AN INVESTIGATION OF VOWEL FORMANT POSITION AND JUDGMENTS OF ROUGHNESS

A Thesis

Presented to

the Faculty of the Department of Speech

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Diane Krieger Linklater

August 1968

ACKNOWLEDGMENTS

The writer wishes to express her sincere appreciation to Dr. Ronald W. Wendahl for his continuous encouragement and endeavors as her thesis advisor, for his direction in the formulation of the experimental design, and for his instruction in procedural techniques.

Appreciation is also extended to Dr. Donna Fox and Dr. Joseph P. Schnitzen for their valued criticisms and constant aid in the writing of this thesis. The writer wishes to acknowledge her gratitude for the special interest and unfailing support extended by Thelma Krieger and Jacqulyn Mannix.

Diane Krieger Linklater

AN INVESTIGATION OF VOWEL FORMANT POSITION AND JUDGMENTS OF ROUGHNESS

An Abstract of A Thesis
Presented to
the Faculty of the Department of Speech
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

---

by
Diane Krieger Linklater
August 1968

ABSTRACT OF THESIS

The purpose of this study was to investigate the importance of the first formant frequency position upon the perceptual judgment of vocal roughness. Vowel-like stimuli were synthesized using a laryngeal analog and vowel filters. The basic experimental variable in the study was the position of the center frequency of the first formant.

The hypothesis upon which the study was generated was that the higher the first formant frequency, the greater would be the judgments of roughness. The results supported the hypothesis. Those vowels produced with low first formant frequencies were judged as less rough than those with higher formant frequencies. Clinical interpretations from the study are drawn which state that diagnostic procedures should include high first formant vowels.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

When a patient with a voice disorder phonates a prolonged vowel, several factors contribute to the perception of his voice quality. At least four factors must be described to circumscribe the important parameters which contribute to the perception of the vocal deviation. These are: (a) jitter (rapid variations from mean period); (b) glottal wave form; (c) shimmer (random glottal amplitude variations; and, (d) the transfer characteristics of the vocal tract.

This study was designed to investigate the effects of the transfer function of the vocal tract upon the perception of vocal roughness. More specifically, this study was designed to test the effect of first formant frequency changes on the perception of roughness. The study represented an attempt to expand on a chain of studies which have been done in an attempt to quantify terms used to describe a quality of voice which has been labeled by over two hundred terms. (1)

The term "roughness" will be used throughout this paper, unless in direct quotes, according to the June 1962

Rehabilitation Codes (33) In this publication the term roughness was posed to include all categories of voice production which previously had been described as harsh and hoarse.

Although the final determination of vocal roughness rests in the listener's ear, there appear to be two independent laryngeal factors which will create the sensation of roughness: (a) an aperiodic vibratory pattern (jitter); and, (b) a vibratory pattern which may be within normal limits of periodicity but which is characterized by glottal openings random with respect to area of opening (shimmer). The perception of roughness at the glottal level can be modified at the supraglottal level. Sherman and Linke (34) reported that vowels produced with a high tongue constriction were judged as less harsh than vowels produced with less constriction. They concluded that "since high vowels are shorter in duration than low vowels, the assumption that short vowels are in general perceived as less harsh than long vowels seems reasonable." Brubaker (3) and Rees (31) presented results in agreement with those of Sherman and Linke although they had not considered the question of duration.

Vowels which are characteristically produced with the tongue high in the mouth tend to result in low first formants, as shown by Fourier analysis. Vowels produced with low tongue positions tend to be associated with high first formant frequencies. The formants associated with the various vowels can be electronically simulated. By this means precise control can be maintained over a simulated glottal source and a simulated vocal tract.

Recent pilot data, gathered from a concurrent study of simulated vowels, support the conclusions of Sherman and Linke in part. Vowels with low first formant frequencies were perceived as less rough than vowels simulated with high first formants. In this study, the first three formants were systematically varied to simulate exact vowel formants. Thus, the spectral power density was distributed over three varying formants. These data indicated that duration was not a significant factor since the results were obtained with vowel duration held constant (6).

Cursory listening caused this writer to believe that the primary perceptual modification of laryngeal roughness was first formant determined. Therefore, the purpose of this study was to investigate the effect of first formant frequency changes and their relationship to listener judgments of roughness. The specific hypothesis was that the higher the position of the first formant frequency, all other factors being held stable, the greater would be the judgments of roughness.

I. REVIEW OF RELATED LITERATURE

Various studies have defined vocal roughness in the time domain. One of the first investigators to describe the psychoacoustic aspects of roughness was Carhart (4,5) who studied the spectra of tones produced by a cushion-pipe model larynx. This mechanical larynx, vibrating under certain conditions, produced a rough sound which Carhart associated with an inharmonic spectrum. He suggested that this roughness was analogous to a sound produced when each vocal fold vibrates at a different rate.

Simultaneous with the work of Carhart was a group of studies reported by Fairbanks and his students (15,16 17,18). These studies represented attempts to describe voice quality by an analysis of the rapid variations in glottal excitation patterns. These investigators found aperiodicity in wave-form pattern to be typical of the vocal productions of seven- and eight-year-old children and adolescent males. Curry (12), studying male adolescents, and Duffy (13), who studied vocal productions of eleven-, thirteen-, and fifteen-year-old females, reported the same results as Fairbanks, et al.

Later investigators found that the presence of rapid variations in glottal excitation patterns was not confined to normal or adolescent populations. Bowler (2), as late as 1964, equated harsh voice quality with extreme and abrupt changes in fundamental frequency. He assigned the term "frequency breaks" to these octave changes. Lieberman (25,26) used the term "perturbation" to describe the aperiodicity he found characteristic of both normal and abnormal laryngeal populations. His work agreed in part with the suggestions of earlier investigators that a degree of aperiodicity is found in all populations of speakers. Lieberman was the first to suggest that aperiodicity could most often be expected during times of rapid speech wave-form changes. He suggested that aperiodicity in normal speaking groups was most commonly associated with transitions between vowels and consonants and inflection patterns of conversational speech.

