71-17,058

WHITEHEAD, Robert Lowell, 1943-
SOME SPECTROGRAPHIC AND PERCEPTUAL FEATURES OF NORMAL, VOCAL FRY, AND SIMULATED ABNORMALLY ROUGH VOWEL PHONATIONS.

The University of Oklahoma, Ph.D., 1970
Speech Pathology

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SOME SPECTROGRAPHIC AND PERCEPTUAL FEATURES OF NORMAL, VOCAL FRY, AND SIMULATED ABNORMALLY ROUGH VOWEL PHONATIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

Robert Lowell Whitehead

Oklahoma City, Oklahoma

1970

SOME SPECTROGRAPHIC AND PERCEPTUAL FEATURES OF NORMAL,

VOCAL FRY, AND SIMULATED ABNORMALLY

ROUGH VOWEL PHONATIONS

APPROVED BY

Floyd W. Emanuel

Walter L. Ellmer

Donald T. Counihan

John W. Keys

Donald Parker

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. Floyd W. Emanuel, director of this study, for his guidance, encouragement, and criticism throughout the planning and completion of the investigation. Appreciation is also expressed to the members of the dissertation committee, Dr. Donald T. Counihan, Dr. Walter L. Cullinan, Dr. Glenda J. Ochsner, and Dr. Roy B. Deal, Jr., for their helpful suggestions during the course of this study.

Additional acknowledgment is made to Dr. Roy B. Deal, Jr. and to Dr. Donald E. Parker, Department of Biostatistics and Epidemiology, University of Oklahoma Medical Center, for their assistance in the statistical analysis of the data. Thanks are due also to the individuals who participated in this study as subjects and judges.

The author also wishes to express his gratitude to his wife, Beth, and to his parents, Mr. and Mrs. L. H. Whitehead, for their constant encouragement, understanding, and support throughout this period of graduate study.

This study was supported in part by an American Speech and Hearing Foundation Traineeship Award.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF TABLES--Continued

LIST OF ILLUSTRATIONS

SOME SPECTROGRAPHIC AND PERCEPTUAL FEATURES OF NORMAL, VOCAL FRY, AND SIMULATED ABNORMALLY ROUGH VOWEL PHONATIONS

CHAPTER I

INTRODUCTION

The human voice is evaluated perceptually with respect to the basic parameters of loudness, pitch, and quality (2, 23, 33, 83, 92). There is normally a high degree of relationship between the perceived loudness and the acoustic intensity and between the perceived pitch and the acoustic frequency of a voice sample; thus, the empirical validity of an individual's vocal pitch or loudness percept may be evaluated by comparing it with the appropriate acoustic correlate. In general, however, acoustic voice features which bear a similarly high degree of relationship to perceived vocal quality have been difficult to isolate. Investigations which help to provide a clear exposition of the acoustic bases of perceived voice quality are presently needed, therefore, to facilitate the verification of perceptually defined voice quality disturbances. Findings from such investigations may be of basic importance to speech pathologists and other clinicians.

Because of the present limited availability of information concerning the acoustic bases of voice quality perception, descriptions of normal and deviant quality tend to be predicated upon the unique and variable perceptual impressions of individuals in a clinical situation and to reflect individual preferences regarding descriptive vocabulary. As might be expected, such descriptions tend to be vague, indistinct, and overlapping. In some instances, however, a broad classification of qualities helps to minimize the ambiguity associated with voice descriptions. The term "vocal roughness," for example, is useful because it describes a quality associated with normal as well as with hoarse or harsh voices (45, 65). Moreover, it appears that perceived vocal roughness is related to specific acoustic features.

Sonagraphic studies (37, 57, 93) have indicated that elevated spectral noise levels are associated with excessive vowel roughness. Because of the relatively broad filter bandwidth (45-Hz) employed and the rapidity of the analysis, however, vowel sonagrams do not lend themselves well to spectral noise level measurements. Recently, therefore, instruments capable of a very narrow-band (3-Hz) acoustic analysis and a relatively slow analysis time have been used (16, 25, 45, 65) to produce vowel spectra in which noise levels can be meaningfully measured. These studies employing narrow-band acoustic filtering have indicated that measured vowel spectral noise levels relate highly and positively to listener judgements of the roughness associated with normal, simulated abnormally rough, and clinically hoarse vowel phonations.

With regard to a further delineation of possible relationships between vowel spectral features and perceived vowel roughness, it is pertinent to consider vocal fry phonation. Vocal fry is described as a peculiar, low-pitched type of phonation which vaguely resembles the sound of a motor boat or rapidly popping corn (20, 55). Considerable roughness is also perceptually associated with fry phonation (20, 83). In the past, some authors have considered vocal fry phonation to be abnormal because of the excessive roughness they perceived to be associated with it (54, 83). This view has since been revised on the basis of research findings indicating that fry phonation differs from other types of unusually rough phonation with regard to the acoustic wave and physiological features which underly its perception (49, 54). Further, it has been observed that vocal fry phonation is commonly heard in speech which is perceived to be normal in quality (29). Many questions remain at this time regarding the acoustic nature of vocal fry phonation and, particularly, regarding the spectral features which may be associated with the roughness of fry phonations. Apparently, an extensive spectrographic study of vocal fry phonation has not been completed previously.

It seems appropriate, therefore, to investigate the possibility that vowel roughness perceptually associated with fry phonations is related to spectral noise levels as it is for vowels produced in other phonatory modes, and to consider the relative roughness of fry phonations on a quality continuum. It is thought that findings bearing on such questions would not only contribute to the fund of basic information regarding speech sound production and perception available to speech scientists, but might have clinical implications as well. The broad purpose of this study, therefore, was to investigate the acoustic spectral noise levels associated with normal, vocal fry, and simulated abnormally rough vowels produced by the same subjects under essentially identical, controlled, experimental conditions. A further purpose was to investigate possible relationships between the vowel spectral noise levels and perceived vowel roughness.

CHAPTER II

REVIEW OF LITERATURE

The objective assessment of vocal quality is of interest to speech and medical clinicians because of their responsibility in the rehabilitation of individuals presenting quality disturbances. Commonly, the evaluation of voice quality is predicated upon the perceptual judgement of a clinician. The value of such judgements may be limited in some instances, however, because the validity of the clinician's unverified perceptual impression is questionable. Additional objective methods for voice quality assessment are currently needed, therefore, to supplement and to verify clinical judgements. It appears likely that when such methods are evolved they will be based in part upon the findings of investigations which have delineated the significant relationships between acoustic voice features and perceived vocal quality. Recently, therefore, numerous studies have been designed to investigate acoustic features which may be related to the perception of specific voice qualities.

Because their perceptual, acoustic, and physiologic similarities and differences are incompletely understood, there is considerable current research interest in vocal fry and in vocal roughness. Recently reported findings (51) suggest that listeners are able to differentiate reliably between vocal fry and abnormally rough phonations.

Certain acoustic wave differences between fry and non-fry phonations have also been reported (30, 49, 91), but acoustic measurement data regarding the spectrographic features of vocal fry, normal, and abnormally rough vowels are sparse. Because such data are potentially useful, the present study employed narrow-band spectrography to investigate further the acoustic features of vocal fry and vocal roughness. Literature reviewed as background to this study is reported under the headings: (a) Perceptual Features of Vocal Fry and Vocal Roughness; (b) Acoustic Features of Vocal Fry and Vocal Roughness; and (c) Physiologic Features of Vocal Fry and Vocal Roughness.

Perceptual Features of Vocal Fry and Vocal Roughness

Vocal Roughness

Recently, vocal roughness has been associated with voices which are essentially normal as well as with those which are abnormal in quality. Emanuel and Sansone (16) and Sansone (65), on the basis of investigations of vowels produced by adult male subjects, reported that normal productions of different vowels tend to differ with respect to perceived roughness. In general, they found that normal productions of the low vowels /ɑ/ and /æ/ were characterized by greater roughness than the high vowels /i/ and /u/. Similar findings have been reported by Lively (45) for adult female subjects.

When vocal roughness is severe, listeners may conclude that it constitutes a voice quality disorder. "Hoarseness," "harshness," "raspiness," "huskiness," and similar terms are commonly used to designate severe and abnormal vocal roughness (74). The descriptions of these aberrant qualities which are subsumed under the general term "abnormal vocal roughness," tend to be somewhat vague. Curtis (13), Van Riper and Irwin (83), Sherman and Linke (68), and others have described harsh quality as an unpleasant, noisy, rasping sound which is associated with excessive strain in phonation. Others (3, 38, 82) have described hoarseness as a combination of breathiness and harshness. Moore (52) has described "dry," "wet," and "rough" types of hoarseness. Thurman (75) reported, however, that "hoarseness," "harshness," and "other terms" are often used to identify qualities which are indistinguishable perceptually. Thus, it is questionable that most listeners can differentiate reliably among the various subtypes of excessive roughness, at least when isolated vowel phonations are evaluated. The term "abnormal vocal roughness" is increasingly accepted by investigators, therefore, as a meaningful label for a "group" of quality disturbances which are often differentiated by writers, but which may not be differentiated reliably by most listeners.

Investigators report that abnormal vocal roughness tends to vary in severity, not only across subjects producing the same isolated vowel, but also across different isolated vowels phonated by the same subject. For example, when Sansone (65) investigated isolated vowels produced with simulated abnormal vocal roughness by adult males, he found that low vowels tended to be perceived as more rough than high vowels. Similar findings have been reported by Lively (45) for vowels phonated with simulated abnormal vocal roughness by adult females and by Hanson (25), Rees (64), and Sherman and Linke (68) for vowels produced by subjects presenting clinically significant abnormal vocal roughness. When vowels are phonated in syllabic contexts, however, the consonant environment tends to affect the degree of vowel harshness perceived. Rees (64), for example, investigated vowels produced in various CVC contexts. She noted that less harshness was perceptually associated with vowels in voiceless than in voiced consonant environments, and in stop plosive than in fricative environments.

To summarize, the literature suggests that vocal roughness is associated with normal as well as with hoarse or harsh vocal quality. The degree of roughness perceptually associated with phonations appears to depend in part upon the vowel produced and, if the vowel is in context, upon the syllabically associated consonants. The general term "abnormal vocal roughness" appears to be gaining acceptance among investigators in part because listeners may be unable to differentiate reliably among the various subtypes of excessive roughness; i.e., hoarseness, harshness, raspiness, and huskiness, when isolated vowels are evaluated. Possibly, such differentiations could be made more readily for connected speech than for isolated vowel samples.

Vocal Fry

Ambiguous and unclear written descriptions of perceptually defined vocal qualities are common in the literature, and an adequate description of the quality associated with vocal fry phonation appears particularly elusive. Moser (55) observed that while vocal fry phonation is easy to demonstrate, the quality associated with it is extremely difficult to describe. "You may recognize it," he said, "as the sound produced by many youngsters in imitating a motor boat, but to me it more nearly resembles the sound of vigorously popping corn." Cleeland (7) described vocal fry quality as "a very low-frequency ticking, a rhythmic beat, or scraping noise." Voelker (84) called it "a rattling, rumbling, cracking, or ticker-like substitute for vocalization." Hollien, Moore, Wendahl, and Michel (29) described fry as having a perceptually distinct "popping" or "pulsing" quality.

In the past, many writers have indicated that vocal fry is a mode of phonation which results in a markedly harsh or rough vocal quality. On the basis of their earlier review of the literature, for example, Moore and von Leden (54) noted that "harshness" was the term used most frequently to describe the vocal quality associated with fry phonation. It has been reported (54), moreover, that Bowler included numerous vocal fry samples in his frequently cited study of ostensibly "harsh" voices. As is suggested by Van Riper and Irwin's (83) guarded comment that fry phonation "contributes to" harsh quality, doubts have arisen concerning the appropriateness of implying that the quality associated with fry phonation is equivalent perceptually to the quality termed "harshness," "hoarseness," or "excessive roughness." Such doubts were given impetus when Moore and von Leden (54) observed by high-speed laryngeal photography that distinctive vocal fold movements, unlike those associated with clinically hoarse phonations, characterize vocal fry phonations.

On the basis of their early photographic studies, Moore and von Leden (54) suggested that vocal fry phonation might be called "dicrotic dysphonia" to indicate that it is a voice disturbance associated with particular physiologic features. The view that vocal fry phonation constitutes a voice disorder has, however, also been questioned of late. On the basis of recent investigations, Hollien, Moore, Wendahl, and Michel (29) concluded that it is inaccurate to regard vocal fry as a voice disturbance. These investigators noted that if one were to phonate solely in vocal fry the resulting voice would undoubtedly be perceived as abnormal, but that vocal fry is often heard in undisturbed phonations and that such usage is apparently normal. They hypothesized that vocal fry phonation is "a normal mode of laryngeal operation that results in a distinctive acoustic signal." Hollien and his associates suggested that their hypothesis might be tested in part by ascertaining whether listeners could perceptually differentiate vocal fry and harsh phonations.

Michel (49) subsequently investigated the possibility that vocal fry and harsh productions could be differentiated perceptually. Ten adult males presenting harsh voices, and ten normal-speaking adult males affecting vocal fry, each recorded a passage of connected speech and a sustained vocalization of the vowel /ɑ/. When these samples were evaluated perceptually, Michel found that both sophisticated and untrained listeners could discriminate reliably between the vocal fry and harsh voice samples when instructed to use "differences in quality" as the sole basis for their discriminatory judgements. These findings are of interest in view of the reports that fry phonations evidence considerable harshness. Possibly, the perceptual identification of vocal fry *per se* is based upon acoustic features other than those associated with the perception of vocal roughness *per se*; thus, similarly-rough fry and non-fry phonations may be differentiated perceptually. It may be suggested further that the acoustic features essential to the perception of vocal roughness are commonly associated to a considerable degree with fry phonations; thus, fry productions are often described as being harsh in quality.

In summary, descriptions of vocal fry quality in the literature appear to be somewhat ambiguous and unclear. This may be attributable in part to rapidly changing concepts of the nature of vocal fry as a phonatory phenomenon. First, vocal fry was described as an abnormal mode of phonation which evoked the perception of harsh vocal quality. More recently, however, it has been suggested that, while substantial harshness or roughness is associated with fry phonation, vocal fry phonation *per se* is not abnormal. Moreover, there is evidence that vocal fry and other types of unusually rough phonation can be differentiated perceptually. Little specific information regarding the degree of roughness associated with fry phonations is currently available, however, and such information would appear to be needed.

Acoustic Features of Vocal Fry and Vocal Roughness

Vocal Roughness

In some respects, the acoustic features of normally and abnormally rough phonations are similar. Michel (49), for example, found that abnormally rough phonations tend to evidence a fundamental vocal frequency which is within the modal range for the subject's age and sex. Similarly, Lieberman (44) studied the fundamental periodicity of pathological larynges and concluded that harshness and hoarseness tend to be associated with fundamental vocal frequencies within the subject's modal register.

Abnormal vocal roughness tends to be associated, however, with vocal frequency perturbations, i.e., small but rapid variations in the fundamental frequency of phonation, which are larger than those associated with normal phonation. Michel (49) and Lieberman (44) have noted that such "pitch" perturbations tend to be larger for harsh or hoarse than for normal phonations. Moore and Thompson (53) found that differences in the periods of consecutive acoustic cycles are generally greater in severely hoarse than in mildly hoarse phonation. Coleman (8) also reported that small random changes in fundamental vocal frequency occur less frequently in normal than in hoarse phonation.

Vocal roughness has also been investigated by relating the amount of time the acoustic wave of phonation is aperiodic to the total phonation time for the sample. According to Michel (49), a wave is aperiodic when there is "a lack of recognizable repeating wave-forms." He investigated the amount of such aperiodicity in standardized passages of connected speech spoken with harsh, vocal fry, and normal vocal qualities. The total time for each sample was first measured from phonellegraphic records of the signal. Subsequently, the amount of aperiodicity, i.e., that portion of the total signal which lacked recognizable repeating cycles was determined, and a ratio computed: the time of aperiodicity divided by total phonation time. It was thus possible for Michel to specify the proportion of total phonation time a phonatory signal was characterized by aperiodicity so extreme that evidence of cyclic vibration could not be discerned. On the basis of this study, Michel reported that normal sustained vowel phonation is aperiodic approximately two percent of the time, while harsh phonation is aperiodic approximately seventeen percent of the time.

Spectrographic studies of the human voice have provided information of importance in understanding vocal roughness. On the basis of a sonagraphic vowel analysis, Thurman (75) reported that vowel formant bandwidth changes and formant frequency shifts are associated with hoarseness. These changes were not consistent within or across subjects, however. Isshiki, Yanagihara, and Morimoto (36) and Yanagihara (93, 94), studying sonagrams of hoarse phonations, found that as the severity of hoarseness increased there was a spread of spectral noise components from the high frequency portions of the spectrum into the lower frequencies.

Using a constant bandwidth wave analyzer, Sansone (65) recently obtained 3-Hz bandwidth frequency-by-amplitude acoustic spectra of vowels produced by adult males both normally and with simulated abnormal vocal roughness. He found that spectral noise levels for abnormally rough vowel productions were characteristically elevated above those for normal productions, and that vowel spectral noise levels were highly and positively correlated with vowel roughness ratings. With respect to differences among vowels, Sansone found that both spectral noise levels and perceived roughness tended to be greater for the low than for the high vowels. Similar results were obtained by Lively (45) who studied the spectral noise levels associated with the normal and simulated abnormally rough vowel phonations of adult females, and by Hanson (25) who studied the phonations of subjects presenting clinically rough voices. Hanson also reported significant positive relationships between the spectral noise levels obtained for the vowels and the roughness ratings obtained for the sentences produced by his subjects.

In summary, acoustic features of abnormal vocal roughness appear to differ in degree rather than in kind from those for normal vocal roughness. Abnormally rough phonations tend to be characterized by a fundamental vocal frequency within the subject's modal range, larger-than-normal pitch perturbations, a large percentage of unmeasurable phonation or wave aperiodicity with respect to sample duration, and elevated spectral noise levels.

Vocal Fry

In contrast to vocal roughness, vocal fry is associated with relatively low fundamental vocal frequencies. McGlone (48) reported a fundamental frequency range for fry phonations from 10.9 to 52.1 Hz. Michel (50) reported fundamentals for fry phonations ranging from 7 to 78 Hz. Similar findings have been reported by Timcke, von Leden, and Moore (77). Hollien and Wendahl (30) reported that fry phonations evidence a train of discrete acoustic excitations of pulses with repetition rates below any level that might be expected for either normal or harsh phonation. They defined a "pulse" as "any of a variety of glottal waveforms of brief duration separated by varying periods of no excitation." It may be noted that Bowler (4) reported a high incidence of extremely low fundamental frequencies for "harsh" phonations, but Moore and von Leden (54) reexamined Bowler's "harsh" samples and reported that they were, in fact, largely fry productions.

Because low fundamental frequencies characterize vocal fry phonations, Hollien and Michel (28) hypothesized that fry phonations may constitute a distinct, low-pitched vocal register. These investigators defined a "register" as "a series or range of consecutive (vocal) frequencies of similar quality," and they suggested that there is "little or no overlap in fundamental frequencies between adjacent registers." To test this hypothesis, Hollien and Michel investigated the frequency limits of modal, falsetto, and vocal fry phonations for twelve male and eleven female subjects. Their results suggested that modal, falsetto, and vocal fry phonations are each associated with a distinct range of fundamental frequencies. Thus, they concluded that vocal fry phonations constitute a low-frequency vocal register.

