University of Tennessee, Knoxville

TRACE: Tennessee Research and Creative Exchange

Masters Theses

Graduate School

8-1997

Vocal jitter in trained and untrained female voices

Christopher Somers Gaskill

Follow this and additional works at: https://trace.tennessee.edu/utk_gradthes

Recommended Citation

Gaskill, Christopher Somers, "Vocal jitter in trained and untrained female voices. " Master's Thesis, University of Tennessee, 1997.
https://trace.tennessee.edu/utk_gradthes/10526

This Thesis is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Masters Theses by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a thesis written by Christopher Somers Gaskill entitled "Vocal jitter in trained and untrained female voices." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Speech Pathology.

Patrick J. Carney, Major Professor

We have read this thesis and recommend its acceptance:

Carl Asp, Dolly C. Davis

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Christopher Somers Gaskill entitled "Vocal Jitter in Trained and Untrained Female Voices." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Speech Pathology.

Patrick J. Carney, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of the Graduate School

Vocal Jitter in Trained and Untrained Female Voices

A Thesis
Presented for the
Master of Arts Degree
The University of Tennessee, Knoxville

Christopher Somers Gaskill
August 1997

ACKNOWLEDGMENTS

There are many people without whose generous assistance this project would not have been possible. I thank my advisor Dr. Patrick J. Carney for his commitment to excellence in this project, his thoroughness with every step in the process and his provision of valuable time and departmental resources. I also am indebted to my other committee members, Dr. Carl Asp and Mrs. Dolly Davis, for their vital input and continual support toward the success of this project.

I owe thanks to several other faculty and staff members who graciously offered me their time, resources and encouragement. Thanks to Dr. Lori Swanson for the use of her laboratory and equipment to collect my data. Dr. Sam Burchfield offered valuable advice for my equipment selection and experimental set-up, as well as continual moral support. Dr. James Thelin provided me with genuine encouragement throughout this project and expert assistance in preparing my project proposal for institutional review. Max Dawson was invaluable in his patient and timely help with all of my equipment needs. Kay Gross was a constant source of optimism and faith in the ultimate success of this project, and efficiently managed the technical details of completing my degree.

Above all, I thank God for His faithfulness and for providing me with the strength and resources to reach my goal.

ABSTRACT

This study sought to determine the differences in vocal fundamental frequency perturbation, or jitter, between two groups of female subjects. Ten trained female singers and ten female speakers with no formal training or singing experience sustained an /ɑ/ vowel for four seconds at two levels of fundamental frequency--the subject's speaking fundamental frequency (SFF) and an octave above SFF--and three levels of intensity: 60, 70 and 80 dB SPL measured at a microphone-to-mouth distance of 12". For each vowel token, intensity and frequency were monitored to maintain experimental criteria, and the vowel samples were digitally analyzed to obtain vocal jitter expressed as relative average perturbation (RAP) in percent. Mean jitter values were compared across frequency, intensity and group factors using an analysis of variance technique.

The trained singers demonstrated no significant change in mean jitter from SFF to one octave above SFF, or from 70 dB to 80 dB. In contrast, the untrained singers demonstrated a significant increase in vocal jitter with frequency, and a significant decrease in jitter with intensity. Both subject groups had increased jitter magnitudes and variability at the 60 dB condition, especially at the lower fundamental frequency, which is consistent with the conclusions of other investigators. The major conclusion of this study is that the trained singers produced stable phonation regardless of frequency or intensity, while the untrained subjects' phonation stability varied significantly with changes in frequency and intensity. This finding suggests a difference between trained and untrained voices in the vocal dynamics of frequency and intensity control.

