THE EFFECT OF PROLONGED VOCALIZATION ON THE VOCAL SPECTRUM

by 680

RUSSELL ERNEST HEDGE

B. S., The University of Kansas, 1968

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

Approved by

Robert S. Brooks
Major Professor

TABLE OF CONTENTS

Problem

Attempts to describe specific variations of the voice have led to the derivation of confusing terms. Where similar terms appear in different texts dealing with the study of voice (Berry and Eisenson, 1956; Curtis, 1956; Moore, 1957; West, Ansberry and Carr, 1960; Fairbanks, 1960; Levin, 1962; Murphy, 1964; Greene, 1964; Van Riper and Irwin, 1965; Anderson, 1965), their definitions are usually different in one or more important ways. Conversely, different terms will describe similar changes in vocal quality. Harsh, hoarse, husky, raspy, gutteral, breathy, rough and strident are among the terms which have been used interchangeably. Their exact definition has been a matter of discussion and debate (Yanagihara; 1967b). The interpretation of the data that are available for the study of voice disorders is confused by uncertainty with regard to percisely what vocal characteristics are being considered.

Van Riper (1965) characterized the problem of describing voice quality as follows:

> That the disorders of voice quality are difficult to describe is indicated not only by the names which we listed earlier in our classification but also by the names we omitted. Voices have been called thick, thin, heavy, sweet, round, brilliant, hard, metallic, and rich, as well as poor. The terms we have used are not much better, but at least they do not confuse auditory perceptions with those of taste and touch. The science of experimental phonetics has not yet been able to provide a better classification for variations in timbre.

The speech clincian can describe the voice with whatever adjectives of phrases he feels most suitable (Villarreal, 1949). The quality of a voice and the degree to which that quality is present "is determined chiefly on the basis of the clincian's subjective perception (Yanagihara, 1967b)." Therefore, two observers could describe a voice disorder as two different vocal qualities. The human observer is characterized by varibility, subjectivity and unbelievable complexity as a receiver and analyzer of the speech-sound signal.

The variance in the acoustical properties of voice-pitch/time, quality and loudness - which signifies a defective voice need to be objectively defined if the speech clincian is to become more specific in his selection of terms as labels for particular voice qualities. Although the experienced listener can usually detect a change in vocal quality, he cannot quantitatively define how the specific parameters of the vocal output have changed. For this reason Ladefoged (1964) says, "...instrumental phonetics may be a very powerful aid and of great use in providing objective records on the basis of which we may verify or amend our subjective impressions." Instruments which supply quantifiable acoustical data, even though these data would be interpreted with a certain degree of subjectivity, could be of great use in defining and measuring changes in vocal quality.

Statement of Problem

Breathiness, harshness or hoarseness might be the expected sequela of prolonged vocalization. Breathy quality results when the vocal folds vibrate, but fail to approximate medially to sufficiently interrupt the airflow from the trachea and lungs (Fairbanks, 1960). Thus, the airflow is continuous. Acoustic analysis of breathy vocal quality reveals a rather broad-band noise superimposed on the periodic vocal tone (Zemlin, 1968). The tone generated by the vocal folds is accompanied by strong frictional noise components, limited in vocal intensity and lowered in pitch. The audible effect produced has been described by words such as fuzzy, veiled, hoarse and whispered (Murphy, 1964). Breathiness can result from laryngeal inflammation due to poor muscular tone caused by vocal abuse. Breathy quality is one variation of vocal dysfunction and is a common accompaniment in hoarse and harsh voices.

Terms used to connote harshness include these: strident, coarse, grating, rasping, metallic and gutteral (Murphy, 1964). Curtis (1956) described harsh voice quality as "an unpleasant rough, rasping sound." Spectrograms, state Fairbanks (1960), reveal that "irregular, aperiodic noise in the vocal-fold spectrum is the distinguishing feature of harshness." The individual who has a harsh voice, continues Fairbanks, will "often overuse the extremely low pitches in their vocal areas, where maximum intensity is relatively low." The term 'strident', reports Curtis (1956), "is sometimes used to describe harsh tones of high pitch." Phonation is often initiated with "glattal attacks" and the existence of vocal fry, the result of ventricular fold vibration, is not unusual with harsh voice quality (Murphy, 1964). Perceived harshness varies among isolated vowels and vowels affected by certain consonant environments (Sherman and Linke, 1952; Rees, 1958).

Harshness is generally considered to be the result of excessive muscular tension, during voice production, throughout the entire laryngeal structure. Curtis (1956) reports that laboratory research has verified laryngeal fatigue in individuals with harsh voices "if they try to talk for any substantial length of time." Sherman and Jensen (1962), however, noted that there is a degree of harshness in normal voices. An acceptable amount of harshness was present in the voices of their subjects judged "not deviant in quality" prior to continuous oral reading. The degree of harshness decreased as the subjects with normal voices continued to read. The authors contributed the reduction to "certain muscular adjustments made by the subject which enabled him to continue for the entire period with decreasing physical discomfort and with increasing efficiency." A second experimental group was selected for Sherman and Jensen's study because their voices were "severe in degree of harshness" before oral reading. Except for the three "most severely harsh subjects", perceived harshness was not increased by continued vocal use.

Hoarseness is one of the most frequently used terms to describe the acoustic symptoms of voice pathology caused by underlying laryngeal disturbances. Yanagihara (1967a) states that considerable emphasis has been placed on hoarseness as being the "cardinal symptom of laryngeal diseases, and is often a sign of importance as a symptom of extralaryngeal involvement as well." Terms considered to be synonymous with hoarseness are husky, harsh, breathy, rough, rasping and strident, Fairbanks (1960) describes hoarseness as a combination "of the features of harshness and breathiness. The harsh element predominates in some hoarse voices, the breathy element in others, and the same kinds of variations may be heard within a given voice." Yanagihara (1967a) analyzed the acoustic properties of hoarseness with the spectrograph and determined that hoarseness was the interaction of three factors: (1) noise components in the main formants of each vowel, (2) high frequency noise components above 3000 Hz, and (3) the loss of high frequency harmonic components.

Hoarseness is the term used by Jackson and Jackson (1937) to characterize the condition known as myasthenia laryngis. One of Jackson's six causes of myasthenia laryngis is "muscular fatigue from prolonged, though not violent use" of the voice. According to Jackson and Jackson (1942), the greatest cause of laryngeal disease is excessive use of one of its normal functions--phonation. The authors mention the great variation in the amount of abuse the larynx of different individuals will stand, but each larynx has its limit. To go beyond this limit means thickening of the cords, and a thickened cord means a hoarse voice. The thickened cord is a poor vibrator and hence additional effort is required of the vocal muscles for phonation. Eventually, the vocal muscles are strained and weakened until they can no longer maintain the normal level of phonation. The result is myasthenis laryngis.

Each of the above terms has been derived a priori to describe the effect of prolonged vocalization. Sherman and Jensen's investigation (1962), alone, represents an empirical approach. These authors sought to induce harshness in the voices of their subjects by continuous oral reading. The listeners who judged the data perceived a significant decrease in the degree of harshness as oral reading continued. The reduction in harshness was evidently due to an alteration of the parameters of the voice which was not discerned by listeners. The induced effects of prolonged vocalization could be recorded and analyzed by scientific instruments.

The purpose of the present study was the preliminary investigation of the effects of prolonged vocalization on the acoustic spectrum of the speech signal.

Summary

In this chapter, a need was implied for a more objective means of defining and measuring the parameters of the pathological voice. Proposed was the objective specification of variations in the acoustic spectrum of speech signals resulting from prolonged vocalization. The purpose of the present pilot study was stated to structure the direction of this investigation.

Review of the Literature

While the purpose posed for the present investigation is relatively simple and straightforward, it is predicated upon the rather complex function of the larynx in the production of the normal voice, the physiology of abnormal laryngeal function related to the defective voice, and the method of analysis of voice. Therefore, a review of some of the literature pertinent to these areas is in order.

Mechanisms of Normal Phonation

Under normal circumstances, phonation is initiated when air forced from the lungs reaches the adducted vocal folds. Air pressure is built up beneath the vocal folds until the folds are forced apart and the lung air is released. The vocal folds, due to their elasticity and the reduction in air pressure, close and the cycle of vibration is repeated. "This alternating flowing and stopping, or slowing of the breath", states Moore (1957), "creates pressure changes in the air of the external atmosphere, and hense to the ear of the listener". To the listener, this phonated tone must satisfy certain criteria to be acoustically acceptable. According to Curtis (1956), the four primary elements constituting the adequate voice are: (1) appropriate loudness, (2) a pitch level appropriate to the sex and age of the individual, (3) a pleasant quality, and (4) the flexibility to make loudness and pitch changes. "An abnormality in the size, shape, tonicity, surface conditions and/or muscular control of the phonating and resonating mechanisms," Moore reports (1957), can result in "defects of pitch, loudness, and quality." A defect of one of these elements is a vocal quality disorder.