Michel (49) also reported that vocal fry tends to differ from other modes of phonation with respect to the magnitude of acoustic frequency perturbations. He found significant differences among the perturbation means for sustained normal, vocal fry, and rough phonations. He observed also that fry productions were characterized by the greatest amount of random variation around a mean fundamental frequency, followed by harsh and then by normal or modal register phonations. Because other investigations had suggested a positive correlation between pitch perturbation in a phonatory sample and perceived vocal roughness, Michel's finding that vocal fry productions were characterized by larger perturbations than harsh productions required explanation. Michel accounted for this finding as follows:

> Although a greater perturbation factor was observed in vocal fry than in normal or harsh voices, the pulse-like character of vocal fry was maintained. It is therefore postulated that even though this perturbation, or irregularity of adjacent period lengths, is associated with vocal fry, probably it is not an important factor for the recognition and identification of vocal fry. Thus it is proposed that vocal fry may be either a regular or slightly irregular train of pulses and still be judged as vocal fry. In the harsh voice, however, there is no over-riding pulse-like signal to which the ear can respond. Accordingly, it is proposed that this leaves the perturbation as one of the predominate percepts.

Thus, Michel suggested that even though vocal fry phonations may evidence larger perturbations than harsh phonations, the perception of vocal fry is related to the pulse-like nature of the acoustic signal, while the perception of harshness is related to the perturbation of the signal.

While abnormally rough phonations are associated with considerable signal aperiodicity with respect to the total signal duration, Michel (49) found that vocal fry phonations lacked such aperiodicity. Hollien and Wendahl (30) have also reported that the vocal fry acoustic wave tends to be "regular" rather than aperiodic.

Characteristically, successive cycles of the acoustic voice wave associated with vocal fry phonations are highly damped (9, 91). This wave feature appears to be uniquely associated with vocal fry and, thus, tends to differentiate vocal fry phonations from normally and abnormally rough phonations. To investigate the acoustic wave features associated with the perception of vocal fry, Wendahl, Moore, and Hollien (91) utilized a programmable electrical laryngeal analog coupled to a vowel synthesizer to produce the vowel /ɑ/ with fry quality. They reported that the chief requisite for the perception of vocal fry was an almost complete damping of the acoustic wave between successive excitations. Coleman (9) subsequently investigated the amount of acoustic wave decay necessary for the perception of vocal fry. Studying the vowels /ɑ/ and /u/, Coleman found that vocal fry was perceived when, for each cycle, the difference between the maximum energy of the acoustic wave and that remaining at the time of reexcitation was 42 dB for /ɑ/ and 44 dB for /u/.

In contrast to vocal roughness, the spectral features of vocal fry have apparently not been investigated. Michel (49), however, has offered an hypothesis regarding the amount of spectral noise to be expected in fry phonations. On the basis of his findings regarding the amount of signal aperiodicity associated with vocal fry, harsh, and normal phonations, he suggested that harsh phonations would be characterized by higher levels of spectral noise than normal or vocal fry phonations.

In summary, while vocal fry phonation is characterized by extremely low fundamental vocal frequencies, abnormally rough phonation is associated with fundamental frequencies within the modal register. Larger pitch perturbations are reportedly associated with vocal fry phonations than with normally or abnormally rough phonations. While "signal aperiodicity," as it is defined by Michel (49), is reportedly absent in vocal fry phonations, abnormally rough phonations are characterized by a large percentage of aperiodicity. Further, an acoustic waveform in which successive cycles are highly damped is associated with vocal fry phonations, but not with normally or abnormally rough phonations. Spectral noise levels exceeding normal levels are associated with abnormally rough phonations. Spectral noise data are currently unavailable for vocal fry phonations, however, and would appear to be needed.

Physiologic Features of Vocal Fry and Vocal Roughness

Vocal Roughness

While normal phonation is characterized by slight variations in the duration and amplitude of vocal fold movements in successive vibratory cycles (54), abnormal vocal roughness is associated with relatively large variations in the vocal fold vibratory pattern. Moore and Thompson (53) found that variations in the vibratory frequency of the vocal folds are related to the perceived severity of hoarseness.

Yanagihara (94) has noted that the amplitude of vocal fold movements is often much less in hoarse than in normal phonation. Using high-speed laryngeal photography, Timcke, von Leden, and Moore (76) observed that when voice quality is abnormally rough, the normal time relationships of opening, closing, and closed phases of vocal fold vibration may be disturbed, and that in some instances the closed phase may be diminished or absent. A distinct asymmetry of vocal fold movements accompanying abnormally rough phonation has also been reported by von Leden, Moore, and Timcke (87).

Isshiki and von Leden (35) have suggested that asymmetric vocal fold movements during phonation tend to produce an imperfect modulation of the air stream at the level of the glottis, a turbulent flow of air, and a rough vocal quality. Phonatory air flow rates were obtained by these investigators for thirty-two subjects presenting diverse laryngeal diseases. They reported flows substantially in excess of normal for hoarse subjects unable to achieve glottic closure during phonation, but some of their hoarse subjects evidenced expiratory flows which were near or within normal limits. They attributed the "normal" flow rates for hoarse subjects to high glottal resistance, i.e., to hyperfunction of the laryngeal musculature during phonation.

To generalize, abnormal vocal roughness has been associated physiologically with disturbances in the normal time relationships of intra-cycle phases of glottic valving, asymmetric vocal fold movements, and variations in the frequency and amplitude of consecutive glottic vibratory cycles. Possibly, the roughness associated with normal modal register phonations reflects similar variations of lesser magnitude. The air flow rates associated with abnormally rough phonations vary from normal to excessive, depending on the degree and type of laryngeal disturbance with which the roughness is associated.

Vocal Fry

Studies of vocal fry phonation suggest a unique pattern of vocal fold vibration. Using high speed laryngeal photography, Moore and von Leden (54) observed that during each cycle of fry phonation the folds separate, move toward approximation but do not close, separate again, then approximate and remain in contact for a relatively long period of time. This "double vibration" pattern has also been observed by others (30, 91), but Hollien and Wendahl (30) found that a single opening and closing of the folds followed by a period of no excitation was common in fry phonation. Thus, it has been suggested (30) that the perception of vocal fry quality is not dependent on the number of "pulses" or glottal abductions per cycle, but rather on a sufficiently long closed phase during each cycle to allow for an almost complete damping of vocal tract acoustic excitation.

Some concepts of the physiological processes underlying fry phonation have been predicated on acoustic speech analyses. On the basis of investigations concerned with the acoustic features of vocal fry, Hollien *et al*. (29) speculated that in fry phonation the adducted vocal folds are relatively thick and compressed; the ventricular folds are somewhat adducted; and, the inferior surfaces of the ventricular folds come into contact with the superior surfaces of the true folds. They also speculated that vocal fold vibration in fry phonation is initiated and maintained by relatively low subglottal pressures and that air flow is probably less for fry than for other types of phonation.

Murry (56) subsequently investigated aerodynamic features of vocal fry phonation. He simultaneously recorded intratracheal air pressure, air flow rate, and sound pressure during sustained phonations of two vowels, /ɑ/ and /i/, at three vocal fry and two modal range fundamental frequencies. Murry reported that subglottal air pressures accompanying vocal fry phonations were greater than those accompanying low-frequency modal range phonations, and that the expiratory air flow rates associated with vocal fry phonations were less than those associated with modal range phonations. Murry's latter finding is consistent with findings reported by McGlone (48) who also studied air flow rates during vocal fry phonations. Using five male and five female subjects, McGlone found that expiratory air flow rates associated with vocal fry phonations were smaller than those associated with modal or falsetto phonations. He noted, however, that increases in the fundamental frequency of vocal fry phonations, like those for modal register phonations, were not linear with respect to changes in the rate of phonatory air flow.

Additional information regarding the physiology of vocal fry phonation has been presented by Hollien, Damste, and Murry (27). These investigators studied relationships between vocal fold length and the repetition rate of phonatory pulses during vocal fry. Lateral laryngeal X-rays were taken for eleven males while the subjects attempted to produce vocal fry phonations at three specified repetition rates: 20 Hz, 35 Hz, and 50 Hz. Hollien and his associates reported that vocal fold length during vocal fry phonation was not related to the phonatory repetition rate; the vocal folds were shorter for vocal fry phonations than for even the lowest frequency of phonation in the modal register; and, the mechanisms which govern frequency changes in the modal register, i.e., a systematic lengthening of the vocal folds with pitch increases, do not govern frequency changes in vocal fry phonation.

To summarize, vocal fry phonation tends to be associated with a single opening and closing of the vocal folds followed by a period of no excitation, and with a unique "double" vibratory pattern where, for each cycle, the vocal folds separate, move toward approximation, separate again, approximate and remain in contact for a relatively long period of time. Subglottic air pressures accompanying vocal fry phonations are reportedly greater than those for low-frequency modal register phonations, while air flow rates for fry phonations are less than those for modal register phonations. The vocal folds tend to be shorter during vocal fry than during modal register phonations. In contrast to the changes in vocal fold length which accompany pitch increases in modal register phonations, increases in the pitch of vocal fry phonations do not appear to be associated with a systematic lengthening of the vocal folds.

Summary

The literature suggests that a degree of vocal roughness characterizes normal as well as abnormal vocal quality. It has been reported that substantial harshness or roughness is associated with vocal fry phonations, but there is little specific information available regarding the roughness of fry phonations relative to the roughness of normal or abnormally rough phonations. The literature also suggests that normal, vocal fry, and abnormally rough phonations may be differentiated perceptually. Further, it appears that vocal fry and non-fry phonations tend to differ with respect to acoustic wave and physiologic features.

Concerning acoustic spectral features associated with the roughness of normal and hoarse or abnormally rough phonations, it is reported that as vowel noise levels increase there is a related increase in perceived vowel roughness. The level of spectral noise associated with vocal fry phonations has apparently not been investigated. Additional information regarding spectral noise level and roughness-rating relationships would appear useful in understanding the acoustic nature of normal, vocal fry, and abnormally rough phonations, and may have clinical implications. This study, therefore, was designed to investigate the spectral noise levels associated with vocal fry, simulated abnormally rough, and normal vowel phonations and the relationship of these noise levels to the perceived roughness of the phonations.

CHAPTER III

DESIGN OF THE INVESTIGATION

The purpose of this study was to investigate acoustic spectral noise levels associated with vowels phonated normally, with vocal fry, and with simulated abnormal vocal roughness, and possible relationships between the spectral noise levels and vowel roughness. Twenty normal-speaking adult males individually phonated five selected vowels normally, with vocal fry, and with simulated abnormal vocal roughness at one intensity. A magnetic tape recording was made of each vowel production. The recorded productions were presented in random order to a panel of eleven judges who rated each for roughness. To provide an index of vowel roughness, medians of the judges' ratings for each production were obtained. The mode of phonation represented by each vowel phonation was also evaluated by the judges. The recording of each production was subsequently analyzed to produce a narrow-band (3-Hz) frequency-by-amplitude acoustic spectrum. To provide a quantitative index of vowel spectral noise, the lowest observable peak of energy in each of seventy-nine successive 100-Hz spectral sections from 100 to 8000 Hz was measured in each vowel spectrum. The fundamental frequency of each recorded test vowel phonation was also obtained. These data were then examined with respect to specific research questions. The research questions and the methods employed in this study are discussed in the following sections.

Research Questions

The following research questions concerning selected vowels produced by adult male subjects with vocal fry, simulated abnormally rough, and normal phonation were investigated.

1. What is the relative roughness of each vowel?
2. What are the spectral noise levels associated with each vowel?
3. What relationships obtain between the measures of spectral noise and the roughness ratings for each vowel?

Where appropriate, the findings pertinent to each of the research questions were compared across the three phonatory modes.

Subjects

Twenty normal-speaking adult males, selected chiefly on the basis of their ability to perform the experimental task, served as subjects. The investigation was limited to adult males to provide homogeneity of the subject sample with regard to vocal pitch. Each potential subject was evaluated by a trained speech pathologist to insure that those selected presented normal voice quality and speech. Subjects ranged in age from twenty-three to thirty-four years.

Speech Sample

Subjects individually sustained at one intensity each of five vowels; first normally, then with vocal fry, and then with simulated abnormal vocal roughness. The vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/ were studied to permit investigation of vowel effects on perceived vocal roughness and on acoustic spectral noise levels. Each production was sustained for seven seconds at 75 (± 1 dB) re: 0.0002 dyne/cm^2 (SPL) at a mouth-to-microphone distance of six inches.

Abnormal vocal roughness was simulated by normal-speaking subjects because it was considered desirable to minimize subject differences in the three phonatory modes. Moreover, findings reported by Bowler (4), Sansone (65), and Lively (45) suggest that judges generally do not distinguish perceptually between simulated abnormal and clinical vocal roughness. In addition, Hanson has noted that the acoustic spectral features of abnormal vocal roughness associated with laryngeal pathology are similar to those for simulated abnormal vocal roughness. Thus, it was thought that the present findings might be useful in understanding clinical vocal roughness.

Instrumentation

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

Description

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

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

Its design specifications indicated that the frequency response of the PZT microphone was flat (± 1 dB) from 20 to 8000 Hz when at a

70° angle of incidence to the sound source. The sensitivity of the microphone was -60.3 dB re: 1v/microbar. The sound level meter was designed to indicate the sound pressure level at its PZT microphone with an average signal-to-noise ratio in octave bands from 20 to 10,000 Hz of at least 66 dB. The tape recorder had a flat frequency response ($\pm$ 2 dB) from 40 to 12,000 Hz with a signal-to-noise ratio of at least 65 dB at a tape speed of 15 inches per second.

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

Wave analyzing system. The experimental vowels were each reproduced from tape loops on the tape recorder described above, and were introduced as complex electrical signals into a graphic wave analyzer (General Radio, Type 1910-A) to produce an acoustic spectrum. The analyzer's frequency accuracy to 50,000 Hz was $\pm$ 0.5% of frequency-dial reading, plus 5 Hz. When used in its 3-Hz bandwidth mode, the instrument functioned as a continuously tunable narrow-band filter with the intensity of the frequency components in a complex signal at least 30 dB down at $\pm$ 6 Hz, and at least 60 dB down at $\pm$ 15 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 and coupled the analyzer to a graphic level recorder to permit automatic recording of the level of

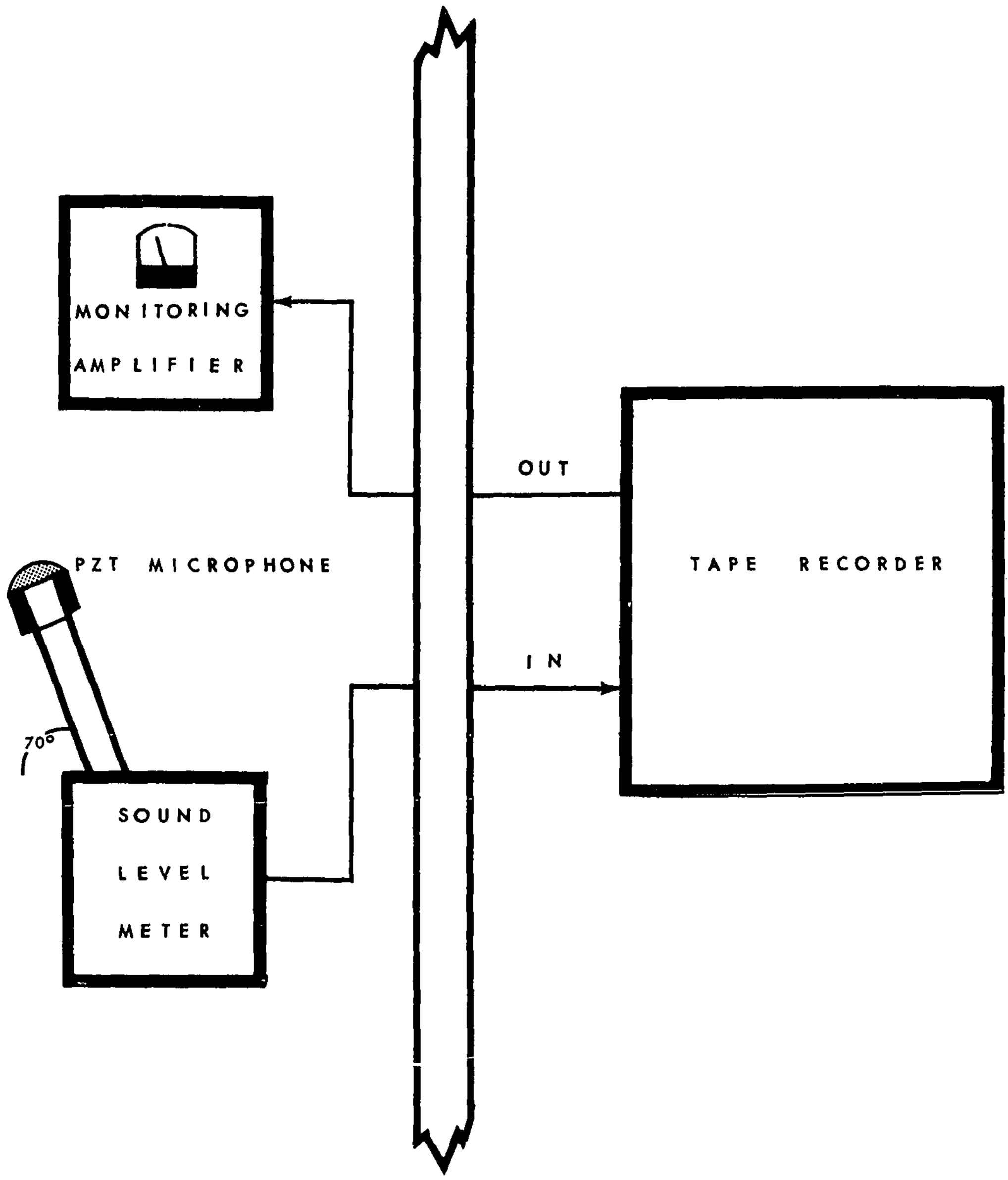

Figure 1.—Simplified diagram of the audio recording system.

components in the complex signal under analysis. The movement of the recorder's chart paper was synchronized with the wave analyzer's frequency-tuning dial. The voltage output of the wave analyzer was proportional to the intensity of the frequency components in a 3-Hz band of the complex signal under analysis and served as an electrical input to the graphic level recorder. The recorder was equipped with an 80 dB input potentiometer designed for accuracy within ± 1% of full scale decibel value. The level recorder's output was proportional to the logarithm of changes in its input and, hence, was linear in decibels. A simplified diagram of the wave analyzing system is presented in Figure 2.

Playback system. The playback system used to present the vowel samples to judges for rating consisted of a dual-channel magnetic tape recorder (Ampex, Model AG 440), an amplifier (Sherwood, Model S9900A), and a loud-speaker (Altec, Model 844A).

Fundamental vocal frequency analyzing system. The system used to determine the fundamental vocal frequency of each test vowel sample consisted of a magnetic tape recorder (Ampex, Model AG 440), a wave analyzer (General Radio, Type 1910-A), and a universal counter (TSI, Model 361). A simplified diagram of this system is presented in Figure 3.

Calibration system. Components employed in instrument calibration included a pure tone oscillator (Hewlett-Packard, Model ABR 200) which drove a loud-speaker (Altec, Model 844A), a sound level meter (General Radio, Type 1551-C), and a manufacturer-calibrated condenser microphone assembly (Bruel and Kjaer, Type 2603). A simplified diagram of the calibration system is presented in Figure 4.

