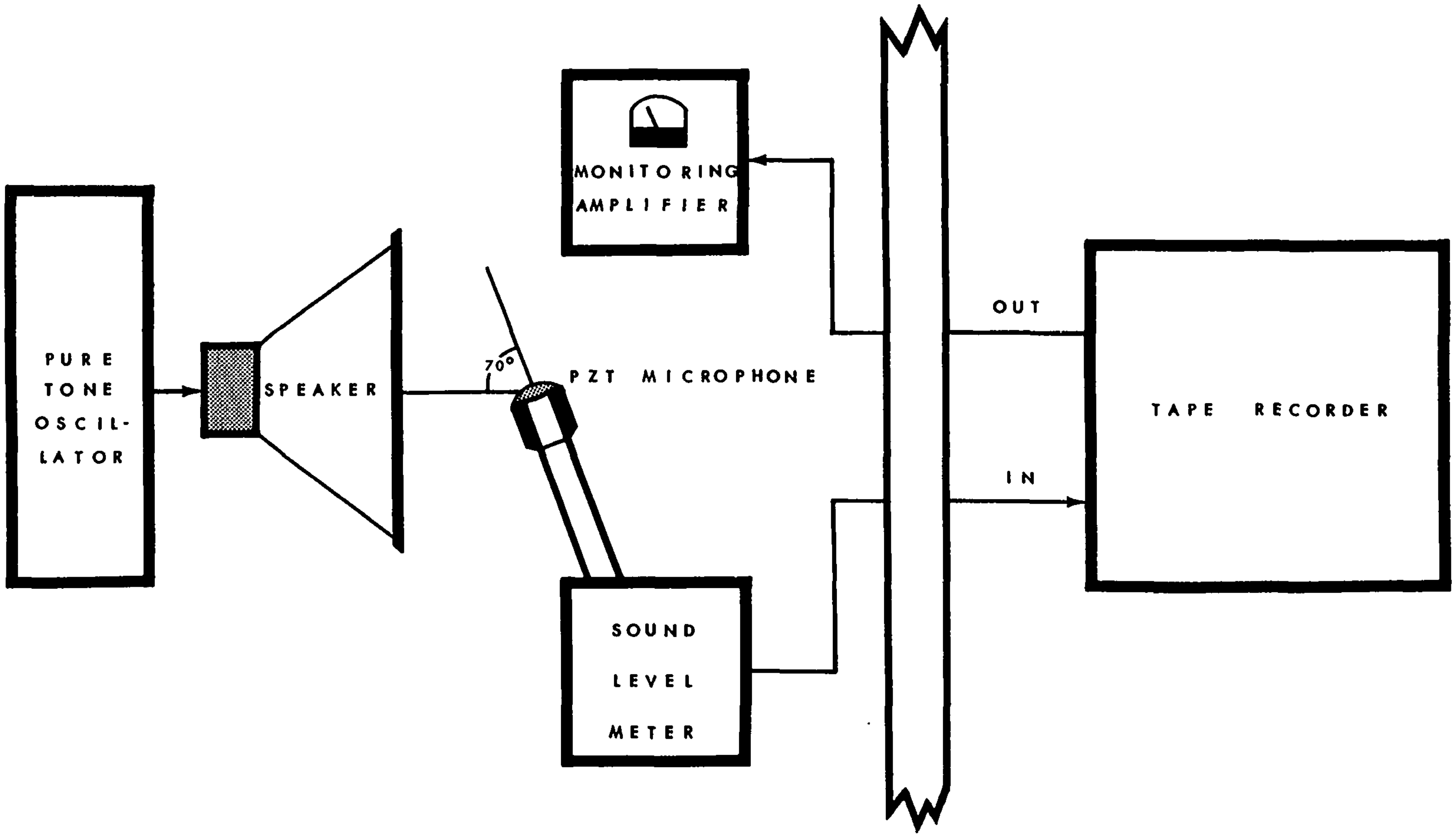

Figure 3.--Simplified diagram of the calibration system.

level-meter microphone was placed at a 70^{0} angle of incidence to, and two feet in front of, the speaker. The intensity of the tone was adjusted to 75 dB SPL. The sound-level-meter output was connected directly to the tape recorder input, and the recorder was adjusted for a -2 dB deflection of its VU meter in response to the 75 dB SPL input. The recorder output was led to the monitoring amplifier, and the deflection of the amplifier's voltmeter was marked as the level each subject was required to maintain during experimental vowel production. A recording was then made of the reference tone. As the reference tone was played back, the audio recorder's reproduce level was adjusted to match the record level. With this adjustment, vowels producing a 75 dB SPL indication on the vocal-intensity-monitoring voltmeter produced a deflection to -2 dB on the recorder's record VU meter. When played back, the recorded vowels produced a deflection to -2 dB on the recorder's reproduce VU meter.

Wave analyzer. The graphic wave analyzer was adjusted to insure minimal carrier frequency intensity at low frequencies, and accurate frequency representation on the analyzer chart paper. The 75 dB SPL 1000 Hz calibration tone was then played into the wave analyzer and the gain of the analyzer was adjusted for a 75 dB SPL pen excursion on the graph paper. The instrument was thus calibrated to record the intensity of complex wave components in decibels SPL.

As evidenced by its accurate plotting of the fundamental and harmonics of a pulse train of known repetition rate produced by the pulse generator assembly, frequency calibration of the wave analyzer appeared to be satisfactory from 0 Hz to 8000 Hz. The frequency response of the coupled audio recording and wave analyzing systems, excluding the microphone, was also checked utilizing a series of pure tones produced by the oscillator, and was found to be flat (±2 dB) from 50 Hz to 12000 Hz. The microphone used in this study was designed for a flat frequency response (±1 dB) from 20 Hz to 8000 Hz. The microphone frequency response was checked against the flat (±.5 dB from 20 Hz to 10000 Hz) response of a calibrated condenser microphone, and was found to be within design specifications.

Procedures

Recording Procedures

All vowel samples were recorded in an acoustically-isolated room with a low ambient noise level. Each subject was first familiarized with the experimental procedure and was then seated in an examination chair. The chair's headrest was adjusted for his comfort, and a headstrap was employed to minimize changes in his position with respect to the microphone. The microphone was placed at a 70° angle of incidence to, and six inches in front of, the subject's mouth.

Subjects sustained, for seven seconds, rough and normal productions of each test vowel at 75 dB SPL (±1 dB). This intensity level was selected because it was noted in preliminary trials that it was a comfortable level for production of both rough and normal vowels. Signal lights were employed to inform the subjects to begin and terminate the test vowel phonations.

The order of vowels was randomized for each subject within normal and rough conditions. The subject was instructed to produce the test vowels first normally and then with vocal roughness. This was done to prevent possible vocal abuse associated with roughness simulation from affecting the normal vowel productions. Instructions read to the subjects are presented in APPENDIX A.

Each test phonation was carefully monitored by the experimenter. If the appropriate vowel was not produced, vocal roughness was not suitably effected, or the experimental intensity was not maintained, the trial was repeated until a satisfactory performance was achieved.

Rating Procedure

The vowels produced by each subject were randomized for presentation to judges. Eleven judges, all graduate students in Communication Disorders, evaluated the vocal roughness associated with each vowel. The judges were seated in a semi-circle nine feet from and facing the loud-speaker. The judgments were made in an acoustically-isolated test room and the recorder used to reproduce the vowels was located in an adjoining control room. Judges listened to each vowel and rated independently the degree of roughness perceived in each. A five point equal-appearing-intervals scale was used in which "1" represented least severe and "5" most severe vocal roughness.

A preliminary rating of all vowel productions was made by the experimenter. Four vowel productions, two representing each of the rating scale extremes, were selected. To provide the judges a common reference for the roughness extremes, these vowels were played several times before rating began. A copy of the instructions given to the judges is presented in APPENDIX B.

The listening session was two and one-half hours long. The vowels were presented in five series each consisting of fifty vowels with ten minute rest periods between each series. The final fifty-vowel series consisted of productions randomly selected from those presented earlier, and were used to evaluate judge reliability. A Pearson _r_ was computed to measure the association between the medians of the eleven judges' first and second ratings of each of these fifty vowels. The _r_ obtained was .96. Percentages of inter- and intra-judge vowel roughness rating agreement for the fifty vowel productions, within 1 scale value, are presented in APPENDIX C. The lowest percentages were generally associated with judge one. The lowest percentage, 80%, was obtained when the vowel ratings made by judge one were compared to those made by judge eight. The percentages for all other judges were equal to or greater than 94%. The intra- and inter-judge reliability indicated by these data appeared adequate for this study. Median scale values (40) of the judges' first ratings for each vowel production were then computed.

Spectral Analysis Procedure

Tape loops, two seconds in duration (tape speed 15 ips), were constructed from the magnetic tape recordings of each test vowel produced by each subject under both normal and rough phonatory conditions. The loops were constructed from the portion of the vowel recording displaying the most uniform intensity as monitored on the recorder's VU meter. The vowel loops were played separately into the wave analyzer to produce a 3-Hz bandwidth graphic spectral analysis of each vowel. The analyzer was operated at a paper speed of .5 inches per minute and a writing speed of 20 inches per second for this analysis to provide adequate data resolution while minimizing writing stylus overshoot.

As a further procedure, a spectrum was also made of recorded test-chamber noise. As illustrated in Figure 4, low noise levels attributable to the instrumental systems and test chamber were evident at all frequencies across the 0 Hz to 8000 Hz spectral frequency range under these conditions. Much lower noise levels were registered at all spectral frequencies, however, when the microphone in the test chamber was connected directly to the wave analyzer and the recording system was bypassed. It appeared, therefore, that the recorder was a major source of the system noise. Repeated spectrographic analyses of recorded room noise indicated that system noise levels would not vary appreciably at different times during the day.

To obtain quantitative spectral data relevant to the research questions, the lowest observable peak level-recorder stylus marking in each 100-Hz section of the 3-Hz bandwidth spectrum of each vowel phonation was measured. Seventy-nine measures, one for each 100-Hz section from 100 Hz to 8000 Hz, from the spectrum of each vowel were obtained. Due to stylus marking overlap, the lowest observable peak in a 100-Hz spectral section, in some instances, may not have been the true low peak. An index of spectral noise levels in dB SPL was, however, provided by this measure. For every measurement, rough vowel noise levels exceeded system noise levels. Occasionally, however, normal

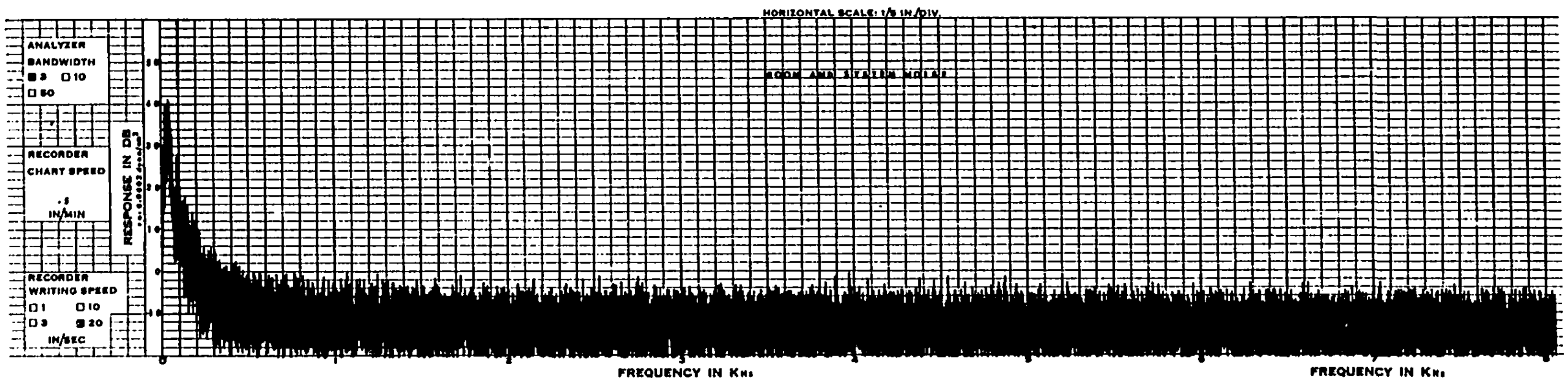

Figure 4.--Spectrum of test chamber and instrument noise.

vowel noise levels did not exceed system levels in the frequencies above 5000 Hz. When this occurred, the high peak system noise level in that 100-Hz spectral section was recorded as the highest noise level possible for the normal vowel production at that measurement point. Thus, high-frequency noise levels for normal vowels may have been elevated slightly because of a system noise artifact. This artifact, if present, may have diminished spectral noise level differences between normal and rough vowel productions in the high-frequency spectral range.