Tone generators. The normal laryngeal mechanism is involved in determining the pitch and intensity of the vocal tone (Moore, 1957; Levin, 1964; Greene, 1964; Luchsinger and Arnold, 1965; Van Riper and Irwin, 1965; Zemlin, 1968). During phonation, the vocal folds completely or nearly completely approximate medially to block the air coming from the lungs. The vocal folds can be likened to the "safety valve of a boiler," contends Moore (1957), "in a continuing attempt to hold constant intratracheal air pressure." When the intratracheal air pressure becomes great enough to separate the vocal folds, the puffs or rushes of air escaping through the glottis are the primary source of the vocal sound (van den Berg, 1958; Levin, 1964; Van Riper and Irwin, 1965). Zemlin (1968) notes that there is a direct relationship between the extent of adduction (by contraction of the interarytenoid, lateral cricoarytenoid, and posterior cricoarytenoid muscles) and the amount of intratracheal air pressure required to force the folds apart. Rubin (1960) emphasizes that the perfect phonation of any pitch depends upon the exact adjustment or balance between muscular tension and breath pressure. The air pressure between the vocal folds is reduced and the vocal folds are sucked together as the velocity of the air flowing through the glottis rapidly increases (Zemlin, 1968). This is accounted for by a aerodynamic law known as the Bernoulli effect, (van den Berg, 1958; Greene, 1964; Zemlin, 1968). The reduction in the subglottal air pressure, the Bernoulli effect, plus the elasticity of the vocal fold tissue, cause the vocal folds to snap back to their adducted position.

To date there is disagreement among proponents of two particular theories of voice production; the myoelastic-aerodynamic theory and the neurochronaxic theory (van den Berg, 1958). The neurochronaxic theory postulates that each vocal fold vibration is innervated by a new nerve impulse. The nerve impulse is transmitted from the brain to the laryngeal muscles by the recurrent branch of the vagus nerve (Greene, 1964). The myoelastic-aerodynamic theory, alluded to above, postulates that the vocal folds are moved together by the air stream from the lungs and the trachea (van den Berg, 1958; Greene, 1964). The frequency of vibration based on well-established aerodynamic principles, is dependent upon the length of the vocal folds in relation to their effective mass and stiffness (Moore, 1957; van den Berg, 1958; Greene, 1964; Zemlin, 1968). Each of the intrinsic and extrinsic laryngeal muscles is actuated to a delicate relationship with the other laryngeal muscles. The innervation of these laryngeal muscles is sustained so the vocal folds are held in a closed position. The subglottal air pressure increases until it separates them and keeps them vibrating. Van den Berg (1958) summarized his review of these two theories of voice by stating, "It is shown that the myoelastic-aerodynamic theory provides a straightforward explanation of all known phenomena of voice production, whereas there is no experimental evidence for the neurochronaxic theory and it is unable to explain a large number of phenomena."

Zemlin (1968) discusses, in some detail, the normal vibrating cycle of the vocal folds. The information comes largely from the analysis of high-speed motion pictures of the larynx. A complete vibratory cycle is composed of three phases: an opening phase, a closing phase and a closed phase. The opening phase occupied approximately 50 percent of the vibratory cycle. The closing phase occupied 37 percent of the total cycle. And the closed phase occupied the remaining 13 percent of the cycle. The vocal folds are separated from beneath. The lower edges of the folds leading the upper edges in an undulating manner. The folds close, again, with the lower edges leading the upper edges. The typical horizontal mode of vocal vibration begins posteriorly and the opening moves anteriorly. The posterior portion is the last to come together. Any change in the frequency or the mode of vocal fold vibration will alter both the pitch and spectral characteristics of the voice.

Fink (1962) analyzed the Bell Telephone Laboratories' "High Speed Pictures of the Human Vocal Cords" to determine the axis of vibration of the human vocal folds. The vibration of the vocal folds is not a simple harmonic motion. The surfaces adjoining the vocal ligament are involved in a complex motion that generates a multiple of sound frequencies, harmonically related to the frequency of the opening of the glottis. The tone produced by the glottal puffs is analogous to a siren. The opening and closing frequency of the glottal chink determines the pitch of the tone. The many harmonics, produced by the glottic surfaces, combine with the fundamental frequency and the entire complex is transmitted into the cavities above the vocal folds as the vocal tone. The axis of vocal fold vibration is paramedial for tone production, rather than at the midline of the vocal folds. For this reason, contact between the vocal folds is maintained for a much smaller fraction of the vibratory cycle than if the axis of vibration was in the midline. This arrangement minimizes laryngeal trauma to the closing vocal folds and reduces the amount of energy required to maintain vibration of the vocal folds.

The larynx is a complex structure capable of producing tones at any one of a vast number of pitches and intensities. Fairbanks (1960) proposed that each individual speaks at the pitch level (fundamental frequency) at which his voice is most efficient for speech. This 'natural level' is one-fourth of the way from the bottom of the total singing range of pitches that the individual's vocal mechanism can produce. Thurman

(1958) examined experimentally the procedure for estimating natural, or optimum, pitch level for 35 subjects and did not find a pitch level where intensity was more efficiently produced. Perkins and Yanagihara (1968) disagree with the concept of a 'natural' or 'habitual' pitch level for an individual vocal mechanism. They recorded 14 peaks of glottal efficiency scattered over low frequencies ranging from 135 Hz to 280 Hz. The peaks were considered to be too widely dispersed to support the idea of "optimum pitch level" and would seem to justify more a concept of optimum pitch range.

Pitch. Pitch level is usually changed by a delicate alteration in the dimensions of the vocal folds. Pitch change is primarily the result of modification in the glottal tension and mass in relation to the length of the vocal folds, and not the effect of an increase in intratracheal air pressure (Zemlin, 1968). Moore (1957) discusses the physical relationships of pitch:

> The heavier and more massive the cords, other factors constant, the slower they will vibrate and the lower will be the pitch of the voice. Conversely, the greater the elasticity of the cords, the quicker they will tend to return to their position of rest when disturbed by the breath stream; hence the faster the rate of movement and the higher the pitch. In two sets of vocal folds having the same cross section and the same tension, but of different length, the longer cord will vibrate more slowly. On the other hand, the lengthening-stretching-of a given set of vocal cords

> causes a decrease in cross-sectional area and an increase in elasticity. The latter changes offset the increased length and the frequency increases. If one or both cords have their mass increased by a growth, swelling, or other causes, the frequency decreases and the pitch is lower.

Hollien (1960a, 1960b) and Hollien, et. al. (1960, 1960, 1962) have studied by means of larynogoscopic photography, lateral x-ray and laminagraphic procedures, the anatomical and physiological factors associated with vocal pitch. Hollien (1960a) photographed the vocal folds of four groups selected so that they differed from one another as radically as possible in pitch level. The groups were composed of (1) six males with very low voices; group LM, (2) six males with very high pitched voices; group HM, (3) six females with very low pitched voices; groups LF, and (4) six females with very high pitched voices; group HF. The subjects were chosen on the basis of pitch range, age, absence of speech or voice problems, and their ability to produce specified vocal tones easily. The laryngoscopic photography procedure allowed the researcher to make measurements of the length of each subject's vocal folds under five conditions. One with the subject's vocal folds abducted and four of each subject's vocal folds during phonation at the 10, 25, 50 and 85 percent points to the nearest semitone of the subject's total pitch range. Hollien (1960a) concluded that the folds are very near maximum length in the abducted position, and substantially shorter when adducted for phonation. As the subject raised the fundamental frequency of phonation, the vocal folds systematically lengthen.

Hollien (1960b) presented data showing the relationship between certain measures which could be considered indices of laryngeal size and vocal pitch. A comparison was made between the data generated by the two sexes and among the individuals of the same sex of the groups in the study reviewed above (Hollien, 1960a). A standard lateral x-ray procedure was used to make four measurements of laryngeal dimensions (two antero-posterior, one vertical, and one area) to establish indices of laryngeal size. The results indicated a correlation between the laryngeal size and pitch level: the smaller the dimensions of the larynx, the higher the pitch. An additional finding was that while the laryngeal size differences between the high pitched male subjects and the low pitched females were equal to or very nearly equal to the size differences between the two male and female groups, the pitch differences were much less than those between the other groups. This suggested that when comparisons are made between the data of the two sexes, these are factors related to laryngeal size which are not completely correlated with the corresponding difference in pitch level.

In one of their studies, Hollien and Moore (1960) investigated the relationship to pitch of smaller variations in the length of the vocal folds. Six male subjects phonated the musical tones of C, E, and A within each octave from the lowest to the highest musical tones sustainable in their pitch range. Measurements were made from laryngoscopic photographs. They demonstrated that the length of the vocal folds at various pitches never exceeds the length give for the abducted position. The pitch seemed to increase systematically with the lengthening of the folds in a 'stair step' manner. The general vocal fold length appeared to bear a moderate relationship to the pitch level, but did not correlate with the absolute fundamental frequency being phonated. The changes in the length of the vocal folds during pitch change would seem to indicate that changes in mass and tension also play a role in the pitch-changing mechanism.