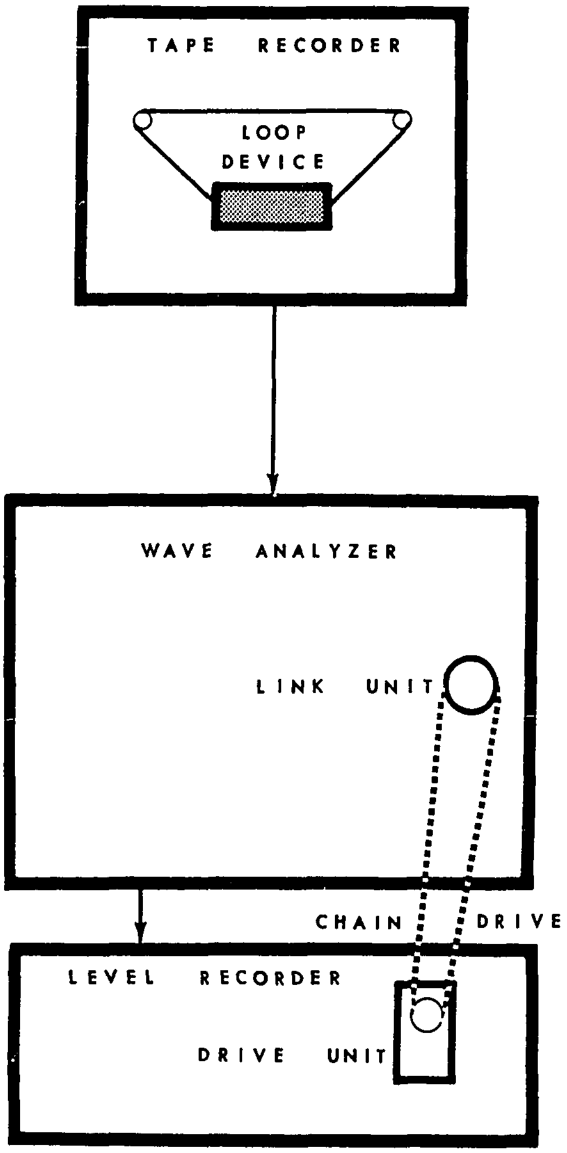

Figure 2.--Simplified diagram of the wave analyzing system.

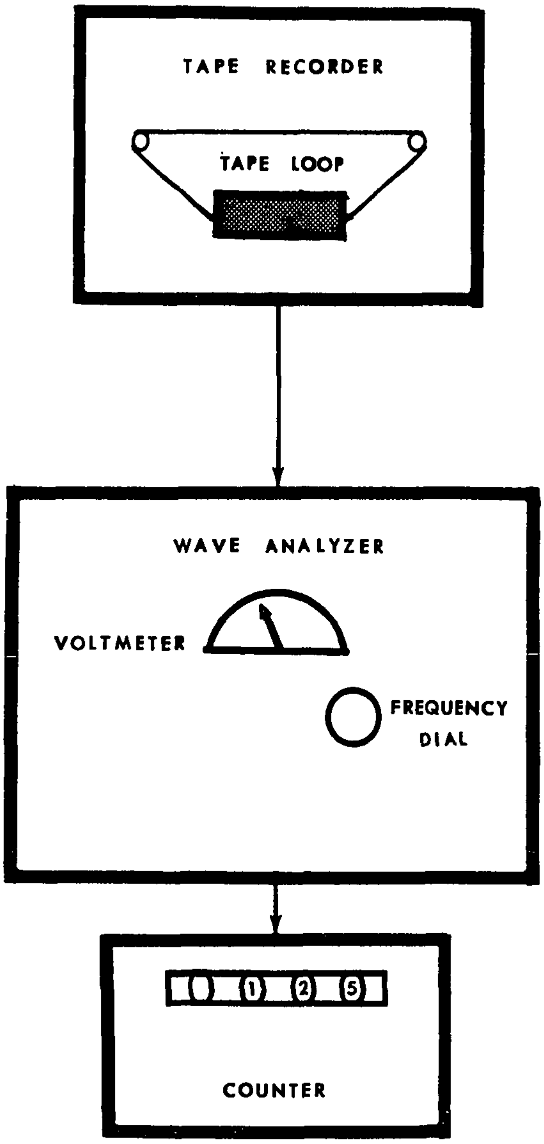

Figure 3.—Simplified diagram of the fundamental vocal frequency analyzing system.

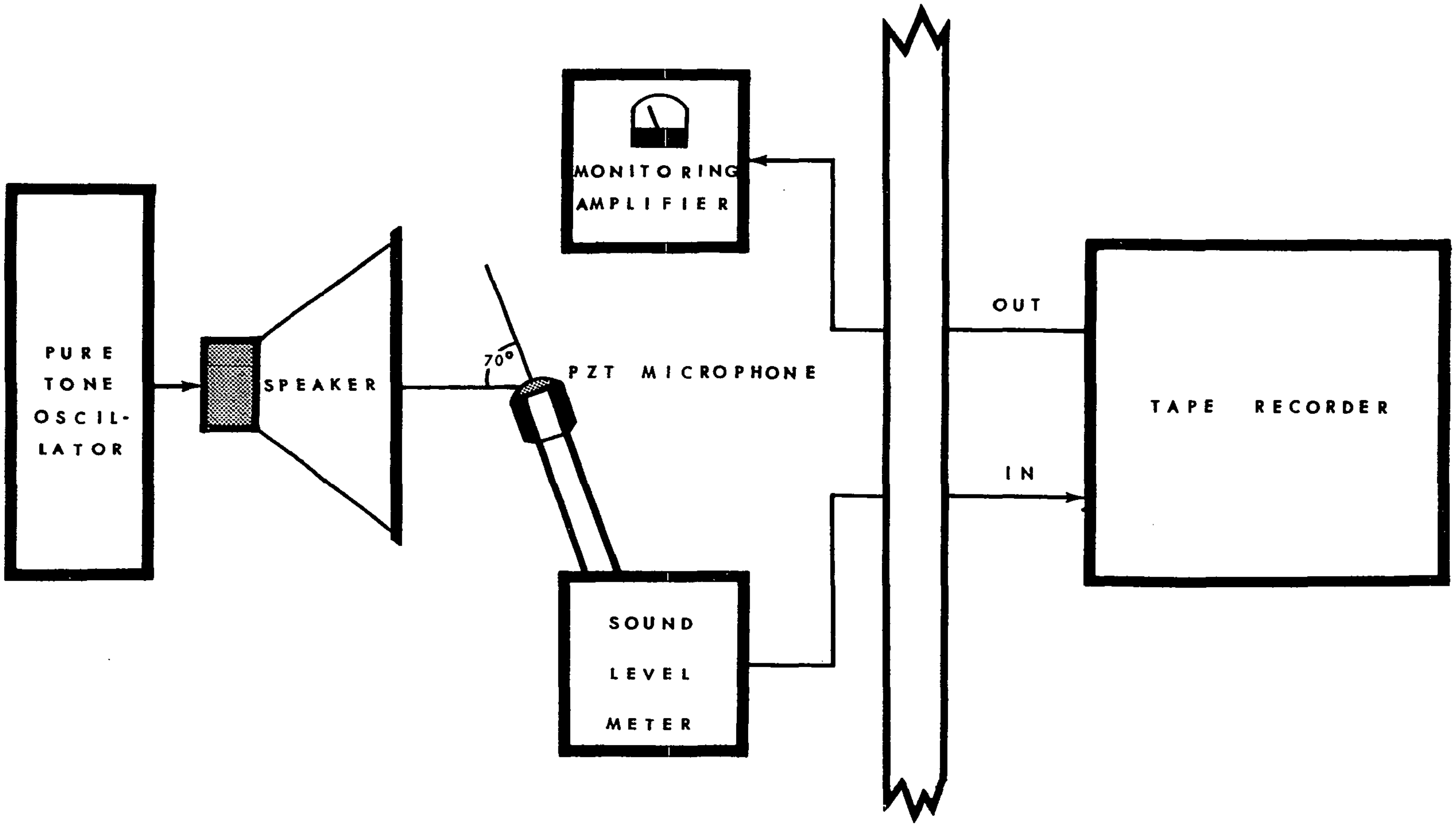

Figure 4.—Simplified diagram of the calibration system.

Calibration

Audio recording system. The magnetic tape recorder was inspected and aligned by an audio engineer immediately prior to data collection. The vocal-intensity-monitoring amplifier's voltmeter was calibrated to indicate when the subject's vocal intensity had reached 75 dB SPL. To calibrate this meter, a 1000-Hz reference tone produced by the pure tone oscillator was led to the loudspeaker. The sound level meter's PZT microphone was placed at a 70^{o} angle of incidence to and two feet in front of the loudspeaker in an acoustically isolated room. The intensity of the pure tone was adjusted until it produced a 75 dB SPL sound level meter deflection. The sound level meter's output was then connected directly to the input of the tape recorder, and the recorder was adjusted for a -2 dB deflection of its VU meter in response to the 75 dB SPL input. The output of the recorder was then led to the monitoring amplifier and the amplifier's input potentiometer was adjusted for a 15 dB deflection on the amplifier's voltmeter in response to the 75 dB SPL input. This deflection of the amplifier's voltmeter was marked as the intensity level each subject was to maintain during test vowel production. The reference tone was then recorded and played back, and the audio recorder's reproduce level was adjusted to match its record level. Thus, speech samples producing a 75 dB SPL indication on the vocal-intensity-monitoring amplifier's voltmeter produced a -2 dB deflection on the recorder's record VU meter. When recorded and played back, the speech samples produced a -2 dB deflection on the recorder's reproduce VU meter.

The PZT microphone used in this study was designed for a flat (± 1 dB) frequency response from 20 to 8000 Hz. Immediately before and after collection of the experimental data, the PZT microphone's frequency response was checked against the flat (± .5 dB from 20 to 10,000 Hz) response of a calibrated condenser microphone and was found to be accurate within the manufacturer's specifications.

Wave analyzing system. Before each use, the graphic wave analyzer was adjusted for minimal carrier frequency intensity at low frequencies and aligned for frequency analysis accuracy within design specifications. After this initial adjustment, intensity calibration was effected by introducing a recorded 75 dB SPL 1000-Hz reference tone into the wave analyzer. The gain of the analyzer and the pen excursion of the graphic level recorder were then adjusted for a 75 dB SPL indication on the graph paper. To insure stability of the wave analyzer's frequency calibration, a daily check was made of its response to a series of reference tones of known frequency produced by the pure tone oscillator.

Procedures

The experimental procedures in this study included recording the subject's productions of the test vowels, analyzing the recorded productions to obtain their fundamental frequencies and frequency-by-amplitude spectra, and presenting the recorded productions to judges for rating.

Recording

All vowel samples were recorded in an acoustically-isolated, two-room testing suite with a low ambient noise level at the Speech and Hearing Center, University of Oklahoma Medical Center. The test room contained the subject's chair, the sound level meter with its attached

PZT microphone, the vocal-intensity-monitoring amplifier, and the signal lights used to control the initiation and termination of each test vowel phonation. The adjoining control room contained the magnetic tape recorder and the cam timer which controlled the activation of the signal lights.

Initially, each subject was familiarized with the experimental procedures and was then seated in the examination chair. The chair's headrest was adjusted vertically for comfort and a headstrap was employed to minimize changes in the subject's position with respect to the microphone during recording. The microphone was placed at a 70° angle of incidence to and six inches in front of the subject's mouth. The monitoring amplifier was positioned to allow the subject to observe readily the intensity of his phonations. The investigator remained in the test room throughout each recording session to monitor the intensity of each vowel production and to cue the subject with printed cards bearing the vowel to be phonated. A copy of the instructions read to the subjects is presented in APPENDIX A.

After being familiarized with the speech material, the subject practiced phonating each test vowel at 75 dB SPL while observing the monitoring amplifier's voltmeter. The subject also practiced timing his phonations with the signal lights until he was able to sustain each vowel for seven seconds while maintaining the required intensity. Upon completion of the training, the normal, vocal fry, and simulated abnormally rough test vowel productions were recorded. The order of vowel presentation was randomized for each subject. The test vowels were produced first normally, then with vocal fry, and then with simulated vocal roughness. This procedure eliminated from normal and fry productions the influence of vocal abuse associated with roughness simulation. Each vowel phonation was carefully monitored by the investigator. If the subject did not produce the appropriate vowel, did not maintain the required intensity, or did not suitably effect vocal fry or abnormal vocal roughness, the trial was repeated until an acceptable performance was achieved.

Spectral Analysis

Tape loops were constructed from the magnetic tape recordings of each normal, vocal fry, and rough vowel produced by each subject. The loops were constructed from a central portion of the vowel recording displaying a uniform intensity of 75 dB SPL ($\pm$ 1 dB) as monitored from the recorder's VU meter. Each loop was two seconds in duration (tape speed 15 ips); thus, initial and terminal vowel inflections were omitted from the analysis. To obtain a 3-Hz bandwidth frequency-by-amplitude spectrum of each vowel, the vowel loops were individually played into the graphic wave analyzer. The graphic level recorder component of the wave analyzer was operated at a paper speed of 0.5 inches per minute and a writing speed of 20 inches per second to produce the vowel spectra. These settings insured adequate resolution of data analyzed in the 3-Hz bandwidth mode and minimized writing stylus overshoot. The time required to produce the spectrum of an individual vowel production under the described conditions was thirty-two minutes.

To evaluate the level of test room and instrumental system noise during data collection, recordings of test chamber noise were made at various times during the day. Tape loops constructed from these recordings were analyzed to produce 3-Hz bandwidth room noise spectra.

To provide an estimate of system noise levels, the high peak of energy in each 100-Hz spectral section from 100 to 8000 Hz was measured in each spectrum. The noise levels attributable to the instrumental systems and test chamber were low and approximately equal at all frequencies above 100 Hz throughout the total spectral frequency range; thus, they appeared to be negligible. There was also negligible noise-level variation in the spectra of test-chamber noise recordings made at different times during the day.

As a quantitative index of vowel spectral noise levels, the lowest observable peak graphic level recorder stylus marking in each 100-Hz section of each vowel spectrum was measured in dB SPL. Seventy-nine measures, one for each successive 100-Hz spectral section from 100 to 8000 Hz, were obtained from the spectrum of each vowel. Measures for the first 100-Hz section of each spectrum were omitted because system noise was greater in that frequency range than in any higher range. Stylus marking overlap, in some instances, may have precluded measurement of the true low noise-level peak in a 100-Hz spectral section; however, measurement of the lowest observable peak provided a numerical index of vowel spectral noise levels.

To assess the reliability of the spectral analysis procedure, three consecutive spectra were made from one tape loop. Spectral noise levels averaged over the frequency range 100 to 8000 Hz did not vary more than ± .2 dB across the three spectra. Differences among noise level means for comparable 1000 Hz segments of the spectra ranged from ± .2 to ± 1.08 dB. Thus, the vowel spectral analysis procedure appeared to be sufficiently reliable.

Fundamental Vocal Frequency Analysis

The vowel tape loops were also used in determining the fundamental vocal frequency of the normal, vocal fry, and rough phonations. The loops were played individually into the graphic wave analyzer which was operated in its 3-Hz bandwidth mode. The analyzer was then hand-tuned upward in frequency from 0 Hz until a large deflection of its voltmeter indicated that the fundamental frequency of the signal had been located. The analyzer's tracking generator output was then coupled to the TSI counter which displayed digitally the fundamental frequency of the vowel production being analyzed.

To evaluate the reliability of the fundamental vocal frequency analysis procedure, all vowel tape loops were analyzed a second time. The obtained fundamentals were found to vary no more than ± 2 Hz across all vowel samples. Thus, the fundamental vocal frequency analysis procedure appeared to be sufficiently reliable.

Roughness Rating

The 300 normal, vocal fry, and rough productions were randomized by means of tape dubbing for presentation to judges. Eleven judges, all graduate students in speech pathology, performed two listening tasks. First, they evaluated the degree of vocal roughness associated with each vowel and, later, they differentiated the normal, vocal fry, and rough productions. The judgements were made in an acoustically isolated room with the judges seated approximately equidistantly from the loudspeaker. The recorder used to reproduce the vowels was located in an adjoining control room. An intercommunication system between the two rooms enabled the judges to indicate when they wished a particular presentation of a vowel sample repeated.

The listening sessions were held on two successive evenings and each was approximately two hours in length. In the initial session, the judges were instructed to rate independently the degree of roughness perceived for each test vowel phonation. A five-point equal-appearing intervals scale in which "1" represented least severe and "5" represented most severe roughness was used. Prior to the rating, the first fifteen vowel samples to be judged were played. The judges were instructed to listen to these samples and to form an independent estimate of the rating scale extremes. A copy of the instructions to judges is presented in APPENDIX B. The speech samples to be rated were then presented in five series of sixty vowel productions and one series of fifty productions, with a brief rest period between series. The final fifty vowel samples, which were the first fifty repeated, were included to evaluate intra-judge reliability. Median scale values of the judges' ratings for each vowel production were then obtained to provide an index of the roughness of each production.

A Pearson $\underline{r}$ of .98 was obtained when the median values of the judges' first and second ratings of the fifty vowel productions in the reliability sample were related. Percentages of intra-judge roughness rating agreement $\pm$ 1 scale value for two ratings of the fifty vowel productions were also obtained. The lowest percentage, 96%, was obtained for Judge 9. Percentages of inter-judge roughness rating agreement $\pm$ 1 scale value for the 300 vowel productions were also obtained. The lowest percentage, 90%, was obtained when the vowel ratings of Judge 9 were compared to those of Judge 4. The results of these procedures are presented in APPENDIX C. The intra- and inter-judge reliability indicated by these data appeared to be adequate.

Mode of Phonation Rating

The appropriateness of the mode of phonation employed for each test vowel production was initially evaluated by the investigator at the time the samples were collected. The validity of the investigator's initial perception was subsequently evaluated, however, by comparing his judgement to that of the judges. In a second listening session, the eleven judges were required to evaluate independently each vowel sample with respect to the mode of phonation it represented. The instructions given to the judges are presented in APPENDIX B. Fifteen vowel samples selected by the investigator, five representing each mode of phonation, were initially played to the judges for practice in differentiating the phonatory modes. The vowel samples were then presented for judgement in five series of sixty productions and one series of fifty productions, with a brief rest period between series. The final fifty productions, which were the first fifty repeated, were included to evaluate intra-judge reliability. The "proper" placement of each vowel production in a given phonatory category was decided on the basis of the assignment made by a majority, i.e., six or more of the judges. APPENDIX D presents the number of judges assigning each phonation to its "proper" category. In no instance did the majority of judges disagree with the investigator's initial opinion regarding the mode of phonation represented by a test phonation. There was unanimous agreement among the judges regarding the mode of phonation represented by 73% of the 300 vowel samples. Ten judges were in agreement regarding 16%, and nine judges regarding 4%, of the 300 samples. Thus, for 93% of the 300 vowel samples nine or more judges agreed with each other and with the investigator's initial opinion regarding the mode of phonation represented. For only seven of the 300 samples did as few as six judges agree with each other and with the investigator regarding the mode of phonation, and these instances occurred only for normal and rough vowel productions. The differences of opinion among judges for these equivocal samples appeared to be attributable to the nature of the samples. That is, disagreements among judges were most frequent for those normal samples which were relatively rough and for those rough samples which were relatively smooth. The judges evidenced comparatively little difficulty in recognizing vocal fry samples. In no instance did fewer than nine judges agree with each other and with the investigator's initial identification of the vocal fry phonations. In general, therefore, agreement among the judges and between the judges and the investigator regarding the mode of phonation represented by each vowel phonation appeared to be adequate. The number of judges "correctly" identifying the mode of phonation of the fifty reliability samples did not diminish from the first to the second evaluation of those samples. Thus, the reliability of the mode-of-phonation evaluations appeared to be adequate. These findings appeared to support the validity of the investigator's initial determination of the mode of phonation represented by the samples.