To determine the reliability of the wave analyzing instrumentation on successive runs of the same tape loop, three vowel spectra produced from one loop were measured. Measures of noise levels in 100-Hz sections were averaged over 1000-Hz segments, from 0 Hz to 8000 Hz, of each of the three spectra. The means for comparable segments did not vary more than ± 1 dB across the three spectra. The mean of the 100-Hz section spectral noise levels were then averaged over the 8000-Hz range for each spectrum. These means did not vary more than $\pm .7$ dB across the three spectra. The reliability of the wave analyzing instrumentation was, therefore, considered satisfactory for this study.

CHAPTER IV

RESULTS AND DISCUSSION

Results

This study investigated vowel spectral noise levels and judges' ratings of vowel roughness. Twenty normal-speaking adult males individually sustained normal and simulated rough productions of each of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/ at one vocal intensity. Tape recordings of each production were randomized and played to a panel of eleven judges who rated the vowels for roughness on a five point equal-appearing-intervals scale. The tape recordings were also analyzed to produce a 3-Hz bandwidth frequency-by-intensity spectrum of each vowel's acoustic components. As an index of vowel spectral noise levels, the low peak level of noise components in each 100-Hz section, from 100 Hz to 8000 Hz, of each vowel spectrum was measured in dB SPL. The vowel spectral noise levels and vowel roughness ratings were then compared.

Ratings

Table 1 shows, for eleven judges, median roughness ratings for the vowels produced by each subject. This table reveals that each vowel phonated with simulated vocal roughness evidenced a higher median scale rating than its

TABLE 1

MEDIAN ROUGHNESS RATINGS FOR EACH NORMAL AND ROUGH VOWEL PRODUCTION

	Vowels									
	/u/		/i/		/ʌ/		/ɑ/		/æ/	
Subjects	N	R	N	R	N	R	N	R	N	R
1	1.42	5.00	1.19	4.95	1.00	4.95	1.58	4.95	1.32	4.95
2	1.00	3.25	1.29	4.95	2.06	4.42	3.00	3.90	1.60	4.81
3	1.00	4.08	1.11	3.86	2.00	3.92	1.42	4.58	2.60	4.95
4	1.88	4.06	1.11	3.86	1.46	4.06	1.89	2.67	2.00	3.58
5	1.29	2.60	2.11	3.42	1.71	3.58	2.00	3.19	2.29	3.60
6	1.00	3.71	1.81	4.58	1.05	3.94	1.05	4.00	2.20	4.14
7	1.11	3.05	1.11	4.19	1.29	4.13	1.80	4.81	2.40	4.81
8	1.00	3.11	1.11	3.29	1.19	3.08	1.11	2.89	2.00	4.14
9	1.05	3.06	1.19	3.38	1.29	3.71	1.11	3.42	1.42	3.80
10	1.00	3.89	1.95	5.00	2.00	3.89	1.19	5.00	2.58	5.00
11	1.05	3.63	1.00	4.00	2.58	3.14	1.86	4.71	1.42	4.89
12	1.00	3.00	1.75	3.42	1.86	3.19	1.29	3.94	2.29	4.00
13	2.29	4.81	1.89	4.42	2.00	4.89	2.00	4.89	1.86	4.71
14	1.19	4.29	1.75	4.89	1.42	3.42	1.29	3.42	1.86	4.78
15	1.00	3.42	1.00	3.08	2.00	2.95	1.19	3.19	1.94	4.14
16	1.00	4.42	1.11	4.25	1.89	4.95	1.75	2.63	2.29	3.25
17	1.05	4.20	1.29	5.00	1.42	3.40	1.86	4.29	2.05	4.95
18	1.05	2.71	1.92	2.86	1.89	3.00	1.71	2.71	2.19	3.19
19	1.29	2.60	1.42	2.67	1.42	2.80	1.42	3.14	2.29	3.71
20	1.11	2.75	2.29	2.81	2.25	3.20	2.11	2.92	2.88	3.75

normal counterpart. The subjects were successful, therefore, in their efforts to simulate relatively rough vowels. Table 2 presents median roughness ratings for each normal and rough vowel averaged over all subjects. Considering the

TABLE 2

AVERAGE MEDIAN ROUGHNESS RATINGS FOR EACH NORMAL AND ROUGH VOWEL PRODUCTION

Vowel	Average Median Roughness Rating Normal	Rough
/u/	1.19	3.58
/i/	1.47	3.90
/ʌ/	1.63	3.76
/ɑ/	1.69	3.65
/æ/	2.07	4.26

average median ratings for normal productions, it may be seen that the high vowels /u/ and /i/ tend to be rated less rough than the other test vowels, with /u/ rated less rough than /i/. The low vowel /ɑ/ is rated slightly more rough than the mid vowel /ʌ/. The low vowel /æ/ is rated most rough. The degree of roughness simulated for each vowel was not controlled, but average median scale ratings for the rough productions also tend to differ. Considering the averages for rough productions, the greatest scale value separation is evident between the vowels /u/ and /æ/, with /u/ rated less rough than /æ/. The average median ratings for /ʌ/, /ɑ/, and /i/ are between those for /u/ and /æ/.

Spectral Noise Levels

Figures 5 and 6, showing spectra of normal and rough /ɑ/ productions respectively, exemplify the spectra obtained in this study. Figure 5 shows that the normal /ɑ/ spectrum is characterized by identifiable harmonics throughout much of the analyzed frequency range. The harmonics are obscured by noise at the high frequency end of the spectrum. Where harmonics can be identified, noise components are seen between them. Figure 6 presents the spectrum of the rough /ɑ/ produced by the same speaker under identical experimental conditions. When it is compared to the normal /ɑ/ spectrum, it may be seen that noise components are generally elevated in the rough vowel production and that the harmonics are obscured except in the very low frequency range. It is also evident that the identifiable harmonics in the rough production tend to be somewhat diminished in amplitude with respect to those in the normal production. Further, for both normal and rough /ɑ/ productions, spectral noise is most prominent in the lower frequencies and diminishes toward higher frequencies. Features similar to these for /ɑ/ were observed in the spectra of all test vowels.

Noise levels in each vowel spectrum were estimated by measures of the lowest energy peak in each 100-Hz spectral section from 100 Hz to 8000 Hz. Selected functions of the spectral measure also considered for each vowel spectrum were: (1) the mean of spectral noise measures from 100 Hz to 2600 Hz, (2) the mean of measures from 2600 Hz to 5100 Hz, (3) the mean of measures from 5100 Hz to 8000 Hz, (4) the mean of measures from 100 Hz to 5100 Hz,

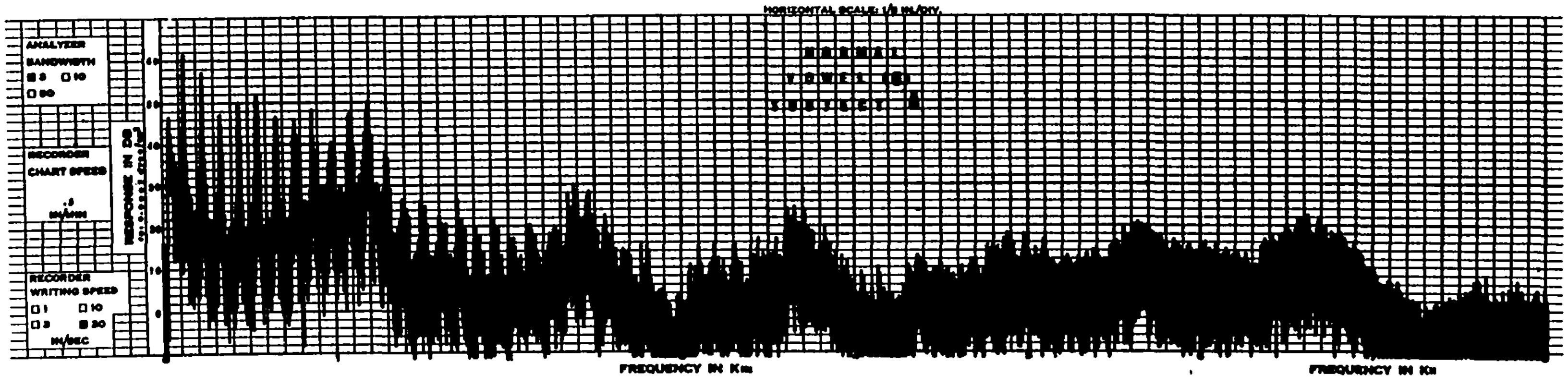

Figure 5.--Spectrum of a normal /a/.

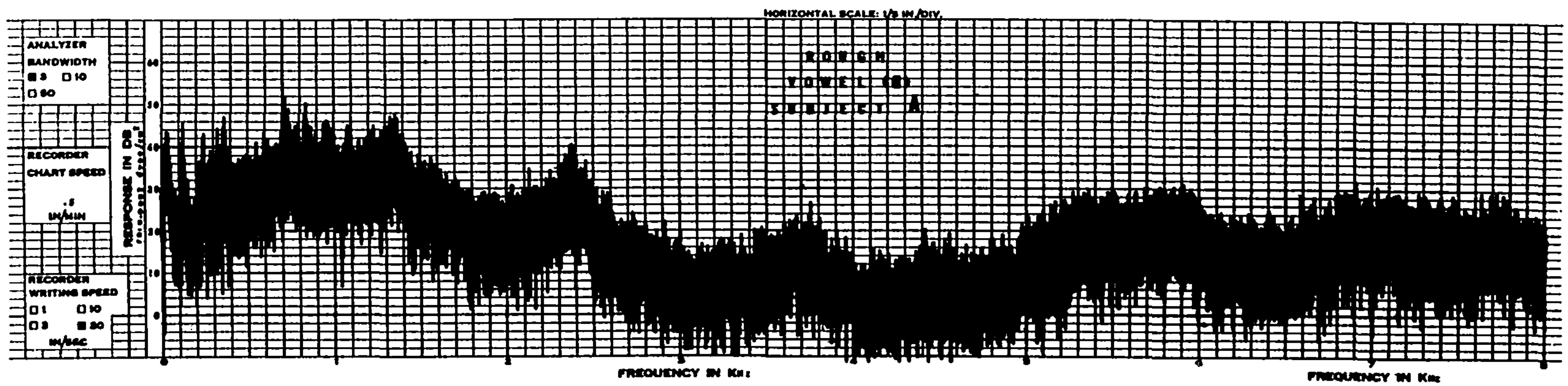

Figure 6.--Spectrum of a rough /a/.

and (5) the mean of measures from 100 Hz to 8000 Hz. Selection of these functions permitted investigation of mean spectral noise levels in the low-frequency range including the vowel formants most critical to vowel identity (47, 48), in the intermediate-frequency range immediately above the major formant frequencies, in the high-frequency range where the elevation of spectral noise in rough phonation has been observed previously (44, 76, 77), in the combined low- and intermediate-frequency range including all major vowel formant frequencies (35), and in the total frequency range considered in previous spectrographic investigations of vocal roughness (64, 77). It was thought that spectral noise level means for these spectral frequency ranges might relate differently to perceived vowel roughness.

To facilitate a detailed presentation of the spectrographic findings, the total spectral frequency range studied (100 Hz to 8000 Hz) is referred to as the TSR; a spectral noise level is referred to as an SNL; and, segments of the total spectral frequency range (TSR) are referred to as SS's. The spectral segments (SS's) studied are referred to as segment one (S-1), 100 Hz to 2600 Hz; segment two (S-2), 2600 Hz to 5100 Hz; segment three (S-3), 5100 Hz to 8000 Hz; and segment four (S-4), 100 Hz to 5100 Hz.