According to a laminagraphic x-ray study, by Hollien and Curtis (1960), the cross-sectional area (mass) or thickness of the vocal folds decreased as the pitch was raised. The investigators concluded from their measurements that there was a general relationship between vocal fold thickness and absolute frequency that transcended differences in the laryngeal anatomy between the pitch groups: LM, HM, LF and HF defined above. The decrease in cross-sectional vocal fold area was greatest at the lower frequencies of the subject's range and becomes proportionately less as the frequency rises.

Hollien and Curtis (1962) followed Hollien's (1960a) earlier procedure to study the (a) relationship between the elevation of the vocal folds and increases in vocal pitch, and (b) the relationship between the upward tilting of the vocal folds and the rise in vocal pitch. A laminagraphic x-ray procedure provided for coronal cross-sectional views of each subject's vocal folds. Measurements of the data obtained indicated that there is a progressive elevation of the vocal folds from low to high vocal pitches. There is also a tendency for the vocal folds to tilt, that is, the superior borders of the folds to slope upward toward the midline. This tilting becomes progressively greater with successive increases in vocal pitch for those in the falsetto register.

Zemlin (1968) deduces from the data of Hollien and Curtis (1960) that the vocal fold thickness (an index of mass) is never reduced below one-half the cross-sectional area of the lowest pitch of phonation. He speculates, thus, that pitch change cannot be accounted for by the reduction in mass alone. Tension must play an important part in the pitch-changing mechanism. Zemlin concludes, "...it is not unreasonable to suppose that an increase in tension of the vocal folds is the sole agent responsible for pitch increases and that the accompanying length and thickness change is simply the result of the elastic tissue of the vocal folds yielding to the marked increase in tension. This concept remains to be proven."

The appearance of the vocal folds changes as the pitch is raised. Van Riper and Irwin (1965) report that at low pitches the vocal folds seem to be relaxed and flaccid, and their edges are rounded and thick. As the pitch is raised, the opening and closing seems to be done with the stratified epithelium of the edges alone, and the folds appear to be thin and rigid. The opening between the folds, notes Zemlin (1968), changed from a narrow triangle in shape to a varible slit as the pitch increased, and only the medial edges undergo vibration.

Intensity. The adequate voice, according to Curtis (1956), is dependent upon the appropriateness of the pitch and loudness. The physical factors which are believed to determine the pitch level of the voice have been discussed in the previous paragraphs, with particular emphasis on the functioning of normal vocal folds. During normal phonation, the vocal folds are an important factor in the variations in intensity of the voice. Moore (1957) attributes the loudness of the voice to the pressure of the released pulsations. Van Riper and Irwin (1965) state that the greater the amplitude of the vibration of the vocal folds, the greater will be the pulsations released. The increase in the amplitude of the movement of the air molecules by the additional pressure expended at the vocal folds, causes a greater excursion of the eardrum, and hence a louder sound. Moore (1957) states, "Greater pressure is acquired through the delicate balance between increases resistance of the vocal cords and greater air flow."

Isshiki (1964) studied the relationship between the voice intensity (sound pressure level), the subglottal pressure, the air flow rate, and the glottal resistance. On a single subject he made simultaneous recordings of the sound pressure level of the voice, the subglottal pressure, the air-flow rate and the volume of air utilized during phonation. Isshiki found that the flow rate remained unchanged or even decreased slightly at very low frequency phonation, while the glottal resistance increased with the vocal intensity. In contrast to this, the flow rate on high frequency phonation was found to increase greatly, while the glottal resistance remained unchanged as the voice intensity increased. On the basis of the data, Isshiki concluded, the voice intensity at the very low pitches was controlled by the larynx. As the pitch was raised, there was less and less laryngeal control of voice intensity and increased expiratory muscle control until at very high pitches, voice intensity is controlled entirely by the flow rate. Charron (1965) tested Isshiki's conclusions by investigating glottal and breathing activity during phonation at various pitches and intensities. He employed laryngoscopic photography and electromyography to relate glottal area variations with the activity of the musculature of exhalation during phonation. His results strongly supported the conclusions of Isshiki (1964).

Zemlin (1968) reports Fletcher's comparison of the modes of vocal fold vibration of three subjects during phonation at moderate intensity and at five and ten decibels above the moderate level. The investigation of the internal laryngeal activity was accompanied with high-speed motion-picture photography. High-speed films of the larynx during a cresendo were also obtained. Two conclusions were apparent: the duration of the closed phase increases as intensity increases, and the glottal area remains essentially constant.

Rubin (1963) explained that the mechanisms of vocal pitch and intensity are so interrelated that to isolate one from the other, except for the most elementary considerations, is virtually impossible. In his study, Rubin demonstrated that vocal intensity may be raised by: (1) increasing air flow at constant cordal resistance (pitch), or (2) increasing cordal resistance at constant air flow. It was concluded that vocal loudness is determined by the balance between air flow and thyroarytenoid-cricothyroid tension. Air flow is increased directly by an increase in subglottal pressure and this results in an increase in sould pressure.

Thus, the above review underscores the normal laryngeal function in the production of an adequate voice. The vocal folds, caused to vibrate by the intratracheal air pressure, determine the pitch and intensity of the voice. Any change in the condition of one or both the vocal folds which prevents them from interrupting the air flow from the lungs for normal phonation will result in a quality disorder.

Vocal Abuse

Vocal abuse may be defined as the improper use of the voice as a result of too high a pitch, excessive air pressure against the under surfaces of the bands (West, Ansberry and Carr, 1957), excessive talking, prolonged vigorous use of the voice such as screaming and shouting (Moore, 1957), abrupt initiation of tone and production of strained sounds in play activity. Greene (1964) recognizes vocal abuse to be the continuation of vocal strain due to psychological problems of the individual. In the case where vocal abuse results in nodules or polyps on the vocal folds, she feels as does Heaver (1962), "that the patient uses the voice box as a natural, biological means of expressing a surcharge of hostile, aggressive impulses."

Wilson (1961) states that vocal nodules, contact ulcers, non-specific laryngitis, polyps, polypoid laryngitis and weakness of hypofunction of laryngeal structures are among the results of vocal abuse. These conditions interfere with the normal phonatory functioning of the vocal folds, and thus produce vocal quality disorders.

Levin (1962) notes that those patients who abuse their voices severly, "such as the barker at the county fair, the auctioneer and the top-sergeant are least likely to complain about their voices no matter how terrible they sound." Abuse is frequent in children who persist in yelling and shouting. Adults who work in noisy surroundings and must strain their voices to be heard will develop the pathologies of vocal abuse. However, Levin states, that the quantitative load a voice has to carry is less important than is generally assumed. The salesman in a busy store, the switchboard operator and the teacher must use their voices constantly for many hours. Levin concludes that normal vocal mechanisms can stand prolonged use very well.

Prolonged vocalization is considered to be one form of vocal abuse. As mentioned above, it is considered by various authors to be the cause of breathiness, harshness or hoarseness. But, there does not appear to be research explicitly designed to define the effect prolonged vocalization has on the voice. Sherman and Jensen (1962), however, studied the effect of oral-reading time on individuals with harsh voices and normal voices. The subjects were adult males, 15 with harsh voices and 15 with normal voices. Sherman and Jensen noted that there was a degree of harshness perceived in their subjects with normal voices. A certain level of harshness was present in the voices of subjects considered "not deviant in quality", while the exceeding of this acceptable level meant a "harsh voice". The subjects read continuously for one and one-half hours. Tape recordings of the subjects reading a standard passage were made prior to the oral reading, after 45 minutes of continuous oral reading, at the end of the reading period, and 30 minutes after the oral reading ended. The recordings were randomized and played backwards to a panel of observers who judged the degree of harshness for each sample on a seven-point equal-appearing interval scale. The observers ratings were used to obtain a degree of harshness for each sample. Sherman and Jensen concluded from their study "that the prediction may not be made that vocal abuse, if it is present in harsh voice production, produces physiological changes in the larynx sufficient to result in an increase of perceived harshness during one and one-half hours of oral reading." While there appeared to be no change in the degree of harshness in the harsh subjects, the degree of harshness perceived in the normal subjects decreased. For the normal voices, differences in perceived harshness were significant between first and second, first and third, and third and fourth readings. Perceived harshness consistently decreased during the period of oral reading and after the period of silence returned to approximately the original level of severity. These results would seem to indicate that prolonged vocalization does in some way affect the voice.

Methods of Appraisal

The experienced human ear is considered to be the best tool for recognizing vocal quality disorders. The diagnosis of vocal behavior, however, would appear to be much more accurate if it was observed and measured both visually and aurally. The sound spectrograph permits visual presentation of the three dimensions of acoustical structure of voice, frequency, intensity and time. For this reason, it has been used in the study of vocal quality disorders. It seems probable that this instrument could be used to study the effects of vocal abuse on the acoustic parameters of the voice, detected or possibly not detected by the human ear.

Thurman (1954) designed an experiment to establish phonographically recorded scales of severity for six voice quality disorders: breathy, nasal, hoarse, harsh, thin and strident. The sound spectrograph was employed to make acoustic analyses of certain of the recorded voice samples. Thurman discovered that in thin and breathy voices the second formant tended to be higher than normal, and in hoarse and harsh voices the first formant tended to be lower than in normal voices. In thin and breathy voices, the rise of the second formant was found to be indicative of judged severity. Differentiation between quality types was shown to be impracticable except as indicated above from the data collected by this sonagraphic technique.