CHAPTER IV

RESULTS AND DISCUSSION

Results

For this investigation, twenty normal-speaking adult males individually phonated each of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/ normally, with vocal fry, and with simulated abnormal vocal roughness at one intensity. These vowel productions were tape-recorded, and the recordings were subsequently analyzed to determine the fundamental frequency and the 3-Hz frequency-by-amplitude acoustic spectrum of each test phonation. Spectral noise levels associated with individual test vowel phonations were then measured in dB SPL over the frequency range 100 to 8000 Hz. The vowel recordings were also randomized and individually rated for roughness on a five-point scale of equal-appearing intervals by eleven judges. The median of the judges ratings for each production was obtained to provide a single-valued index of the roughness of each test phonation. The median roughness ratings and the spectral noise levels for the vowel phonations were then related. The findings are reported in detail in the following sections.

Fundamental Vocal Frequency

Table 1 presents separately for each test vowel and each of the twenty subjects the fundamental frequency obtained for each normal,

TABLE 1

FUNDAMENTAL VOCAL FREQUENCY FOR EACH NORMAL (N), VOCAL FRY (VF), AND ROUGH (R) VOWEL PRODUCTION

	Vowels								
	/u/			/i/			/ʌ/		
Subject	N	VF	R	N	VF	R	N	VF	R
1	108	50	162	108	56	122	102	56	128
2	119	47	172	106	49	155	122	40	150
3	125	75	132	125	52	121	122	58	133
4	96	60	108	98	49	101	97	44	98
5	107	66	108	116	68	101	100	58	116
6	130	55	139	127	72	145	129	44	135
7	133	66	183	131	59	166	116	54	170
8	99	67	144	98	69	103	109	48	101
9	98	72	130	98	74	112	95	60	113
10	122	69	146	153	50	132	125	56	122
11	128	49	169	131	53	148	134	50	138
12	105	64	140	108	67	105	98	60	125
13	112	72	130	96	66	129	98	46	121
14	105	71	157	113	51	142	107	56	117
15	120	63	139	122	57	183	120	47	143
16	139	67	131	137	46	136	132	48	131
17	116	64	138	103	66	162	112	36	138
18	122	51	188	128	50	200	123	39	153
19	117	58	121	130	52	128	123	45	121
20	116	70	129	113	52	128	116	56	132

TABLE 1--Continued

Subject	/ɑ/ N	/ɑ/ VF	/ɑ/ R	/æ/ N	/æ/ VF	/æ/ R
1	106	48	119	98	49	112
2	101	43	153	103	30	143
3	121	32	114	122	41	114
4	99	30	101	99	40	96
5	96	57	114	98	47	97
6	118	45	126	116	35	144
7	122	56	163	133	54	147
8	106	59	100	106	56	107
9	95	66	105	100	61	108
10	134	44	132	114	48	124
11	124	48	134	130	49	136
12	108	49	132	101	47	104
13	101	44	111	96	41	121
14	121	48	113	101	47	105
15	127	50	111	116	43	123
16	126	40	138	130	47	139
17	120	40	137	115	45	122
18	120	47	161	123	38	169
19	113	48	119	128	48	129
20	109	46	121	114	47	119

vocal fry, and simulated abnormally rough phonation. This table shows that for each subject the fundamental for the vocal fry phonation of each vowel was lower than that for the normal and rough production. Table 1 also shows that a higher fundamental tended to be associated with the rough than with the normal phonation of each test vowel. These trends are also seen in Table 2 which presents separately for normal, vocal fry, and simulated abnormally rough productions, the fundamental frequency means and standard deviations, and the range of fundamentals for each test vowel. The means are over the twenty subjects. Table 2 shows that for each test vowel the range of relatively low fundamentals for vocal fry productions did not overlap the range of relatively high fundamentals for either normal or rough productions. There was, however, considerable overlap in the ranges for normal and rough productions of each vowel.

Table 2 reveals further that the fundamental frequency means tended to vary directly with tongue height in vowel production. When either simulated rough or vocal fry productions are considered, for example, the test vowels may be ranked in descending order of average fundamental frequency: /u/, /i/, /ʌ/, /ɑ/, and /æ/. A similar trend obtains for the normal test vowel productions with only one slight reversal, i.e., the mean for /i/ was higher than that for /u/. The results of randomized complete-block analyses of variance and Duncan's New Multiple Range Tests employed to test the significance of these trends are summarized in APPENDIX E. These tests revealed that, although all the observed trends toward differences in fundamental frequency were not significant, vowel fundamentals at the extremes of the range for each mode of phonation were demonstrably different. For both abnormally

TABLE 2

FUNDAMENTAL VOCAL FREQUENCY MEANS AND STANDARD DEVIATIONS, AND THE RANGE OF FUNDAMENTALS FOR NORMAL, VOCAL FRY, AND SIMULATED ABNORMALLY ROUGH VOWELS PRODUCED BY TWENTY ADULT MALES

	Normal			Vocal Fry			Rough		
Vowel	FVF Mean	SD	Range	FVF Mean	SD	Range	FVF Mean	SD	Range
/u/	116	12.03	96–139	63	8.45	47–75	143	22.37	108–188
/i/	117	15.56	96–153	58	8.90	46–74	136	27.33	101–200
/ʌ/	114	12.53	95–134	50	7.38	36–60	129	17.21	98–170
/ɑ/	113	11.45	95–134	47	8.44	30–66	125	18.47	100–163
/æ/	112	12.43	96–133	46	7.09	30–61	123	18.90	96–169

rough and vocal fry productions, for example, the high vowels /u/ and /i/ evidenced significantly ($p < 0.05$) higher fundamentals than the low vowels /ɑ/ and /æ/. For normal productions, the fundamental for /i/ was significantly ($p < 0.05$) higher than that for /æ/.

Roughness Ratings

Table 3 presents the median of the eleven judges' roughness ratings for each of the five vowels produced normally, with vocal fry, and with simulated abnormal vocal roughness by each of the twenty subjects. This table shows that, for each subject, vocal fry and simulated abnormally rough productions of each test vowel received higher median roughness ratings than normal productions. The range of median ratings shown in Table 3 was, for normal productions, from 1.05 to 3.38; for fry productions, from 3.19 to 5.00; and, for rough productions, from 2.29 to 5.00. The relatively large range associated with the simulated abnormally rough productions may be attributable to the fact that the degree of vowel roughness subjects simulated was not controlled.

Table 4 presents the median roughness ratings for normal, vocal fry, and rough productions of each vowel, averaged over the twenty subjects. This table shows, when the means for normal productions are considered, that the high vowel /u/ tended to be rated least rough and the low vowel /æ/ most rough. The scale separation between these extreme vowels is .55. Regarding vocal fry productions, Table 4 shows that the high vowel /i/ tended to be rated least rough and the low vowel /æ/ most rough. Because of the very restricted range of ratings assigned to fry productions, however, the scale separation between the extreme vowels is only .28. Regarding the simulated abnormally rough

TABLE 3

MEDIAN ROUGHNESS RATINGS FOR EACH NORMAL, VOCAL FRY, AND ROUGH VOWEL PRODUCTION

	Vowels								
	/u/			/i/			/ʌ/		
Subject	N	VF	R	N	VF	R	N	VF	R
1	2.08	4.95	4.71	1.29	4.95	4.71	1.42	4.89	4.42
2	1.60	5.00	4.71	1.42	4.89	4.95	1.42	4.95	4.95
3	1.19	4.00	4.59	2.06	4.00	4.59	1.19	4.89	4.89
4	2.11	4.95	4.59	2.71	4.89	4.89	1.80	4.95	4.89
5	1.11	3.71	4.00	1.81	4.06	4.00	1.86	4.42	4.00
6	2.40	4.71	4.81	1.19	4.14	3.80	1.19	3.89	3.40
7	1.11	4.60	4.00	1.11	5.00	3.42	1.86	4.81	4.81
8	2.25	4.29	3.75	2.59	3.19	4.08	1.19	4.89	4.19
9	1.19	4.00	3.75	1.42	4.11	3.80	1.29	4.29	4.59
10	1.05	4.20	4.89	2.00	4.59	4.89	1.67	4.59	4.33
11	1.42	4.81	4.59	1.19	4.71	2.19	1.60	4.95	2.42
12	2.19	4.89	5.00	2.00	4.81	5.00	2.20	4.95	5.00
13	1.42	4.14	4.25	1.94	4.59	3.59	1.86	4.95	4.95
14	1.59	4.19	4.71	2.14	4.95	3.67	2.59	5.00	4.89
15	1.29	5.00	4.42	2.06	4.95	4.59	1.42	5.00	3.75
16	1.05	4.81	3.86	1.42	4.95	4.33	1.11	4.95	4.19
17	1.80	4.71	4.14	2.06	4.11	3.11	1.71	4.89	3.25
18	1.11	4.95	3.40	1.19	4.95	4.33	1.11	4.42	3.59
19	1.42	4.59	5.00	2.20	4.06	4.89	1.75	4.89	4.14
20	1.86	4.11	4.08	1.11	3.80	2.86	1.29	4.29	2.29

TABLE 3--Continued

	Vowels					
	/ɑ/			/æ/		
Subject	N	VF	R	N	VF	R
1	1.42	4.95	4.00	2.06	4.89	4.00
2	1.29	4.89	4.95	1.86	4.95	5.00
3	2.25	4.00	4.59	1.75	4.71	4.81
4	2.59	5.00	4.81	2.25	4.59	4.81
5	1.42	4.40	3.71	2.29	4.42	4.00
6	1.29	4.81	4.08	2.00	4.89	4.25
7	2.00	4.89	3.59	2.42	4.12	3.59
8	2.19	4.19	4.81	2.42	4.81	4.81
9	1.92	4.00	3.88	2.20	4.89	4.08
10	2.42	5.00	4.81	2.60	4.59	4.81
11	1.71	4.89	2.40	1.42	4.81	3.86
12	3.38	4.71	4.81	2.29	4.81	4.95
13	1.05	4.95	4.00	2.00	4.89	3.88
14	1.71	5.00	4.71	2.59	5.00	4.81
15	1.60	4.71	4.14	2.00	5.00	4.89
16	1.81	4.06	3.88	1.86	4.71	4.71
17	1.19	4.95	3.00	2.00	4.71	3.40
18	1.29	5.00	4.00	1.60	5.00	3.86
19	1.95	4.71	4.59	3.00	4.71	4.89
20	1.62	4.25	3.08	1.59	4.89	4.42

TABLE 4

MEDIAN ROUGHNESS RATINGS FOR NORMAL, VOCAL FRY, AND ROUGH PRODUCTIONS OF EACH VOWEL AVERAGED OVER TWENTY SUBJECTS

	Average Median Roughness Rating		
Vowel	Normal	Vocal Fry	Rough
/u/	1.56	4.53	4.36
/i/	1.75	4.48	4.08
/ʌ/	1.58	4.74	4.15
/ɑ/	1.81	4.67	4.09
/æ/	2.11	4.76	4.39

productions, Table 4 shows that the high vowel /i/ tended to be rated least rough and the low vowel /æ/ most rough. The scale separation between these extreme vowels is only .31. Table 4 reveals, therefore, that for all three modes of phonation the low vowel /æ/ tended to evidence greater roughness than the other test vowels. Least roughness tended to be associated with either /u/ or /i/, depending on the mode of phonation; thus, least roughness was associated with a high vowel for all three modes of phonation. It may also be noted in Table 4 that when ratings for vocal fry and simulated abnormally rough productions of the test vowels are compared, the average rating for fry productions of each vowel tended to be higher than that for the simulated abnormally rough productions. In general, however, for each test vowel the average ratings for fry and for simulated abnormally rough productions tended to be similar and to be relatively high on the roughness scale with respect to those for normal productions.

Spectral Noise Levels

Example acoustic spectra for normal and fry, and for rough and fry phonations of one of the test vowels are presented in Figures 5 and 6 respectively. These spectra are for the vowel /æ/ as produced by Subject 9. In general, the spectral features observed for /æ/ were also observed for the other test vowels. Figure 5 shows that the spectrum for the normal /æ/ production is characterized by identifiable harmonics over approximately the lower one-half (100 to 4000 Hz) of the analyzed frequency range (100 to 8000 Hz). The harmonics are obscured by the noise at the high-frequency end of the spectrum. Where harmonics are identifiable, noise components may be seen between them. In contrast, Figure 6 shows that, in the rough vowel spectrum, noise components are elevated and harmonics are obscured by noise except in the low-frequency range. Identifiable low-frequency harmonics in the spectrum of the rough production appear to be diminished in amplitude with respect to those in the spectrum of the normal production.

Inspection of the spectrum for the vocal fry production in either Figure 5 or 6 reveals that spectral noise levels are markedly elevated across the frequency range analyzed. With respect to spectral noise levels, the rough and fry production spectra in Figure 6 appear more similar than the normal and fry production spectra in Figure 5. Figure 6 also shows that even fewer harmonics are discernable in the fry than in the rough spectrum. Moreover, the low-frequency harmonics which may be delineated in the fry spectrum appear to be more greatly diminished than those in the rough spectrum.

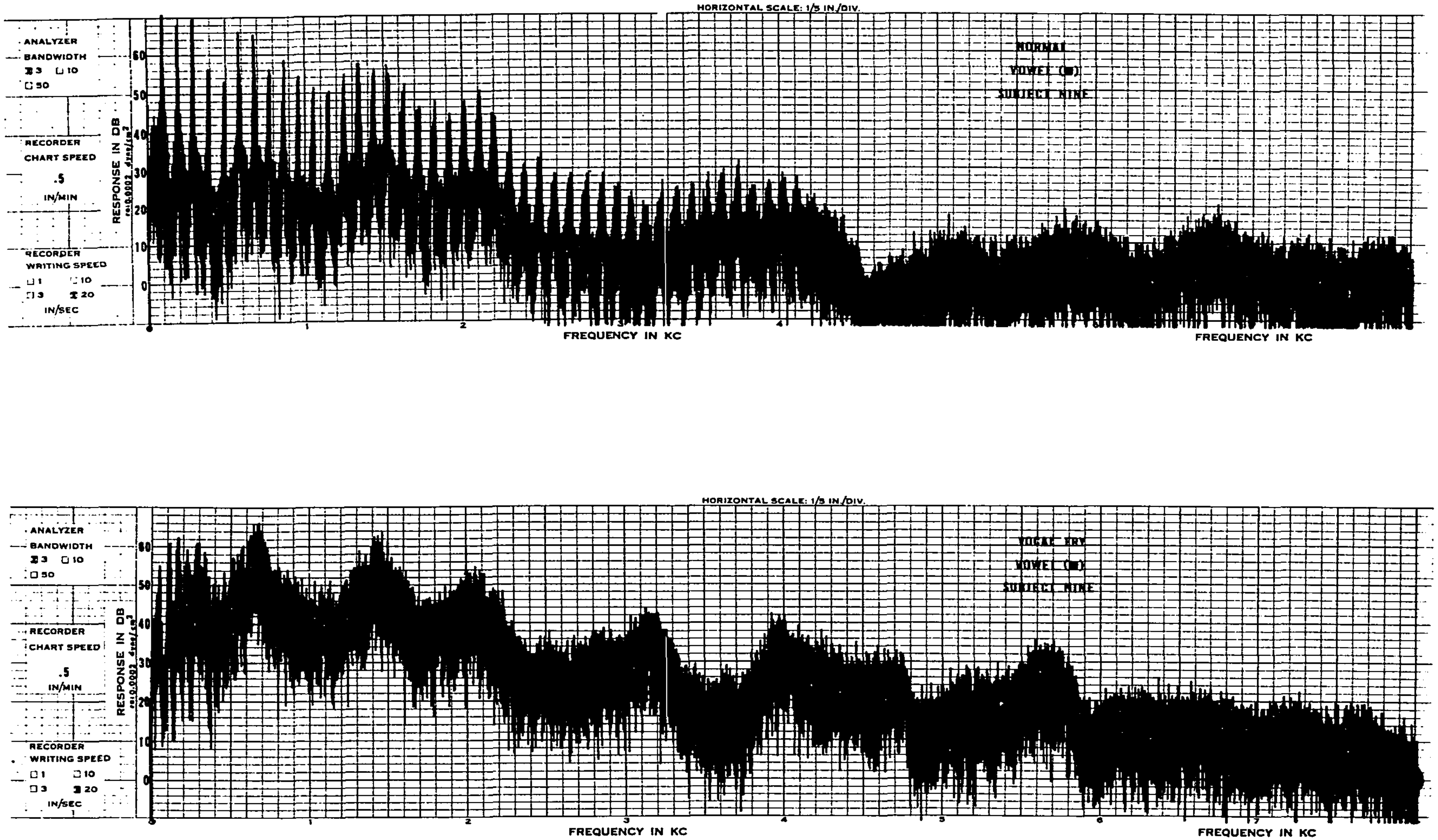

Figure 5.—(Top) Spectrum of a normal /æ/; (Bottom) Spectrum of a vocal fry /æ/.