Table 3 presents mean SNL's and standard deviations for rough and normal productions of each test vowel. The means are over all subjects, over the TSR, and separately, over each of the spectral segments (SS's). It may be seen in this table that the mean SNL's for rough productions of each vowel are higher than those for its normal production and that this trend holds for the TSR

TABLE 3

NORMAL AND ROUGH VOWEL SPECTRAL NOISE LEVEL (SNL) MEANS AND STANDARD DEVIATIONS FOR TWENTY MALE SUBJECTS, AND DIFFERENCES BETWEEN THE NORMAL AND ROUGH VOWEL SNL MEANS (SNLD'S)

Spectral Segment S-1 (100 Hz to 2600 Hz)

Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD
/u/	10.9	2.4	26.4	6.2	15.5
/i/	14.9	2.8	27.7	5.0	12.8
/ʌ/	23.9	3.4	35.9	3.6	12.0
/ɑ/	25.4	2.6	36.6	3.7	11.2
/æ/	29.8	3.2	39.3	2.8	9.5

Spectral Segment S-2 (2600 Hz to 5100 Hz)

Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD
/u/	-.2	3.7	11.6	6.0	11.8
/i/	19.1	5.7	29.5	3.9	10.4
/ʌ/	11.5	4.9	23.3	6.1	11.8
/ɑ/	12.4	4.5	24.3	5.0	11.9
/æ/	18.1	4.5	28.9	4.1	10.8

Spectral Segment S-3 (5100 Hz to 8000 Hz)

Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD
/u/	-3.4	3.4	9.2	8.4	12.4
/i/	9.2	4.4	18.4	7.9	9.2
/ʌ/	2.1	4.0	15.0	6.6	12.9
/ɑ/	2.5	5.1	16.1	7.8	13.6
/æ/	6.3	4.9	19.8	6.3	13.5

TABLE 3--Continued

Spectral Segment S-4 (100 Hz to 5100 Hz)					
Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD
/u/	5.3	3.1	19.0	6.1	13.7
/i/	17.0	4.3	28.6	4.5	11.6
/ʌ/	17.7	4.2	29.6	4.8	11.9
/ɑ/	18.9	3.5	28.5	4.4	9.6
/æ/	23.9	3.8	34.1	3.5	10.2

Total Spectral Range (100 Hz to 8000 Hz)					
Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD
/u/	2.5	3.2	15.7	6.9	13.2
/i/	13.0	4.7	25.2	5.6	12.2
/ʌ/	12.5	4.2	24.7	5.4	12.2
/ɑ/	13.4	4.1	24.4	5.5	11.0
/æ/	18.1	4.2	29.3	4.4	11.2

and for each SS. It is also evident that mean SNL's for rough and normal productions of /i/ vary considerably both in absolute magnitude and in relative magnitude with respect to the means for other vowels, across the spectral segments. The rough and the normal vowels exclusive of /i/ could be ranked, in most SS's, with respect to increasing mean SNL's: /u/, /ʌ/, /ɑ/, and /æ/. Reversals in this order occur only among the S-4 (100 Hz to 5100 Hz) and the TSR (100 Hz to 8000 Hz) SNL means for rough vowels. In each of these spectral ranges, rough and normal productions considered separately, the SNL mean for /ʌ/ slightly exceeds the mean for /ɑ/, but the means for /u/ and /æ/ remain at the low and high extremes of the order respectively. Within spectral segment S-1 (100 Hz to 2600 Hz), each of the normal test vowels considered, vowels could be ranked with respect to increasing mean SNL's: /u/, /i/, /ʌ/, /ɑ/, and /æ/. A randomized complete block analysis of variance (58) in which subjects were treated as blocks and vowels as treatments was employed to determine whether the normal vowel S-1 SNL means were significantly different. A Duncan's New Multiple Range Test (58) was also employed to locate differences among the means detected by the analysis of variance. A .05 significance level was set for these analyses. A summary of these statistical analyses is presented in APPENDIX D. The results of these tests indicate that, with the exception of those for /ʌ/ and /ɑ/, all normal vowel S-1 SNL means are significantly different from each other at the .05 level.

Examination of the differences between SNL means for normal and rough productions of each vowel, presented in Table 3 reveals that the SNLD's were generally similar across each of the SS's and the TSR. These differences are illustrated for the vowel /ʌ/ in Figure 7. Figure 7 shows a plot of individual vowel SNL's in each 100-Hz section of the TSR averaged over all subjects, normal and rough /ʌ/ productions considered separately. Differences between SNL means for normal and rough productions of each vowel, similar to those for /ʌ/, were observed for all of the test vowels.

The standard deviations for the normal vowel SNL means are shown in Table 3. It may be seen that for the frequency ranges studied the greatest normal vowel SNL variability is associated with S-2 (2600 Hz to 5100 Hz) and S-3 (5100 Hz to 8000 Hz), while least variability is associated with S-1 (100 Hz to 2600 Hz). It may also be seen that standard deviations for the normal /u/ are generally smaller than those for the other vowels, regardless of the spectral segment considered. SNL variability for /u/ is least in S-1. Standard deviations for mean SNL's in each SS were also obtained for each of the rough vowels. It may be seen in Table 3 that the SNL variability for rough productions generally exceeds that for normal productions of each test vowel in each of the SS's. The greater variability for rough vowels may reflect the fact that the degree of roughness simulated for each vowel by each subject was not controlled.

Spectral Noise Level and Roughness Rating Relationships

To study the relationships between vowel mean SNL's and the median judgments of vowel roughness, scatter diagrams for each of the five experimental vowels were plotted. All the diagrams suggested a positive relationship between

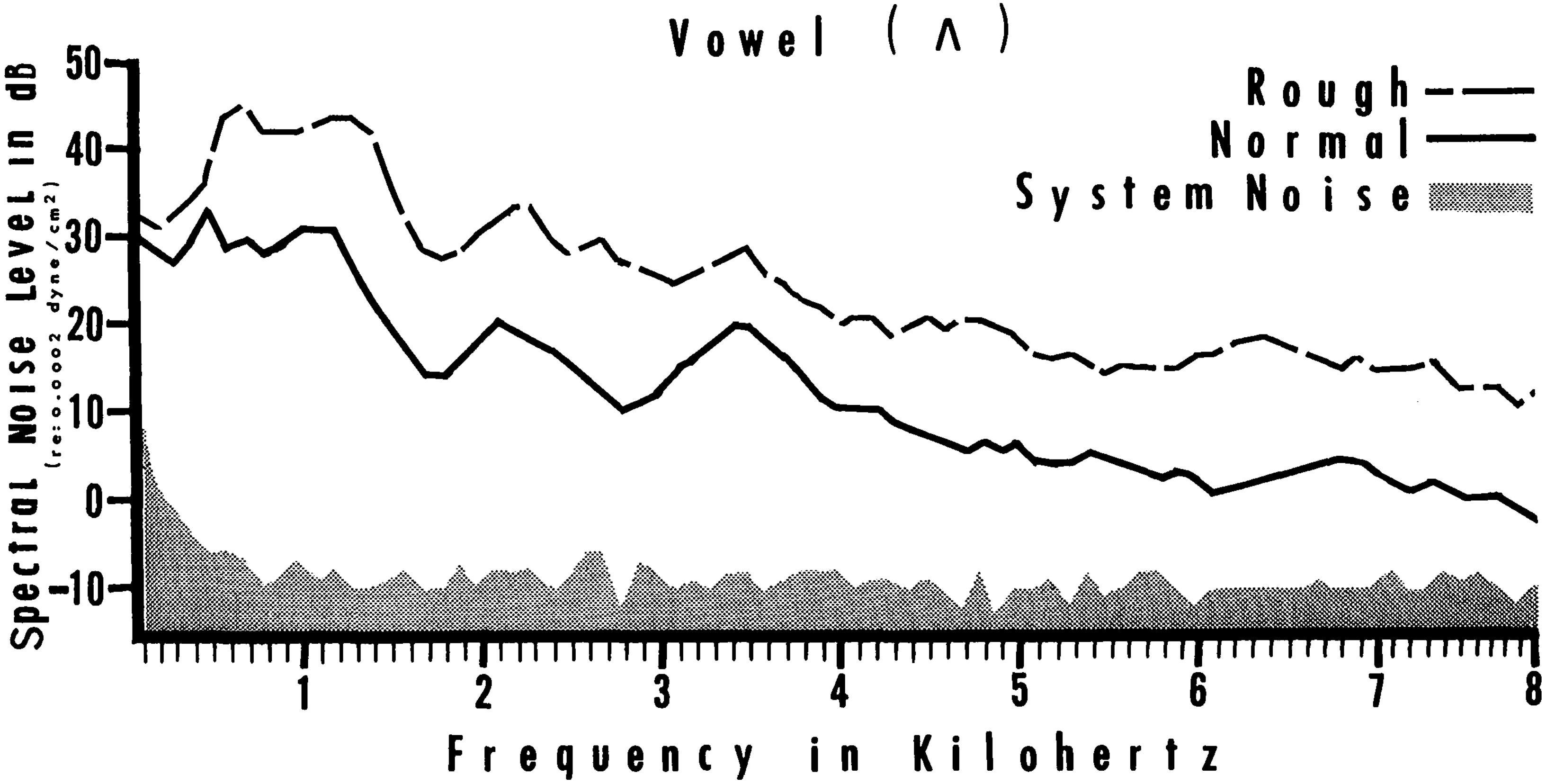

Figure 7.--Noise levels in each 100-Hz spectral section averaged over twenty male subjects for normal and for rough productions of the vowel /ʌ/.

mean spectral noise levels in individual SS's and the TSR and roughness ratings for each of the experimental vowels. In general, as the roughness of each vowel increased, its spectral noise level tended to increase. This relationship was most evident, however, in the low spectral frequencies where data point scatter tended to be less than in higher frequencies. Mean spectral noise levels and perceived vowel roughness appeared to be most closely associated when S-1 SNL means were compared to the roughness judgments for each vowel. The scatter diagrams for S-1 (100 Hz to 2600 Hz) for each test vowel are presented in Figures 8 through 12. The data points in the diagram for each vowel represent SNL's averaged over S-1 and median roughness ratings for each subject's normal and rough vowel productions.

To investigate further the degree of association between mean spectral noise levels and rated roughness for each vowel, a simple correlation statistic (Pearson r) (58) was employed. The experimental significance level was set at .05 for this correlation. Table 4 presents correlation coefficients indicating the degree of association between mean spectral noise levels for the TSR and individual SS's and the median roughness ratings for each of the test vowels. Coefficients obtained for S-4 (100 Hz to 5100 Hz) and the TSR (100 Hz to 8000 Hz) however, are not statistically independent of the coefficients for S-1, S-2, and S-3 and, thus, were not tested for significance. Each of the coefficients for S-1, S-2, and S-3 are greater than .73 and all are statistically significant at the .05 level. All the test vowels considered individually, the coefficients were largest for S-1 (100 Hz to 2600 Hz), ranging from .90 to .92. However, the