Sawyer (1955) found that the spectrograph is applicable for distinguishing between the efficient and the inefficient voice. From the spectrograms of low-pitched male subjects, he concluded "that the characteristics that distinguish vocal efficiency (a normal voice) from inefficiency (a deviant voice) are lower frequency for the fundamental, more energy in formant one, more consistent appearance of formants three, four and five and a greater regularity and distinctness of the acoustic patterns."

Gunn (1960) investigated the characteristics of six quality disorders in sung vowels by comparing the voice quality ratings of judges with narrow-band spectrograms. The principal problem was to find the relationship between perception of a particular voice quality and variance in the following vowel characteristics on the spectrograms: (1) the fundamental frequency; (2) the frequency position; relative intensity and bandwidth of formants one, two and three; (3) the relative intensity at 250, 500, 1000, 2000, 2400, and 2800 Hz. The relationship between the voice quality ratings and the selected acoustic measures was tested by means of the Pearson product moment coefficients. The study concluded that the differences.in voice qualities investigated appear to be related to peculiar shifts in frequency position and energy distribution among the vowel formants. There is a positive relationship between listener ratings of the extent of the voice quality defect and the frequency position of formants one and two. The intensity seen at formants one and three increased and decreased directly with the judged severity of the quality deviation. The perceived intensity changes of the harmonics in the region of 2800 Hz increased with the severity as judged by the listeners and were directly related to observable characteristics of the spectrograms as follows: (1) head quality was directly related to the bandwidth of formants two and three; (2) nasality was associated a reduction in intensity of formant one combined with the increase in formant bandwidth; and (3) throaty quality can be recognized by shifts in the frequency position of formant three in the direction of the low frequencies.

The audible characteristics of the deviant voice were portrayed visually by the sound spectrograph in Gunn's study. While the distinction between different acoustic phenomena was not quantitatively approached, three parameters applicable to quantifying voice quality were used successfully. The qualities heard were seen to be shifts in frequency position and energy distribution among the vowel formants. The formant levels at which these aspects appeared were different for the particular voice quality where the same formant was affected in different voice qualities. The manner and/or the extent of change was also a distinguishing factor.

O'Brien (1960) studied, by means of subjective judgment and acoustical analysis, what kind of voice quality is produced by a larynx which has undergone any of the tissue changes classified as chronic non-specific laryngitis. Secondary goals were the exploration of the usefulness of the sound spectrograph for research on voice quality and the provision of objective data on which the definitions of some terms used to define voice qualities might be based. The subjects were nine males who had had a diagnosis of chronic non-specific laryngitis, and nine males who had not. Their voices were recorded while they read four sentences. Ten expert judges rated the voices on over-all quality and on breathiness, huskiness, hoarseness, harshness, raspiness, denasality, nasality, metallic tone, stridency, muffled tone and throatiness. Spectrograms were made of one sentence of the recorded samples, and oscillograms were made of one word. Four judges ranked the spectrograms on two physical features.

The spectrograms indicated that there is a greater tendency for the distribution of acoustic energy throughout the frequency range to shift abruptly every three or four vocal cord cycles. On spectrograms, these abrupt changes in the spectrum appear as shape changes in the darkness of the formants. The degree to which these shifts in spectrum occur bears a moderate degree of correlation with all the quality faults studied, except metallic tone and stridency, the same two qualities that did not distinguish the groups. O'Brien decided that no feature of the spectrograms or the oscillograms seemed to be associated with any particular voice quality characteristic.

Narrow-band spectrograms were studied to achieve the conclusions of Gunn's (1960) and O'Brien's (1960) studies. The frequency scanning filters of the spectrograph relay the intensity of the sound to the stylus and the stylus burns the spectrographic paper proportionate to the intensity magnitude. The differences were observable, but would be difficult to quantify. Amplitude quantification might best be performed upon contour spectrograms in a manner similar to Kersta's procedure (1962, 1965).

Nessel (1960) compared the spectra of pathologically-altered voices with those of normal voices, by means of a sound frequency spectrograph of high selectivity and wide range (sound-tracking principles according to Gruetzmacher). The main vowels were recorded on tape in a room with low reverberation. Analyses were carried out on the tape loops. In all, 478 spectrograms were evaluated for the characteristics of different types of pathological conditions of the voice. The main conclusion was that while the spectrum of normal voices contain practically no acoustic energy above 5 kc, the vast majority of cases of 'hoarseness' are marked by spectral criteria above 5 kc. The presence of additional noise components with definite frequency location in the upper range of the sound-frequency spectrum represent the distinguishing features.

Luchsinger and Faabrog-Anderson (1966) found the sound spectrograph was capable of detecting changes in the parameters of the voice not detected by the human ear. The researchers studied the registration of halting tones during phonation at constant pitch with electromyography of the cricothyroid muscle and spectra analysis of three singers. The sound traces were recorded at a tape speed of 19 cm/sec and played back on a spectrograph recording pitch and intensity by an ink-writer on continuously moving paper. In the beginning of the phonation a transitory, unsteady adjustment was observed, both in spectrograms and in the electromyograms, but was not heard by the human ear.

Isshiki, Yanagihara and Morimoto (1966) endeavored to establish an objective method for the diagnosis of hoarseness in parallel with the mechanism of hoarse voice production. The hoarse voice was discussed from two points of view--noise component in relation to harmonic component, and frequency variation. The sound spectrograph was used to study this relationship.

Four classifications of hoarse voice were determined according to the frequency region and the intensity of noise component relative to the harmonic component. The aperiodicity of the fundamental frequency was found to be closely related to the degree of noise components in the voice since an increase of noise component in voice would naturally intensify the frequency variation.

Yanagihara (1967a) demonstrated that the acoustical analysis of the voice, with the spectrograph, in conjunction with air flow measurement during phonation may lead to a better understanding of pathophysiological mechanisms in the production of hoarseness. Ultrahigh speed cinematographic analysis of the glottal area function, sound spectrographic analysis of voice, and measurement of the air flow rate during phonation were carried out on ten subjects having hoarseness. The results suggest that abnormal variation in the glottal area function has close correlation to the abnormal findings in the spectrogram and air flow rate.

In another study, Yanagihara (1967b) demonstrated with the spectrograms of subjects with pathological hoarse voice quality, a high correlation between the increase of noise components and the loss of harmonic components with judged severity as rated by experienced listeners. The severity of hoarseness perceived was found to be the result of these factors: (1) the noise component in the main formant of each vowel; (2) high frequency noise components above 3000 Hz; and (3) the loss of high frequency harmonics.

Yanagihara supplemented the data revealed by the sound spectrographic analysis with a synthetic study. Band-filtered noise was added to vowels produced by a normal male voice and recorded on an endless tape loop. Before mixing, the intensity level of the band-filtered noise was attenuated to a predetermined value that caused listeners to perceive hoarseness in the mixed tone. Thus the relative intensity levels of each vowel and the corresponding band noise were kept constant at the final stage of the synthesis. The quality and degree of hoarseness of synthetic sounds-vowels-band noise mixtures were judged by six otolaryngologists in a sound proof room. The results supported the conclusions determined by spectrographic analysis.

Werner-Kukuk, von Leden and Yanagihara (1968) performed objective measurements of the laryngeal function and voice on a patient with extensive leukoplakia and a bilateral carcinoma of the vocal cords before, during and after the course of cobalt 60 teletherapy. Ultrahigh speed motion pictures were used in comparison with measures of air flow rate and maximum phonation time achieved basically with a pneumotachograph and tape recorder and spectrograms used to evaluate the changes in the patients voice. The results of the aerodynamic and acoustic investigation were compared with the clinical findings and ultrahigh speed photographic studies. An analysis of this investigation demonstrates the effect of radiation on laryngeal physiology and stresses the value of aerodynamic and spectrographic studies for the diagnosis and prognosis of laryngeal disease. In some cases, the effects of teletherapy were evident in the aerodynamic and spectrographic material when no change seemed apparent in the laryngeal photographs.

Greiner, Dillenschneider, and Conraux (1968) reported that a combination of stroboscopic analysis with sound spectrography permitted a more accurate description of the lesions of the larynx due to trauma. They concluded that the sound spectrograph provided a more percise and objective control of the treatment of laryngeal lesions, and a valid means to compare changes in function as vocal re-education progresses.

Hecker and others (1968) induced stress in subjects by having them perform a task involving the addition of numbers. The subjects was required to read six meters and announce the correct sum of his readings together with a test phase. The experimenter could vary the amount of time the meter display was presented to the subject, and sometimes the time allowed the subject was reduced to the point where the subject failed. An incorrect response was not rewarded. For each of 10 subjects, numerous verbal responses were obtained while the subject was under stress and while he was relaxed. Contrasting responses containing the same test phase were assembled into paired-comparison listening tests. Listeners could identify the stressful responses of some subjects with better than 90 percent accuracy and of others only at chance level. The test phases from contrasting responses were analyzed with respect to the level of the subject's speech and the fundamental frequency. The effects of stress on other parameters were examined by comparing wide-band and narrow-band spectrograms of all test phases with the control condition. The results indicate that task-induced stress can produce a number of characteristic changes in the acoustic speech signal. Most of the effects of stress that were noted are related to the manner in which the glottal pulses are generated in the larynx. Stress influenced the amplitude of these pulses (level), the average rate at which the pulses are generated (fundamental frequency), the contour of fundamental frequency during an utterance, the shape or frequency spectrum of each pulse, the regularity in shape of successive pulses, and the initiation of glottal vibration following a voiceless interval. Other effects of stress are related to articulation. The durations of phonetic segments can be altered, and the precision with which articulatory targets for vowels and for consonants are reached can be affected. The utterances produced by the individual under stress usually exhibit only some of these effects.