Fairbanks, et al (16) defined abrupt wave-to-wave changes in fundamental frequency as "voice breaks." They stated that these were typically one octave in extent and occurred in both upward and downward directions. Fairbanks (14), commenting on the failure of judges to perceive the voice breaks in children, stated "the answer may be found in the relationship between the typical location of the breaks . . . and the mode pitch level of the individual in question." He hypothesized that breaks down from, or up to, the individual's modal frequency would not result in the perception of voice breaks.

The question of whether octave voice breaks would be perceived as rough in quality was investigated by Wendahl (48). Synthesized voice breaks of one octave were presented to listeners whose reactions were reported as "the quality of the tone was an interesting auditory experience but neither unpleasant nor rough." This suggested that a reevaluation of the role of octave voice breaks in the perception of roughness was necessary.

Wendahl (17) studied persons with known laryngeal pathologies judged to have severe voice quality deviations. He did not find a single voice break of one octave in oscilliographic analyses of isolated vowels showing aperiodicity. Coleman (9) used photophonellographic procedures to measure the number and extent of voice breaks and the degree of judged hoarseness. He reported that the phonations of patients suffering from extreme hoarseness were characterized by vocal aperiodicities of a far less degree than the suggested one octave changes.

He suggested that the adolescent male quality deviation and other similar phenomena exhibited by younger males and females were not comparable to the breaks found in his subjects. He concluded that perceived roughness was related to "frequency changes less than one octave in extent occurring randomly on a cycle-to-cycle basis and perceived as a voice quality deviation."

Thompson (37) using ultra high-speed photography and phonellegraphic techniques, found that consecutive cycles varied more in time for phonations in voices judged to be harsh. He reported that the extent of the variation did not correspond in magnitude to those reported by Bowler (2) and Fairbanks (15,16,17,18).

Subsequent to these reports, Thompson and Moore (38) studied a population with known pathological larynges whose voice quality was judged as rough. Their results concurred with those of Coleman (9), Wendahl (48), and Thompson (37). The consensus was that comparatively small wave period variations were associated with voices judged to be rough.

Flanagan (19) and Stevens and House (36) suggested that normal vocal fold vibration, or laryngeal tone, was "quasi-periodic." Lieberman (25) found that the introduction of a small degree of aperiodicity into the fundamental frequency of synthetic speech enhanced its normal quality.

Cooper, Peterson, and Fahringer (11) maintained the trend of describing voice quality in the time domain. They attributed vocal roughness to the presence of abrupt, random, cycle-to-cycle period variations which they termed "jitter."

Although there was general agreement that jitter was present in normal as well as rough vocal productions, there remained some question about the amount of irregularity necessary for the perception of vocal roughness. Investigators had been limited to studying the vocal productions of human subjects in which all the possible variations in speech were occurring and uncontrollable. These variables included articulation, modal pitch level, invlection, amplitude variations caused by the glottal source, and modifications of the glottal source by the transmission characteristics of the vocal tract.

Wendahl (45,46,47) began a systematic study of jitter as a psychoacoustic correlate of roughness. Using an electrical laryngeal analog, LADIC, to generate the stimuli, he found that saw-tooth waves were judged to be rough when wave lengths of consecutive cycles varied as little as plus or minus one Hertz around a median frequency of 100 Hz. He found that amounts of jitter were directly related to the roughness judgments by listeners. Wendahl's findings supported the tentative suggestions of Coleman (9) and Thompson and Moore (38) regarding the effect of smaller variations than those noted by earlier investigators.

Jitter, or aperiodicity of vocal fold vibratory patterns, may not supply the entire answer to the cause of perceived vocal roughness. Relatively few studies have explored the effects of random amplitude variations of the glottal wave on judgments of roughness.

Moore and von Leden (30) related gross amplitude variations in successive openings of the vocal folds to deviant voice quality. Coleman (9) found in records obtained from a pathological larynges population, that both jitter and amplitude variations occurred. Although he was not able to quantify or equate amplitude differences with perceived roughness, he suggested that these were primary factors in the roughness judgments.

Moore (28),through the use of ultra high-speed photography, demonstrated asynchrony of vibratory patterns between the two folds. (See Figure 1) For this specific patient, the right fold was opening at a slower rate than the left. A close look at area of openings vs. amplitude of openings is shown schematically in Figures 1-A and 1-B. Both time and amplitude variations were present in the voice of this patient.

Figure 1-A Theoretical openings of the vocal folds showing asynchrony of opening.

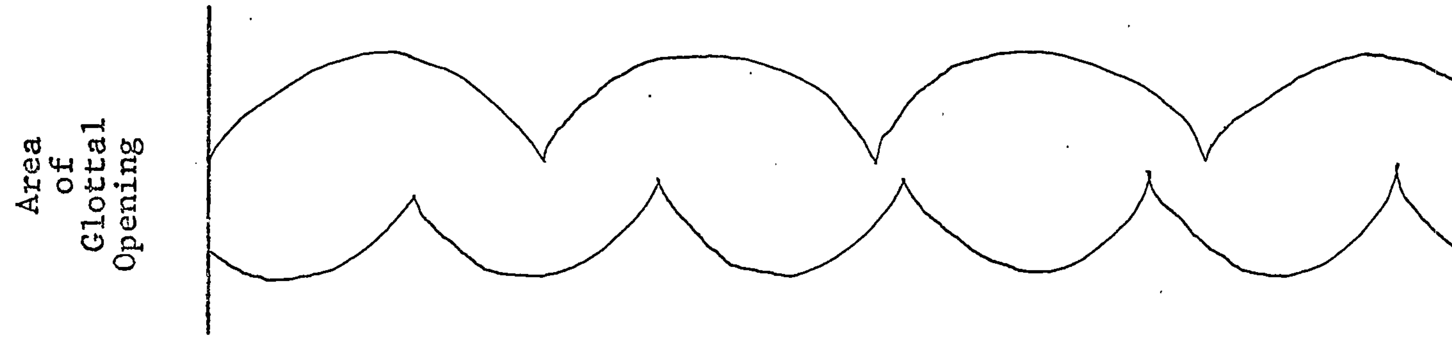

Figure 1-B Theoretical representation of the expected sound pressure level which would result from this vibratory pattern.