TABLE OF CONTENTS

CHAPTER ONE: OVERVIEW OF STUDY
I. INTRODUCTION 1
II. REVIEW OF THE LITERATURE
Definition and Measurement of Jitter 2
Jitter Measurement Techniques 3
Factors Affecting Jitter
Jitter as a Function of Gender and Fundamental Frequency 5
Jitter as a Function of Vowel 8
Jitter as a Function of Age 9
Jitter as a Function of Intensity 10
Differences Between Trained and Untrained Voices 14
III. PURPOSE 21

CHAPTER TWO: METHODS
I. SUBJECTS 22
II. PROCEDURES 23

CHAPTER THREE : RESULTS
I. RELIABILITY
Intrasubject Reliability 26
Jitter Measurement Reliability 26
Signal Digitization Reliability 27
II. RESULTS 27
First Analysis of Variance 31
Group Difference in Mean Jitter 31
Group Difference in Mean Jitter with Changes in Intensity 31
Group Difference in Mean Jitter with Changes in Frequency at Each Intensity 31
Second Analysis of Variance 35
Group Difference in Mean Jitter 38
Group Difference in Mean Jitter with Changes in Intensity 38
Group Difference in Mean Jitter with Changes in Frequency at Each Intensity 41

CHAPTER FOUR: DISCUSSION 42

CHAPTER FIVE: SUMMARY AND CONCLUSIONS 53

REFERENCES 56

APPENDICES.....66

VITA.....70

LIST OF FIGURES

1. Mean jitter for each experimental condition.....29

2. Mean jitter as a function of frequency comparing the three different experimental intensity levels of 60, 70 and 80 dB SPL.....33

3. Mean jitter as a function of the three experimental intensities (60, 70 and 80 dB SPL) comparing the 10 trained versus 10 untrained voices.....34

4. Mean jitter as a function of frequency comparing the medium (70 db) and loud (80 dB) conditions at SFF and one octave above SFF for all subjects in both groups.....37

5. Mean jitter as a function of the two experimental frequencies (SFF and one octave above SFF) comparing trained versus untrained voices, without the data for the soft condition (60 dB) used for analysis.....39

6. Mean jitter as a function of intensity comparing 10 trained versus 10 untrained voices for the medium (70 dB) and loud (80 dB) conditions.....40

7. Mean jitter for each experimental condition, illustrating the reduction in variability in mean jitter as intensity increases.....51

CHAPTER ONE: OVERVIEW OF STUDY

I. INTRODUCTION

Fundamental frequency perturbation, or jitter, is defined as small cycle-to-cycle differences in vocal fundamental frequency, and is a natural acoustic feature of all human phonation (Titze, 1994). Many factors, including fundamental frequency, intensity and different vowels have been investigated for their effects on jitter. Investigators have reported a consistent inverse relationship between jitter and vocal intensity, with larger perturbations occurring at low intensities (Gelfer, 1995; Glaze, 1990; Jacob, 1968; Orlikoff & Kahane, 1991; Pabon, 1991). Jitter also tends to decrease with increasing fundamental frequency (Horii, 1979; Orlikoff & Baken, 1990). An interaction between frequency and intensity has also been reported, with the highest levels of jitter found at low-frequency, low-intensity conditions (Gelfer, 1995; Jacob, 1968; Pabon, 1991).

Jitter is often defined as an index of vocal fold vibration stability (Baken, 1987; Fukazawa et al., 1988; Heiberger & Horii, 1982; Koike, 1973; Lieberman, 1963; Ludlow et al., 1983). Although jitter has been investigated as an indicator of the instability in phonation caused by various vocal pathologies (Fukazawa et al., 1988; Koike et al., 1977; Laver, 1992; Ludlow et al., 1987; Ludlow et al., 1983; Murry & Doherty, 1980; Wolfe et al., 1991), little is known about what changes may occur in jitter as a result of vocal training. It is possible that vocal training may increase the stability of vocal fold vibration and thereby reduce the level of jitter in the voice. The purpose of this study was to investigate the the difference between trained and untained voices in the amount of vocal jitter. Specifically, the amount of vocal jitter for the two subject groups was compared at two different frequency and three different intensity levels.