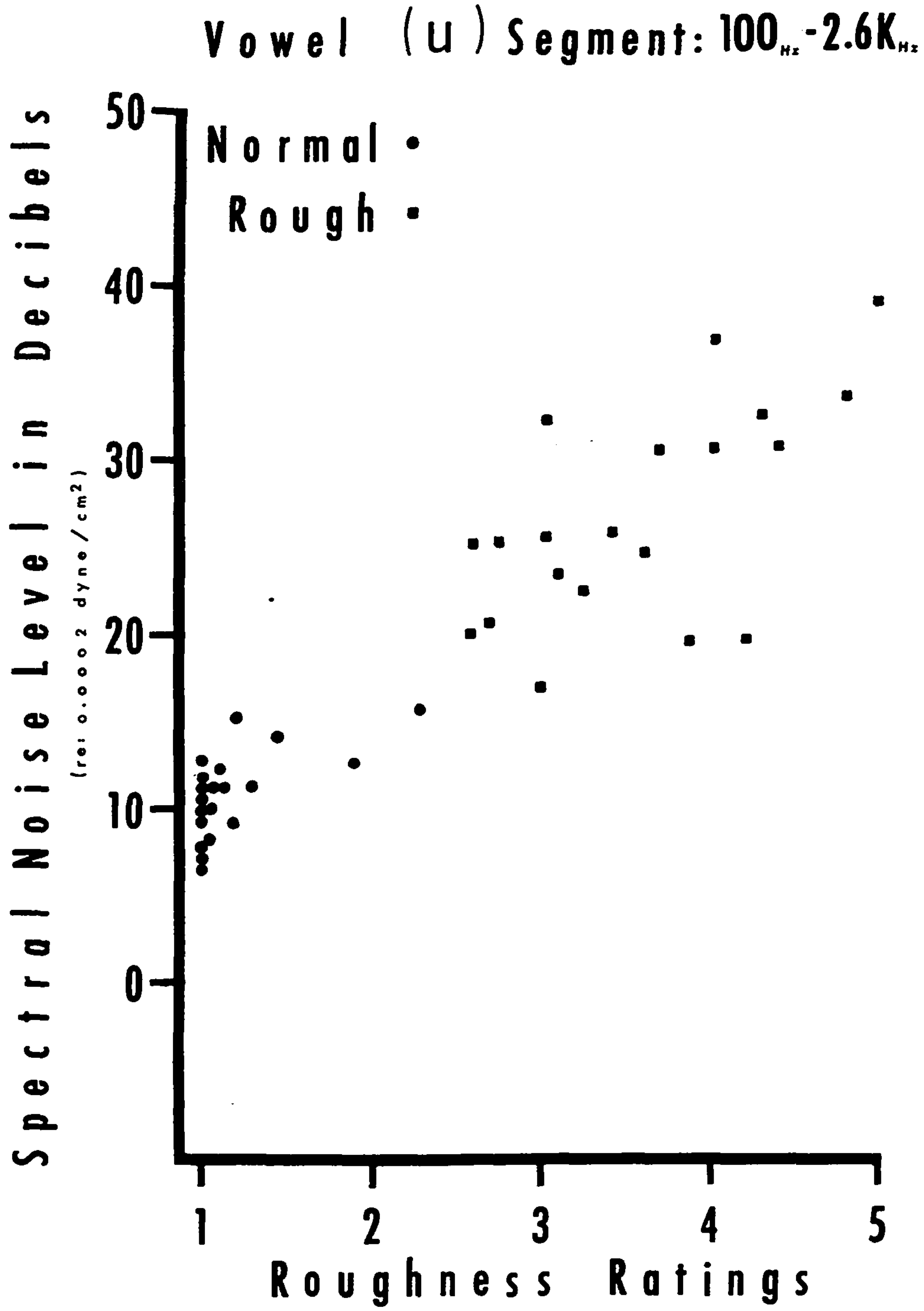

Figure 8.--Spectral segment S-1 (100 Hz to 2600 Hz) spectral noise level means and median roughness ratings over eleven judges for twenty male subjects' normal and rough productions of the vowel /u/.

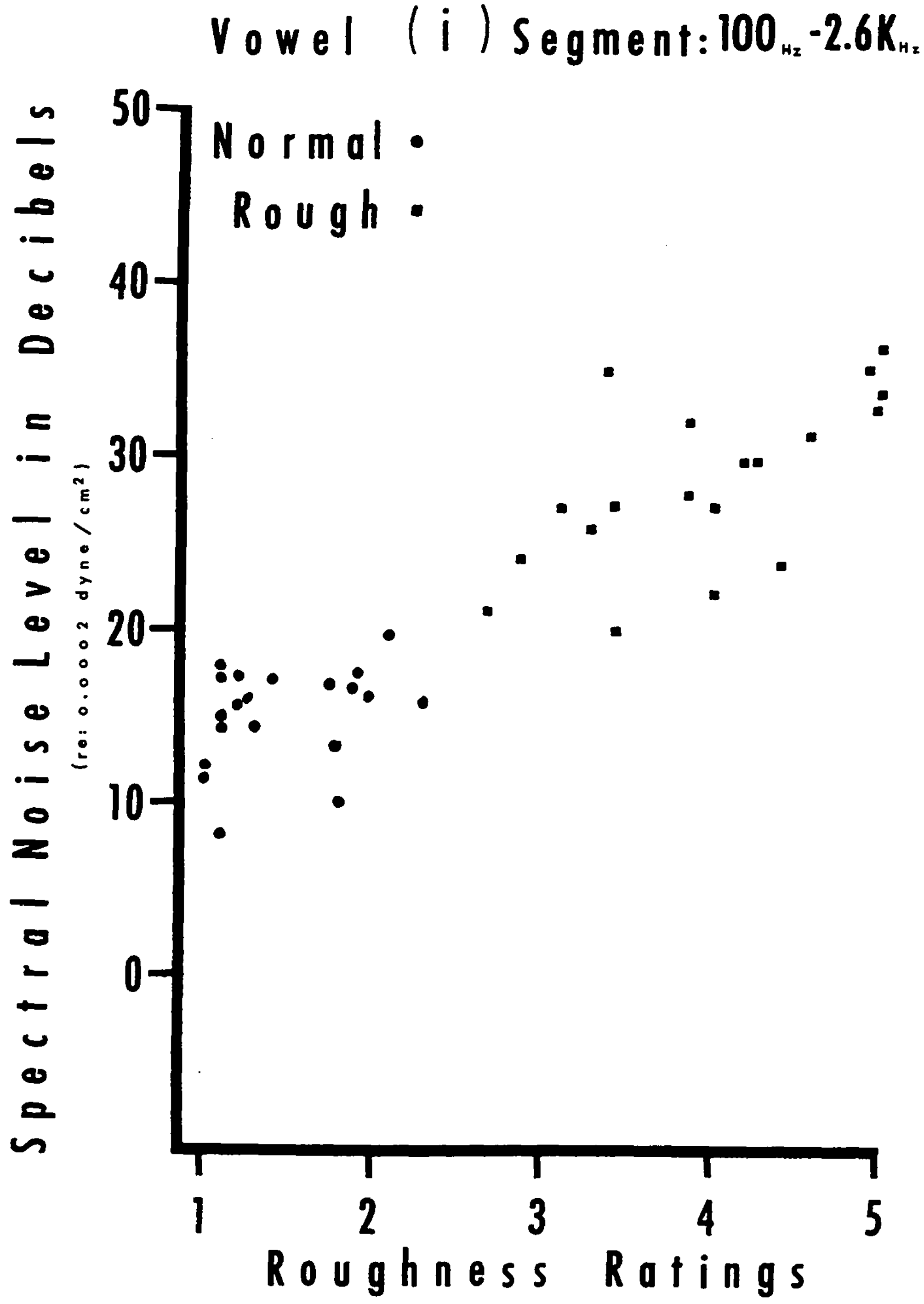

Figure 9.--Spectral segment S-1 (100 Hz to 2600 Hz) spectral noise level means and median roughness ratings over eleven judges for twenty male subjects' normal and rough productions of the vowel /i/.

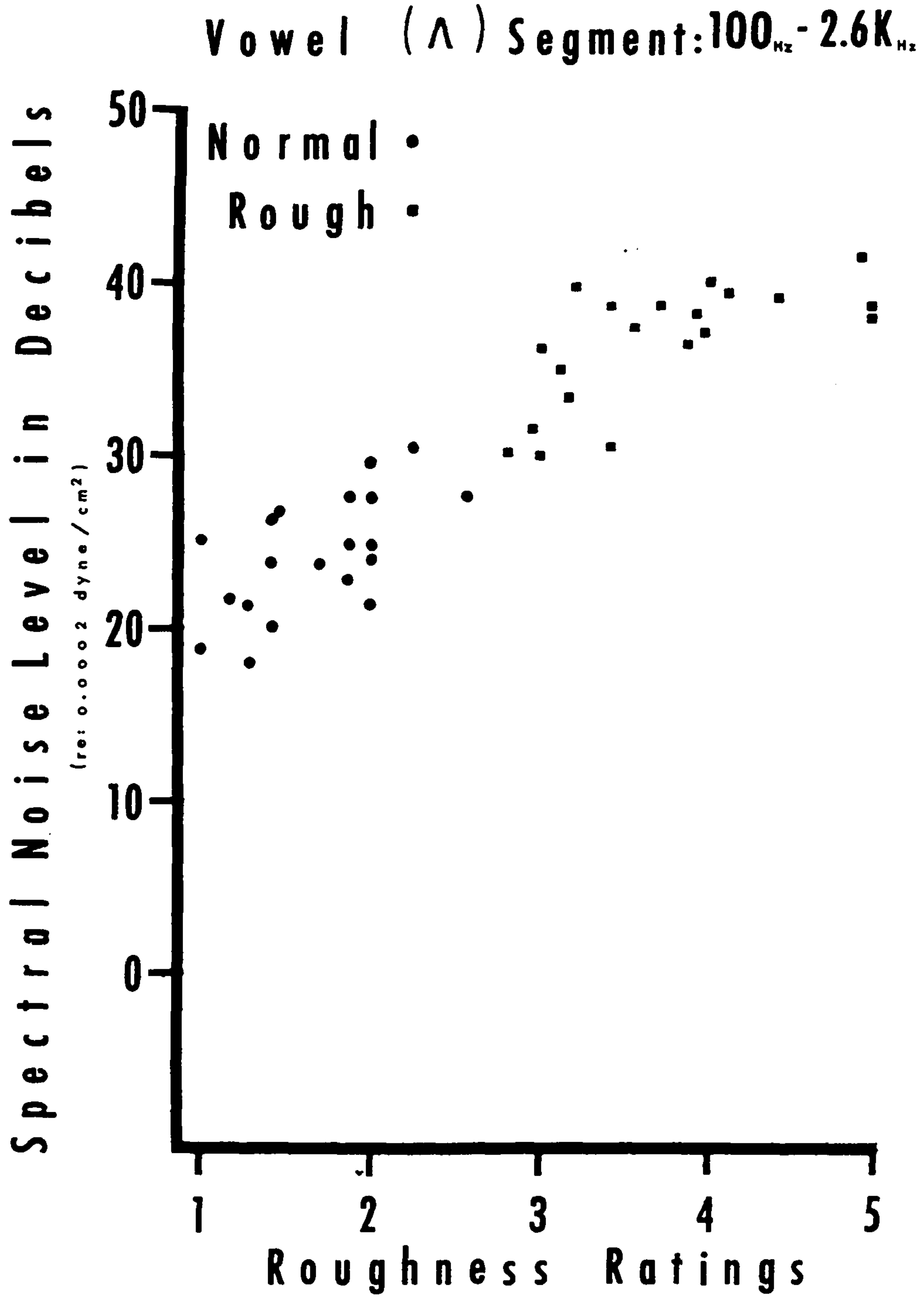

Figure 10.--Spectral segment S-1 (100 Hz to 2600 Hz) spectral noise level means and median roughness ratings over eleven judges for twenty male subjects' normal and rough productions of the vowel /ʌ/.