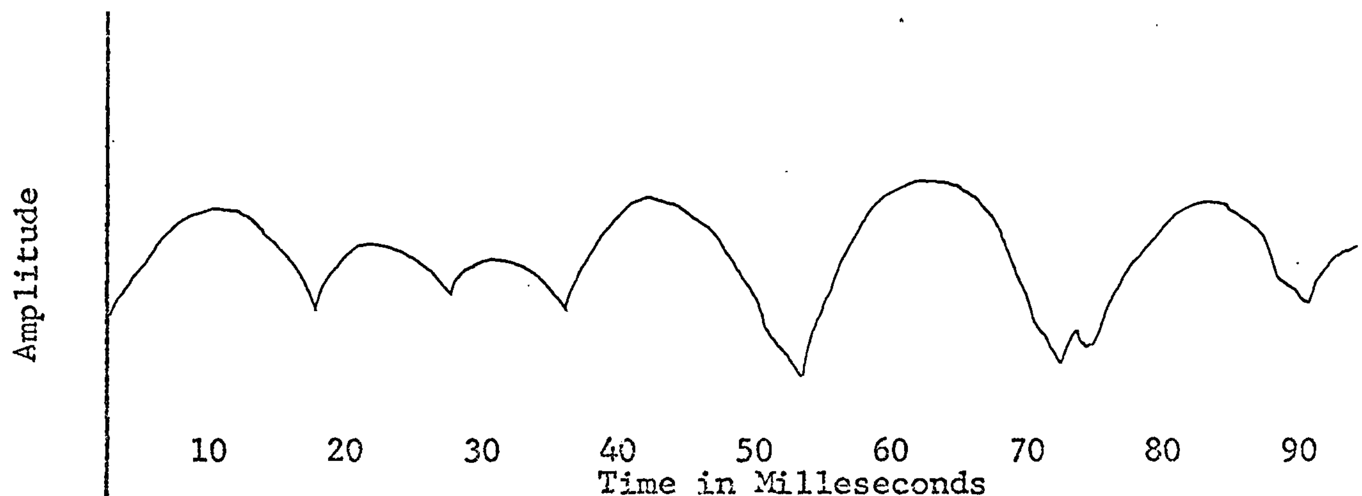

FIGURE 1

SCHEMATIC DRAWING OF ASYNCHRONY OF VOCAL FOLD VIBRATORY PATTERNS

The tentative data from both Coleman (9) and Moore (28) suggested the need to study the acoustic parameter of amplitude variations. Through the use of computer techniques Wendahl (46) investigated the relationship between wave amplitude variations and judgments of roughness. He used the term "shimmer" to describe these rapid and random amplitude variations among successive glottal impulses. He found that the presence of shimmer in saw-tooth waves resulted in judgments of roughness which could be scaled on the dimension of auditory roughness in the same manner as jitter.

Sonneson (35), using glottographic techniques, where a photosensor was placed above the larynx and a light source on the neck below the folds, demonstrated that a patient with a hemiparesis of one fold had a vibratory pattern of one large fold opening, followed by a small opening. This patient, who exhibited no jitter, was reported to sound very rough. Thus Sonneson, using human subjects, verified Wendahl's earlier work.

This parameter of shimmer may answer the question presented when vocal roughness was reported in the absence of concomitant jitter. The consensus implied that the perception of vocal roughness rested upon at least two glottal factors - time and amplitude (jitter and shimmer).

Another aspect, that of vocal fry, has often been equated with vocal roughness. Several investigators (8,20,21,27,30,49) suggested that vocal fry showed more aperiodicity than a voice judged to be rough. These studies associated vocal fry with both a very low fundamental frequency and the damping time of the vocal tract. These suggestions might explain much of the early literature that related roughness to low pitch. The low frequency vibrations found at the end of phonations for nearly all subjects, both normal and pathological, constitute segments of vocal fry which may be conceived of as a normal wave-period fluctuation in connected speech. Much of the early literature reporting octave breaks may be related to the presence of vocal fry. That vocal fry, exhibiting larger amounts of jitter was not perceived of as rough suggested that the ear may operate differently at very low frequencies than at average speaking levels.

Various studies (19,22,29) suggested the importance of noise components and the loss of high frequency harmonics in the perception of vocal roughness. Thurman (39) found that vowels produced by rough voices showed some tendency toward lower positions of the first and second formants. His results were questionable since he could not demonstrate the relationship between consistent lowering of vowel formants and increasingly severe harsh phonations.

Van den Berg (40) related rough voice quality to a greater concentration of energy in the higher harmonics. Laguaite (24) reported a preponderance of low frequency distributions in hoarse populations. Fairbanks (15) found well-defined first formant regions and pulse-like signals in the upper formants in harsh voices. Yanagihara (50) suggested that the loss of high frequency harmonics resulted in increased judgments of hoarseness.

There are many factors which may mitigate the perception of vocal roughness. Some have been completely unstudied, others have been given partial consideration. Among those factors about which no information is available at present, is the effect of the opening and closing quotients of the vocal folds. Coleman (8) has recently recorded data relative to this area, but results from his study are not available at this time.

CHAPTER II

PROCEDURE

I. HYPOTHESIS

The specific hypothesis of this study was that the position of the first formant of an isolated vowel was the primary determinant in the perception of vocal roughness. Roughness is generated at the larynx and modified by the various transfer functions of the vocal tract. Vowels which have greater constriction will be judged as less rough than vowels with a small amount of constriction.