II. REVIEW OF THE LITERATURE

Definition and Measurement of Jitter

Human phonation is a dynamic and constantly variable process and has been investigated in many different ways (Fant, 1960; Moore & von Leden, 1958; Scripture, 1906). Lieberman (1961) examined oscillographic displays of speech and reported almost constant changes in fundamental frequency during speech. In 1963, he suggested the use of what he referred to as a "pitch perturbation factor" for the detection of certain laryngeal pathologies after observing increased aperiodicity in the pitch of speakers with vocal fold lesions. This factor was defined as the percentage of perturbations that were 0.5 msec or greater. The magnitude of this perturbation factor increased with certain vocal pathologies (Lieberman, 1963). He concluded that vocal lesions interfered with the biomechanical and aerodynamic forces of phonation at the level of the glottis and increased the aperiodicity of vocal fold vibration. Other investigators have subsequently confirmed that increased amounts of perturbation in the voice signal are associated with a rough quality (Askenfelt & Hammarberg, 1986; Fukazawa et al., 1988; Horii, 1979; Koike, 1973; Koike et al., 1977; Ludlow et al., 1987; Murry & Doherty, 1980; Wolfe et al., 1991).

Titze (1994a) defined jitter as a perturbation in fundamental frequency, which is a momentary change, in contrast to vocal tremor and vibrato, which are long-term fluctuations in frequency and amplitude. A perturbation is defined as a disturbance that leaves the ultimate output in a steady state. A fluctuation, however, means that the output constantly oscillates between two distinct states, never converging upon a single steady state. A fluctuation in fundamental frequency will be perceived by a listener as two distinct

frequencies, while frequency perturbation will be perceived as roughness or a harsh quality.

Jitter Measurement Techniques

Investigators have typically used sustained vowels, rather than speech as in Lieberman's studies (1961, 1963), for the measurement of acoustic perturbations of the voice (Horii, 1979). Intentional and unintentional frequency variations found in normally inflected speech and at the moments of voice onset and offset confound the accurate measurement of minute, unintentional changes in vocal output (Baken, 1987; Karnell, 1991b).

Measurements of jitter rely on highly accurate measurements of fundamental frequency, and recent improvements in computer technology have made frequency and jitter measurement using digital equipment the preferred method of most researchers (Deem et al., 1989; Gould & Korovin, 1994). One method for analyzing vocal jitter involves recording the airborne acoustic signal from the subject's voice and either taping the sample or directly digitizing it for analysis by a computer program designed to extract fundamental frequency and jitter values.

Two other methods of fundamental frequency (F_O) extraction and jitter measurement are the accelerometer and the electroglottograph (EGG) (Baken, 1987). These both employ the use of either a contact microphone or electrodes placed externally at the thyroid cartilege. Some researchers have used the EGG or accelerometer to extract F_O and measure jitter because of the simple waveform it provides (Casper, 1984; Haji et al., 1986; Horii, 1982; Koike, 1973; Koike et al., 1977; Orlikoff, 1995; Orlikoff & Baken, 1990; Orlikoff & Kahane, 1991; Sussman & Sapienza, 1994). However, others have concluded that these methods are unsuitable for clinical and research purposes due to limitations

of the reliability of electrode or contact microphone placement and variation of neck tissue density (Bough et al., 1996; Karnell et al., 1995).

In comparing EGG and accelerometer measurements to acoustic methods, very slight differences have been reported (Orlikoff, 1995; Horii, 1982). Other investigators have shown that jitter values are not significantly different using either acoustic or non-acoustic methods (Haji et al., 1986; Koike et al., 1977; Sussman & Sapienza, 1994). Given the limitations of EGG and accelerometer meaurements and the advent of digital acoustic recording techniques, most investigators now prefer acoustic methods for jitter measurement (Bielamowicz, 1996; Brown et al., 1989, 1990; Gelfer, 1995; Glaze et al., 1990; Hecker & Kreul, 1970; Heiberger & Horii, 1982; Jacob, 1968; Ludlow et al., 1987, 1983; Pabon, 1991; Ramig & Rengel, 1983; Sabol et al., 1995; Sorensen, 1983, 1984; Wilcox & Horii, 1980). Futhermore, several investigators have reported that direct digitization of acoustic voice samples provides the most accurate perturbation values, and suggest using directly digitized versus tape recorded voice samples for acoustic perturbation analysis (Deem et al., 1989; Doherty & Shipp, 1988; Gelfer & Fendel, 1995; International Association..., 1992; Perry et al., 1996).