Hecker explains that the acoustical characteristics associated with a given effect of stress for one individual may be quite different from those for another individual. For one person, an increase in a particular variable of the speech signal is indicative of stress; and for another person a decrease in the same variable is equally significant. As long as such effects occur in a consistent manner in most of the utterances of a given individual under stress, they could be used to predict whether other utterances by the same individual were also produced under stress. While the effects were consistent in some subjects, they were sporadic in still other individuals and very inconsistent in others.

Summary

In this chapter, an attempt has been made to define the rather complex function of the nonpathological larynx in the production of the vocal spectrum. The change in the physiological parameters of the larynx, due to vocal abuse, were noted with particular interest in the effect prolonged vocalization had on the acoustic sound signal. The use of the sound spectrograph, in studies examining the pathological and nonpathological voice, were reviewed to examine its application for interpreting acoustic data.

Method

For the purpose of generating tentative conclusions regarding the effects of prolonged vocalization on the visual display of the acoustic spectrum of speech, a small group of subjects was selected for study, methods of investigation devised, and the necessary materials and instrumentation arranged. These procedures are described in detail in the present chapter.

Subjects

Five adult males, 18 to 22 years of age, were selected among students enrolled in introductory speech classes at Kansas State University. Subjects were chosen who had no voice disorder, reported no history of laryngeal pathology, and who had received no formal training in speech or singing. Each subject was judged by the experimenter to have a normal voice. Additionally, each subject was required to exhibit ease in oral reading. The subject's were instructed not to abuse their voices during the three days prior to their participation in the experiment.

Method

The method utilized resembles that followed by Sherman and Jensen (1962) to determine whether perceived harshness varies as a function of oral reading time. The method of the present study was designed to demonstrate the effect of prolonged vocalization in the visual display of the acoustic spectrum of the speech signals of subjects with previously normal voices. This approach required that each subject read aloud continuously for two hours at his normal loudness level. Recordings of the subject's voice could then be made as he spoke a standard passage, a well-known sentence and six sustained vowel sounds (1) following one-half hour of silence prior to when the subject started the continuous and prolonged reading; (2) after every 30 minutes of oral reading; and (3) after one-half hour of silence following the two hours of prolonged vocalization. The procedure provided for two temporally consecutive speech samples after the initial one-half hour of silence and just before the beginning of the oral reading. The first sample was for the purpose of representing the voice after rest and before abuse. The second was a control. This provided a measure of the variance in the vocal spectrum known to exist between consecutive readings of the same material by an individual to which changes which were the result of the prolonged vocalization could be compared. These recording conditions were labled respectively as follows: (1) Pre-abuse A, (2) Pre-abuse B, (3) 30 minutes abuse, (4) 60 minutes abuse, (5) 90 minutes abuse, (6) 120 minutes abuse, and (7) Post abuse.

Narrow-band spectrograms were selected as the media of visual display for the recorded samples of the vowels, the sentence and selected sections of the standard passage.

Reading material. The reading material for the prolonged vocalization was selected and handled in a manner that paralleled Sherman and Jensen (1962). This reading material was selected to satisfy certain criteria: adequate length to occupy two hours of continuous oral-reading time, reading difficulty within the range of the subjects' reading abilities, and sufficiently interesting to retain the attention of the subjects. For this purpose, anecdotal material from Reader's Digest was pica-typed, double-spaced, and placed in a loose-leaf notebook.

Experimental speech samples. The standard passage sample used in this investigation was the Rainbow Passage (Fairbanks, 1960). The sample sentence was: You will make that line send (Koenig, Dunn and Lacy, 1946). The vowels selected for samples were /i/, /e/, /ɛ/, /æ/, /u/, and /a/. The order in which the speech samples were recorded on tape was: (1) the first paragraph of the standard passage; (2) the sentence; and (3) the sustained vowels. The subject was familarized with the standard passage and the sentence and was taught to produce the isolated vowels prior to the beginning of the initial one-half hour of silence.

Instruments

Recordings of the speech samples were made on an Ampex AG 500 tape recorder coupled to an Electro-Voice Model 664 microphone in an IAC Model 1202-ACT double walled room. The tape-recorded speech samples were analyzed with a Sona-graph Model

6061-A Sound Spectrograph on type B/65 sonagram paper. An Ampex Model 620 speaker unit was coupled to the monitor output of the tape recorder to permit the examiner to monitor the recorded speech sample as he manipulated the recording circuits of the Sonagraph.

Procedures

The subject was taken into the IAC room by the experimenter. He was familarized with the Rainbow Passage, the sentence, and taught to sustain the isolated vowels. The subject was then positioned before the microphone and the experimenter adjusted the record level while the subject practiced the standard passage. The face of the microphone was placed within two inches of the subject's mouth at an angle of incidence of 90°. This position was also maintained during the oral-reading and recording procedure. The recording levels were adjusted during the practice period so that the VU meter of the tape recorder peaked at zero for each subject during the practice reading period. This was taken as the subject's normal loudness level. Each subject monitored his vocal intensity level throughout the oral-reading and recording session by maintaining the needle of the VU meter approximately at zero. This procedure enabled the subjects to sustain a relatively constant vocal output level. The examiner was seated to the left, one foot behind the reader to instruct him to alter his voice level when necessary.

Each subject was informed of the steps of the procedure. The subject was instructed to remain silent for one-half hour. Recordings of the sample speech material were then made twice immediately before beginning the reading of the anecdotal material from Reader's Digest. After each half hour of continuous reading, the sample speech material was again recorded. A final recording of the standard passage, the sentence and the sustained vowels was taken after one-half hour of silence following the two hour period of oral reading. The subjects were asked to report how their throats felt and how their voices sounded to them at various times during the experiment.

The Sona-graph was employed to display for measurement, the speech parameters that were affected by the continuous oral reading. The principal problem was to determine the relationship between prolonged vocalization and variance in the following characteristics of the vocal spectrum: (1) the fundamental frequency; (2) the distribution of acoustic energy; (3) the presence of noise components in the main formants of each vowel; (4) noise components in the high frequency range; (5) loss of high frequency harmonic components; (6) shifts in the main formants of the vowels; and (7) any other acoustic parameter observed to be affected by prolonged vocalization. Narrow-band spectrograms were prepared of selected sections of the standard passage, the sentence and the sustained vowels. For a given subject, spectrograms representing the control condition (two recorded samples taken before oral reading) were critically compared with spectrograms of the same material recorded at each of the intervals described above. Differences in the spectrograms which exceeded the variance in the control spectrograms were attributed to changes in the mechanisms of phonation resulting from prolonged vocalization.

Results and Discussion

The purpose of continuous oral reading was to induce and measure change in the vocal spectrum of subjects with normal voices. The sound spectrograph was employed, in this preliminary investigation, to detect some of the speech parameters that may be affected by prolonged vocalization.

Measurements of Fundamental Frequency

Table 1 presents the results of the fundamental frequency measures tabulated to determine whether the subject's responses representing the control condition differed in this parameter following prolonged vocalization. The purpose of Table 1 was to demonstrate the increase in fundamental frequency after two hours of prolonged vocal use. Measures taken at 60 minutes and

TABLE 1

Mean Fundamental Frequency Measures

Trials	Subjects				
	A.L.	S.D.	L.R.	C.M.	J.O.
Pre-abuse A	192.5	166.0	128.5	106.0	99.0
Pre-abuse B	192.0	161.0	129.0	106.0	99.0
30 min. abuse	200.7	178.6	146.0	102.5	98.0
120 min. abuse	233.0	185.0	159.0	112.0	104.0
Post-abuse	215.7	178.0	125.0	115.0	100.5

90 minutes were omitted since they represented consistent intermediate rises in fundamental frequency between the 30 and 120 minute measures. Narrow-band spectrograms were prepared and used to measure the mean fundamental frequency of the six isolated vowels. The results for the isolated vowels were critically compared with the mean fundamental frequency of the same vowels in the context of the standard passage and the sentence of the experimental sample being examined and found to be representative. The measurements of the fundamental frequency were taken by (1) reading the frequency of the fifth harmonic and dividing the result by five, and (2) dividing 1 K Hz by the estimated number of harmonics in that frequency range. The 1 K Hz frequency range, which is a distance of one-half inch on a spectrogram, was measured vertically from the spectrograms' baseline. Where the 1 K Hz measure fell between two horizontal striations (harmonics), the distance between the last harmonic within the measured frequency range and the upper limit of the measure was approximated to the nearest tenth. The results were the same for the two methods used to determine the fundamental frequency. In a few instances the harmonic structure was either obscured by noise or too indefinite for measurement. However, there was no problem experienced determining the fundamental frequency. No measurement of fundamental frequency was determined from one speech sample alone. In Table 2, the means representing the control condition were subtracted from the means

TABLE 2

Mean Fundamental Frequency Differences

Trials	Subjects				
	A.L.	S.D.	L.R.	C.M.	J.O.
Pre-abuse B minus Pre-abuse A	0	-5	0.5	0	0
30 min. abuse minus Pre-abuse A	8.2	12.6	17.5	-3.5	-1.0
120 min. abuse minus Pre-abuse A	40.5	19.0	30.5	6.0	5.0
Post-abuse minus Pre-abuse A	23.2	12.0	-2.5	9.0	1.5

representing the condtions following prolonged vocalization after 30 minutes, 120 minutes, and the final half hour of silence.