II. ASSUMPTIONS

As basic to the rationale of this study the following assumptions were made: (a) Roughness judgments are vowel dependent; (b) high vowels, with greater constriction would be judged as less rough than low vowels, with less constriction, regardless of stimuli duration; and, (c) the reason for this judgment would be the position of the first formant.

III. METHODOLOGY AND PROCEDURE

General Procedure

The basic stimuli were generated on an electrical laryngeal analog to digital computer, LADIC, so that each acoustic factor could be isolated and controlled. The complete operation of LADIC has been described in several publications (45,46,47) and will not be described again beyond the statement that a basic saw-tooth wave-form was chosen which could be programmed to change adjacent period intervals between 16.6 milleseconds and 1.11 milleseconds at will. The choice of jitter program was dictated by this published data. The program was one which had been scaled by psychophysical methods and one known to produce a tone which would be judged rough by most listeners. The complete program can be found in Table I.

The output of LADIC was fed into a series vocal tract simulator and stored on magnetic tape. The tape was spliced into a master tape in a paired-comparisons design and played to a group of twenty listeners in a quasi-sound-treated recording studio. The listeners were instructed to listen to the pairs of sounds and to state which of each pair sounded more rough.

TABLE I

WAVE PERIODS AND RELATIVE AMPLITUDES FOR THE MODEL AND JITTER PROGRAMS

MODEL PROGRAM				JITTER PROGRAM 300 CENTS	
Pulse Number	Relative Amplitude in db	Frequency	Periods in Milleseconds	Frequency	Periods in Milleseconds
1	-30	100	10.0	100	10.0
2	-27	100	10.0	100	10.0
3	-24	100	10.0	100	10.0
11	(zero)	100	10.0	119	8.46
12		101	9.9	82	12.22
13		102	9.8	121	8.24
14		103	9.7	84	12.04
15		104	9.6	124	8.05
16		105	9.5	85	11.75
17		106	9.4	126	7.92
18		107	9.3	87	11.45
19		108	9.2	129	7.75
20		109	9.1	88	11.31
21		110	9.0	132	7.63
22		111	8.9	91	11.07
23		112	8.8	132	7.63
24		113	8.7	93	10.79
25		114	8.6	136	7.36
40		129	7.8	105	9.57
41		130	7.7	155	6.54
42		131	7.6	106	9.25
43		132	7.6	157	6.35
44		133	7.5	108	9.23
45		134	7.4	159	6.27
46	---	---	---	---	----

Specific Procedures

Generation of stimuli: LADIC was programmed to present a 300 cent jitter program. This represented a deviation in frequency from the model program in the amount of plus or minus 300 cents or three semitones. An example of this deviation is Pulse 11, where the model program called for a wave period corresponding to 100 Hz. This frequency was raised by approximately 300 cents, or to 119 Hz. The program called for Pulse 12 to be lowered by 300 cents or to 82 Hz.

The experimental design required that thirty-five channels of LADIC be set at 300 cents jitter superimposed on a rising inflection from 100 to 134 Hz. The output wave-form was a saw-tooth with a slow rise and rapid decay. The program called for variations in adjacent time period intervals of between 16.6 milleseconds and 1.11 milleseconds. The first ten channels, representing 100 milleseconds in time, were set at a steady state frequency of 100 Hz while the amplitude was systematically modulated from -30 db to 0 db. LADIC was set to present only one cycle of each frequency before advancing to the next channel. The forty-sixth channel was set at zero amplitude to produce a five second silent interval. The program allowed LADIC to recycle after this silent period.

The output of LADIC was fed into a vowel filter, POVO, which was set to the following characteristics: A first formant frequency of 500 Hz; a second formant frequency of 1,500 Hz; and, a third formant frequency of 2,500 Hz. The half-power bandwidths were set to 50, 75, and 100 Hz respectively. Following this procedure, the bandwidths were allowed to vary freely with frequency. POVO was then set to have center formant frequencies for the second through fifth formants of 1,500, 2,500, 3,500, and 4,000 Hz respectively. These center frequencies were held constant throughout the experiment. The only major variable was the position of the first formant which was systematically varied from 200 to 800 Hz in 100 Hz steps.

The jitter program was played through POVO under each of the experimental conditions. Although there were differences in formant bandwidth and formant amplitude resulting from the variation of the first formant, these factors were not isolated or controlled. These differences will be discussed in greater detail in the discussion section.

Construction of the Master Tape: The master tape consisted of a jittered signal of 300 cents for each of the seven conditions of first formant frequency. Each of the seven conditions was recorded a number of times on lengths of magnetic tape by an Ampex PR 10 AM recorder set to run at fifteen inches per second. These segments were later cut into shorter lengths and spliced into a paired-comparisons design in which every segment was paired with every other segment. Each segment was ordered as follows: a stimulus of approximately 333 msec.; a silent interval of 500 msec.; a stimulus of approximately 333 msec.; and, a five second silent interval preceding the subsequent pair of stimuli.

The final tape consisted of two stimuli demonstrating "rough" and"smooth." These were simulated by programming LADIC to present a 100 Hz steady state through the vowel filters which were set for the vowel /a/ - smooth; and, a 100 Hz jittered signal - rough. Following these demonstration items were ten pair of practice items, twenty-four pair of test items, and three pair of items included for test-retest reliability measures. None of the practice items were used in the master tape.

Playback Procedure: Playback was achieved through paired KLH Model 10 electrostatic speakers in a partially sound-treated room. Sound output was adjusted to a comfortable listening level and was maintained constant throughout the experiment.

Listeners: For this experiment, twenty subjects served as volunteer listeners. All were undergraduate students enrolled in the Department of Speech at the University of Houston. None had a history of hearing loss.

Instructions to the Listeners: The descriptive adjectives of "smooth" and "rough" were not verbally defined, but rather demonstrated by presenting the 100 Hz steady-state vowel and the 100 Hz jittered signal generated from LADIC. Neither stimuli was included in the experimental program. Complete instructions given to the listeners may be found in the appendix.