There are two major classifications of jitter indices: absolute and frequency-related (Baken, 1987). Absolute jitter indices do not compensate for the fundamental frequency of the vocal signal, while frequency-related jitter indices compensate for the negative jitter-frequency correlation. This second type of jitter index is obtained by dividing by the mean period of the signal, resulting in a measure of jitter as a ratio or percent. This allows more accurate comparison of jitter magnitudes between different fundamental frequencies of a given signal.

Absolute jitter indices include perturbation factor (Lieberman, 1963) and directional perturbation factor (Hecker & Kreul, 1971; Sorensen & Horii, 1984). Some frequency-related jitter measures include jitter ratio (Horii, 1979; Jacob, 1968), jitter factor (Hollien et al., 1973) and relative average perturbation (RAP) (Koike, 1973). Each of these indices compensates for the negative jitter-frequency correlation by essentially computing a ratio (which takes various forms) of mean perturbation to mean period. In his discussion, Baken (1987) emphasized that Hollien, Michel and Doherty (1973) and Horii (1979) reported that any such compensation with a ratio of jitter to frequency is, at best, a compromise; no completely "uninfluenced" jitter index seems possible. Nonetheless, investigators have typically reported frequency-related jitter values, such as those mentioned above, or simply "percent jitter" (e.g. Brown et al. 1989, 1990; Gelfer, 1995; Horii, 1979, 1982; Ludlow et al., 1983; Orlikoff & Kahane, 1991; Pabon, 1991). In addition, two commonly used commercial voice analysis systems employ Koike's (1973) RAP (Bielamowicz et al., 1996; Karnell, 1991a).

Factors Affecting Jitter

Jitter as a Function of Gender and Fundamental Frequency

Many researchers have reported that male voices have higher levels of jitter than female voices (Casper, 1984; Higgins & Saxman, 1989; Ludlow et al., 1987; Milenkovic, 1987; Nittrouer, 1990; Sussman & Sapienza, 1994). Data reported from single-gender studies with male voices (Horii, 1982; Ramig & Rengel, 1983; Wilcox & Horii, 1980) and female voices (Brown et al., 1989, 1990) also indicate that male voices exhibit higher jitter than female voices. A few researchers, however, have reported increased jitter in female voices over male voices (Deem et al., 1989; Sorensen & Horii, 1983).

Higher jitter in male than female voices is consistent with the finding that the longer periods (lower frequencies) found in men's lower voices, are associated with higher degrees of frequency variability. Investigators have reported that greater amounts of short-term variability in fundamental frequency, or jitter, are associated with longer fundamental periods (Horii, 1979; Koike, 1973; Lieberman, 1963). Stated in terms of frequency, or the inverse of period, jitter tends to decrease with increasing fundamental frequency. Several researchers have reported that higher amounts of jitter are associated with low vocal fundamental frequencies (Gelfer, 1995; Jacob, 1968; Koike, 1973; Lieberman, 1963; Pabon, 1991; Verstraete et al., 1993).

Given the fundamental frequency differences in male and female voices, the relationship between jitter and gender must also be examined in terms of fundamental frequency. The interrelationship between gender and frequency creates some difficulty in jitter measurement. There is some indication that using an jitter index such as jitter ratio or percent jitter overcorrects for the jitter-frequency relationship (Horii, 1979), especially for higher (e.g., female) voices (Orlikoff & Baken, 1990).

Orlikoff and Baken (1990) reported jitter data as a function of frequency for both male and female voices. Measurements were made at specific musical semitone intervals above and below each subject's self-selected comfortable fundamental frequency. They reported no significant difference in absolute jitter between males and females across the range of experimental frequencies. Their analysis revealed a small but significant negative correlation between mean absolute jitter and F_0 for the subjects as a whole. However, when male and female jitter data were analyzed separately, the

male F_0's (below 150 Hz) did show a linear relationship with jitter, while the female F_0's (above 150 Hz) did not.