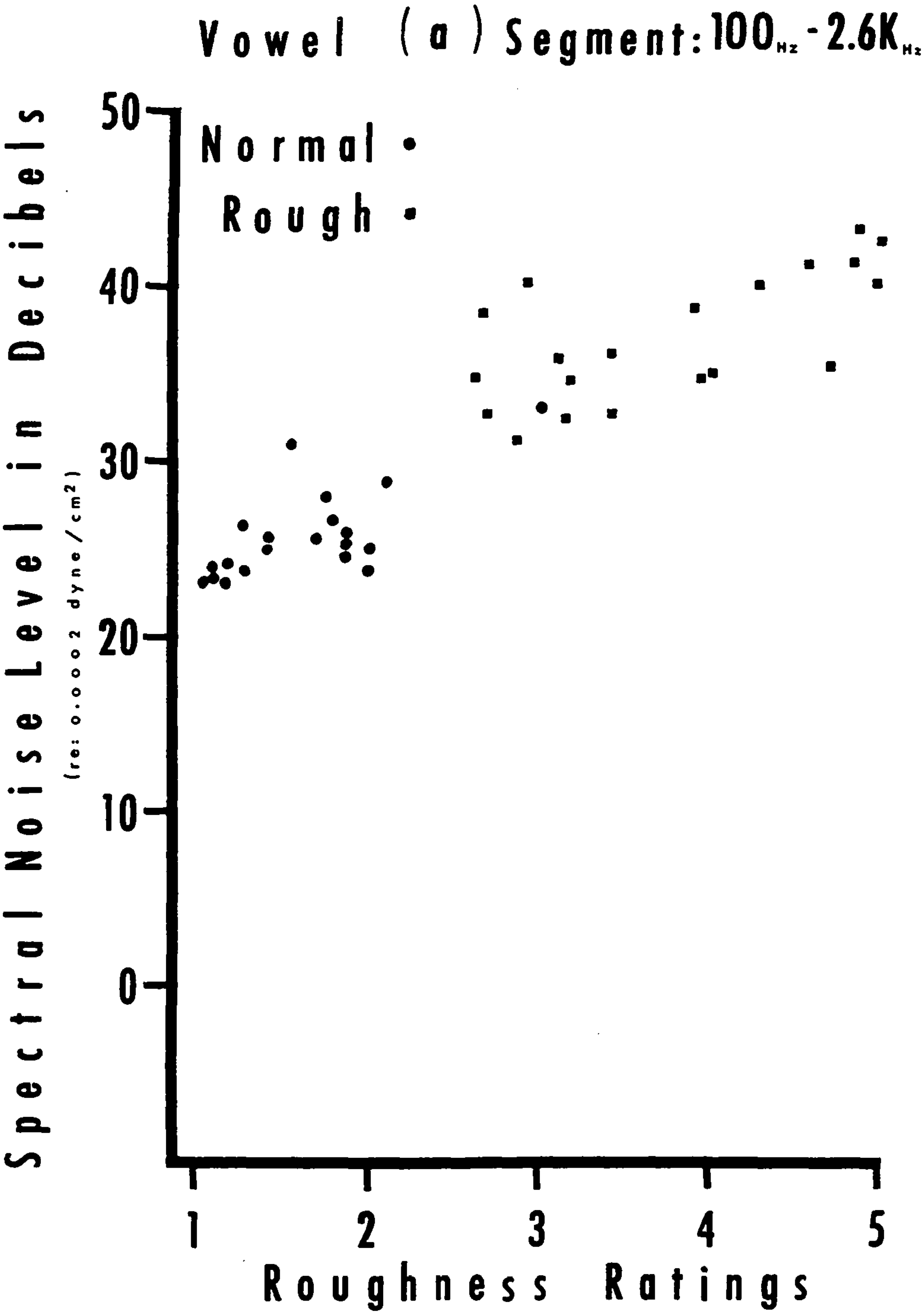

Figure 11.--Spectral segment S-1 (100 Hz to 2600 Hz) spectral noise level means and median roughness ratings over eleven judges for twenty male subjects' normal and rough productions of the vowel /a/.

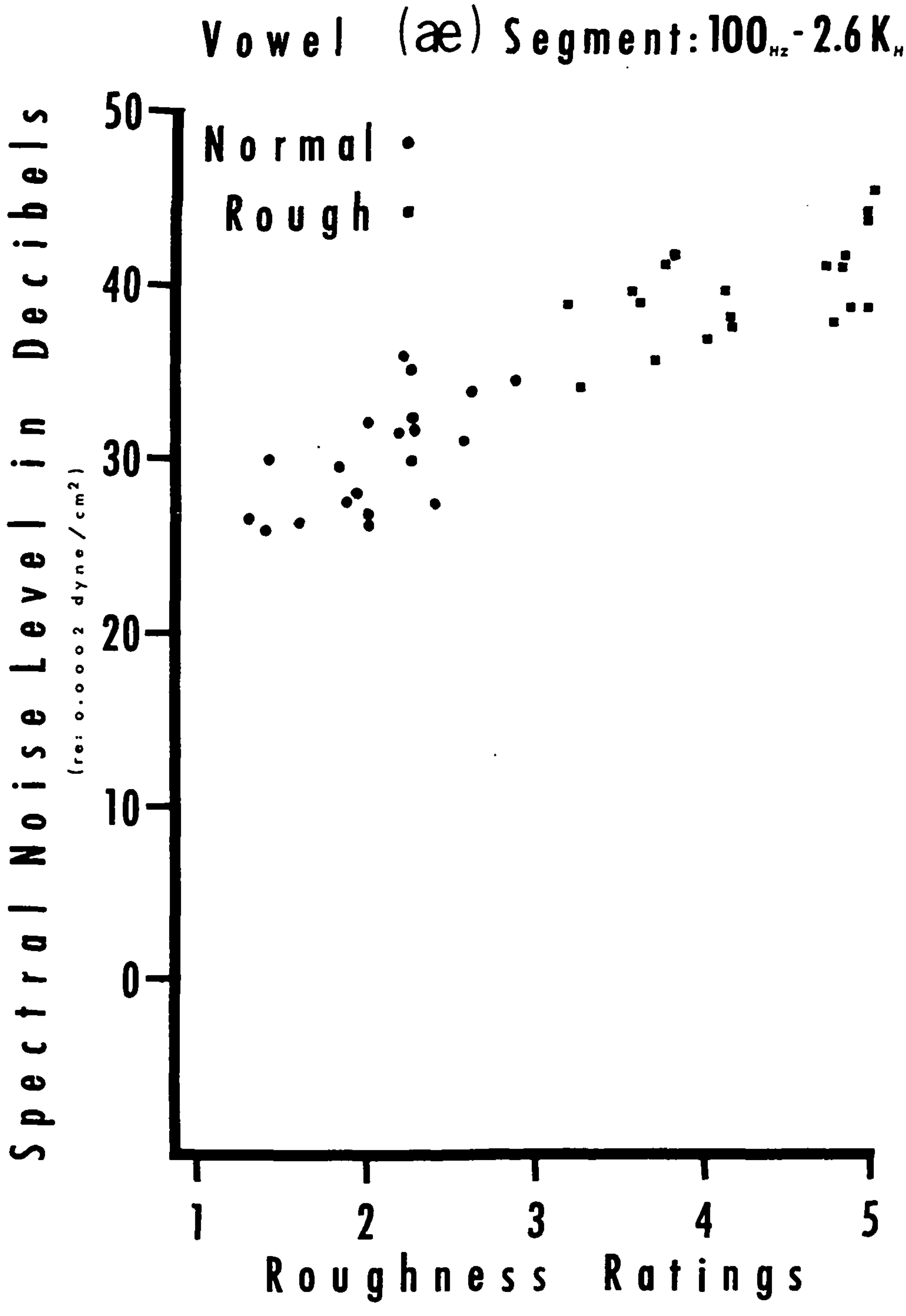

Figure 12.--Spectral segment S-1 (100 Hz to 2600 Hz) spectral noise level means and median roughness ratings over eleven judges for twenty male subjects' normal and rough productions of the vowel /æ/.

TABLE 4

CORRELATION COEFFICIENTS FOR MEAN SPECTRAL NOISE LEVELS AND ROUGHNESS SEVERITY RATINGS FOR EACH TEST VOWEL

	Correlation Coefficients*				
Vowel	S-1 (100 Hz to 2600 Hz)	S-2 (2600 Hz to 5100 Hz)	S-3 (5100 Hz to 8000 Hz)	S-4 (100 Hz to 5100 Hz)	TSR (100 Hz to 8000 Hz)
/u/	.91	.83	.79	.90	.91
/i/	.90	.76	.80	.90	.92
/ʌ/	.91	.78	.80	.89	.90
/ɑ/	.92	.80	.77	.89	.89
/æ/	.90	.74	.79	.86	.88

* All coefficients for S-1, S-2, and S-3 are significant at .05 level as determined by analyses of variance.

highest correlation for /i/, .92, was obtained for the TSR, while correlations for /u/, were of the same magnitude for both S-1 and the TSR. Coefficients for the TSR and S-4 (100 Hz to 5100 Hz) were slightly less than those for S-1. Coefficients for S-2 (2600 Hz to 5100 Hz) and S-3 (5100 Hz to 8000 Hz) were the lowest obtained. Because the correlations were uniformly high for S-1 (100 Hz to 2600 Hz), a more detailed inspection of relationships between the S-1 SNL's and vowel roughness-ratings was made.

For each vowel production, a multiple linear regression analysis (58) was performed relating the SNL in each 100-Hz section of S-1 (100 Hz to

2600 Hz) to the median roughness rating for that production. A significance level of .05 was set for this analysis. Table 5 presents the multiple correlation coefficients for each of the test vowels. The coefficients obtained in this analysis tend to be higher than those obtained when spectral segment SNL means and roughness ratings for each vowel were compared. Table 5 shows that the

TABLE 5

CORRELATION COEFFICIENTS FOR THE MULTIPLE REGRESSION BETWEEN SPECTRAL NOISE LEVELS IN EACH 100-Hz SPECTRAL SECTION FROM 100 Hz TO 2600 Hz AND ROUGHNESS RATINGS FOR EACH TEST VOWEL

Vowel	Correlation Coefficients*
/u/	.98
/i/	.98
/ʌ/	.98
/ɑ/	.97
/æ/	.98

* All coefficients significant at .05 level as determined by analyses of variance.

multiple linear regression correlation coefficients for the vowels /u/, /i/, /ʌ/, and /æ/ are each .98 and the coefficient for /ɑ/ is .97. All of these coefficients were statistically significant at the .05 level. The magnitude of these coefficients indicates a high degree of linear relationship between the S-1 (100 Hz to 2600 Hz) 100-Hz section SNL's and the median roughness ratings for all the test vowels. Because these coefficients were high and significant, the median roughness ratings for each test vowel production could be predicted from its S-1 100-Hz section SNL's. The regression equation (58) used for the prediction was:

$$Y = B_0 + B_1X_1 + B_2X_2 + \ldots . B_{25}X_{25}$$

where Y equals the roughness prediction, B_0 the Y intercept estimated by the regression analysis, B_{1-25} the regression coefficients estimated by the regression analysis, and X_{1-25} each of the S-1 100-Hz section SNL's for each vowel production.

Table 6 shows judges' median roughness ratings, roughness ratings predicted by the regression equation, and residuals (the observed rating minus the predicted rating) for twenty subjects' individual normal and rough productions of the vowel /ɑ/. Residuals for this vowel were the largest obtained and are presented to show the magnitude of the greatest residuals with this equation. It may be seen in this table that the roughness predictions for only three of forty vowel productions deviate more than .50 scale value from the median roughness ratings for those productions. Inspection of similar data for the other test vowels revealed that, for /u/ and for /æ/, roughness predictions for only two productions of each vowel differed more than .50 scale value from the median roughness ratings for those productions. For each of the vowels /i/ and /ʌ/, one roughness prediction deviated more than .50 scale value from the median roughness rating. The remaining residuals for /u/, /i/, /ʌ/, and /æ/ were relatively small.

TABLE 6

MEDIAN ROUGHNESS RATINGS FOR ELEVEN JUDGES, ROUGHNESS RATINGS PREDICTED BY THE REGRESSION EQUATION, AND RESIDUALS FOR TWENTY SUBJECTS' NORMAL AND ROUGH PRODUCTIONS OF THE VOWEL /ɑ/

	Normal			Rough		
Subject	Roughness Rating	Prediction	Residual	Roughness Rating	Prediction	Residual
1	1.58	1.76	-.18	4.95	4.69	.26
2	3.00	2.94	.06	3.90	4.10	-.20
3	1.42	1.47	-.05	4.58	4.11	.47
4	1.89	1.80	.09	2.67	3.14	-.47
5	2.00	2.12	-.12	3.19	3.46	-.27
6	1.05	1.34	-.29	4.00	3.54	.46
7	1.80	1.82	-.02	4.81	4.20	.61*
8	1.11	1.44	-.33	2.89	2.48	.41
9	1.11	1.12	-.01	3.42	3.68	-.26
10	1.19	1.15	.04	5.00	5.47	.47
11	1.86	1.88	-.02	4.71	4.23	.48
12	1.29	1.51	-.22	3.94	3.63	.31
13	2.00	2.00	.00	4.89	4.86	.03
14	1.29	0.88	.41	3.42	3.35	.07
15	1.19	1.14	.05	3.19	3.42	-.23
16	1.75	1.58	.17	2.63	2.82	-.19
17	1.86	1.67	.19	4.29	3.90	.39
18	1.71	1.85	-.14	2.71	3.35	-.64*
19	1.42	1.21	.21	3.14	2.80	.34
20	2.11	1.97	.14	2.92	3.83	-.91*

* Residual > .50 scale value.