Differences between the average fundamental frequencies among four experimental samples (the two experimental samples read before anecdotal reading, after one-half hour of reading, after two hours of reading, and after 30 minutes of silence) were readily apparent for each subject. Except for C. M. the fundamental frequency for each subject increased during oral reading and then decreased after the half hour of vocal rest. It is of interest that the fundamental frequency of subject C. M. decreased following the initial half hour of anecdotal reading, as did subject J. O.'s, but unlike subject J. O., his fundamental frequency reached its highest level after the period of silence. The fundamental frequency of subjects A. L., C. M. and S. D. did not return to the baseline level after the post abuse silence period. In comparison, subject L. R. and J. O. spoke at their approximate baseline level during the post abuse period. It appears that the effect of prolonged vocalization is dependent upon the individual's laryngeal mechanism. Both subject A. L.'s and subject S. D.'s initial pitch levels were higher than the other subjects. Hollien (1960b) indicated the correlation between laryngeal size and pitch level; the smaller the larynx the higher the pitch level. It is possible that slight changes in the length, tension, and/or mass of the vocal folds of smaller larynges could result in larger shifts of the fundamental frequency. Indeed, the size of the larynx may be a factor directly related to the degree of pitch level change due to prolonged vocalization and the observable residual of the effect after a half hour of vocal rest.

The increase in mean fundamental frequency, as shown in Figure 1, coincides with the reduction of "harshness" in the normal voices of Sherman and Jensen's study (1962). Contrary to Sherman and Jensen's prediction, the group of subject's with normal voices decreased in perceived harshness during the period of oral reading and after the period of silence returned to approximately the original level of severity. It should be

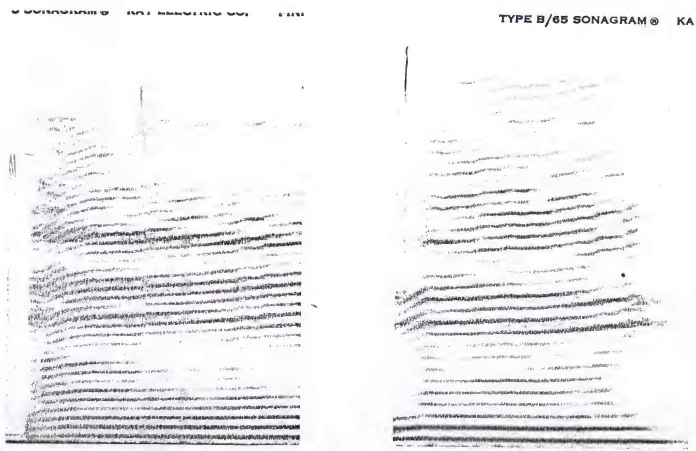

Fig. 1. Spectrograms of vowel /i/ spoken by subject A.L. in control condition Pre-abuse A (left) and experimental condition 120 minutes abuse (right).

noted that the fundamental frequencies of the present study, except for subject C. M., decreased after a half hour of silence. This might explain the increase in degree of harshness indicated by Sherman and Jensen (1962). It is notable that nowhere in the literature is there a direct examination of the relationship between fundamental frequency and the severity of harshness. Fairbanks (1960), however, has suggested that the therapist "experiment with connected reading at a slightly higher pitch level" than that ordinarily used by the harsh individual. "This will not harm his voice," states Fairbanks (1960), "and it often reduces harshness immediately." The results of the present investigation would seem to give some support to this rationale.

Under the procedure of the present investigation, a subject was required to maintain a constant level of intensity. It may have been that the fundamental frequency would have remained the same or have been lowered had the subject been allowed to speak at a confortable effort level.

Rees (1958) suggested direct examination of the relationship between fundamental frequency, duration and relative power of vowels in speech samples from persons with harsh voices. Her study was patterned after the investigation by Sherman and Linke (1952), in which the writers demonstrated that judgements of severity of harshness are vowel dependent. High vowels were perceived by listeners to be less harsh than low vowels. Rees (1958) postulated from Sherman and Linke's data that one of the acoustical characteristics responsible for the reduction of harshness in high vowels was their high fundamental frequencies. Rees' study also demonstrated that harshness was perceptually diminished for vowels in consonant environments with high fundamental frequencies. Since it has been postulated by Rees (1958) that harshness may be reduced for higher vowels and vowels with certain consonant environments because they have higher fundamental frequencies, then it would seem logical to hypothesize that connected speech would be affected in the same manner. In Sherman and Jensen's study (1962), the standard passage was judged to decrease in harshness as oral reading continued. This could conceivably have been due to an increase in fundamental frequency similar to that which occurred following prolonged vocalization in the present study.

The fundamental frequency rise possibly is due to the increased contraction of those muscles of pitch elevation. The vocal folds are thus tensed and elongated for the production of high-pitched tones. Rises in pitch may be accomplished by increases in subglottal pressure, though the present study did not examine this parameter. It is possible that the increase in muscular tension was an adjustment to vocal fatigue by the laryngeal mechanism. The subjects of the present study commented that their throats felt "tired" and "strained" and that "it was an effort for them to speak" during and after the oral reading period. To them, their voices were weak, unsteady and tended to break in certain spots (as shown in Figure 2). There was a tendency for the subjects to clear their throats constantly. They reported pain on swallowing, pain in the sides of their necks, and that their throats became sore. A possible interpretation of this information could be that the voice was produced by a fatigued and strained laryngeal mechanism--the result of prolonged vocalization. Phonation, under this condition, would thus appear to be contraindicated if the laryngeal mechanisms are to return to their normal state as rapidly as possible.

Sherman and Jensen (1962) assumed that the reduction in perceived harshness might be accounted for by muscular adjustments as a response to the speaking situation which worked against the factor of vocal misuse. This does not appear to be the case. As indicated by the results reported in Table 3 for subject S.D. (these data appear to be representative of the upward shift in fundamental frequency for the other subjects), there would have appeared to have been a continual increase in muscular tension in order to support a rise in fundamental frequency which could possibly mean a corresponding increase in vocal misuse. West, Ansberry and Carr (1957) noted that vocal abuse may be defined as the improper use of the voice as a result of too high a pitch. As a result of prolonged vocalization, each subject in the present study appears to have been induced to speak at continually higher frequencies above his normal pitch

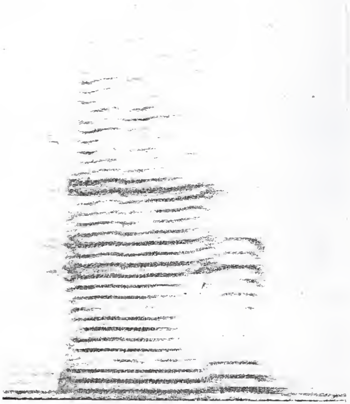

Fig. 2. Spectrogram of voice break at end of vowel /i/ spoken by subject A.L. following 60 minutes abuse.

TABLE 3

Fundamental Frequency Values (in Hertz) of Six Vowels

Trials	Vowels					
	/i/	/e/	/ɛ/	/æ/	/u/	/o/
Pre-abuse A	185	167	159	156	164	154
Pre-abuse B	189	167	159	156	164	152
30 min. abuse	196	189	169	167	182	169
60 min. abuse	204	192	176	176	189	176
90 min. abuse	196	196	192	185	196	179
120 min. abuse	200	190	189	176	185	172
Post-abuse	192	172	169	169	196	169

level. Following the half hour of vocal rest the subjects reported that their voices sounded "husky", "raspy", "rough", and "hoarse". The subjects were again interviewed a few hours following the experiment and they perceived their vocal quality as even worse. If such is the case, the effect of vocal abuse, which may be defined as prolonged vocalization at too high a pitch, would begin to appear as a characteristic audible vocal quality disorder sometime after the conclusion of continuous oral reading. The results of the present investigation indicated a downward trend in fundamental frequency with rest. Unfortunately, post experimental samples were obtained at 30 minutes rest only, and only two subjects were observed to return to baseline levels.