CHAPTER III

RESULTS AND DISCUSSION

Although the original group of listeners used in this study consisted of twenty listeners, a total of only fifteen listeners' responses were used in the final tabulation of data. The responses of five listeners were removed from the data because of their failure to meet the reliability criterion of giving the same response to a minimum of two of the three test-retest items included in the experimental tape.

The total number of votes for each stimulus was summated. The results of this summation appear in Table II.

TABLE II

RELATIONSHIP BETWEEN CENTER FREQUENCY OF THE FIRST FORMANT AND JUDGED ROUGHNESS

Center Frequency of First Formant in Hertz	200	300	400	500	600	700	800
Summed Roughness Judgments	17	23	40	41	57	62	75

These data are presented in graphic form in Figure 2.

From these data it is apparent that the judgment of roughness was influenced by the center frequency position of the first formant. Both the amount of jitter and the duration of each signal were held constant at 300 cents and 333 milleseconds respectively. The number of judgments of roughness rose as the first formant was increased in frequency. While there was a positive and approximately linear relationship between judgments of roughness and increased values of first formant frequencies, the data from this study cannot be extrapolated to state that the first formant center frequency is the only major variable or, ceteres parabus, the significant variable.

Although the center frequencies of the second through fifth formants were held constant throughout the experiment, no attempt was made to control differences in formant amplitude resulting from the variation of the first formant. As the first formant was increased in frequency in this experiment, it was brought closer to the center frequency of the higher fixed formants. As the first formant frequency approached the higher formants, their amplitudes would be expected to increase. The rationale for this prediction was as follows: When any

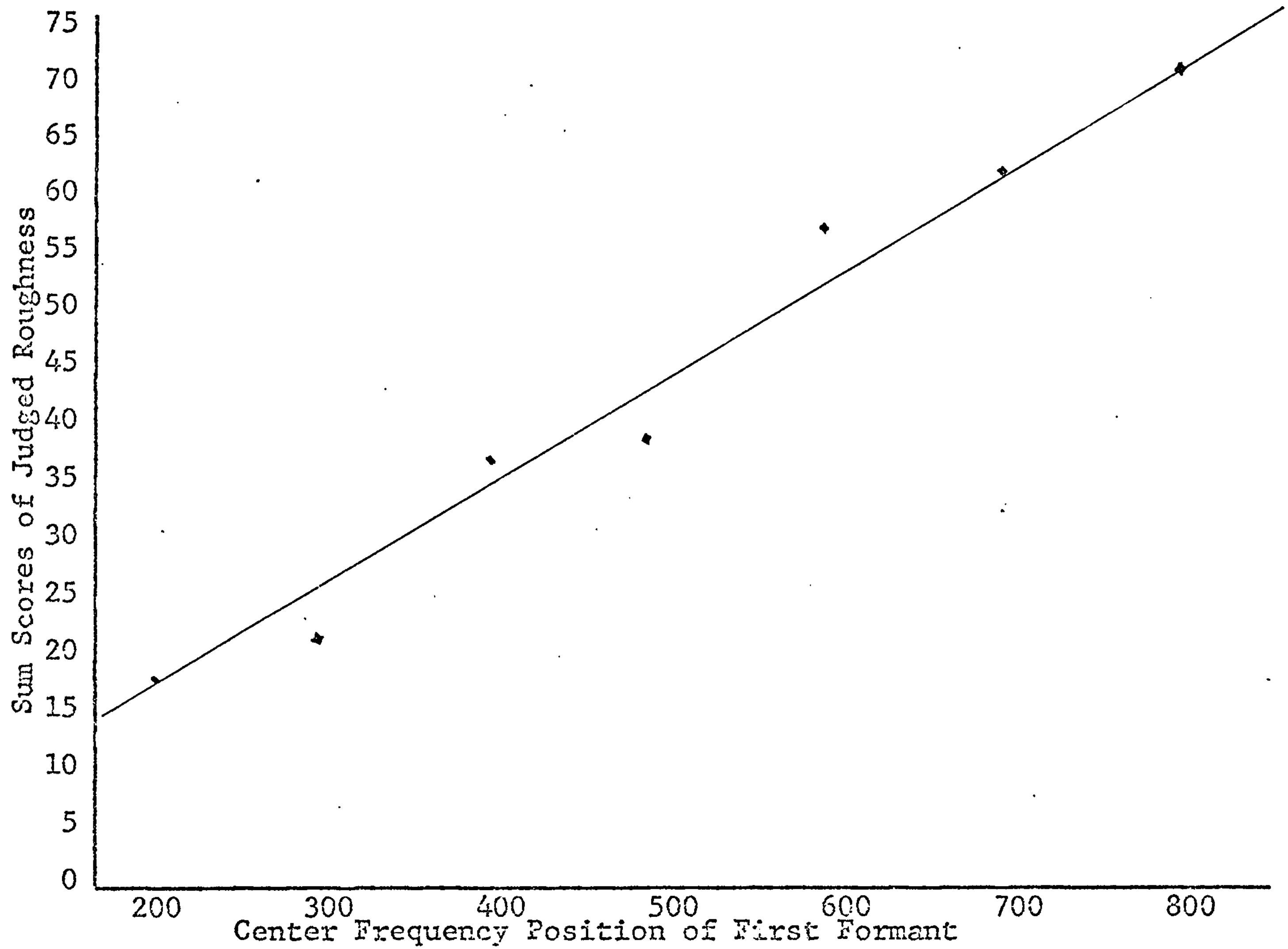

FIGURE 2

JUDGED ROUGHNESS AS A FUNCTION OF FIRST FORMANT FREQUENCY POSITION number of filters are connected in a series, their frequency response is asymptotic. The tails of one filter will contribute to the amplitude of adjacent filters. This statement is illustrated in Figure 3.