Discussion

Vowels phonated with simulated vocal roughness uniformly evidenced higher median scale roughness ratings than their normal counterparts. Generally, however, the judges did not seem to be aware that the roughness they heard in many vowel productions was simulated by normal-speaking subjects. Following the judgment session, some of the judges reported that they thought the roughest vowels were produced by speakers with clinical voice problems. Bowler (5) has also reported that his listeners did not distinguish simulated and "real" vocal harshness. Because they appear to be perceptually similar, it may be that simulated and clinical vocal roughness are related to similar acoustic features and underlying physiological mechanisms.

The present findings indicate that normal productions of vowels tend to differ in perceived roughness. In general, normal vowels produced with higher tongue positions were evaluated as less rough than those with lower tongue positions. A similar tendency was evident for vowels produced with simulated roughness, though the severity of this roughness was not controlled. Similar findings for vowels judged for harshness have been reported by Rees (51) and by Sherman and Linke (55). It may be that the degree of roughness which listeners evaluate as being within normal limits is different for different vowels and dependent, in part at least, upon relative tongue height in the vowel's production.

The findings indicate that both normal and rough vowels evidence noise components over a wide spectral range. While it has been speculated in the past (7) that vowel noise components should be expected only in the presence of voice abnormality, such predictions appear to have been made in the absence of detailed information regarding the quasi-periodic nature of vocal fold vibration in normal phonation. The fact that noise components in vowels have seldom been reported in studies of rough or normal phonation may be attributed, in part, to the design of instruments used for acoustic analysis. Early instruments performing a Fourier series analysis of complex waves apparently did not reveal vowel noise components. Generally, more recent instruments have lacked the sensitivity and narrow frequency selectivity which appears essential to a clear spectrographic presentation of vowel noise components.

Previous investigations (74, 75) have indicated that the presence of jitter and shimmer in synthesized complex waves is associated with the perception of roughness in the signal. Similarly, rapid, random variations in fundamental vocal frequency are reported to be related to the perception of vowel roughness (38, 42). A possible relationship between vowel spectral noise levels and variations in the wave periodicity for both normal and rough vowel waves is suggested. Moreover, because variations in the periodicity of vowel waves are demonstrably related to variations in glottic valving (37, 38, 42), a possible relationship between vowel spectral noise levels and perturbations in glottic valving is also suggested. It may also be that the elevation of vowel SNL's observed in this study is attributable largely to glottic valving perturbations. Glottic perturbations may be reflected in expiratory air stream disturbances which generate noise. It has been suggested previously by others (29, 76) that imperfect modulation of the expiratory air stream at the glottis during hoarse phonation may produce air flow turbulences causing noise components in vowel spectra.

The finding, in this study, that rough productions of each vowel are characterized by an elevation of mean SNL's over those for normal productions is generally consistent with the findings of other recent investigations (44, 76, 77). The present spectra of simulated rough vowels appear similar, moreover, to frequency-by-amplitude spectra of hoarse vowels reported by Nessel (44). Nessel observed, for hoarse vowels, an elevation of noise components both in the formant ranges and in higher frequencies, and a diminution of harmonic energy in the formant ranges. Similar findings based on Sonographic analyses of hoarse vowels have been reported by Yanagihara (76, 77). It appears, therefore, that the spectral features associated with simulated vocal roughness may be similar to those associated with clinical hoarseness.

The observation that harmonics tend to diminish in amplitude with an increase in spectral noise in vowels is consistent with the findings of other investigations (44, 77). This observation suggests a possible trading relationship between vowel SNL's and harmonic levels.

The findings that the low-frequency spectral segment S-1 (100 Hz to 2600 Hz) evidenced the least normal vowel SNL variability is of interest. Vowel harmonic and noise energy is found predominantly in the lower spectral frequencies. The location of spectral energy prominances, or vowel formants, in the lower spectral frequencies is apparently important in the perception of vowel identity (48). It may be, therefore, that the subjects tended to control acoustic energy distribution in the low spectral frequencies to preserve vowel identity, thus limiting SNL variability in this region. Among the normal vowels, the vowel /u/ evidenced the least SNL variability. The variability of SNL measures for /u/ was least in S-1 (100 Hz to 2600 Hz). Possibly, therefore, normal /u/ SNL measures in S-1 would provide a standard to which similar data for rough voiced speakers might be compared clinically.

The relative magnitudes of SNL's and roughness ratings for each test vowel phonated normally and with simulated vocal roughness in this study and the harshness ratings for vowels studied by Rees (51) may be compared. Using harsh-voiced speakers, Rees found that vowels considered in the present study were ordered with respect to increasing severity of harshness: /u/, /i/, /ʌ/, /æ/, and /ɑ/. With respect to increasing roughness, normal vowels in this study were ordered: /u/, /i/, /ʌ/, /ɑ/, and /æ/. Further, the rough and normal test vowels for this study were ordered with respect to increasing mean SNL's for S-1 (100 Hz to 2600 Hz): /u/, /i/, /ʌ/, /ɑ/, and /æ/. The similarity in the order of vowels rated for roughness and harshness and the S-1 SNL's for those vowels suggests that listener judgments of both harshness and roughness in vowels may be related to low-frequency spectral noise levels.

Rees (51) and others have observed that the order of vowels with respect to increasing harshness seems to be inversely related to tongue height in vowel production as indicated in the conventional vowel triangle (33). Similarly, in this study, there appears to be a possible relationship between vowel tongue height and vowel SNL's. In general, vowels produced with relatively low tongue positions evidenced greater spectral noise than those with high tongue positions, within normal and within simulated rough phonatory conditions. It may be too that the vowel noise levels are related to the overall configuration of the supraglottic cavity since tongue position greatly influences this configuration (35). Moreover, it is thought (11, 35) that the relative maximum amplitudes of vowel harmonic components at various spectral frequencies are largely determined by frequency-selective acoustic damping in the vocal tract and that this damping is related to supraglottic cavity configuration. For normal vowel productions in particular, the present vowel noise levels in the low-frequency spectral range tend to be relatively high when harmonic amplitudes are relatively high and vice versa. Thus, it appears that the relative amplitudes of vowel noise components at various spectral frequencies may also, to some extent, be affected by the acoustic damping of the vocal tract. Possibly, this helps to explain why SNL means for discrete spectral frequency ranges associated with different vowels produced at the same intensity tend to be different.

A primary purpose of this study was to investigate possible relationships between vowel spectral noise levels and listener judgments of vowel roughness. Previous studies (76, 77) indicate that vowel spectral noise elevation and decreases in harmonic energy accompany perceived increases in hoarseness. The present findings appear consistent with these previous reports. In the present study, vowel spectral noise levels were found to increase as vowel roughness increased. Yanagihara (77) has suggested that four types of hoarseness may be differentiated on the basis of the level and frequency location of noise in Sonographic spectra. To the extent that vocal roughness may be assumed to vary in its severity along a continuum, however, it would seem desirable to study its acoustic correlates in ways which provide measurements of relevant acoustic features along a continuum. In this regard, the measures of noise in narrow-band frequency-by-amplitude vowel spectra made in this study appear to offer advantages over measures of vowel noise in Sonographic spectra. Yanagihara obtained, for vowels produced by hoarse speakers, a .65 correlation between the four types of spectrograms and judges' perception of hoarseness severity. Correlations obtained in this study between vowel SNL's in 100-Hz sections of S-1 (100 Hz to 2600 Hz) and vowel median roughness ratings were higher, being .98 for four, and .97 for one, of the five test vowels.

Previous investigators (30, 76, 77) have generally considered a wide spectral frequency range, e.g., 80 Hz to 8000 Hz, in relating Sonographic features to judgments of hoarseness severity. High correlations between perceived vowel roughness and spectral noise levels were obtained in this study, however, when only the 100 Hz to 2600 Hz spectral range was considered. This suggests that acoustic information which cues the perception of vocal roughness may be redundant in the vowel spectrum.

Multiple linear correlation coefficients between 100-Hz section SNL's in S-1 (100 Hz to 2600 Hz) and roughness ratings were high for all test vowels. The relationship between the S-1 100-Hz section SNL's and listener judgments of roughness was, therefore, nearly linear for the range of roughness studied for all test vowels. A regression equation was used to predict roughness ratings for individual productions of each test vowel from S-1 100-Hz section SNL measures for each production with small residuals. If further studies reveal that the SNL and roughness-rating relationship is sufficiently linear outside of the roughness scale employed in this study, the regression equation may be employed to predict listener judgments for extremely low or high vowel SNL's.

It has been reported previously (37, 38) that measures of variations in the periodicity of vowel waves, i.e. of pitch perturbation, might provide an objective clinical index of voice disturbance in hoarse subjects. The present data suggests that measures of vowel spectral noise may be similarly useful. Moreover, the relative easewith which noise levels were displayed and measured in the present spectra suggests that these measures may be simpler, and thus more readily applied clinically, than pitch perturbation measures. Lieberman (38) observed that an accurate determination of pitch perturbation was not possible when hoarseness was severe and the acoustic wave was "filled in." If isolated vowels are considered, spectral noise measures similar to those made in this study would appear feasible even for severely hoarse subjects.

To the extent that simulated rough vowels are acoustically similar to clinically rough vowels, the present equation might be used to predict roughness ratings for clinically rough vowels. For example, spectral noise measures may be made from the spectrum of a vowel produced by a person presenting clinical vocal roughness. It might then be predicted how the judges for this study would rate the voice sample. Initially, predicting vowel roughness from vowel SNL measures may help to demonstrate the clinical usefulness of SNL measures.
To illustrate, the vowel /u/, phonated by a fifty-five year old white male with

medically diagnosed hyperkinetic dysphonia, was recorded and analyzed in the same manner as the test vowels in this study. This vowel recording was placed in random order with recordings of forty-nine other clinically rough vowels, fifty simulated rough vowels, and fifty normal vowels. These were played to the eleven judges under controlled listening conditions. The judges rated the vowel productions for roughness on a five point equal-appearing-intervals scale. The median scale roughness rating obtained for the dysphonic subject's /u/ production was 2.00. The roughness rating predicted by the equation from the S-1 100-Hz section SNL measures for this vowel was 1.92, with a .08 scale value residual. As relationships between spectral noise and predicted roughness ratings for vowels are better defined, clinicians may choose to utilize the SNL measures alone as an index of vocal roughness, without reference to the judgment predictions. Thus, with further study, spectral noise measures may provide a useful index of voice disturbance.

CHAPTER V

SUMMARY

Findings

The purpose of this study was to investigate spectral noise levels (SNL's) for normal and for simulated rough vowels and possible relationships between the SNL's and perceived vowel roughness. The study was attempted because of the need for measurement data regarding the acoustic features of vocal roughness. Though previous studies (44, 77) have suggested that the elevation of vowel spectral noise components may be associated with the perception of vowel roughness, quantitative data pertaining to the level of these noise components and to the specific relationships between them and perceived vowel roughness are lacking.

Twenty normal-speaking adult males serving as subjects for this study sustained seven second productions of each of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/. The vowels were phonated first normally and then with simulated vocal roughness within ± 1 dB of a monitored intensity. The intensity selected, 75 dB SPL at a mouth-to-microphone distance of six inches, was comfortable for both normal and rough phonations. Each vowel production was tape-recorded and all productions were presented in random order to eleven judges for rating. Each judge rated the vowels for roughness on a five point equal-appearing-intervals scale and the median of the eleven judges' ratings of each test vowel was computed. The recording of each vowel production was also analyzed to produce a graphic 3-Hz bandwidth frequency-by-intensity acoustic spectrum with a range from 0 Hz to 8000 Hz. The analysis was made from a two second portion of each vowel production evidencing a uniform intensity ($\pm$ 1 dB). The low peak of energy recorded in each 100 Hz section of each vowel spectrum was measured in dB SPL as an index of the vowel's spectral noise levels. The median roughness ratings obtained for each test vowel indicated that each simulated rough production was judged more rough than the normal production of the same vowel. The average median roughness ratings for the normal vowel productions indicated that the high vowels /u/ and /i/ tended to be rated less rough than the other test vowels, with /u/ rated less rough than /i/. The mid vowel /ʌ/ was rated more rough than /i/ and slightly less rough than /ɑ/. The vowels /ɑ/ and /æ/ were rated most rough, with /ɑ/ rated less rough than /æ/. The rough vowel productions were similarly ordered with respect to average median judgments.