With respect to several studies made by Wendahl (1963, 1966a, 1966b), harsh voice quality was demonstrated to result from abrupt cycle-to-cycle frequency variations (jitter) or amplitude variations among successive glottal impulses (shimmer). An electrical laryngeal analog was used to generate stimuli which emanated jitter and shimmer. Of particular interest is Wendahl's study (1963) dealing with the degree of aperiodicity required for listener judgments of roughness. The results showed that even very slight frequency variations, as little as $\pm$ 1 Hz around the median fundamental frequency of 100 Hz, sounded rough or harsh. While frequency variations of this magnitude apparently can be perceived by the human ear, they cannot be displayed by the sound spectrograph. This might also explain the increase in perceived harshness heard by Sherman and Jensen's (1962) listeners after vocal rest, but could not be supported by the method of spectrographic study employed in the present investigation.

A second result of Wendahl's investigation (1963) would seem to further verify the proposed relationship between increased pitch and reduced harshness. A programmed amount of jitter was always judged to be less rough at a 200 instead of a 100 Hz median fundamental frequency. Thus, it was anticipated that the higher pitched of two voices, having equal frequency variations, would sound least rough. It is possible, however, that the reduction in harshness is not completely due to a higher fundamental frequency. The assumed increase in vocal fold tension could prevent both jitter and shimmer and in this way reduce audible roughness.

Other Spectral Features

The fundamental frequency is only one of the speech parameters that may be affected by prolonged vocalization. The spectrograms were examined for other spectral features that could be attributed to this form of vocal abuse. It was impossible, from the amount of data collected to recognize consistent shifts in the frequency position, variation in the bandwidth, or energy redistribution among the main formants. Observations of interest were made with some subjects in certain portions of the speech samples, however, that might be demonstrated with more consistency in a larger sample. While the fundamental frequency increased, the position of the main formants appeared to remain relatively constant. The weak, irregular harmonic components present in the high frequency areas above 4000 Hz were intensified by use of the high frequency pre-emphasis circuit of the sound spectrograph. There did not seem to be a change in the amount of high-frequency energy in the glottal pulses.

The quality and personal characteristics associated with voiced vowels largely depend upon the intensity, shape and periodicity of the glottal waves (Werner-Kukuk, von Leden and

Yanagihara, 1968). With the use of a narrow filter setting, the sound spectrograms of the subject phonating the isolated vowels prior to prolonged vocalization appear to have basic and common characteristics regardless of the difference in vowel sounds, which may represent an aspect of vibratory behavior of the vocal folds. Within the range of main formant frequencies, the transverse striations corresponding to fundamental frequency and harmonic frequencies are regularly spaced without showing abrupt and apparent initiation or discontinuation. There may not be any notable additional sound components in the space between each transverse striation.

From inspection of the narrow-band spectrograms, there did not seem to be indication of (1) noise components in the main formants, (2) high frequency noise components above 300 Hz, or, (3) a loss of high frequency harmonic components throughout the speech samples. It appears that hoarseness was not induced during the experimental period according to these criteria. One manifestation of prolonged vocal use that appeared for each subject, however, was an increase in the amount of noise components at the beginning and end of the isolated vowels and the beginning of the sentence (Figure 3 shows this effect). The vowels, in particular, were initiated with the harmonic components replaced or dominated by noise. The effect might be termed a "glottal attack" (Murphy, 1964). Werner-Kukuk, von Leden and Yanagihara (1968) attributed the generation of noise

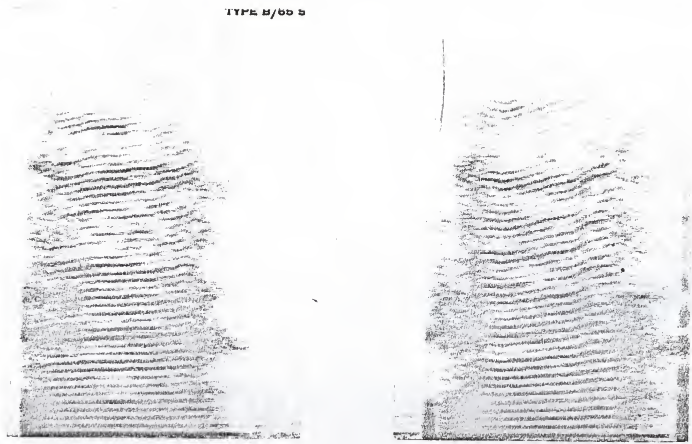

Fig. 3. Spectrograms of vowel /æ/ spoken by subject S.D. in control condition Pre-abuse A (left) and experimental condition 60 minutes abuse (right).

to the imperfect modulation of the air stream and subsequent turbulence in the flow of air at the level of the glottis. The vocal folds, in effect, were apparently not vibrating immediately in response to the air stream for vowel production. There also appeared to be a premature discontinuation of voicing at the end of the vowels. The onset of the harmonic components, following this initial region of noise, were usually aperiodic in frequency at as low as 500 Hz. The aperiodicity appeared to be reduced during the medial one-half of the vowel sound, but reappeared before the end of the vowel, especially if the vowel ended with noise.

Another effect of prolonged vocalization that seemed to occur in the spectrograms of the subjects is described below and illustrated in Figure 4. The normal irregularity experienced in the higher frequency harmonics appeared to shift downward. The fundamental frequency and the harmonic compostion of the main formants did not seem to be affected, but in many cases the components of formants four and five were. It is possible that the vocal quality of the speech signal was changed in that manner.

Summary

The main purpose of the present study was the preliminary investigation of the effects of prolonged vocalization on the acoustic spectrum of the speech signal. The terms normally

Fig. 4. Spectrograms of vowel /ɛ/ spoken by subject A.L. in control condition Pre-abuse B (left) and experimental condition 120 minutes abuse (right).

used to describe voice disorders follow the onset of laryngeal pathology, whereas in this investigation pathology was induced so that the specific effects of prolonged vocalization on the voice could be studied directly.

Sherman and Jensen (1962) had sought to increase the degree of listener perceived harshness during a period of continual vocal use, in the speech signals of both normal and harsh voices. The voices of adult males, with normal and harsh voices, were recorded before, during and a half hour following continuous oral reading. The recordings were judged by listeners as to degree of harshness. There was no evidence that harsh voices, in general, typically show any change in severity of perceived harshness after a period of one and one-half hours of oral reading. Conversely, normal voices decreased in degree of perceived harshness during the same period of oral reading, and returned to their original level of severity at the end of one-half hour of silence following the reading period. The authors concluded, that if vocal abuse was present, it did not produce physiological changes in the larynx which would result in an increase of perceived harshness. The possible cause for the reduction of "perceived harshness" and insight into the laryngeal mechanism of phonation during prolonged use were tentatively established by the present study.

The present investigation employed the sound spectrograph to display for measurement, some of the speech parameters that were affected by continuous oral reading. Five male subjects, 18 to 22 years of age, with normal voices were selected to read orally at their normal loudness level for two hours following an initial half hour of silence. Recordings were made of the subject's voice as he spoke a standard passage, a sentence and sustained six vowels (1) following a half hour of silence prior to oral reading, (2) after every 30 minutes of vocal use, and (3) after one-half hour of silence following the period of prolonged vocalization. Narrow-band spectrograms were then prepared for inspection and analysis from the recordings. An attempt was made to relate the results derived by the Sona-graphic technique with changes in perceived vocal quality.

The results of this exploratory study indicate that there are a number of potential effects of prolonged vocalization on the acoustic speech signal:

There seems to be an increase in the fundamental frequency of each subject's voice during prolonged vocalization and a decrease in fundamental frequency after a half hour of vocal rest. It was postulated from the data of this study that the rise in pitch could possibly result from increased laryngeal muscular tension as an adjustment of the mechanism of vocal misuse. In turn, prolonged phonation at too high a pitch is considered a form of vocal abuse (West, Ansberry and Carr, 1957). It is possible that this increase in fundamental frequency was a manifestation of two hours of continual oral-reading at a constant intensity level. The subjects' fundamental frequencies may have remained the same or become lower had they been allowed to select their own level of vocal intensity. The results suggest a need for further investigation into the physiological and acoustic characteristics of prolonged vocal use.

Additional relevant observations were made as follows, although their appearance was not sufficiently consistent in all the speech samples collected on any one subject to offer more than suggested avenues for further study:

1. On some occasions, for each subject, there were noise components in the initiation and discontinuation of the vowels and in the initiation of the sentence, followed and preceeded by irregularity in the harmonic structure.

2. An increased irregularity in the harmonics in the higher frequency range and in formants four and five was noted inconsistently with each subject.

3. Harshness, hoarseness or breathiness did not appear to be present in the acoustic spectrum of the subjects' speech signals from the limited amount of data collected.

It also seems possible to hypothesize that there may be a correlation found between a rise in fundamental frequency and a decrease in perceived harshness in the normal voice with prolonged vocalization.

REFERENCES

Berg, Jw. van den. Myoelastic-aerodynamic theory of voice production. Journal of Speech and Hearing Research, 1958, 1, 227-242.

Berry, Mildred F., and Eisenson, J. Speech disorders: Principles and practices of therapy. New York: Appleton-Century-Crofts, 1956.

Charron, R. An instrumental study of the mechanisms of vocal intensity. Umpublished Master's Thesis, University of Illinois, 1965.