Furthermore, both male and female mean absolute jitter levels were negatively correlated more with relative F_0 (i.e. semitone level above or below their comfortable pitch) than with absolute F_0. Females had much less change in jitter as a function of F_0 than did men. When jitter ratio (JR) was calculated and analyzed in relation to F_0 for all voices, using JR resulted in a small but significant overcorrection in the levels of jitter in female voices, which is contrary to the purpose of utilizing a jitter ratio.

Orlikoff and Baken (1990) concluded that: (1) overall average jitter values between males and females do not significantly differ when measured over a significant portion of their frequency ranges; (2) jitter is related more to relative fundamental frequency within overall frequency range than to absolute frequency; and (3) the jitter-F_0 relationship is nonlinear over the entire frequency range, with jitter decreasing quasi-exponentially as a function of frequency. Therefore, the jitter in female (higher) voices is influenced much less by increasing fundamental frequency than in male (lower) voices. They suggested confining the use of jitter factor or jitter ratio to male voices, since these indices distort the jitter data in female voices. Newman and Emanuel (1991) also reported a difference in vowel roughness (in terms of signal-to-noise ratio, which is a correlate of jitter) in terms of relative pitch within overall pitch range, in agreement with Orlikoff and Baken's (1990) conclusions.

To summarize then, an adult male and an adult female with normal voices producing a pitch at the same relative point within their overall frequency range should have similar jitter ratios. That is, the amount of

instability relative to the location within their pitch range is equivalent, while absolute jitter (which is a function of absolute F_0) should be higher for the male voice since the actual F_0 is lower than the female's F_0. However, as Orlikoff and Baken (1990) reported, the use of jitter ratios is counter-productive as F_0 increases in female voices. Orlikoff & Baken (1990) suggested that investigators use easily defined and replicable pitches within a subject's overall range for making jitter measurements.

Jitter as a Function of Vowel

A review of the existing literature regarding jitter differences for various vowels reveals inconsistent findings. Some researchers have reported that jitter differs as a function of vowel (Deem et al., 1989; Lieberman, 1963; Sorensen & Horii, 1983 and 1984; Sussman & Sapienza, 1994; Wilcox & Horii, 1980), while others have reported no significant differences in jitter between vowels (Casper, 1984; Gelfer, 1995; Horii, 1982; Orlikoff, 1995). However, the jitter differences reported for different vowels may be due to the different intrinsic pitches of the various vowels, and not due to intrinsic differences among the vowels themselves.

For example, Orlikoff's (1995) study comparing eight English vowels in a group of ten males and ten females reported very consistent amounts of jitter for all the vowels when F_0 and intensity were controlled for each subject. He noted that the previous studies that have reported vowel differences for vocal jitter did not adequately control fundamental frequency. The jitter data were therefore reflecting a frequency dependence based on the intrinsic "comfortable" pitch chosen for each vowel, and not an actual vowel dependence. Orlikoff (1995) also stated that there was little evidence to support any supralaryngeal effects on vocal phonation stability. If jitter is to be

used as a measure of phonatory function, it must by definition be impervious to supralaryngeal effects. Heiberger and Horii (1982) also claim that there is no physiologic reason to expect jitter to be vowel-dependent.

Jitter as a Function of Age

Age is another factor that has been investigated for its effects on vocal jitter. In a study of twenty young (mean age 23.3 years) and twenty older (mean age 69.8 years) adult males with healthy voices, Wilcox and Horii (1980) reported significantly greater average jitter during sustained vowels for the older subject group. They concluded that their data were consistent with previous research that supported the existence of perceptual, structural and functional changes in the aging voice. Other researchers, however (Brown et al., 1989 and 1990; Ramig & Rengel, 1983; Casper, 1984), did not report any differences in vocal jitter as a function of age.

Ramig and Rengel (1983) also studied 48 male subjects, grouping them into three age ranges (25-35, 45-55, and 65-75) and into groups for good and poor physical condition. While there was a significant effect of physical condition on jitter during maximum sustained phonations, there were no significant jitter differences as a function of age or physical condition during sustained phonations of a comfortable duration.