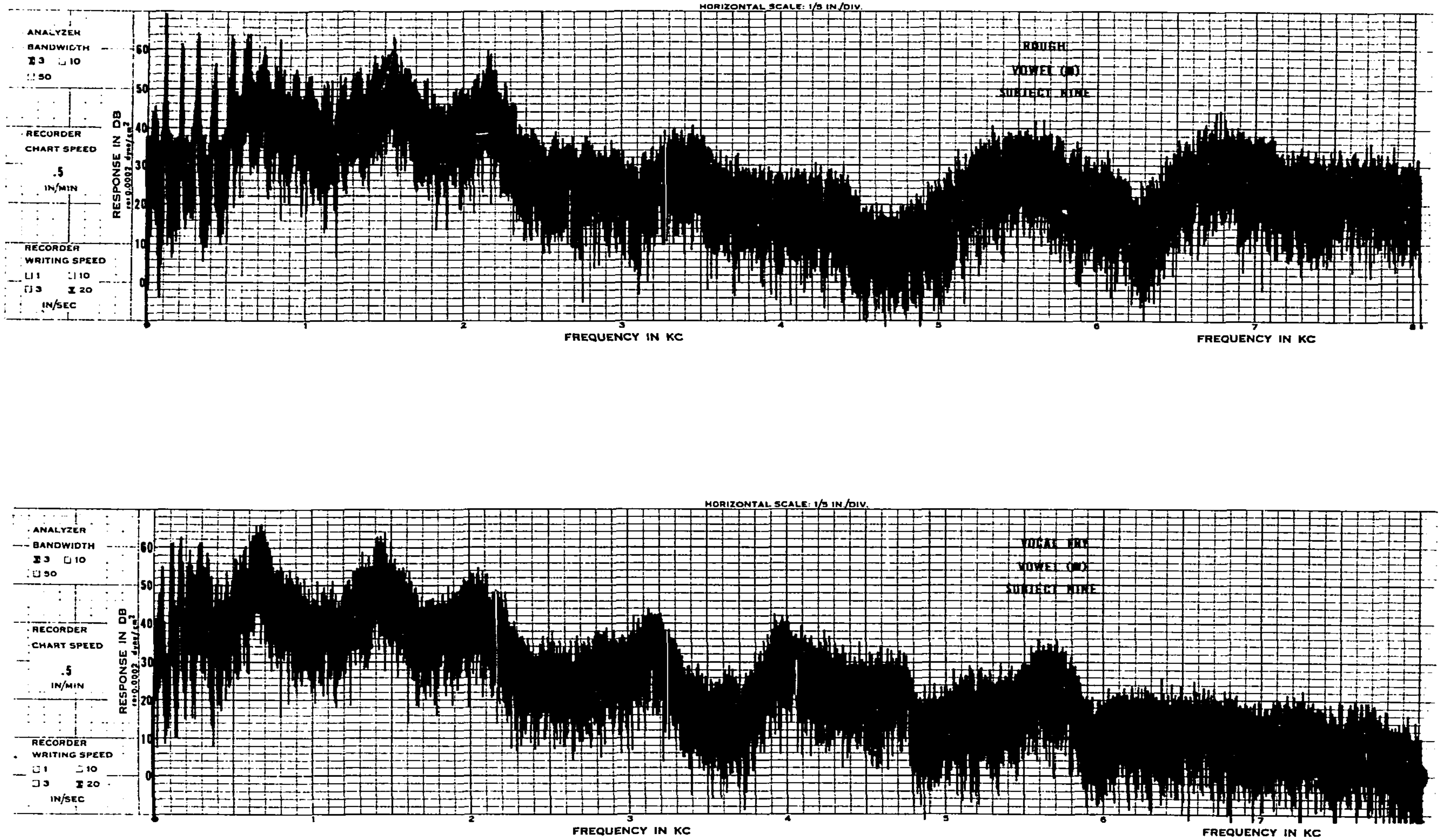

Figure 6.—(Top) Spectrum of a rough /æ/; (Bottom) Spectrum of a vocal fry /æ/.

Figures 5 and 6 also reveal a general tendency for both harmonic and noise levels to reflect the formants of the vowel produced. This tendency has been noted previously (25, 45, 65), and is thought to be attributable to frequency-selective acoustic damping in the vocal tract resonators.

Spectral noise levels (SNLs) for each of the test vowels averaged over all subjects and over the spectral frequency range 100 to 2600 Hz, and the standard deviations associated with the SNL means, are presented in Table 5. The frequency range considered (100 to 2600 Hz) was selected for two reasons. First, others (25, 45, 65) have reported that the variation in individual vowel spectral noise levels across subjects is less for measures in this than in other spectral ranges. Second, the cited previous investigations have demonstrated a very high degree of linear relationship between vowel production SNL measures over this frequency range and the roughness perceptually associated with the productions.

Table 5 reveals that the SNL means for the rough and vocal fry productions tend to exceed those for normal productions of each test vowel. Further, except for /u/, the SNL means for the vocal fry productions of each vowel tend to exceed those for rough productions. For /u/, the mean for rough productions tended to exceed very slightly the mean for fry productions. Table 5 also shows that when normal and fry productions were considered, the vowels could be ranked with respect to increasing SNL means: /u/, /i/, /ʌ/, /ɑ/, and /æ/. In general, this order also held for rough productions, except that the SNL mean for rough productions of /i/ was slightly smaller than that for /u/. Table 5 reveals also that, in general, low vowels tended to be characterized

TABLE 5

NORMAL, VOCAL FRY, AND ROUGH VOWEL SPECTRAL NOISE LEVEL (SNL) MEANS AND STANDARD DEVIATIONS FOR TWENTY MALE SUBJECTS, AND DIFFERENCES BETWEEN THE VOCAL FRY AND NORMAL, VOCAL FRY AND ROUGH, AND ROUGH AND NORMAL VOWEL SNL MEANS (SNLDS)

	Frequency Range 100-2600 Hz				
Vowel	Normal SNL Mean	Standard Deviation	Vocal Fry SNL Mean	Standard Deviation	SNLD Fry-Normal
/u/	13.2	2.9	30.0	4.3	16.8
/i/	16.1	3.4	33.1	3.6	17.0
/ʌ/	25.7	4.7	38.6	2.3	12.9
/ɑ/	26.8	4.3	38.9	2.3	12.1
/æ/	31.4	3.9	43.8	1.9	12.4

	Frequency Range 100-2600 Hz				
Vowel	Rough SNL Mean	Standard Deviation	Vocal Fry SNL Mean	Standard Deviation	SNLD Fry-Rough
/u/	30.2	4.9	30.0	4.3	- 0.2
/i/	30.1	4.8	33.1	3.6	3.0
/ʌ/	37.7	4.5	38.6	2.3	0.9
/ɑ/	37.7	4.7	38.9	2.3	1.2
/æ/	41.2	2.9	43.8	1.9	2.6

	Frequency Range 100-2600 Hz				
Vowel	Normal SNL Mean	Standard Deviation	Rough SNL Mean	Standard Deviation	SNLD Rough-Normal
/u/	13.2	2.9	30.2	4.9	17.0
/i/	16.1	3.4	30.1	4.8	14.0
/ʌ/	25.7	4.7	37.7	4.5	12.0
/ɑ/	26.8	4.3	37.7	4.7	10.9
/æ/	31.4	3.9	41.2	2.9	9.8

by larger SNL means than high vowels, regardless of the phonatory mode employed in vowel production.

Spectral noise level differences (SNLDs) between the means for vocal fry and normal, vocal fry and rough, and rough and normal productions of each test vowel are also presented in Table 5. Table 5 shows that, in general, fry-normal and rough-normal SNLDs tend to be smaller for low than for high vowels. It is also evident in Table 5 that fry-normal and rough-normal SNLDs tend to be substantially larger than fry-rough SNLDs. Thus, the spectral noise levels associated with fry phonations more nearly resemble those associated with simulated abnormally rough than with normal phonations.

Table 5 also presents the standard deviations associated with the normal, rough, and vocal fry SNL means. This table shows that the greatest SNL variability tended to be associated with the rough productions. For the rough and vocal fry productions there was a general trend toward smaller standard deviations for low than for high vowels. A similar trend was not evident, however, for the normal productions.

Differences between the SNLs associated with vocal fry, rough, and normal productions of /æ/ are illustrated in Figure 7. In this figure, individual 100-Hz SNLs averaged over the twenty subjects are plotted separately for normal, vocal fry, and simulated abnormally rough productions of /æ/. Figure 7 shows that the 100-Hz SNL means for both vocal fry and rough /æ/ productions tended to exceed those for the normal /æ/ productions over the total spectral range (100 to 8000 Hz). Similar trends were evident for the other test vowels. Figure 7 also reveals for /æ/ that 100-Hz SNL means for fry productions tended to exceed those for the rough productions from 100 to 1700 Hz. The 100-Hz

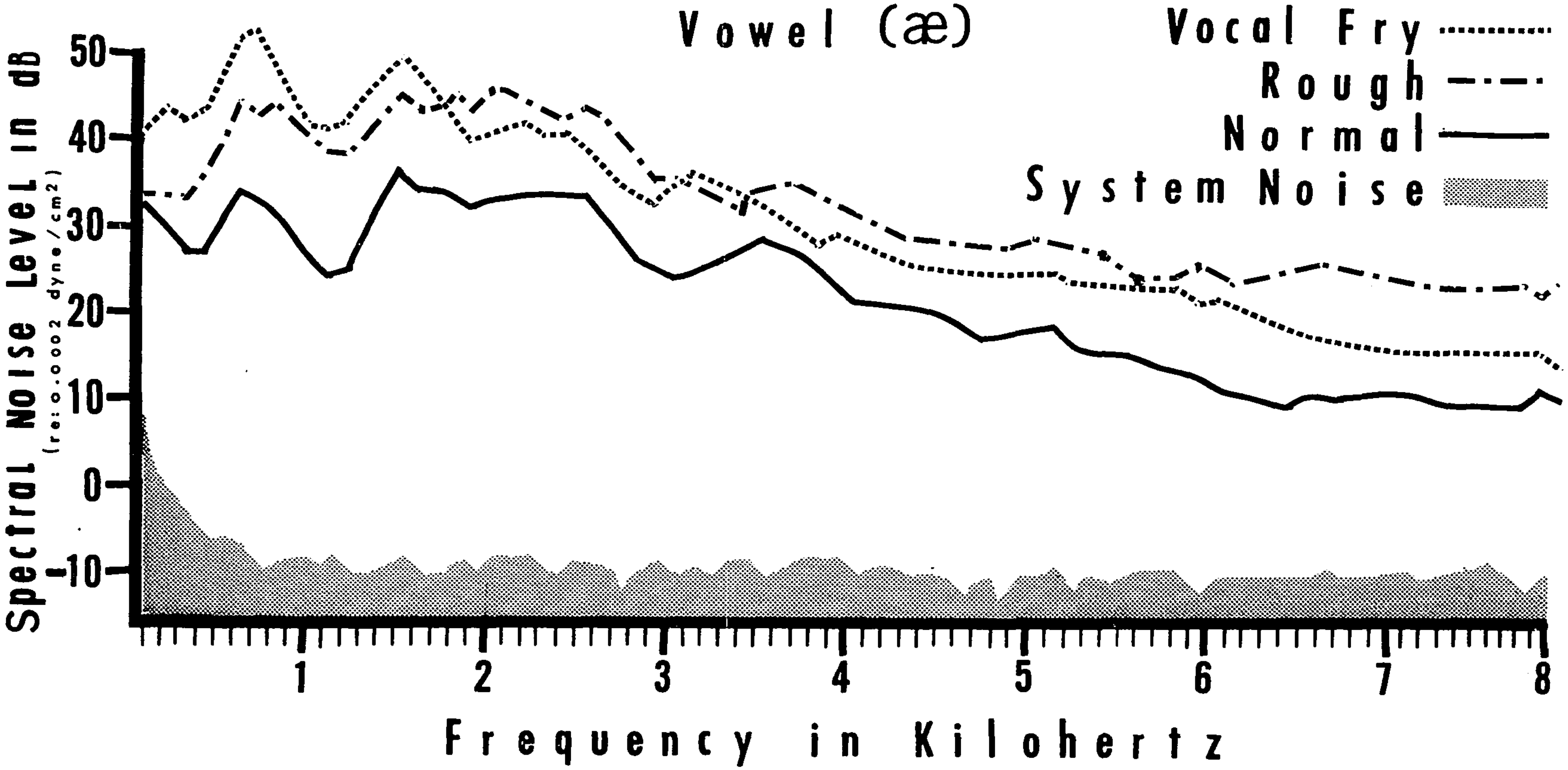

Figure 7.--Noise levels in each 100-Hz spectral section averaged over twenty male subjects for normal, vocal fry, and simulated abnormally rough productions of the vowel /æ/.

means for the rough productions tended to exceed those for the vocal fry productions at higher spectral frequencies, except from 3200 to 3400 Hz.

Spectral Noise Level and Roughness Rating Relationships

To investigate possible relationships between vowel spectral noise levels and roughness severity ratings, scatter diagrams of SNLs averaged over the spectral segment 100 to 2600 Hz and median roughness ratings for the productions of each vowel were plotted. The scatter diagrams for each test vowel are presented in Figures 8 through 12. These figures suggest a positive relationship between the SNL means and the median roughness ratings for each of the test vowels. In general, as the roughness of each vowel increased, its spectral noise level tended to increase.

Table 6 presents correlation coefficients (Pearson r)

TABLE 6

CORRELATION COEFFICIENTS FOR NOISE LEVELS AVERAGED OVER THE SPECTRAL FREQUENCY RANGE 100 TO 2600 Hz AND MEDIAN ROUGHNESS RATINGS FOR THE PRODUCTIONS OF EACH TEST VOWEL

Vowel	Correlation Coefficients *
/u/	0.92
/i/	0.90
/ʌ/	0.89
/ɑ/	0.88
/æ/	0.91

* All coefficients significant ($p < 0.01$)

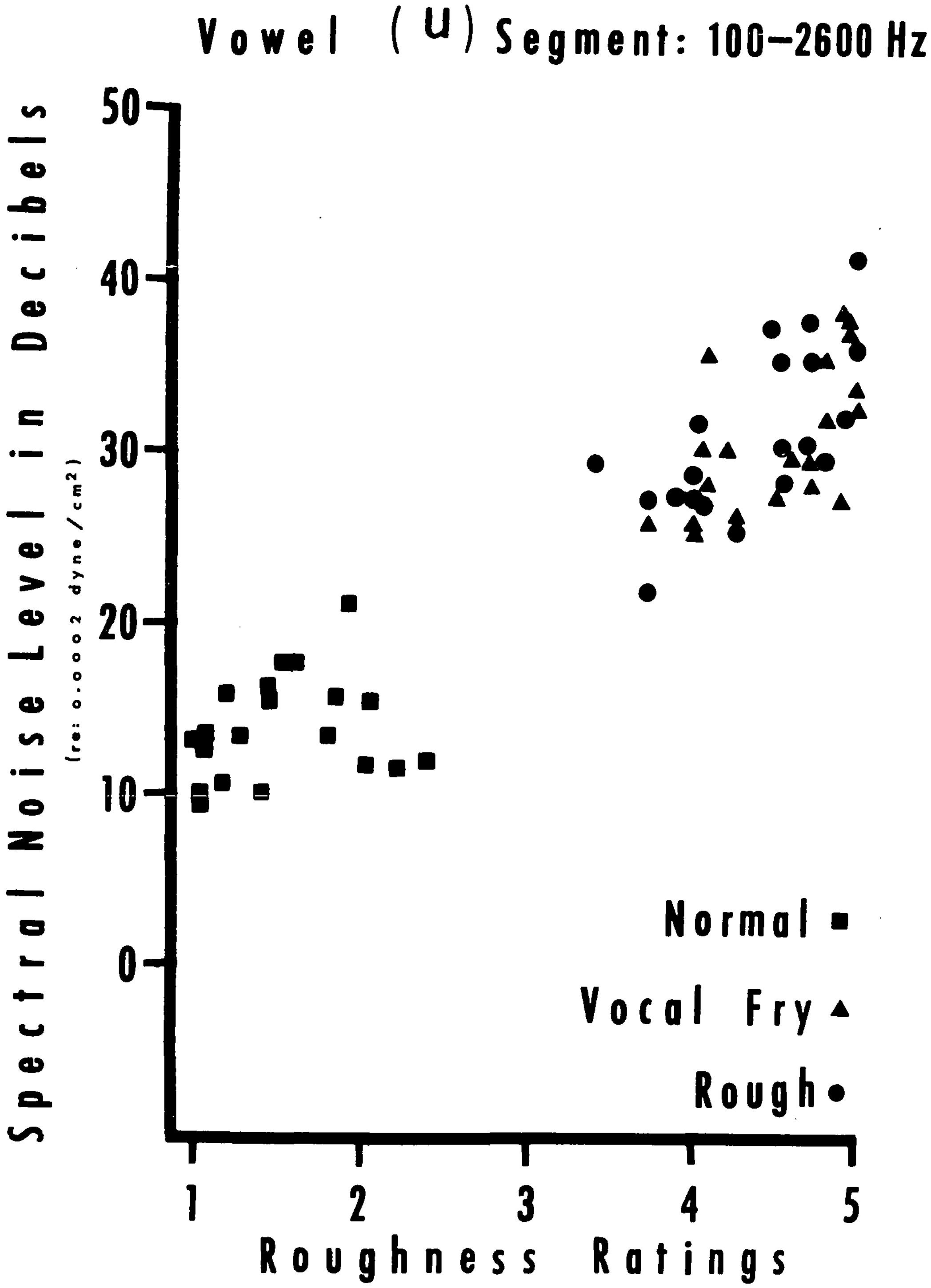

Figure 8.--Spectral range 100 to 2600 Hz noise level means and median roughness ratings over eleven judges for twenty male subjects' normal, vocal fry, and simulated abnormally rough productions of the vowel /u/.

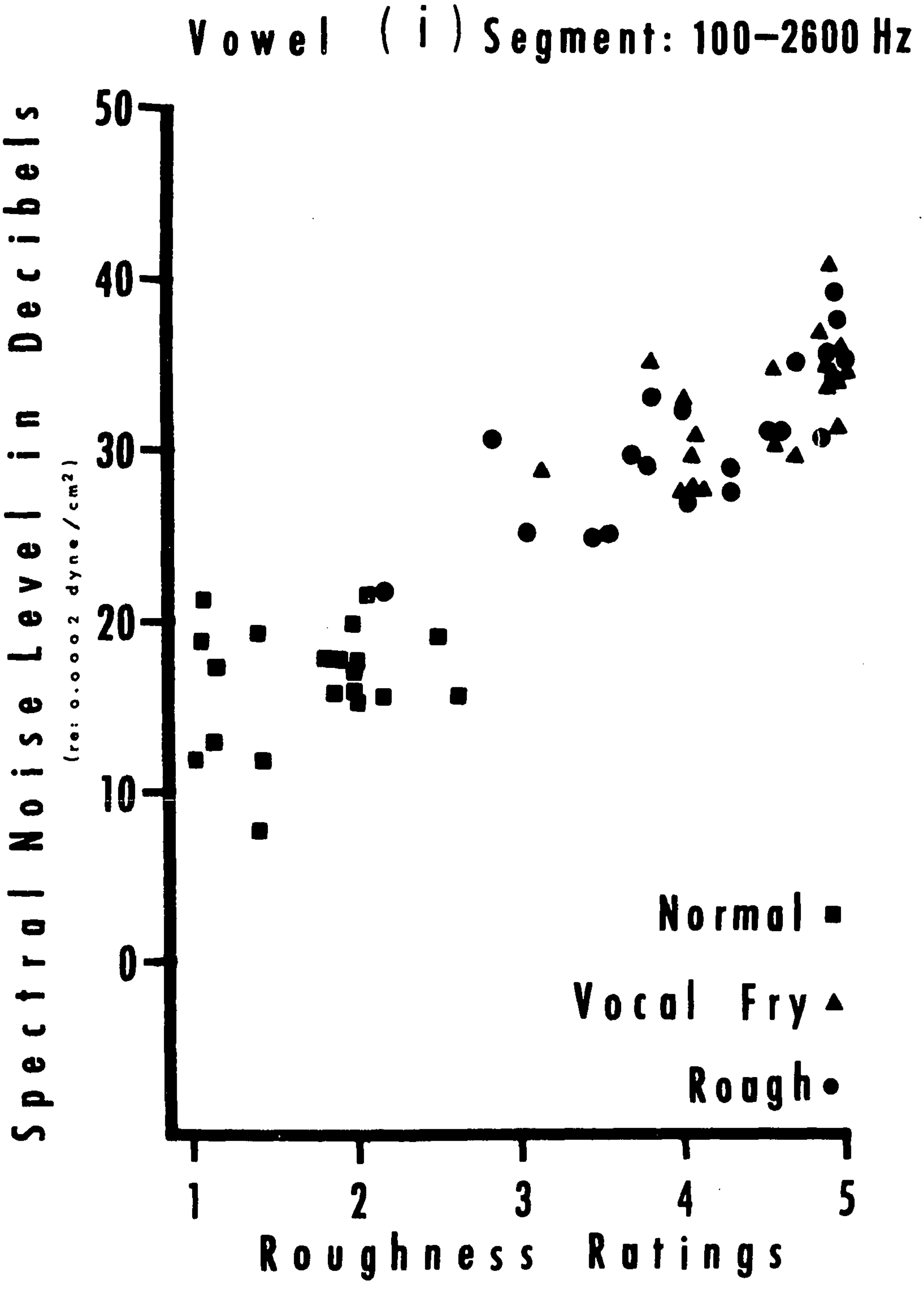

Figure 9.—Spectral range 100 to 2600 Hz noise level means and median roughness ratings over eleven judges for twenty male subjects' normal, vocal fry, and simulated abnormally rough productions of the vowel /i/.