Curtis, J. F. Disorder of voice. In Johnson, W., Brown, S. F., Curtis, J. F., Edney, C. W., and Keaster, Jacqueline. Speech handicapped school children. New York: Harper and Row, 1956.

Fairbanks, G. Voice and articulation drillbook. (2nd ed.) New York: Harper and Row, 1960.

Fink, B. R. Adaptation for phonatory efficiency in the human vocal folds. Annals of Otology, Rhinology and Laryngology, 1962, 71, 79-85.

Greene, Margaret. The voice and its disorders. London: Pitman, 1964.

Greiner, G. H., Dillenschneider, E., and Conraux, C. Stroboscopic and sonagraphic aspects of the traumatized larynx. French Journal of Otology, Rhinology and Larynology, 1968, 17, 237-241.

Gunn, G. H. An acoustical analysis of quality variations in sung vowels. (Doctoral dissertation, University of Michigan) Ann Arbor, Mich.: University Microfilms, 1960. No. 60-6874.

Heaver, L. Psycho-somantic aspects of non-verbal communication. Logos, 1962, 5, 60.

Hecker, M. H. L., Stevens, K. N., Bismarck, G. von, and Williams C. E. Manifestations of task-induced stress in the acoustic speech signal. Journal of the Acoustical Society of America, 1968, 44, 993-1001.

Hollien, H. Vocal pitch variation related to changes in vocal fold length. Journal of Speech and Hearing Research, 1960, 3, 50-56. (a)

Hollien, H. Some laryngeal correlates of vocal pitch. Journal of Speech and Hearing Research, 1960, 3, 52-58. (b)

Hollien, H., and Curtis, J. F. A laminagraphic study of vocal pitch. Journal of Speech and Hearing Research, 1960, 3, 361-371.

Hollien, H., and Curtis, J. F. Elevation and tilting of vocal folds as a function of vocal pitch. Folia Phoniatrica, 1962, 14, 23-36.

Hollien, H., and Moore, P. Measurements of the vocal folds during changes in pitch. Journal of Speech and Hearing Research, 1960, 3, 157-165.

Isshiki, N. Regulatory mechanism of voice intensity variation. Journal of Speech and Hearing Research, 1964, 7, 17-29.

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

Jackson, C., and Jackson, C. L. The larynx and its diseases. Philadelphia: Saunders, 1937.

Jackson, C., and Jackson, C. L. Diseases and injuries of the larynx. New York: Macmillan, 1942.

Kersta, L. G. Voiceprint infallibility. Paper presented at the meeting of the Acoustical Society of America, Seattle, Washington, November 7, 1962.

Kersta, L. G. Voiceprint classification. Bell Telephone Laboratory Report, June 5, 1965.

Ladefoged, P. N. A phonetic study of West African languages. London: Cambridge University Press, 1964.

Levin, N. M. Voice and speech disorders: Medical aspects. Springfield, Ill.: Charles C. Thomas, 1962.

Luchsinger, R., and Arnold, G. E. Voice-speech-language. Belmont, Calif.: Wadsworth, 1965.

Luchsinger, R., and Faaborg-Andersen, N. Phonetic and electromyographic registration of halting tones. Folia Phoniatrica, 1966, 18, 91-97.

Moore, P. Voice disorders associated with organic abnormalities. In L. E. Travis (Ed.), Handbook of speech pathology. New York: Appleton-Century-Crofts, 1957.

Murphy, A. T. Functional voice disorders. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Nessel, E. The spectrum of sound frequencies of the pathologically altered voice. Acta Oto-laryngology, 1960, 157, 1-45. (Supplement)

O'Brien, A. O. Some acoustic concomitants of chronic nonspecific laryngitis. (Doctoral dissertation, Columbia University) Ann Arbor, Mich.: University Microfilms, 1960. No. 60-5835.

Perkins, W. H., and Yanagihara, N. Parameters of voice production: I. Some mechanisms for the regulation of pitch. Journal of Speech and Hearing Research, 1968, 11, 246-267.

Rees, Maryjane. Some varibles affecting perceived harshness. Journal of Speech and Hearing Research, 1958, 1, 155-168.

Rubin, H. J. Vocal pitch and intensity. Laryngoscope, 1963, 73, 973-1015.

Sawyer, G. M. An experimental study of perceived differences in efficient and inefficient voice production in low-pitched male voices by acoustic spectrography. Unpublished doctoral dissertation, University of Southern California, 1955.

Sherman, D., and Jensen, P. J. Harshness and oral-reading time. Journal of Speech and Hearing Disorders, 1962, 27, 172-177.

Sherman, D., and Linke, E. The influence of certain vowel types on degree of harsh voice quality. Journal of Speech and Hearing Disorders, 1952, 17, 401-408.

Thurman, W. L. The construction and acoustic analyses of recorded scales of severity for six voice quality disorders. Speech Monograph, 1954, 21, 219-220.

Thurman, W. L. Frequency-intensity relationships and optimum pitch level. Journal of Speech and Hearing Research, 1958, 1, 117-123.

Van Riper, C. Speech correction: Principles and methods. (4th ed.) Englewood Cliffs, N.J.: Prentice-Hall, 1965.

Van Riper, C., and Irwin, J. V. Voice and articulation. (2nd ed.) Englewood Cliffs, N. J.: Prentice-Hall, 1965.

Villarreal, J. J. Consistency of judgments of voice quality. Southern Speech Journal, 1949, 15, 10-20.

Wendahl, R. W. Laryngeal analog synthesis of harsh voice quality. Folia Phoniatrica, 1963, 15, 241-250.

Wendahl, R. W. Some parameters of auditory roughness. Folia Phoniatrica, 1966, 18, 26-32. (a)

Wendahl, R. W. Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. Folia Phoniatrica, 1966, 18, 98-108. (b)

Werner-Kukuk, E., von Leden, H., and Yanagihara, N. The effects of radiation therapy on laryngeal function. Journal of Laryngology and Otology, 1968, 82, 1-15.

West, R., Ansberry, M., and Carr, A. The rehabilitation of speech. (3rd ed.) New York: Harper and Row, 1957.

Wilson, D. K. Children with vocal nodules. Journal of Speech and Hearing Disorders, 1961, 26, 19-26.

Yanagihara, N. Hoarseness: Investigation of the physiological mechanisms. Annals of Otology, Rhinology and Laryngology, 1967, 76, 472-488. (a)

Yanagihara, N. Significance of harmonic changes and noise components in hoarseness. Journal of Speech and Hearing Research, 1967, 10, 531-541. (b)

Zemlin, W. R. Speech and hearing science. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

THE EFFECT OF PROLONGED VOCALIZATION ON THE VOCAL SPECTRUM

by

RUSSELL ERNEST HEDGE

B. S., The University of Kansas, 1968

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

The main purpose of the present study was the preliminary investigation of the effects of prolonged vocalization on the acoustic spectrum of the speech signal. The terms normally used to describe voice disorders follow the onset of laryngeal pathology, whereas in this investigation pathology was induced so that the specific effects of prolonged vocal use on the voice could be studied directly.

The sound spectrograph was employed in the present investigation to display for measurement some of the speech parameters that were affected by continuous oral reading. Five male subjects, 18 to 22 years of age, with normal voices were selected to read orally at their normal reading loudness level for two hours following an initial half hour of silence. Recordings were made of the subject's voice as he spoke a standard passage, a sentence and sustained six vowels (1) following a half hour of silence prior to continuous oral reading, (2) after every 30 minutes of vocal use, and (3) after one-half hour of silence following the period of prolonged vocalization. Narrow-band spectrograms were then prepared for inspection and analysis from the recordings. An attempt was made to relate the results derived by the Sona-graphic technique with possible changes in perceived vocal quality.

The results of this exploratory study appear to indicate that there are a number of potential effects of prolonged vocalization on the acoustic speech signal:

There seems to be an increase in the fundamental frequency of each subject's voice during prolonged vocalization, and a decrease in fundamental frequency after a half hour of vocal rest. It was posulated from the data of this study that the rise in pitch could possibly result from increased laryngeal muscular tension as an adjustment of the laryngeal mechanism to vocal misuse. In turn, prolonged phonation at too high a pitch is considered a form of vocal abuse. There seems to be an increase in the fundamental frequency of each subject's voice during prolonged vocalization and a decrease in fundamental frequency after a half hour of vocal rest. The results suggest a need for further investigation into the physiological and acoustic characteristics or prolonged vocal use.

Additional relevant observations were made as follows, although their appearance was not sufficiently consistent in all the speech samples collected on any one subject to offer more than suggested avenues for further study:

1. On some occasions, for each subject, there were noise components in the initiation and discontinuation of the vowels and in the initiation of the sentence, followed and preceeded by irregularity in the harmonic structure.

2. An increased irregularity in the harmonics in the higher frequency range and in formants four and five was noted inconsistently with each subject.

3. Harshness, hoarseness or breathiness did not appear

to be present in the acoustic spectrum of the subjects' speech signals from the limited amount of data collected.

It also seemed possible to hypothesize that there may be a correlation found between a rise in fundamental frequency and a decrease in perceived harshness in the normal voice with prolonged vocal use.