The findings indicated that spectral noise was associated with normal as well as rough vowel productions, but rough productions of each vowel tended to evidence spectral noise levels greater than those for the normal productions. For each test vowel, harmonic amplitudes in the spectra of rough productions tended to be somewhat diminished in amplitude with respect to those in the normal productions. For each vowel, the difference between SNL's for normal and rough productions were generally of similar magnitude throughout the total spectral range (TSR) from 100 Hz to 8000 Hz. For both normal and rough productions of each vowel, spectral noise was most prominant in lower spectral frequencies and decreased toward the higher frequencies. Considering both normal and rough vowel productions, an increase in mean SNL's over the TSR appeared to be associated with a decrease in tongue height in test vowel production. The high vowel /u/, in both phonatory conditions, generally evidenced the lowest SNL's while the low vowel /æ/ evidenced the highest. The mid vowel /ʌ/ and the low vowel /ɑ/ were located toward the center on this continuum, with /ʌ/ generally evidencing smaller SNL's than /ɑ/. SNL's for the vowel /i/ appeared the most variable across the TSR. Generally, relatively high SNL's were recorded for /i/ in the very low and high spectral frequency ranges, while relatively low SNL's were recorded for this vowel in the intermediate spectral range. Considering the spectral segment S-1 (100 Hz to 2600 Hz), both normal and rough vowel productions, considered separately, could be ranked with respect to increasing mean SNL's: /u/, /i/, /ʌ/, /ɑ/, and /æ/. With the exception of those for the vowels /ʌ/ and /ɑ/, all normal test vowel S-1 (100 Hz to 2600 Hz) SNL means were significantly (.05) different.

Normal vowel SNL variability was of interest since it seemed possible that this information might be relevant to the use of SNL measures in clinical voice evaluations. It was found that the greatest vowel SNL variability among subjects was associated with vowel SNL means over spectral segment S-2 (2600 Hz to 5100 Hz) and segment S-3 (5100 Hz to 8000 Hz). The vowel SNL means for S-1 (100 Hz to 2600 Hz) evidenced least variability. Among the vowels, /u/ evidenced the least variability. The rough vowels' SNL variability tended to be greater than that for normal vowels for all SS's and the TSR.

A primary purpose of the investigation was to explore possible relationships between normal and simulated rough vowel SNL's and listener judgments of vowel roughness. For each test vowel, correlation coefficients indicating the relationship between the TSR and SS mean SNL's and the median ratings of vowel roughness were equal to or greater than .74. The highest correlations between vowel noise measurement means and roughness judgments were obtained when S-1 (100 Hz to 2600 Hz) SNL means were considered. Data for this spectral segment was analyzed further. For each vowel, multiple correlation coefficients relating SNL's in each 100-Hz section of S-1 (100 Hz to 2600 Hz) and vowel roughness ratings were computed. The correlation coefficients obtained between these variables for each vowel were: /u/, .98; /i/, .98; /ʌ/, .97; /ɑ/, .98; and /æ/, .98. These coefficients indicated a strong and nearly linear relationship between S-1 (100 Hz to 2600 Hz) 100-Hz section SNL's and the median roughness rating for each vowel production. A regression equation was employed, therefore, to predict the roughness ratings for each vowel production from its S-1 100-Hz section SNL's. It was found that the roughness ratings for each vowel production could be predicted from its S-1 100 Hz section SNL's with only small residuals. Roughness predictions for /ɑ/, the vowel evidencing the greatest prediction residuals, were different from the median of judges' roughness ratings as much as .50 scale value for only three of forty productions.

The principal findings of this investigation may be summarized as follows:

1. Normal-speaking adult male subjects appeared able to simulate a vowel quality which was judged to be rough by trained listeners.

2. Normal test vowel productions tended to be ranked with respect to increasing roughness: /u/, /i/, /ʌ/, /ɑ/, /æ/.

3. Measurable spectral noise, above system noise levels, appeared to be associated with normal as well as with simulated rough vowel productions.

4. For the individual vowels studied, spectral noise levels tended to be higher for rough productions than for normal productions when both were phonated at the same vocal intensity.

5. Harmonic amplitudes for rough productions of each vowel tended to be somewhat diminished with respect to those for normal productions.

6. For both normal and simulated rough vowel productions, spectral noise appeared to be most prominent in lower spectral frequencies and decreased toward higher frequencies.

7. With the exception of means for the vowels /ʌ/ and /ɑ/, individual normal test vowel spectral noise levels averaged over S-1 (100 Hz to 2600 Hz) were significantly (.05) different.

8. An increase in mean spectral noise levels for both normal and rough test vowels appeared to be associated with decreasing tongue height in

vowel production.

9. For both normal and rough productions, inter-subject spectral noise level variability for each test vowel appeared less for levels averaged over S-1 (100 Hz to 2600 Hz) than for levels averaged over the other spectral segments studied.

10. For both normal and rough productions, inter-subject spectral noise level variability tended to be less for the vowel /u/ than for the other vowels studied, regardless of the spectral segment considered.

11. For each test vowel, spectral noise levels averaged over the total spectral range and separately over each of the spectral segments correlated highly with the median roughness ratings for that vowel.

12. Individual vowel spectral noise levels averaged over spectral segment S-1 (100 Hz to 2600 Hz) tended to correlate more highly with the median roughness rating for each vowel than spectral noise levels averaged over the other spectral segments considered.

13. High and significant (.05) multiple linear correlation coefficients were obtained between each test vowel's S-1 (100 Hz to 2600 Hz) 100-Hz section spectral noise levels and judges' ratings of the vowel's roughness.

14. A regression equation predicted, with small residuals, each vowel production's median roughness rating from its S-1 (100 Hz to 2600 Hz) 100-Hz section spectral noise levels.

Limitations

The experimental design of the present investigation might be altered profitably in future studies of vowel SNL's. First, because one of the purposes of the present investigation was to explore SNL's in normal and simulated rough vowels produced at the same intensity, subject vocal intensity during vowel production was limited to 75 dB ± 1 dB. Measures of SNL's in vowels produced at other intensity levels might also be of interest particularly with respect to possible clinical application of the data. In some instances, persons presenting clinical vocal roughness may not be able to achieve a constant 75 dB ± 1 dB intensity at a mouth-to-microphone distance of six inches.

Second, previous investigations (73, 74) have suggested that the perceived roughness of an acoustic signal is influenced by the fundamental frequency of the stimulus. Only male subjects were utilized in this investigation. It might be useful also to study vowel SNL and roughness-rating relationships for females presenting higher vocal pitches. Data for female subjects might be relevant to the clinical use of vowel SNL measures.

Third, the spectral analysis procedures employed in evaluating vowels in this study may be useful in the experimental evaluation of clinical vocal roughness. Though similarities between the present spectra of simulated rough vowels and the spectra of vowels produced by persons presenting clinical vocal roughness were observed, it would seem desirable to replicate this study using subjects with clinically rough voices.

Alterations in the instrumentation used in this study might also be advisable in future studies of vowel SNL's. First, because the level of high-frequency acoustic energy in the spectra of normal vowel productions in the present study occasionally did not exceed the system noise, the use of instruments with less internal noise appears desirable.

Second, the wave analyzer level recorder assembly employed in this study provided two charting speeds, one ten times faster than the other. When the fast speed was used, each vowel spectrum required approximately thirteen feet of chart paper, and this speed did not facilitate the vowel SNL measurements sufficiently to justify its use. Use of the slower speed in this study resulted in some level recorder overwrite. The true low peak level of noise in each 100-Hz spectral section may, therefore, have been obscured in some instances. Noise level measures from vowel spectra produced with the slower speed, however, correlated highly with listener judgments of vowel roughness. Charting speeds between those possible in this study seem desirable.

Third, it might be desirable to employ instruments permitting real-time spectral analysis in vowel SNL studies. These instruments could enhance spectral resolution, decrease analysis time, and eliminate the need for tape recording. Analyzers providing these advantages may be the instruments of choice in future studies of vowel spectral noise level measurements.

BIBLIOGRAPHY

1. Bell Telephone Laboratories. High-speed motion pictures of the human vocal cords. New York, 1939.

2. Black, J. W. Acoustic spectra of the successive waves of the vowels. Jour. Acous. Soc. Amer., 1939, 10:204.

3. Black, J. W., Lierlie, D. M. and Tiffin, J. Harmonic analysis of consecutive waves in the vowel "ah" in normal voices before and after surgical removal of the nasal polyps. Jour. Acous. Soc. Amer., 1936, 7:235.

4. Bogert, B. P. and Peterson, G. E. The acoustics of speech. In Travis, L. (Ed.) Handbook of Speech Pathology. New York: Appleton-Century-Crofts, 1957.

5. Bowler, N. Fundamental frequency analysis of harsh voice quality. Sp. Monogr., 1964, 31:128.

6. Brubaker, R. S. and Dolpheide, W. R. Consonant and vowel influence upon judged voice quality of syllables. Jour. Acous. Soc. Amer., 1955, 27:1000.

7. Carhart, R. The spectra of model larynx tones. Sp. Monogr., 1941, 8:76.

8. Coleman, F. Some acoustic correlates of hoarseness. Unpublished Master's Thesis, Vanderbilt University, 1960.

9. Coleman, F. and Wendahl, R. Vocal roughness and stimulus duration. Sp. Monogr., 1967, 34:241.

10. Cooper, F., Peterson, G. and Fahringer, G. Some sources of characteristic vocoder quality. Jour. Acous. Soc. Amer., 1957, 29:183.

11. Fant, G. M. Descriptive analysis of the acoustic aspects of speech. Logos, 1962, 5:3.

12. Fant, G. M. On the predictability of formant levels and spectrum envelopes from formant frequencies. In Lehiste, I. (Ed.) Readings in Acoustic Phonetics. Cambridge, Massachusetts: M.I.T. Press, 1967.

13. Farnsworth, D. W. High-speed motion pictures of the human vocal cords. Bell Tel. Lab. Rec., 1940, 18:203.

14. Fletcher, H. Speech and Hearing in Communication. Princeton, N. J.: D. Van Nostrand Co., 1953.

15. Fourier. Theorie Analytique de la Chaleur, Ch. III. Paris, 1822. Cited by Scripture.

16. Green, M.C.L. The Voice and Its Disorders. (2nd Ed.) Philadelphia: J. B. Lippincott Co., 1964.

17. Hall, H. Recording analyzer for the audible frequency range. Jour. Acous. Soc. Amer., 1935, 7:102.

18. ______ Sound analysis. Jour. Acous. Soc. Amer., 1935, 8:257.

19. Halle, M., Hughes, G. W. and Radley, J. P. Acoustic properties of stop consonants. Jour. Acous. Soc. Amer., 1957, 29:107.

20. Hollien, H. A study of some laryngeal correlates of vocal pitch. Jour. Sp. Hear. Res., 1960, 3:52.

21. ______ Selected vocal characteristics and physical size measurements of pre-adolescent males. N. I. H. Progress Report, Grant NB-02813, 1962.

22. ______ Vocal fold thickness and fundamental frequency of phonation. Jour. Sp. Hear. Res., 1962, 5:237.

23. ______ Vocal pitch variation related to changes in vocal fold length. Jour. Sp. Hear. Res., 1960, 3:150.

24. Hollien, H. and Curtis, J. F. A laminagraphic study of vocal pitch. Jour. Sp. Hear. Res., 1960, 3:361.

25. ______ Elevation and tilting of vocal folds as a function of vocal pitch. Folia Phoniat., 1962, 14:23.

26. Hollien, H. and Moore, G. P. Measurements of the vocal folds during changes in pitch. Jour. Sp. Hear. Res., 1960, 3:157.

27. Isshiki, N. Regulatory mechanism of the pitch and volume of voice. Otorhinolaryng. Clinic Kyoto, 1959, 52:461.

28. ________ Regulatory mechanism of voice intensity variation. Jour. Sp. Hear. Res., 1964, 7:17.

29. Isshiki, N. and von Leden, H. Hoarseness: aerodynamic studies. Arch. Otolar., 1964, 80:206.

30. Isshiki, N., Yanagihara, N., and Morimoto, M. Approach to the objective diagnosis of hoarseness. Folia Phoniat., 1966, 18:393.