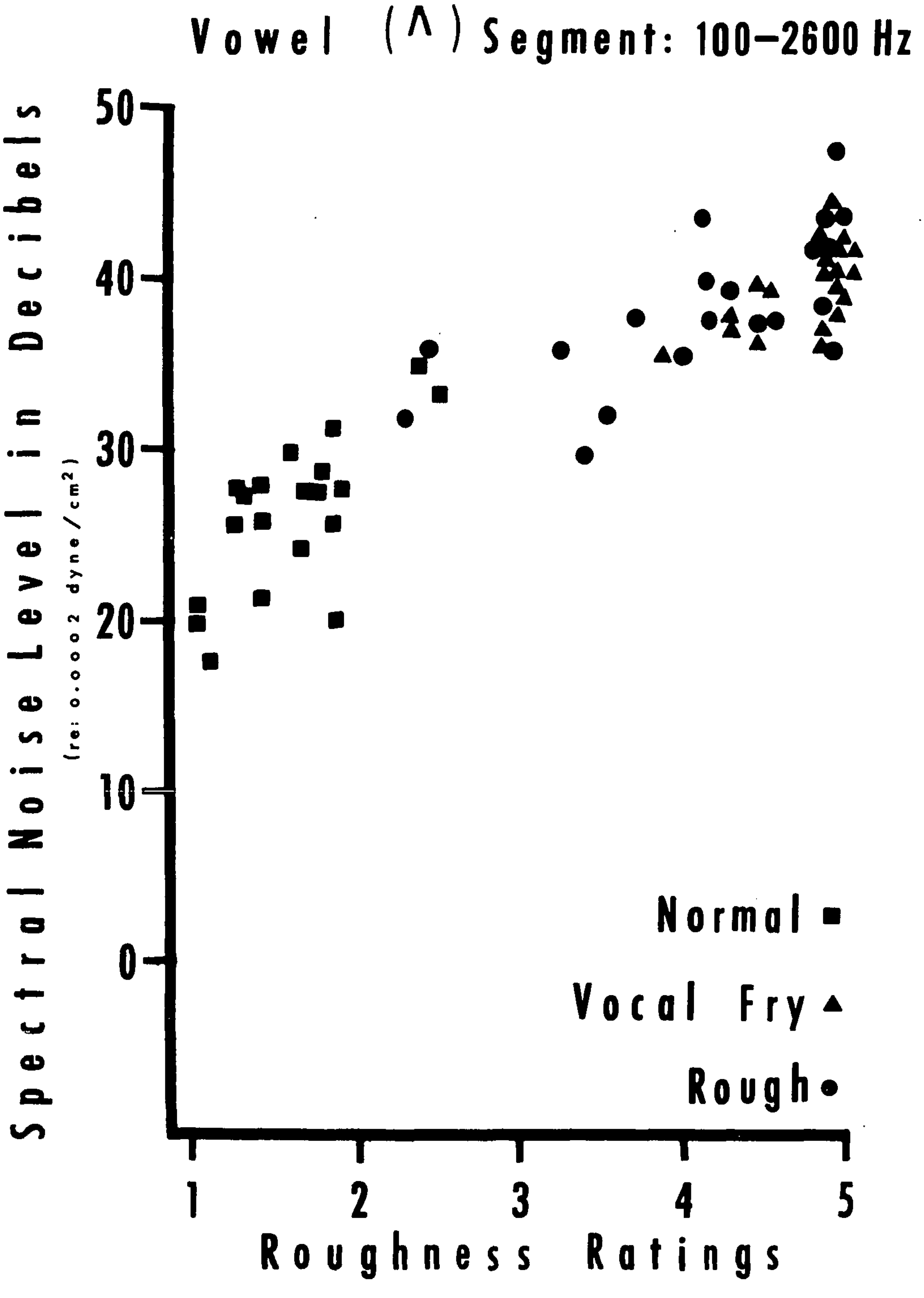

Figure 10.—Spectral range 100 to 2600 Hz noise level means and median roughness ratings over eleven judges for twenty male subjects' normal, vocal fry, and simulated abnormally rough productions of the vowel /ʌ/.

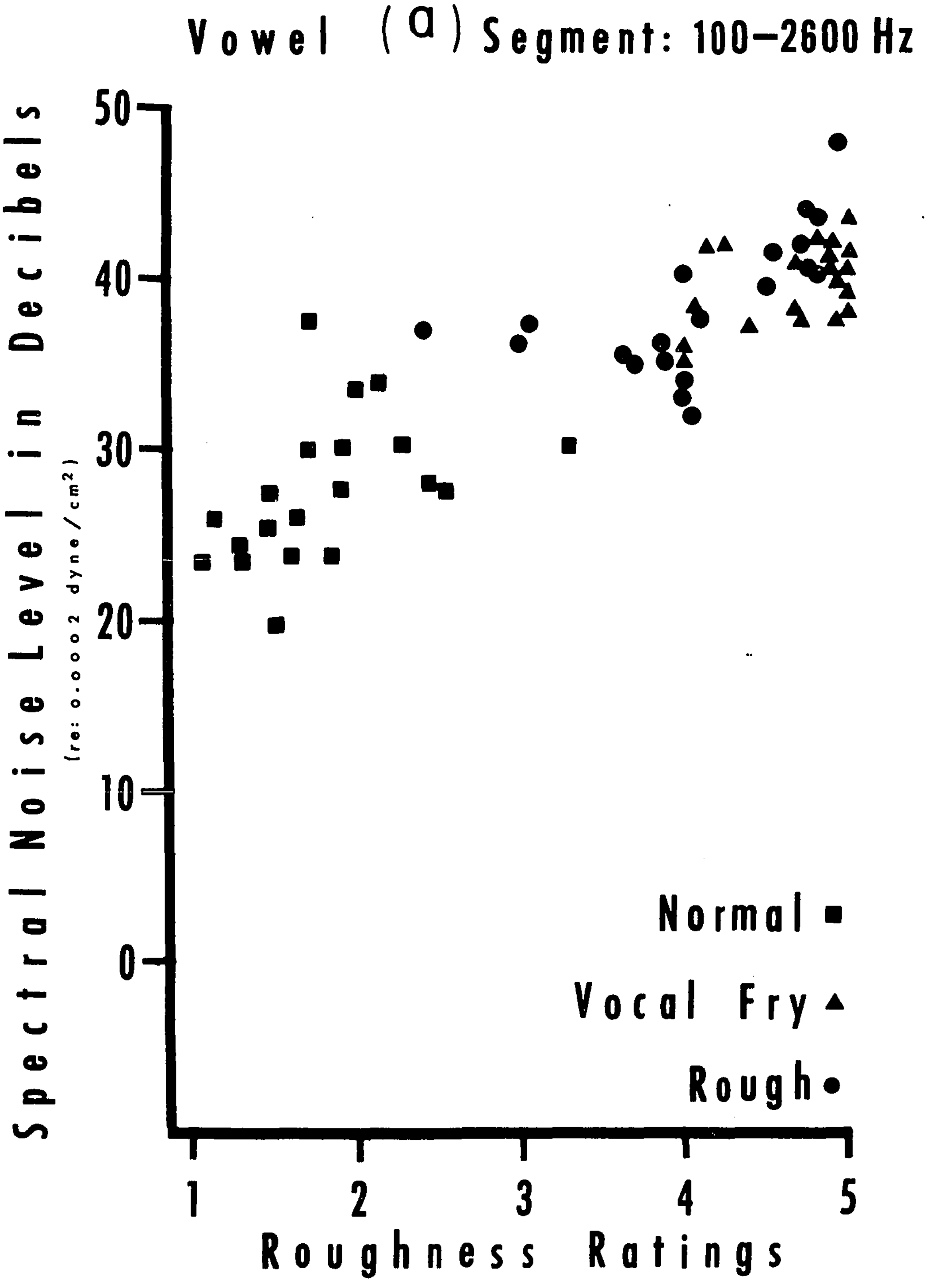

Figure 11.--Spectral range 100 to 2600 Hz noise level means and median roughness ratings over eleven judges for twenty male subjects' normal, vocal fry, and simulated abnormally rough productions of the vowel /ɑ/.

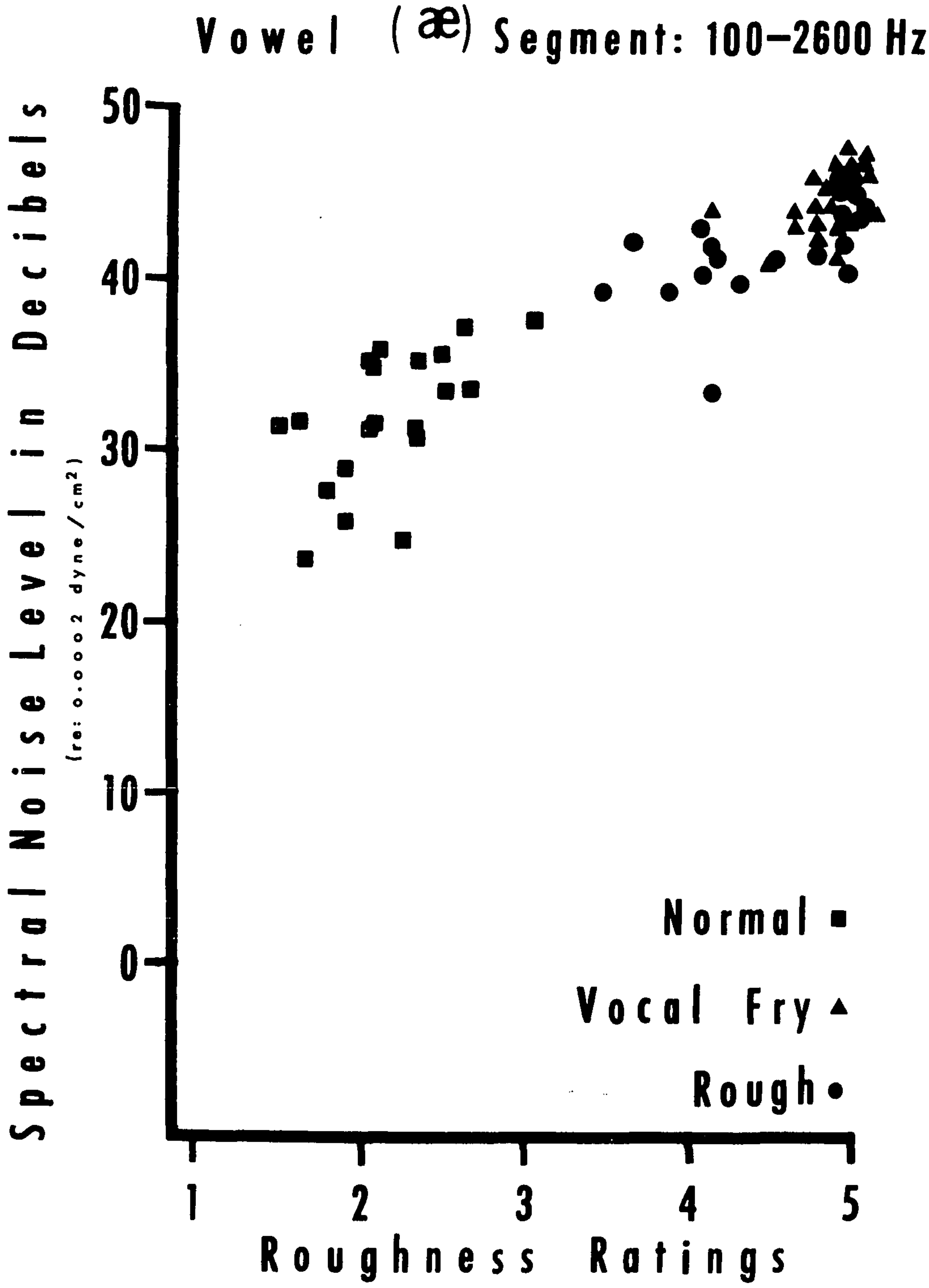

Figure 12.—Spectral range 100 to 2600 Hz noise level means and median roughness ratings over eleven judges for twenty male subjects' normal, vocal fry, and simulated abnormally rough productions of the vowel /æ/.

indicating the degree of association between the SNL means and the median roughness ratings for the productions of each test vowel. Table 6 reveals that the obtained coefficients ranged from .88 for /ɑ/ to .92 for /u/, and each was statistically significant ($p < 0.01$). Because the coefficients were uniformly high across vowels when SNLs for each production were averaged over the 100 to 2600 Hz range, an inspection of relationships between the SNLs in each 100-Hz segment of this range and median roughness rating was made.

Table 7 presents correlation coefficients obtained when a

TABLE 7

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

Vowel	Correlation Coefficients *
/u/	0.97
/i/	0.96
/ʌ/	0.97
/ɑ/	0.95
/æ/	0.95

* All coefficients significant ($p < 0.01$)

multiple regression analysis was performed relating the 100-Hz section SNLs to the median roughness ratings for the productions of each vowel. Table 7 shows that the obtained coefficients ranged from .95 for /ɑ/ and /æ/ to .97 for /ʌ/ and /u/. The coefficients for all five vowels were significant ($p < 0.01$), and tended to be higher than those obtained when the 100 to 2600 Hz SNL means and median roughness ratings for productions of each vowel were related. The magnitude of these coefficients indicates a high degree of linear relationship between 100-Hz section SNLs and the median roughness rating for each of the test vowels. Because the multiple correlation coefficients were uniformly high and significant for all test vowels, it appeared that the median roughness rating for each vowel production might be predicted from the spectral noise levels associated with it.

Three different regression analyses were employed to predict the roughness of the vowel productions. Table 8 presents, for each prediction method, the number of residuals (the obtained roughness rating minus the predicted rating) for each vowel which deviated more than .50 or 1.0 scale value. This table shows that the number of relatively large residuals tended to be greater for Methods II and III than for Method I. Thus, of the three methods employed, Method I appeared to offer advantages with respect to the accuracy of prediction of vowel median roughness ratings when all three modes of phonation were considered. In Method I, the 100-Hz spectral section SNLs from 100 to 2600 Hz and the median roughness ratings for normal, simulated abnormally rough, and vocal fry productions were used in fitting a regression plane for each test vowel. The derived multiple regression equation was then used to predict the median roughness ratings for all productions of each vowel. The Method I multiple regression equation has the form

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

where Y equals the roughness prediction, B_0 the Y intercept determined

TABLE 8

THE NUMBER OF RESIDUALS FOR THREE METHODS OF PREDICTION GREATER THAN .50 AND 1.0 SCALE VALUE FOR TWENTY NORMAL, TWENTY VOCAL FRY, AND TWENTY SIMULATED ABNORMALLY ROUGH PRODUCTIONS OF EACH OF FIVE VOWELS

		Method I		Method II		Method III	
Vowel	Phonatory Mode	Residual>.5 Scale Value	Residual>1 Scale Value	Residual>.5 Scale Value	Residual>1 Scale Value	Residual>.5 Scale Value	Residual>1 Scale Value
/u/	Normal	3	0	1	0	13	0
	Rough	1	0	3	0	8	0
	Fry	2	0	7	3	9	1
/i/	Normal	4	0	1	0	8	2
	Rough	4	0	0	0	11	1
	Fry	4	1	14	10	11	0
/ʌ/	Normal	6	0	2	0	14	4
	Rough	1	0	0	0	11	3
	Fry	4	0	11	3	4	2
/ɑ/	Normal	3	1	4	0	10	3
	Rough	6	0	1	0	6	2
	Fry	3	0	17	16	8	1
/æ/	Normal	6	0	4	0	11	2
	Rough	2	1	4	0	5	1
	Fry	0	0	10	5	1	0

by the regression analysis, B_{1-25} the regression coefficients determined by the regression analysis, and X_{1-25} the successive 100-Hz SNLs from 100 to 2600 Hz for each vowel production. Table 8 reveals that when Method I was employed the predicted median rating for forty-nine of the 300 vowel productions deviated more than .50 scale value from the obtained median rating for those productions. Only three of the forty-nine residuals greater than .50 scale value were larger than 1.0 scale value. The largest residual associated with Method I was 1.36.

In Method II, only the normal and abnormally rough vowel 100-Hz spectral section SNLs from 100 to 2600 Hz and the median roughness ratings were utilized in fitting the regression plane. The derived regression equation was then used to predict the median roughness ratings for the vocal fry as well as for the normal and abnormally rough vowels. Table 8 shows that the number of relatively large residuals for fry productions tended to be greater for Method II than for Methods I or III. This finding suggests that the linear regression for vocal fry vowel productions may differ slightly from that for normal and rough vowel productions.

In Method III, the SNLs averaged over the spectral segment 100 to 2600 Hz and the roughness ratings for normal, vocal fry, and rough productions of each vowel were utilized in fitting a regression line. The derived regression equation was then used to predict the median roughness ratings for all productions of each vowel. When all three modes of phonation are considered, the number of relatively large residuals obtained using Method III tended to exceed that obtained using either Methods I or II. This was thought to be attributable to the fact that the 100 to 2600 Hz SNL mean was a less precise indicator of the spectral noise associated with each vowel production than the individual measures of noise in successive 100-Hz spectral sections. The regression equations and other details of Methods II and III are presented in APPENDIX F.

Discussion

To investigate the spectral noise levels and perceived roughness of normal, vocal fry, and simulated abnormally rough productions of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/, subjects individually produced each of the test vowels at one intensity utilizing each of the three modes of phonation. The appropriateness of the mode of phonation employed for each production was initially evaluated by the investigator, but his percepts were subsequently compared to those of eleven judges. In every instance, a majority of judges agreed with the investigator's initial opinion regarding the mode of phonation represented by a test phonation. These findings appear to support the assumption that the investigator's categorization of the test vowel phonations was valid, at least insofar as the mode of phonation employed in vowel production may be delineated perceptually. It may also be noted that the judges successfully differentiated among the vowel samples with regard to the mode of phonation each represented as they listened to randomized recordings of the samples. This suggests that normal, vocal fry, and abnormally rough phonations can be differentiated perceptually on the basis of audible cues alone.

As a further procedure, the fundamental frequency of each test phonation was evaluated. This analysis revealed that the fundamentals for vocal fry phonations were uniformly much lower than those for the normal or the simulated abnormally rough phonations. Moreover, the range of fundamentals for fry samples did not overlap the range for normal or for rough samples. The fundamentals for the normal and rough productions were generally similar, although a trend toward slightly higher fundamentals for rough than for normal productions was observed. These findings are consistent with those of others (28, 50) who have reported that the fundamental frequency of fry phonations is generally lower than that for normal and harsh phonations; that the range of fundamentals for fry phonations does not overlap the range for normal or harsh phonations; and, that the fundamentals of harsh phonations tend to be within the subject's modal frequency range. Further, the observed differences in the range of fundamentals for normal and for fry phonations appears consistent with an hypothesis that fry phonations represent a distinct vocal register.