In Figure 3, as Filter A is moved closer to Filter B and Filter C, the formant amplitude of all three filters will be increased. This cross effect results from the dependence of Filter A upon the energy at B' and C' for amplitude. This factor could be controlled in future studies and an attempt could be made to evaluate its contribution to vocal roughness perception.

Another factor which entered into the data was the systematic change in formant bandwidth as the first formant frequency was raised. A characteristic of the series filters used in this experiment is an increase in formant bandwidth as center frequency positions are raised. This factor was not isolated in the present experiment since these filters function as an analog of the human vocal tract. In future studies, an attempt might be made to investigate the significance of bandwidth variations in order to evaluate their contribution to the perception of vocal roughness.

The original hypothesis of this experiment will be

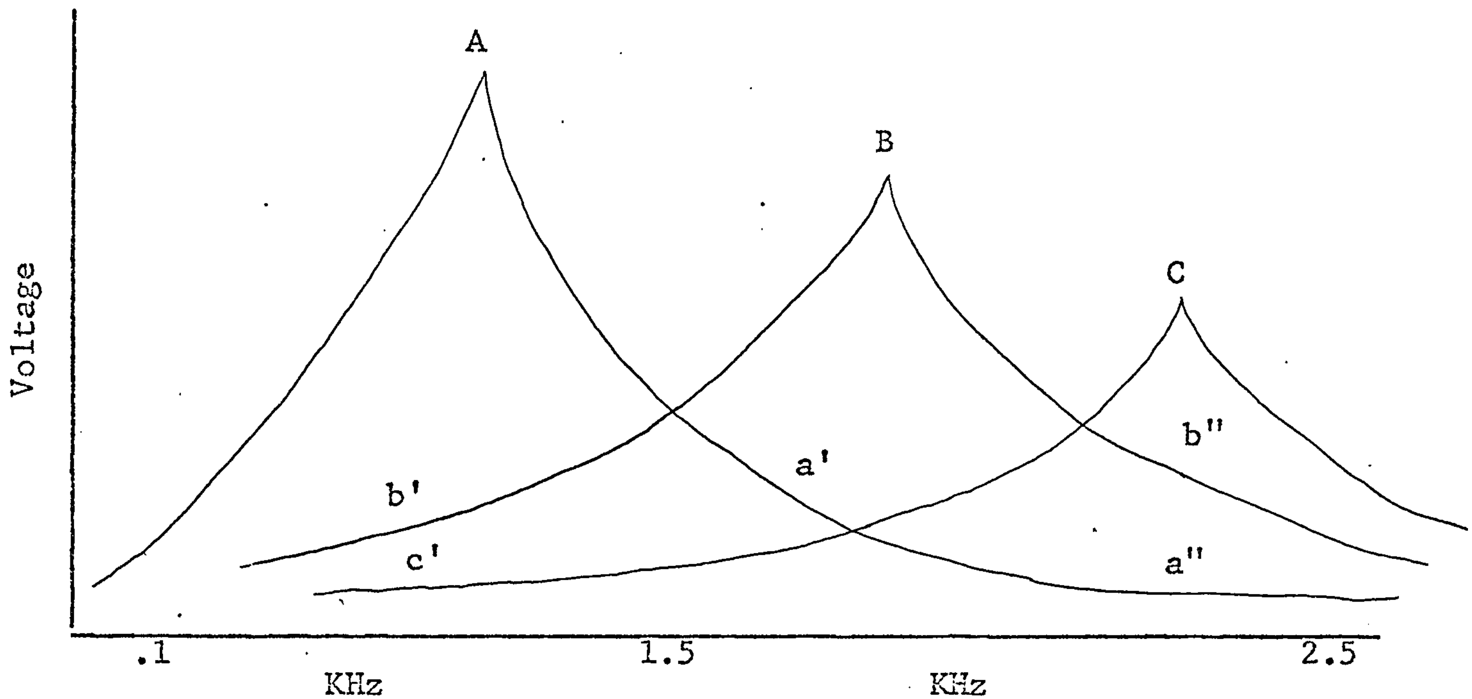

FIGURE 3

THE RELATIONSHIP BETWEEN FORMANT AMPLITUDE AND CENTER FREQUENCY OF ADJACENT FILTERS

*

Upper case letters indicate the amplitude of the center frequency of the first three formants.

Lower case prime numbers show the energy potentially contributed to the amplitude of the center frequency.

accepted, however, until such a time as more detailed investigations of both formant amplitude and bandwidth are undertaken. The hypothesis stated only that the number of judgments of roughness would be increased as the first formant was increased in frequency from 200 to 800 Hz. This hypothesis was supported by the data resulting from this study.

These data tend to support the findings of Sherman and Linke (34) who stated that vowels which are characteristically produced with the tongue high in the mouth are judged to be less rough than those vowels produced with the tongue low in the mouth. To the extent that the first formant frequency position is related to the amount of tongue constriction or height, these data are consistent with the findings of Sherman and Linke. But, the data from this study cannot be interpreted in the way that Sherman and Linke interpreted their results. They concluded that roughness judgments were dependent upon the duration of the vowel and assumed that since high vowels were shorter in duration than low vowels, they would get fewer roughness judgments. They assumed that the reason that the low vowels sounded more rough was that they were generally of longer duration than high vowels.

While Sherman and Linke's assumptions may be generally true for utterances produced by humans, the stimuli used in this experiment were electronically generated with duration held constant for all stimuli. Duration of the vowel was excluded as a variable in this study.

The results of this study have possible clinical implications suggestive of a rationale for an altered approach to the treatment of voice disorders. In a clinical setting, where one is working directly with a person who has a voice disorder, a primary goal is to facilitate the client's perception and discrimination of his vocal roughness. As he comes to recognize the difference between phonations, he is then able to alter his production and judge improvements in voice quality.