31. Joos, M. Acoustic phonetics. Lang. Supp., 1948, 24:1.

32. Judson, L.S.V. and Weaver, A.T. Voice Science, New York: Appleton-Century-Crofts, 1942.

33. Kenyon, J. S. American Pronunciation. Ann Arbor, Michigan: George Wahr, 1940.

34. Koening, W., Dunn, H. K. and Lacy, L. Y. Sound spectrograph. Jour. Acous. Soc. Amer., 1946, 18:19.

35. Ladefoged, P. Elements of Acoustic Phonetics. Chicago: University of Chicago Press, 1966.

36. Laguaite, J. and Waldrop, W. Acoustic analysis of fundamental frequency of voices before and after therapy. Folia Phoniat., 1964, 16:183.

37. Lieberman, P. Intonation, Perception, and Language. Cambridge, Massachusetts: M.I.T. Press, 1967.

38. ________ Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. Jour. Acous. Soc. Amer., 1963, 35:344.

39. Luchsinger, R. and Arnold, G. Voice-Speech-Language Clinical Communicology: Its Physiology and Pathology. Belmont, California: Wadsworth Publishing Co., 1965.

40. McNemar, Q. Psychological Statistics. New York: John Wiley and Sons, Inc., 1962.

41. Meyer, E. Method for very rapid analysis of sounds: sound grating spectroscopy, Jour. Acous. Soc. Amer., 1935, 7:88.

42. Moore, P. and Thompson, C. L. Comments on the physiology of hoarseness. Arch. Otolaryng., 1965, 81:97.

43. Moore, P. and von Leden, H. Dynamic variation of the vibratory pattern in the normal larynx. Folia Phoniat., 1958, 10:205.

44. Nessel, E. Über das tonfrequenzspektrum der pathologisch veränderten stimme. Acta-Otol., 1960, Supp. 157:3.

45. Palmer, J. M. Hoarseness in laryngeal pathology: review of literature. Laryng., 1956, 66:500.

46. Peacher, G. Contact ulcer of the larynx, Part III. Etiological factors. Jour. Sp. Dis., 1947, 12:177.

47. Peterson, G. E. Parameters of vowel quality. Jour. Sp. Hear. Res., 1961, 4:10.

48. Peterson, G. E. and Barney, H. L. Control methods used in study of the vowels. Jour. Acous. Soc. Amer., 1952, 24:175.

49. Pressman, J. J. Physiology of the vocal cords in phonation and respiration. Arch. of Otolaryng., 1942, 35:355.

50. Rees, M. Harshness and glottal attack. Jour. Sp. Hear. Res., 1958, 1:344.

51. ________ Some variables affecting perceived harshness. Jour. Sp. Hear. Res., 1958, 1:155.

52. Rubin, H. and Lehrhoff, I. Pathogenesis and treatment of vocal nodules. Jour. Sp. Hear. Dis., 1962, 27:150.

53. Scripture, E. W. Researches in Experimental Phonetics: Study of Speech Curves, Washington, D. C.: The Carnegie Foundation, 1906.

54. Sherman, D. The merits of backward playing of connected speech in the scaling of voice quality disorders. Jour. Sp. Hear. Dis., 1954, 19:312.

55. Sherman, D. and Linke, E. The influence of certain vowel types on degree of harsh voice quality. Jour. Sp. Hear. Dis., 1952, 17:401.

56. Sivian, L. J. Speech power and its measurement. Bell System Technical Journal, 1929.

57. Sivian, L. J., Dunn, H. K. and White, S. D. Absolute amplitudes and spectra of certain musical instruments and orchestras. Jour. Acous. Soc. Amer., 1931, 2:330.

58. Steel, R. G. and Torrie, J. H. Principles and Procedures of Statistics. New York: McGraw-Hill Book Company, 1960.

59. Steinburg, J. C. Analysis of successive periods of vowel sounds. Jour. Acous. Soc. Amer., 1934, 6:18.

60. ______________ Application of sound measuring instruments to the study of phonetic problems. Jour. Acous. Soc. Amer., 1934, 6:16.

61. Stout, B. Harmonic structure of vowels in singing in relation to pitch and intensity. Jour. Acous. Soc. Amer., 1938, 10:137.

62. Strevens, P. Spectra of fricative noise in human speech. In Lehiste, I. (Ed.) Readings in Acoustic Phonetics. Cambridge, Massachusetts: M.I.T. Press, 1967.

63. Third Regional Workshop on the Rehabilitation Codes and Communicative Disorders. PHS. Grant #B-3676, The National Institute of Neurological Diseases and Blindness, Communicative Disorders Research Training Committee, 1967.

64. Thurman, W. The construction and acoustic analysis of recorded scales of severity for six voice quality disorders. Unpublished Ph.D. Dissertation, Purdue University, 1954.

65. Timcke, R., von Leden, H. and Moore, P. Laryngeal vibration: measurements of the glottic wave, Part I. The normal vibratory cycle. Arch. of Otolaryng., 1958, 68:1.

66. ______________ Laryngeal vibrations: measurements of the glottic wave, Part II. Physiologic variations. Arch. of Otolaryng., 1959, 69:438.

67. Van den Berg, J. W. Myoelastic aerodynamic theory of voice production. Jour. Sp. Hear. Res., 1958, 1:227.

68. ______________ Subglottic pressures and vibrations of the vocal folds. Folia Phoniat., 1957, 9:65.

69. Van Dusen, R. Laboratory study of metallic voice. Jour. Sp. Hear. Dis. 1941, 6:140.

70. Von Leden, H. The clinical significance of hoarseness and related disorders. Jour. Lancet, 1958, 78:50.

71. ________ The mechanism of phonation: search for a rational theory of voice production. Arch. Otolaryng., 1961, 74:660.

72. Von Leden, H., Moore, P. and Timcke, R. Laryngeal vibrations: measurements of the glottic wave, Part III. The pathologic larynx. Arch. Otolaryng., 1960, 71:26.

73. Wendahl, R. Laryngeal analog synthesis of harsh voice quality. Folia Phoniat., 1963, 15:24.

74. ________ Laryngeal analog synthesis of jitter and shimmer: auditory parameters of harshness. Folia Phoniat., 1966, 18:98.

75. ________ Some parameters of auditory roughness. Folia Phoniat., 1966, 18:26.

76. Yanagihara, N. Hoarseness: investigation of the physiological mechanisms. Ann. Oto. Rhino. Lar., 1967, 76:472.

77. ________ Significance of harmonic changes and noise components in hoarseness. Jour. Sp. Hear. Res., 1967, 10:531.

78. Zemlin, W. R. Speech and Hearing Science: Anatomy and Physiology. Champaign, Ill.: Stipes Publishing Co., 1964.

79. Zinn, W. The significance of hoarseness. Ann. Oto. Rhino. Lar., 1945, 54:136.

APPENDIX A

Instructions to Subjects

Instructions to Subjects

In this experiment you will phonate five vowel sounds, at first normally and then while simulating vocal roughness, into the microphone. The vowel sounds you are to produce are the underlined sounds in the words printed on the cards: /i/ as in <u>bee</u>, /æ/ as in <u>cat</u>, /ɑ/ as in <u>hot</u>, /ʌ/ as in <u>but</u>, and /u/ as in <u>boot</u>. You are not to say the entire word, but only the vowel sound that is underlined. The cards will be held so you can see them easily during recording. I will also say each vowel immediately before you speak it.

You should say the vowel sounds loudly enough so that the needle on the meter will peak at the green mark. You will be given two signals from the signal lights. The amber light will come on briefly, indicating that you are to begin to phonate and to peak the needle of the meter steadily at the mark. When the red light comes on, you are to continue to keep the needle steadily at the mark as long as the red light is on. Be very careful to keep the needle on the meter at the mark. Some of the sounds are weak sounds and will have to be spoken loudly to peak at the mark. Some of the sounds are strong sounds and will not have to be spoken as loudly to peak the needle at the mark. You will be given an opportunity to practice peaking the needle on the vowel sounds before actually making the recording.

Produce vocal roughness by phonating while "making your throat tight." A "tight throat" occurs on the initiation of a cough. If you have trouble making your throat tight, start to cough, hold your laryngeal structures in that posture, and phonate. If you wish, I will demonstrate vocal roughness for you. When you are simulating vocal roughness be sure to avoid producing "glottal fry." I will indicate to you if you produce "glottal fry." If you do, we will re-record the vowel. I will also indicate to you if you are not producing the vowel printed on the card. Sometimes while simulating vocal roughness the vowel is distorted. If you do not produce the vowel we will re-record. Are there any questions?

APPENDIX B

Instructions to Judges

Instructions to Judges

You are asked to listen to 250, seven-second sustained vowel samples produced by adult males. The samples are comprised of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/, and represent a range of vocal production from smooth to rough. The vowel samples will be presented to you one at a time, and you are to judge each in relation to a five point scale of severity of vocal roughness. Make your judgments on the basis of the severity of vocal roughness perceived.

Each vowel is to be rated on a scale of equal appearing intervals with scale values from 1 to 5. Scale value 1 represents least severe vocal roughness and 5 represents most severe. Do not attempt to rate vowel samples between any two scale points.

Four vowels representing the extremes of the judgment scale will be presented now to help you locate the extremes of the scale. The first two vowels will represent least severe vocal roughness and the second two will represent most severe. You may listen to these productions as many times as you wish before the judging begins.

The vowels to be judged will now be presented to you in random order. There will be a short interval between productions and each will be preceded by a number announcement.

You are to judge each of the vowel samples in relation to the five point scale of severity of vocal roughness. Record on your response sheet the scale value from 1 to 5 you think each production should be assigned. Because you are asked to scale your perceptions of the severity of vocal roughness, there are no right or wrong scale values. Thus, a scale value you record for a vowel may not be the scale value the person sitting next to you records for that same vowel. For this reason, be sure to make your judgments independently. Record the scale value assigned to each vowel to the right of its number on your response sheet. You may hear each vowel production to be judged as many times as you wish. Notice that you will start at the top of a column and work down. Be sure to record a judgment for every vowel sample. Leave no blank spaces. The vowels will be presented in five segments of fifty vowels each with short rest periods between segments. The instructions and vowel samples representing the scale extremes will be presented again at the beginning of each segment. Are there any questions?

APPENDIX C

Percentages of Inter- and Intra-Judge Roughness Rating Agreement

TABLE 7

PERCENTAGE OF INTER-JUDGE ROUGHNESS RATING AGREEMENT ±1 SCALE VALUE FOR FIFTY VOWEL PRODUCTIONS

	Judge									
Judge	11	10	9	8	7	6	5	4	3	2
1	86	90	88	80	84	96	90	88	100	94
2	100	100	98	100	100	100	100	100	100	
3	96	98	100	98	98	100	100	98		
4	100	100	98	100	100	98	98			
5	96	98	98	100	98	100				
6	98	96	100	98	96					
7	100	98	96	100						
8	98	100	96							
9	98	96								
10	100									

TABLE 8

PERCENTAGE OF INTRA-JUDGE ROUGHNESS RATING AGREEMENT ±1 SCALE VALUE FOR TWO RATINGS OF FIFTY VOWEL PRODUCTIONS

Judge										
1	2	3	4	5	6	7	8	9	10	11
92	100	98	98	98	94	100	100	96	100	100

APPENDIX D

Summary of analysis of variance and Duncan's New Multiple Range Test

TABLE 9

SUMMARY OF THE ANALYSIS OF VARIANCE FOR S-1 (100 Hz to 2600 Hz) SPECTRAL NOISE LEVEL MEANS FOR NORMAL VOWELS

Analysis of Variance

Source	df	ss	ms	F
Subjects	19	242.3	12.8	1.71
Vowels	4	4903.5	1225.9	164.18*
Error	76	561.1	7.4	
Additivity	1	1.0	1.0	.14
Residual	75	560.0	7.5	

* $P < .05$

TABLE 10

DUNCAN'S NEW MULTIPLE RANGE TEST FOR DIFFERENCES AMONG NORMAL VOWEL S-1 (100 Hz to 2600 Hz) SPECTRAL NOISE LEVEL (SNL) MEANS

Vowels	/u/	/i/	/ʌ/	/ɑ/	/æ/
SNL Means	10.9	14.9	23.9	25.4	28.8

(23.9 and 25.4 are underscored by the same line.)

Note: Any two means not underscored by the same line are significantly different at the .05 level.

Any two means underscored by the same line are not significantly different.