This study indicated, in general, that vocal fry and simulated abnormally rough vowel productions tend to receive similar median roughness ratings and that both tend to exceed normal productions in perceived roughness. Moreover, the range of median ratings obtained for the vocal fry productions tended to be smaller than the ranges for either the normal or the rough productions. That is, the ratings for fry productions tended to be restricted toward the upper or "more severely rough" end of the rating scale. These findings may help to explain why some investigators have reportedly (54) included vocal fry samples in their studies of ostensibly harsh voices, and why others (20, 54) have tended to regard vocal fry as an abnormal mode of phonation. It may be noted, however, that while the present findings are consistent with previous observations indicating that substantial harshness or roughness is perceptually associated with fry phonations (83, 54), they do not suggest that abnormally rough vocal quality and vocal fry quality are perceptually equivalent. On the contrary, the observation that vocal fry and simulated abnormally rough non-fry productions receiving similar roughness ratings could be successfully differentiated by the judges for this investigation suggests that vocal fry quality per se differs in a perceptually significant way from abnormally rough vocal quality. These findings may have implications with regard to the design of future investigations of rough vocal quality.

The present findings regarding the roughness of fry samples also appear consistent with previous findings suggesting that the perception of vocal roughness is related to the magnitude of frequency perturbations in the acoustic wave (8, 44, 53), and that vocal fry samples tend to evidence large pitch perturbations relative to normal or abnormally rough productions (49). These observations concerning the perturbation factor suggest that fry samples should tend to evidence considerable roughness when evaluated perceptually. Moreover, on the basis of his analog studies, Wendahl (88, 89) has observed that given high- and low-frequency signals evidencing equal frequency perturbation, greater roughness tends to be perceptually associated with the signal which is relatively low in frequency. Possibly, therefore, the extreme roughness perceptually associated with fry phonations may be in part attributable to their low fundamentals.

The present findings regarding the relative roughness of the individual test vowels indicated that, regardless of the mode of phonation, a trend was evident for least roughness to be associated with a high vowel and most roughness with the low vowel /æ/. Similar findings have been reported previously by Sansone (65) for vowels produced normally and with simulated abnormal vocal roughness by adult males, by Lively (45) for vowels produced normally and with simulated vocal roughness by adult females, and by Hanson (25), Sherman and Linke (68), and Rees (64) for vowels produced by subjects presenting clinically significant vocal hoarseness or harshness. This trend toward greater roughness for low than for high vowels suggests that the degree of roughness associated with vowel phonations tends to be related to vocal tract configuration during phonation. Possibly, the configuration of the supraglottic resonators during high vowel production is such that the acoustic components which contribute to the perception of vocal roughness tend to be more highly damped than they are during low vowel production. Another possibility is that greater laryngeal periodicity is associated with the production of high than low vowels. These hypotheses may be tested in future investigations. The present findings suggest, however, that physiologic conditions which effect differences in the relative roughness of normal and abnormally rough low and high vowel phonations may also obtain for low and high vocal fry vowel phonations.

The findings regarding the spectral noise levels associated with the test phonations indicated that in low-frequency spectral ranges, the noise levels for vocal fry phonations tended to exceed slightly those for abnormally rough phonations, but the reverse was true in higher spectral frequency ranges. Further investigation will be needed to ascertain whether such trends are reliable. In general, however, the level of spectral noise associated with vocal fry phonations was similar in magnitude to that for the simulated abnormally rough phonations, and both vocal fry and abnormally rough productions tended to evidence spectral noise levels which were considerably elevated with respect to those for normal productions. Moreover, the SNLs for vocal fry phonations in this study tended to be larger than those reported by Hanson (25) for clinically hoarse phonations. These findings are of interest in view of the speculation by Michel (49) that considerably more spectral noise should be expected for harsh than for vocal fry phonations. This hypothesis was apparently predicated upon his observation that there was a greater percentage of "aperiodic" phonation associated with harsh than with vocal fry phonations. The present findings do not appear to support Michel's hypothesis. It may be that in formulating his hypothesis Michel underestimated the influence on spectral noise levels of acoustic wave perturbations which are not readily apparent upon visual inspection of the wave. Lively and Emanuel (46) have offered an alternative hypothesis which seems to fit the present findings better. These investigators suggested that the level of spectral noise associated with vowel phonations may be related directly to acoustic wave deviations from perfect periodicity, i.e., to wave perturbations. This hypothesis implies that relatively high spectral noise levels should be expected for phonations which evidence relatively large variations in acoustic wave periodicity over time, even when cycles are discernable in the wave. It may be noted in this regard that Michel's investigation (49) suggested that fry phonations are characterized by frequency perturbations which are larger than those characterizing harsh phonations. Thus, the present findings regarding spectral noise levels for fry phonations and those of Michel regarding the wave characteristics of fry phonations appear consistent with the hypothesis suggested by Lively and Emanuel.

While the level of harmonic components in the spectra of the test phonations for this investigation were not measured, certain trends were noted. For example, the low-frequency harmonics for fry and abnormally rough phonations tended to be diminished with respect to those for normal phonations. Similar findings reported previously by others (45, 65) for normal and simulated abnormally rough productions have been interpreted to suggest that there may be a trading relationship trend between the level of spectral noise and the level of low-frequency harmonics for vowel phonations produced at one intensity. On the basis of the present findings, it appears that a similar relationship may obtain for vocal fry phonations.

Inspection of the spectra obtained for the test phonations raised questions, however, about the level of the higher harmonics for normal, vocal fry, and rough productions. It was observed that while higher vowel harmonics were not clearly discernable because of elevated spectral noise, spectral levels often appeared to be elevated at higher harmonic frequencies. Moreover, these elevations tended to be greater for fry and for abnormally rough than for normal phonations. Similar trends were observed earlier by Emanuel and Whitehead (17) for vowels phonated normally and with simulated abnormal vocal roughness. These investigators suggested that the apparent level of higher vowel harmonics may reflect the noise level at harmonic frequencies more than the signal level at those frequencies. Thus, the apparent elevation of higher vowel harmonics with increasing roughness may be attributable to the elevation of noise in higher spectral frequency ranges.

Possible relationships between spectral noise levels and the degree of roughness associated with the normal, simulated abnormally rough, and vocal fry phonations were also of interest. Previous investigations, similar in design to the present study, have considered such relationships for vowels produced normally and with simulated abnormal vocal roughness by adult males (65) and by adult females (45). This study of males largely replicated one of the cited previous studies, but also provided new information regarding relationships between spectral noise levels and roughness ratings for vocal fry productions. High and significant ($p < 0.01$) correlation coefficients were obtained for each of the five test vowels when spectral noise levels averaged over the frequency range 100 to 2600 Hz and median roughness ratings for each vowel production were related. The magnitude of these coefficients, ranging from .88 for /ɑ/ to .92 for /u/, was similar to that reported earlier by Sansone (65) when only normal and simulated abnormally rough productions were evaluated. With the inclusion of vocal fry samples, therefore, the linear relationship trend between 100 to 2600 Hz vowel spectral noise level means and median roughness ratings remained strong. Thus, the present findings support an hypothesis that spectral noise levels tend to be linearly related to roughness ratings for vocal fry as well as for normal and abnormally rough vowel productions.

As a further procedure, relationships between vowel roughness ratings and spectral noise levels obtained in each 100-Hz segment of each vowel spectrum were investigated using a multiple regression procedure. Very high and significant ($p < 0.01$) correlation coefficients were obtained for each test vowel when this procedure was employed. Thus, using the appropriate linear regression equation, it was possible to predict the roughness ratings for normal, rough, and vocal fry productions of each vowel from the 100-Hz SNLs for each production. In general, small residuals (observed rating minus the predicted rating) were obtained for each vowel production using this procedure (Method I). This finding is similar to that reported by Sansone (65) for normal and simulated abnormally rough vowels produced by adult males. Thus, it appears that spectral noise levels and roughness ratings for vocal fry phonations tend to be related, much as noise levels and roughness ratings for abnormally rough and normal phonations are related.

Somewhat larger residuals were obtained, for fry productions in particular, however, when a regression equation predicated upon the 100-Hz SNL-roughness rating relationships for only the normal and the simulated abnormally rough vowel productions was utilized in predicting the roughness of normal, simulated rough, and fry phonations (Method II). This suggests that the linear regression for vocal fry vowels may differ slightly from that for normal and rough vowels. Inspection of the residuals revealed that, in general, obtained median roughness ratings for fry samples tended to be greater than the ratings predicted for those samples by Method II. It may be, therefore, that the relatively large residuals for the vocal fry productions associated with Method II are attributable to the effect of relatively low fry fundamental frequencies on the perceived roughness of those productions. The relatively large residuals for all vowel samples obtained using a regression equation predicated upon the relationship of SNL means to vowel roughness ratings (Method III) appeared to be due to the loss of precision in estimating vowel SNLs associated with the use of 100 to 2600 Hz SNL means in the prediction equation.

The present findings may be considered with respect to some possible implications. They suggest that vocal fry phonations are in some respects like abnormally rough phonations. Specifically, it appears that the roughness associated with fry phonations is similar in severity to that associated with simulated abnormally rough phonations. Moreover, similar acoustic spectral features, i.e., elevated spectral noise and diminished low-frequency harmonic levels appear to be associated with the perception of the severe roughness for fry and for abnormally rough phonations. Possibly, therefore, similar acoustic wave features underlie the roughness associated with fry and abnormally rough phonations. It may be that a deviation from periodicity is the wave feature most closely associated with perceived vowel roughness, regardless of the mode of vowel phonation.

Vocal fry phonations apparently can be differentiated perceptually from abnormally rough phonations, however, and it seems reasonable to speculate that the acoustic wave features which cue such differentiation are other than those underlying the elevated spectral noise levels for fry phonations. Possibly, as others (9, 29, 49) have suggested, the perception of vocal fry phonation *per se* is predicated upon such factors as a low fundamental frequency and a rapid decay of acoustic intensity in successive cycles of phonation.

With respect to their clinical implications, the following observations regarding the findings seem pertinent. They do not support a conclusion that vocal fry phonation *per se* is abnormal, even though such phonation tends to be quite rough. It would seem that clinical decisions regarding the normality or abnormality of fry phonation ought to be decided on the basis of such factors as: whether or not the fry phonation is maladaptive; whether or not it impairs in some socially significant way the subject's ability to communicate; and, whether or not it is associated with some underlying pathology, disease process, or abusive behavior. It appears that individuals presenting vocal fry phonation which would be regarded as abnormal on the basis of such criteria are rare if, indeed, there are any. On the other hand, available information suggests that vocal fry phonations are generally not indicative of vocal abnormally. This study suggested, for example, that fry phonation is rather easily produced by normal-speaking subjects without noticeable strain or excessive effort. Moreover, fry phonation is reportedly (11, 29) common in connected speech judged to be within normal limits in quality. Thus, it appears that fry phonation is usually normal. The finding that vocal fry and normal or modal register phonations can be differentiated perceptually appears to support an hypothesis that fry phonations constitute a perceptually distinct, low-frequency, vocal register.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this study was to investigate spectral noise levels (SNLs) in narrow-band (3-Hz) acoustic spectra of normal, vocal fry, and simulated abnormally rough vowels produced by adult male subjects, and possible relationships between the SNLs and perceived vowel roughness. Other studies (45, 65) have considered relationships between SNLs and perceived roughness for vowels phonated normally and with simulated abnormal vocal roughness. A unique feature of the present study, however, was investigation of such relationships for vocal fry vowel phonations as well as for normal and simulated abnormally rough phonations.

Twenty normal-speaking adult males served as subjects for this investigation. The subjects individually produced each of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/ first normally, then with vocal fry, and then with simulated abnormal vocal roughness. Each production was sustained for seven seconds at 75 dB (± 1 dB) SPL at a mouth-to-microphone distance of six inches. The vowel productions were recorded on magnetic tape and were individually presented in random order to eleven judges for rating. Each judge first rated the vowels for roughness on a five-point equal-appearing intervals scale in which "1" represented least severe and "5" represented most severe roughness. The median of the eleven judges' ratings was obtained as an index of each vowel production's roughness. In a second listening session, the judges independently evaluated the randomized vowel samples to identify the mode of phonation each represented.

The recording of each vowel production was also analyzed to produce a narrow-band (3-Hz) frequency-by-amplitude spectrum of its acoustic components from 0 to 8000 Hz. This analysis was made from two-second tape loops constructed from a central portion of each recording evidencing a uniform intensity (75 dB SPL $\pm$ 1 dB). As a quantitative index of vowel spectral noise levels, the lowest observable peak of energy in each of seventy-nine successive 100-Hz sections from 100 to 8000 Hz was measured in dB SPL in each vowel spectrum. The vowel tape loops were also individually analyzed to determine the fundamental vocal frequency of the normal, vocal fry, and rough phonations.

The findings revealed that the eleven judges were generally in unanimous or nearly unanimous agreement regarding the mode of phonation represented by each vowel sample. In all instances, the majority of judges agreed with the investigator's initial opinion regarding the mode of phonation represented by each test phonation. The fundamental vocal frequency analysis revealed that the vocal fry productions evidenced uniformly lower fundamentals than either normal or simulated abnormally rough productions, and that the range of fundamentals for fry samples did not overlap the range for normal or for rough samples. It was also found that the fundamentals for high and low vowels were at the extremes of the fundamental frequency range for each mode of phonation and tended to differ significantly ($p < 0.05$).

The findings regarding the median roughness ratings obtained for each test vowel production indicated that each vocal fry and simulated rough production was judged more rough than its normal counterpart. When the means of the median ratings for vocal fry and simulated abnormally rough productions of each test vowel were compared, it was noted that they tended to be similar in magnitude and relatively high on the roughness scale. For all three modes of phonation, the low vowel /æ/ tended to evidence greater roughness than the other test vowels, and least roughness tended to be associated with a high vowel.

With respect to vowel spectral noise levels, this study indicated that normal, vocal fry, and abnormally rough vowel productions tend to evidence noise levels above system noise levels over the spectral range 0 to 8000 Hz. For each test vowel, spectral noise levels tended to be higher for vocal fry and for rough phonations than for normal phonations. The spectral noise levels for the vocal fry and the rough productions of each test vowel were similar in magnitude. Identifiable low-frequency harmonics for vocal fry and rough productions of each vowel tended to be somewhat diminished in amplitude with respect to those for normal productions. For all three phonatory modes, spectral noise levels tended to be relatively high in vowel formant ranges and relatively low in interformant ranges, and an increase in mean spectral noise levels for the spectral range 100 to 2600 Hz appeared to be associated with changes in vocal tract configuration relating to decreasing tongue height in vowel production. That is, for the normal, vocal fry, and rough productions, high vowels were generally characterized by lower spectral noise levels than low vowels. For each test vowel, the greatest SNL variability for the spectral segment 100 to 2600 Hz was associated with the rough productions.

Spectral noise levels averaged over the spectral segment 100 to 2600 Hz were highly, positively ($\geq .88$), and significantly ($p < 0.01$) related to the median roughness ratings for the productions of each test vowel. Thus, as mean spectral noise levels for the test vowel productions increased, the median roughness ratings of the productions also tended to increase. High ($\geq .95$) and significant ($p < 0.01$) multiple correlation coefficients were obtained when the noise level in each 100-Hz spectral section from 100 to 2600 Hz was related to the median roughness rating for the productions of each vowel. Generally, small residuals (observed rating minus the predicted rating) were obtained when a multiple linear regression equation predicated upon the relationship of the 100-Hz section SNLs in the spectral frequency range 100 to 2600 Hz and the observed median roughness ratings for normal, simulated abnormally rough, and vocal fry phonations was used to predict each vowel production's median roughness rating. A greater number of relatively large residuals was obtained, particularly for fry samples, when a regression equation predicated upon the 100-Hz section SNLs and the median roughness ratings for the normal and abnormally rough vowels was used to predict the roughness ratings for the vocal fry, normal, and rough vowel productions. The residuals for all vowel samples were also relatively large when the roughness of each sample was predicted using a regression equation predicated upon the relationship of SNLs averaged over the spectral range 100 to 2600 Hz to vowel roughness ratings.

Some general conclusions and implications of the present study include the following. Vocal fry and abnormally rough phonations appear similar with respect to perceived roughness and with respect to specific acoustic features associated with their roughness, i.e., elevated spectral noise and diminished low-frequency harmonic levels. It is suggested, therefore, that similar acoustic wave features, i.e., deviations from periodicity or perturbations, may underlie the roughness associated with fry and with abnormally rough phonations. Because vocal fry phonations can be reliably differentiated perceptually from abnormally rough phonations, however, it appears that the acoustic wave features which cue such differentiation are other than those underlying the elevated spectral noise levels. As others (9, 29) have suggested, the perception of vocal fry phonation *per se* may be predicated upon such factors as a low fundamental vocal frequency and a rapid decay of acoustic intensity in successive cycles of phonation.

The present study suggests that normal-speaking subjects can produce fry phonations at will and without apparent discomfort or excessive tension. Moreover, it has been observed (29) that fry phonation occurs commonly in speech judged to be normal. These observations, and a lack of clinical evidence to the contrary, support the view that vocal fry phonation *per se* does not constitute a voice disorder, i.e., it is usually normal. On the other hand, the finding that vocal fry vowel phonations can be differentiated perceptually from modal register vowel phonations appears to support the hypothesis of Hollien *et al.* (29) that fry phonations constitute a perceptually distinct, normal, low-frequency vocal register.

BIBLIOGRAPHY

1. Arnold, G.E. Vocal rehabilitation of paralytic dysphonia: II. Acoustic analysis of vocal function. Arch. Otolaryng., 62 593-601 (1955).

2. Berry, M. and Eisenson, J. The Defective in Speech. New York: Appleton-Century-Crofts, Inc. (1942).

3. Berry, M. and Eisenson, J. Speech Disorders: Principles and Practices of Therapy. New York: Appleton-Century-Crofts, Inc. (1956).

4. Bowler, N.W. A fundamental frequency analysis of harsh vocal quality. Ph.D. Dissertation, Stanford University (1957).

5. ___________ A fundamental frequency analysis of harsh vocal quality. Speech Monogr., 31, 128-134 (1964).

6. Brubaker, R.S. and Dolpheide, W.R. Consonant and vowel influence upon judged voice quality of syllables. Jour. Acous. Soc. Amer., 27, 1000-1002 (1955).

7. Cleeland, C.E. Definitions of a voice quality called X: An attempt to isolate a voice quality and define its relations to other voice qualities and basal metabolic rate. Ph.D. Dissertation, University of Denver (1948).

8. Coleman, R. Some acoustic correlates of hoarseness. Master's Thesis, Vanderbilt University (1960).

9. ___________ Decay characteristics of vocal fry. Folia Phoniatr., 15, 256-263 (1963).

10. ___________ Effect of median frequency levels upon the roughness of jittered stimuli. J. Speech Hearing Res., 12, 330-336 (1969).

11. Coleman, R. and Wendahl, R. Vocal roughness and stimulus duration. Speech Monogr., 34, 85-92 (1967).

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

13. Curtis, J.F. Disorders of voice. In. W. Johnson et al. (Eds.) Speech Handicapped School Children (Revised Ed.). New York: Harper and Brothers (1956).