The clinical procedures leading to this goal will vary according to whether one proceeds from Sherman and Linke's assumptions or the data from the present study. If one assumed Sherman and Linke's interpretation that vowel duration is the significant factor in the determination of vocal roughness, he would have the client phonate a vowel over a considerable period of time without respect to the particular vowel phonated. If one agreed with the view expressed by the writer of this study, he would select a vowel for therapy which is characteristically produced with the tongue low in the mouth. In the latter view, the duration of the phonation would be a secondary factor.

CHAPTER IV

SUMMARY AND CONCLUSIONS

This study was designed to investigate the effects of vowel first formant frequency position upon listener evaluations of vocal roughness. The specific hypothesis investigated was that as the first formant of synthetic vowels was raised in frequency, listener judgments of roughness would increase.

The stimuli for this experiment were electronically generated on a special purpose computer designed to simulate the laryngeal source, LADIC. LADIC was set to a program which had been scaled by psychophysical methods and known to produce a tone which would be judged as rough by most listeners. The same basic source was used for all the stimuli included in this study. The output of LADIC was fed into a series vowel apparatus, POVO, which was set to the following characteristics: The second through fifth formants were held fixed at 1,500, 2,500, 3,500, and 4,000 Hz respectively. The duration of all stimuli was held constant at 333 milleseconds. The only major variable was the frequency position of the first formant which was systematically varied from 200 to 800 Hz in 100 Hz steps.

Half-power bandwidths were set to 50, 75, and 100 Hz for the first through third formants. Formant bandwidth and formant amplitude were allowed to vary freely as first formant position was raised.

The stimuli were stored on a high quality AM tape recorder and arranged in a paired-comparisons design in which each stimuli was paired with every other stimuli. The pairs of stimuli were played to listeners through electrostatic speakers in a sound-treated environment. The listeners were allowed to listen to each pair as many times as necessary in order to make a judgment indicating which in each pair sounded the more rough. The original listening group of twenty judges was reduced to fifteen as a result of five judges' failure to meet the reliability criterion.

Analysis of the listeners' judgments resulted in a positive and approximately linear relationship to the original hypothesis. Vowels received increasingly more judgments of roughness as the value of the first formant frequency was systematically raised.

The results of this experiment led to the following conclusions: First formant frequency position of vowels is a primary determinant of judged roughness. This statement was found to hold true when duration of the stimuli was held constant. It follows that clinically in the diagnosis and retraining of persons with rough voice quality, a low vowel such as /a/ or /ae/ would be used. The latter statement is in direct conflict with the interpretation by Sherman and Linke (34).

The following limitations of the data from this study are recognized. Formant amplitude and formant bandwidth varied freely as first formant frequency was varied. These factors were not measured so their importance in the perception of vocal roughness cannot be estimated.

Although these limitations are recognized, neither is considered important to the clinical interpretations drawn from this study. A study of the effects of formant bandwidth and formant amplitude on the perception of vocal roughness would be experimentally interesting. Both of these factors are currently under investigation at the University of Houston and at the University of Florida.

APPENDIX

INSTRUCTIONS TO THE LISTENERS

You are about to hear a series of stimuli which have been paired so that one of the pair is more rough. It will be your task to listen to each pair and to indicate which, either One or Two, sounds more rough. You will have to make a decision for each of the pairs presented to you. You may not leave a blank indicating that you cannot make a decision. You are also not allowed to check both One and Two indicating that they are equally rough in your estimation. If you do either of these on any stimulus, we will have to discard your paper.

You are encouraged to raise your hand and request as many repeats on any pair of sounds you may find necessary in order to make a decision. Please do not feel embarrassed to ask for repeats. Also, do not feel compelled to keep your first answer if somebody else has asked for a repeat and your decision has changed. You are allowed to change any judgment.

The task we are going to give you will range from very easy to very hard. It is up to you to work harder on the difficult items and make the best decision you can. Are there any questions? You will be presented ten practice items and then we will stop for another question period before going on to the main research project.

BIBLIOGRAPHY

1. Anderson, V.P. Improving The Child's Speech. New York: Oxford University Press, 1953.

2. Bowler, N.W. "A Fundamental Frequency Analysis of Harsh Vocal Quality." Speech Monographs, 31: 128-34, 1964.

3. Brubaker, R.S. and Dolpheide, W.R. "Consonant and Vowel Influences upon Judged Voice Quality of Syllables." J. Acous. Soc. America, 27:1000, 1955.

4. Carhart, Raymond. "Infra-Glottal Resonance and a Cushion Pipe." Speech Monographs, 8:76-84, 1941.

5. ____________. "The Spectra of Model Larynx Tones." Speech Monographs, 8: 76-84, 1941.

6. Cole, Belinda. Personal Communication.

7. Coleman, R. F. "Decay Characteristics of Vocal Fry." Folia Phoniatrica, 15:256-63, 1963.

8. ____________. Personal Communication.

9. ____________. "Some Acoustic Correlates of Hoarseness." Unpublished Master's Thesis, Vanderbilt Univ., 1960.

10. Coleman, R. F. and Wendahl, R. W. "Vocal Roughness and Stimulus Duration." Speech Monographs, 34:86-92, 1967.

11. Cooper, F. S., Peterson, E. and Fahringer, G.S. "Some Characteristics of Vocoder Quality." J. Acous. Soc. America, 29:183A, 1957.

12. Curry, E. Thayer. "The Pitch Characteristics of the Adolescent Male Voice." Speech Monographs, 6:48-62, 1940.

13. Duffy, R. J. "The Vocal Pitch Characteristics of Eleven-, Thirteen-, and Fifteen-Year Old Female Speakers." Unpub. Ph.D. Dissertation, State University of Iowa, 1958.

14. Fairbanks, Grant. "Recent Experimental Investigations of Vocal Pitch in Speech." J. Acous. Soc. America, 11: 457-66, 1940.

15. ____________. Voice and Articulation Drillbook. 2nd Ed., New York: Harper Brothers, 1960. pp. 127, 170-83.