14. ____________ Acoustics of speech production and nasalization. In D.C. Spriestersbach and D. Sherman (Eds.) Cleft Palate and Communication. New York: Academic Press (1968).

15. Denes, P.B. and Pinson, E.N. The Speech Chain: The Physics and Biology of Spoken Language. Bell Telephone Laboratories, Inc., Baltimore, Maryland: Waverly Press, Inc. (1963).

16. Emanuel, F.W. and Sansone, F.E. Some spectral features of "normal" and simulated "rough" vowels. Folia Phoniatr., 21, 401-415 (1969).

17. Emanuel, F.W. and Whitehead, R.L. Some relationships between vocal spectral harmonic levels and vocal roughness. J. Speech Hearing Res. (submitted September, 1970).

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

19. ____________ Quoted in H. Hollien and R. Wendahl, Perceptual Study of vocal fry. Jour. Acous. Sco. Amer., 43, 506-509 (1968).

20. Fisher, H.B. Improving Voice and Articulation. Boston: Houghton Mifflin Co. (1966).

21. Fisher, H.B. and Logemann, J.A. Objective evaluation of therapy for vocal nodules: a case report. J. Speech Hearing Dis., 35, 277-285 (1970).

22. Flanagan, J.L. Some properties of the glottal sound source. J. Speech Hearing Res., 1, 99-116 (1968).

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

24. Halle, M. Hughes, G.W. and Radley, J. P. Acoustic properties of consonants. Jour. Acous. Soc. Amer., 29, 107-116 (1957).

25. Hanson, W. Vowel spectral noise levels and roughness severity ratings for vowels and sentences produced by adult males presenting abnormally rough voice. Ph.D. Dissertation, University of Oklahoma (1970).

26. Hanson, W. and Emanuel, F.W. Some spectral noise and vocal roughness relationships for adult males presenting laryngeal pathology. J. Speech Hearing Res. (submitted January, 1970).

27. Hollien, H., Damste, H. and Murry, T. Vocal fold length during vocal fry phonation. Folia Phoniatr., 21, 257-265 (1969).

28. Hollien, H. and Michel, J.F. Vocal fry as a phonational register. J. Speech Hearing Res., 11, 600-604 (1968).

29. Hollien, H., Moore, P., Wendahl, R.W., and Michel, J.F. On the nature of vocal fry. J. Speech Hearing Res., 9, 245-247 (1966).

30. Hollien, H. and Wendahl, R.W. Perceptual study of vocal fry. Jour. Acous. Soc. Amer., 43, 506-509 (1968).

31. Hollinger, P. and Johnson, K.C. Benign tumors of the larynx. AMA Archives of Otolaryngology, 60, 496-509 (1952).

32. House, A.S. and Fairbanks, G. The influence of consonant environment upon the secondary acoustical characteristics of vowels. Jour. Acous. Soc. Amer., 25, 105-113 (1953).

33. Irwin, R. Speech and Hearing Therapy. New York: Prentice-Hall, Inc. (1953).

34. Isshiki, N. Regulatory mechanism of voice intensity variation. J. Speech Hearing Res., 7, 17-29 (1964).

35. Isshiki, N. and von Leden, H. Hoarseness: aerodynamic studies. Arch. Otolaryng., 80, 206-213 (1964).

36. Isshiki, N., Yanagihara, N. and Morimoto, M. Approach to the objective diagnosis of hoarseness. Folia Phoniatr., 18, 393-400 (1966).

37. Iwata, S. and von Leden, H. Voice prints in laryngeal disease. Arch. Otolaryng., 91, 346-351 (1970).

38. Johnson, W., Darley, F.L. and Spriestersbach, D.C. Diagnostic Methods in Speech Pathology. New York: Harper and Row (1963).

39. Joos, M. Acoustic phonetics. Lang. Suppl., 24, 1-65 (1948).

40. Koike, Y. Vowel amplitude modulations in patients with laryngeal diseases. Jour. Acous. Soc. Amer., 45, 839-844 (1968).

41. Ladefoged, P. Elements of Acoustic Phonetics. Chicago: University of Chicago Press (1966).

42. Laguaite, J. and Waldrop, W. Acoustic analysis of fundamental frequency of voices before and after therapy. Folia Phoniatr., 16, 183-192 (1964).

43. Lieberman, P. Intonation, Perception, and Language. Cambridge, Mass.: M.I.T. Press (1967).

44. Lieberman, P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. Jour. Acous. Soc. Amer., 35, 344-355 (1963).

45. Lively, M.A. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult females. Ph.D. Dissertation, University of Oklahoma (1969).

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

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

48. McGlone, R.E. Air flow during vocal fry phonation. J. Speech Hearing Res., 10, 299-304 (1967).

49. Michel, J.F. Vocal fry and harshness. Ph.D. Dissertation, University of Florida (1964).

50. ___________ Fundamental frequency investigation of vocal fry and harshness. J. Speech Hearing Res., 11, 590-594 (1968).

51. Michel, J.F. and Hollien, H. Perceptual differentiation of vocal fry and harshness. J. Speech Hearing Res., 11, 439-443 (1968).

52. Moore, P. Voice disorders associated with organic abnormalities. In L. Travis (Ed.) Handbook of Speech Pathology. New York: Appleton-Century-Crofts, Inc. (1957).

53. Moore, P. and Thompson, C.L. Comments on the physiology of hoarseness. Arch. Otolaryng., 81, 97-102 (1965).

54. Moore, P. and von Leden, H. Dynamic variation of the vibratory pattern in the normal larynx. Folia Phoniatr., 10, 205-238 (1958).

55. Moser, H. Symposium of unique cases of voice disorders: A presentation of a case. J. Speech Hearing Dis., 7, 173-174 (1942).

56. Murry, T. Subglottal pressure measures during vocal fry phonation. Communication Sciences Laboratory Quarterly Report, University of Florida, 7, 3-4 (1969).

57. Nessel, E. Über das tonfrequenzspektrum der pathologisch veränderten stimme. Acta Otolaryng. Suppl., 157, 3-45 (1960).

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

59. Palmer, J.M. Hoarseness in laryngeal pathology: review of literature. Laryngoscope, 66, 500-516 (1956).

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

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

62. Proceedings of the Workshop on Nomenclature of Communication Disorders. PHS Grant No. B-3676, Bethesda, Maryland (1962).

63. Rees, M. Harshness and glottal attack. J. Speech Hearing Res., 1, 344-349 (1958).

64. ________ Some variables affecting perceived harshness. J. Speech Hearing Res., 1, 155-168 (1958).

65. Sansone, F.E. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult males. Ph.D. Dissertation, University of Oklahoma (1969).

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

67. Sherman, D. The merits of backward playing of connected speech in the scaling of voice quality disorders. J. Speech Hearing Dis., 19, 312-321 (1954).

68. Sherman, D. and Linke, E. The influence of certain vowel types on degree of harsh voice quality. J. Speech and Hearing Dis., 17, 401-408 (1952).

69. Shipp, T. and Huntington, D. Some acoustic and perceptual factors in acute-laryngitic hoarseness. J. Speech Hearing Dis., 30, 350-359 (1965).

70. Smith, W.R. and Lieberman, P. Studies in pathological speech production. Project 5628, Data Sciences Laboratory Air Force Cambridge Research Laboratories, Office of Aerospace Research, United States Air Force, L.G. Hanson Field, Massachusetts (1964).

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

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

73. Stevens, K.N. and House, A.S. Development of a quantitative description of vowel articulation. J. Acous. Soc. Amer., 27, 484-493 (1955).

74. Third Regional Workshop on the Rehabilitation Codes and Communicative Disorders. PHS Grant No. B-3676, The National Institute of Neurological Diseases and Blindness, Communicative Disorders Research Training Committee (1967).

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

76. Timcke, R., von Leden, H. and Moore, P. Laryngeal vibrations: measurements of the glottic wave, Part I. The normal vibratory cycle. Arch. Otolaryng., 68, 1-19 (1958).

77. ______ Laryngeal vibrations: measurements of the glottic wave, Part II. Physiologic variations. Arch. Otolaryng., 69, 438-444 (1959).

78. Van den Berg, J.W. Myoelastic aerodynamic theory of voice production. J. Speech Hearing Res., 1, 227-244 (1958).

79. ______ Subglottic pressures and vibrations of the vocal folds. Folia Phoniatr., 9, 65-71 (1957).

80. Van Dusen, R. A laboratory study of the metallic voice. J. Speech Dis., 6, 137-140 (1941).

81. Van Riper, C. Speech Correction: Principles and Methods. (2nd Ed.) New York: Prentice-Hall (1947).

82. ______ Speech Correction: Principles and Methods. (3rd Ed.) Englewood Cliffs, N.J.: Prentice-Hall, Inc. (1954).

83. Van Riper, C. and Irwin, J.V. Voice and Articulation. Englewood Cliffs, N.J.: Prentice-Hall, Inc. (1958).

84. Voelker, C. Phoniatry in dysphonia ventricularis. Annals of Otology, Rhinology, and Laryngology, 44, 471-473 (1935).

85. Von Leden, H. The clinical significance of hoarseness and related disorders. J. Lancet, 78, 50-53 (1958).

86. Von Leden, H. and Koike, Y. Detection of laryngeal disease by computer technique. Arch. Otolaryng., 91, 3-10 (1970).

87. Von Leden, H., Moore, P. and Timcke, R. Laryngeal vibrations: measurements of the glottic wave, Part III. The pathologic larynx. Arch. Otolaryng., 71, 26-45 (1960).

88. Wendahl, R. Laryngeal analog synthesis of harsh voice quality. Folia Phoniatr., 15, 241-250 (1963).

89. __________ Some parameters of auditory roughness. Folia Phoniatr., 18, 26-32 (1966).

90. __________ Laryngeal analog synthesis of jitter and shimmer. Folia Phoniatr., 18, 98-108 (1966).

91. Wendahl, R., Moore, P. and Hollien, H. Comments on vocal fry. Folia Phoniatr., 15, 251-255 (1963).

92. West. R., Kennedy, L. and Carr. A. The Rehabilitation of Speech. New York: Harper and Brothers (1947).

93. Yanagihara, N. Significance of harmonic changes and noise components in hoarseness. J. Speech Hearing Res., 10, 531-541 (1967).

94. ____________ Hoarseness: investigation of the physiological mechanisms. Ann. Oto. Rhino., Laryng., 76, 472-488 (1967).

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

96. Zinn, W. The significance of hoarseness. Ann. Oto. Rhino. Laryng., 54, 136-138 (1945).

APPENDIX A

Instructions to Subjects

Instructions to Subjects

In this experiment you will phonate five vowel sounds, at first normally, then with vocal fry, and then while simulating abnormal vocal roughness, into the microphone. The vowel sounds you are to produce are the underlined sounds in the words printed on the cards: /u/ as in boot, /i/ as in feet, /ʌ/ as in up, /ɑ/ as in father, and /æ/ as in hat. 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 black 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.

I will indicate to you if you are not producing the vowel printed on the card. Sometimes while simulating vocal roughness or phonating in vocal fry, 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: Vowel Ratings

You are asked to listen to 350, 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 productions 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 judgements 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. The vowel samples may vary according to parameters other than roughness; however, you are asked to ignore these variations. Restrict your attention to the degree of roughness perceived.

Prior to the rating of the vowels, the first fifteen vowel samples to be judged will be played to familiarize you with the range of vocal roughness. You are asked to listen to these samples and to form an independent estimate of the rating scale extremes.

The vowels to be judged will 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. As 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 judgements 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 judgement for every vowel sample. Leave no blank spaces. Are there any questions?

Instructions to Judges: Differentiation of Vowel Samples

You are asked to listen to 350, seven-second sustained vowel samples produced by adult males. The samples are comprised of the vowels /u/, /i/, /ʌ/, /ɑ/, and /æ/, and represent three modes of phonation: normal, vocal fry, and abnormally rough. The vowel samples will be presented to you one at a time and you are to evaluate each vowel independently and to identify the mode of phonation it represents.

To acquaint you with the three types of phonation to be differentiated, five vowel samples representing normal phonation, five vowel samples representing vocal fry phonation, and five vowel samples representing abnormally rough phonation will be played. You may listen to these productions as many times as you wish before the judging begins.

The vowels to be judged will 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 identify each vowel sample with respect to the mode of phonation it represents. On your response sheet, for every vowel sample, record either an "N" for normal phonation, an "F" for vocal fry phonation, or an "R" for abnormally rough phonation. Record the mode of phonation assigned to each vowel to the right of its number on your response sheet. Be sure to make your identification independently. You may hear each vowel production to be identified as many times as you wish. Notice that you will start at the top of the column and work down. For every vowel sample be sure to record a phonatory mode. Are there any questions?

APPENDIX C

Percentages of Inter- and Intra-Judge Roughness Rating Agreement

TABLE 9

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

	Judge									
Judge	11	10	9	8	7	6	5	4	3	2
1	99	95	93	97	98	98	98	98	95	98
2	98	98	93	98	94	92	95	93	97	
3	95	97	97	96	94	95	96	92		
4	97	93	90	94	97	98	96			
5	99	95	95	97	99	97				
6	96	93	94	95	97					
7	96	94	93	96						
8	98	98	94							
9	92	93								
10	98									

TABLE 10

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
100	100	100	98	100	100	98	98	96	100	100

APPENDIX D

The Number of Eleven Judges Concurring with the Investigator's Identification of Normal, Vocal Fry, and Rough Samples of Five Vowels Produced by Twenty Subjects

TABLE 11

THE NUMBER OF ELEVEN JUDGES CONCURING WITH THE INVESTIGATOR'S IDENTIFICATION OF NORMAL, VOCAL FRY, AND ROUGH SAMPLES OF FIVE VOWELS PRODUCED BY TWENTY SUBJECTS

	Vowels														
	/u/			/i/			/ʌ/			/ɑ/			/æ/		
Subject	N	VF	R	N	VF	R	N	VF	R	N	VF	R	N	VF	R
1	11	11	11	11	11	11	11	11	10	11	11	11	9	11	11
2	10	11	11	11	11	11	10	11	11	11	11	11	11	11	11
3	11	9	11	10	9	11	11	11	11	10	11	11	7	11	11
4	11	11	11	9	11	11	10	11	11	10	11	11	8	11	11
5	10	11	11	10	11	6	9	11	11	11	11	8	11	11	11
6	8	11	11	11	11	11	11	11	10	11	11	11	10	11	11
7	11	10	11	11	11	10	11	11	11	10	11	10	6	11	11
8	10	11	10	8	11	11	11	11	11	8	11	11	9	11	10
9	11	11	6	11	10	11	11	11	11	11	11	11	11	11	11
10	11	10	11	8	11	11	10	11	11	6	11	11	6	11	11
11	11	11	10	11	11	6	9	11	7	11	11	9	11	11	11
12	8	11	11	10	11	11	6	11	11	7	11	11	10	11	11
13	11	11	10	10	11	11	9	11	11	11	11	11	10	11	11
14	11	10	11	10	11	11	7	11	11	7	11	11	11	11	11
15	10	11	11	10	11	11	11	11	11	11	11	11	10	11	9
16	11	11	9	10	11	11	10	11	11	11	11	11	11	11	11
17	10	11	11	10	10	10	10	11	11	11	11	11	11	11	10
18	10	11	11	10	11	11	11	11	11	10	11	11	10	11	11
19	10	11	11	10	11	11	11	11	11	10	11	11	8	11	11
20	11	11	8	11	9	7	11	11	7	11	11	9	11	11	11

APPENDIX E

Summary of Analyses of Variance and Duncan's New Multiple Range Tests

TABLE 12

SUMMARY OF AN ANALYSIS OF VARIANCE FOR THE FUNDAMENTAL VOCAL FREQUENCIES OF NORMAL VOWELS

Analysis of Variance				
Source of Variation	df	ss	ms	F
Subjects	19	12767.75	671.99	17.03
Vowels	4	308.36	77.09	1.95 ≁
Residual	76	2998.84	39.46	

≁ Not Significant ($p > 0.05$)

TABLE 13

DUNCAN'S NEW MULTIPLE RANGE TEST FOR DIFFERENCES AMONG NORMAL VOWEL FUNDAMENTAL VOCAL FREQUENCIES (FVF)

Vowels	/i/	/u/	/ʌ/	/ɑ/	/æ/
FVF	117	116	114	113	112

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

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

TABLE 14

SUMMARY OF AN ANALYSIS OF VARIANCE FOR THE FUNDAMENTAL VOCAL FREQUENCIES OF VOCAL FRY VOWELS

Analysis of Variance				
Source of Variation	df	ss	ms	F
Subjects	19	3121.31	164.28	4.05
Vowels	4	4345.06	1086.26	26.77 *
Residual	76	3083.74	40.57	

* $p < .05$

TABLE 15

DUNCAN'S NEW MULTIPLE RANGE TEST FOR DIFFERENCES AMONG VOCAL FRY VOWEL FUNDAMENTAL VOCAL FREQUENCIES (FVF)

Vowels	/u/	/i/	/ʌ/	/ɑ/	/æ/
FVF	63	58	50	47	46
			______	______	
				______	______

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

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

TABLE 16

SUMMARY OF AN ANALYSIS OF VARIANCE FOR THE FUNDAMENTAL VOCAL FREQUENCIES OF ROUGH VOWELS

Analysis of Variance

Source of Variation	df	ss	ms	F
Subjects	19	32810.10	1726.84	13.407
Vowels	4	5535.06	1383.76	10.743 *
Residual	76	9788.95	128.80	

* $p < .05$

TABLE 17

DUNCAN'S NEW MULTIPLE RANGE TEST FOR DIFFERENCES AMONG ROUGH VOWEL FUNDAMENTAL VOCAL FREQUENCIES (FVF)

Vowels	/u/	/i/	/ʌ/	/ɑ/	/æ/
FVF	143	136	129	125	123

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

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

APPENDIX F

Details of Methods II and III

Method II

The regression equation used in Method II was identical to that used in Method I and has the form

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

where Y is the roughness rating predicted, B_0 the Y intercept determined by the regression analysis, B_{1-25} the regression coefficients determined by the regression analysis, and X_{1-25} the successive 100-Hz SNLs from 100 to 2600 Hz for each vowel production. The two methods differ in that the SNLs for vowel samples representing all three modes of phonation were used in estimating the betas for Method I, but in Method II only the SNLs for the normal and rough phonations were used. Thus, the greater number of relatively large residuals associated with Method II was attributable to the fact that the vocal fry samples were not included in fitting the regression plane. The largest residual associated with Method II was 2.56.

Method III

The regression equation used in Method III has the form

$$Y = B_0 + B\overline{X}$$

where Y is the predicted roughness rating, B_0 the Y intercept, and $\overline{X}$ the mean of the 25 SNL measurements for each test vowel production. The number of relatively large residuals associated with Method III as compared to Method I was expected because in Method III: (a) SNLs averaged over the 100 to 2600 Hz spectral range and the median roughness rating for each normal, vocal fry, and simulated abnormally rough production were used in fitting a regression line for each test vowel; and, (b) the correlation coefficients obtained (Table 7) when the 100 to 2600 Hz SNL means and the median roughness ratings for all productions of each vowel were related were smaller than those obtained (Table 8) when 100-Hz section SNLs and the median roughness ratings for all productions of each vowel were related. The largest residual associated with Method III was 1.37.