16. Fairbanks, Grant, Wiley, J. H., and Lassman, Frank. "An Acoustical Study of Vocal Pitch in Seven- and Eight-Year Old Boys." Child Dev., 20:63-9, 1949.

17. Fairbanks, Grant, Herbert, E.C., and Hammond, M. J. "An Acoustical Study of Vocal Pitch in Seven- and Eight-Year Old Girls." Child Dev., 20:71-8, 1949.

18. Fairbanks, Grant and Pronovost, Willibert. "An Experimental Study of the Pitch Characteristics of the Voice During the Expression of Emotion." Speech Monographs, 6:89-104, 1939.

19. Flanagan, J. "Some Properties of the Glottal Sound Source." J. Speech & Hearing Research, 1:99-116, June, 1958.

20. Hollien, H., Moore, P., Wendahl, R. W., and Michel, John. "On the Nature of Vocal Fry." J. Speech & Hearing Research, 9:245-47, 1966.

21. Hollien, H., and Wendahl, R. W. "Perceptual Study of Vocal Fry." J. Acous. Soc. America. 43:506-10, March 1968.

22. Isshiki, N. and von Leden, H. "Hoarseness-Aerodynamic Studies." Arch. Otolaryngology, 80:206-13, 1964.

23. Johnson, Wendell. Speech Handicapped School Children. New York: Harper Brothers, 1948.

24. Laguaite, J. and Waldrop, W. F. "Acoustic Analysis of Fundamental Frequency of Voice Before and After Therapy." New Zealand Speech Therapist Journal, 18: 23-5, 1963.

25. Lieberman, P. "Perturbations in Vocal Pitch." J. Acous. Soc. America, 35: 344-53, 1963.

26. Lieberman, P. "Some Acoustic Measures of the Fundamental Periodicity of Normal and Pathological Larynges." J. Acous. Soc. America., 35:344-53, 1963.

27. Michel, John. "Pitch Characteristics of Adult Males, Vocal Fry and Harshness." Progress Reprt, Nat'l Institute of Neurological Disorders and Blindness, National Institute of Health, October, 1964.

28. Moore, Paul. "Theoretical Concepts in Phonation." Paper presented at Communication Sciences, Short Course on Voice Function, University of Florida, Gainesville, Florida, May , 1967.

29. ____________. "Voice Disorders Associated with Organic Abnormalities." Handbook of Speech Pathology. L. E. Travis, Ed., New York: Appleton-Century-Crofts, 1957.

30. Moore, P. and von Leden, H. "Dynamic Variation of the Vibratory Pattern in the Normal Larynx." Folia Phoniatrica, 10:205-38, 1958.

31. Rees, M. "Some Variables Affecting Perceived Harshness." J. Speech & Hearing Research, 1:155-88, 1958.

32. Rehabilitation Codes. Five-Year Progress Report 1957-1963. Special Project RD-788, Office of Vocational Rehabilitation, New York, 1964.

33. Rehabilitation Codes. Proceedings of the Workshop on Nomenclature of Communication Disorders. PHS Grant No. B-3676, Bethesda, Maryland, 1962.

34. Sherman, D. and Linke, E. "The Influence of Certain Vowel Types on Degree of Harsh Voice Quality." J. Speech & Hearing Disorders, 17:401-8, 1952.

35. Sonnesson, Bertil. "Photo-Glottographical Studies of the Vibrating Larynx - Part I." Paper presented at Communication Sciences, Short Course on Voice Function, University of Florida, Gainesville, Florida, May, 1967.

36. Stevens, K. and House, A. "An Acoustical Theory of Vowel Production and Some of its Implications." J. Spc. & Hear. Research, 4:303-20, 1961.

37. Thompson, Carl L. "Wave Length Perturbations in Phonation of Pathological Larynges." Progress Report, National Institute of Neurological Diseases and Blindness, NIH, December, 1962.

38. Thompson, Carl L. and Moore, Paul. "Comments on the Physiology of Hoarseness." Arch. Otolaryngology, 81:97-102, January 1965.

39. Thurman, W. L. "The Construction and Acoustic Analysis of Recorded Scales of Severity for Six Voice Quality Disorders." Unpub. Ph.D. Dissertation, Purdue University, 1963.

40. Van den Berg, J. "On the Role of the Laryngeal Ventricles in Voice Production." Folia Phoniatrica, 7:57-69, 1955.

41. Van Dusen, C. "A Laboratory Study of the Metallic Voice." J. Spc. & Hearing Disorders, 6:137-40, 1941.

42. von Leden, H., Moore, P., and Timke, R. "Laryngeal Vibrations: Measurements of the Glottic Wave, Part III, The Pathological Larynx." Amer. Medical Assoc. Arch. Otoloaryngology, 71:16-35, 1960.

43. Wendahl, Ronald W.. " A Photophonellographic Analysis of Hoarse Voice Quality." Proceedings of Fourth International Congress of Phonetic Science. Helsinki, 1961. pp. 307-10.

44. ______________. "Laryngeal Analog Synthesis of Glottograms." Proceedings of the Fifth International Congress of Phonetic Science. Munster, Germany, 1964, pp. 569-73.

45. ______________. "Laryngeal Analog Synthesis of Harsh Vocal Quality," Folia Phoniatrica, 15:241-50, 1963.

46. ______________. "Laryngeal Analog Synthesis of Jitter and Shimmer, Auditory Parameters of Harshness." Folia Phoniatrica, 18: 98-108, 1966.

47. ______________. "Some Parameters of Auditory Roughness." Folia Phoniatrica, 18:26-32, 1966.

48. _______________. " The Synthesis of Harsh Voice Quality." Paper presented at the American Speech and Hearing Association Convention, New York, 1962.

49. Wendahl, R. W., Moore, P., and Hollien, H. "Comments on Vocal Fry." Folia Phoniatrica, 15:251-55, 1963.

50. Yanagihari, Naoki. "Significance of Harmonic Changes and Noise Components in Hoarseness." J. Speech and Hearing Research, 10:531-41, 1967.