Brown, Morris and Michel (1989 and 1990) analyzed sustained phonations by young and aged females, while Casper (1984) reported vocal jitter data for both males and females in three age groups, similar to Ramig and Rengel (1983). None of these researchers reported any significant age effect for measurements of vocal jitter. While previous researchers have concluded that there are significant changes in the human vocal mechanism with advancing age (Wilcox & Horii, 1980; Ramig & Rengel, 1983), there are

insufficient data at this time to support a significant effect of age on vocal jitter.

Jitter as a Function of Intensity

Several researchers have reported a consistent negative correlation between jitter magnitude and vocal intensity level, with jitter decreasing as intensity increases (Gelfer, 1995; Glaze et al., 1990; Jacob, 1968; Orlikoff & Kahane, 1991; Pabon, 1991). There is also evidence that a significant interaction between intensity (SPL) and fundamental frequency (F_0) exists, with the highest levels of jitter reported at the low-SPL, low-F_0 condition (Gelfer, 1995; Jacob, 1968; Pabon, 1991). Since the jitter-intensity relationship is of particular interest in this study, three of these studies will be reviewed below in more detail.

Jacob (1968) reported normative data on jitter in normal voices to use as a standard of comparison when analyzing pathological voices. Fifteen male and fifteen female adults with normal voices sustained the vowel /ɑ/ at three frequencies at each of three intensities. The frequencies selected were at the 25th, 50th and 75th percentiles of each subject's overall modal frequency range. The intensity levels selected were the softest and loudest phonations each subject could produce at each of the three frequencies, and the intensity at the midpoint between these extremes.

All recordings were made in a sound-treated booth and the voice signals were tape recorded for later analysis. Subjects matched their productions with a reference frequency, and were instructed to sustain the vowel as steadily as possible. The subjects were also instructed to monitor their intenisty with a sound-level meter for each specified intensity level. A minimum of five seconds of phonation time was recorded for each vowel.

Jacob (1968) reported the highest level of jitter ratio (10.86) at the lowest experimental intensity, with a moderate trend for jitter ratios to decrease with increasing intensity. Above the moderate intensity level, the changes reported in jitter were small. He also reported a similar relationship between jitter ratio and fundamental frequency percentile, as well as a significant interaction between frequency and intensity, with the highest jitter levels reported at the condition with the lowest frequencies and intensities. Jacob (1968) hypothesized that with vocal training, his subjects (who reported specific difficulty sustaining the soft phonation) might have been able to better control phonation at the soft intensities more accurately, and therefore produce sustained vowels with less jitter.

Orlikoff and Kahane (1991) investigated the relationship between intensity and both jitter and shimmer, which is a correlate to jitter indicating the cycle-to-cycle variability in amplitude of vocal vibration, or perceptually, loudness of the phonation. They used ten adult males with normal voices for the study, measured jitter in percent using an EGG method. Each subject sustained /ɑ/ at three intensities: 60-68 dB (soft), 70-78 dB (moderate) and 80-88 dB (loud). Subjects monitored their loudness level with a sound-level meter and were instructed to try not to elevate pitch in producing the loud phonation. Phonations of at least ten seconds were obtained, with all but the initial and final one second of phonation used for analysis.

Orlikoff and Kahane (1991) reported mean percent jitter across intensity levels of 0.436%, which is comparable to values reported for normal adult males at or near a "comfortable" F_0 (Hollien et al., 1973; Horii, 1982; Ramig & Rengel, 1983). Their analysis of the jitter data as a function of intensity revealed a significant ($r = -0.87$) negative correlation between jitter and

intensity, with jitter decreasing with increasing intensity. In addition, the differences in jitter between each SPL level were statistically significant ($p = .001$). There were no significant differences in F_0 used by the subjects at the three intensity levels. These data follow the trend seen in Jacob's (1968) data for his male subjects, but are lower in jitter magnitude overall. This is due most likely to the very different measurement methods employed by the two studies.

Orlikoff & Kahane (1991) explained their results in terms of complex relationships between glottal airflow and vocal fold tension, stating that measures of perturbation like jitter are indicators of the relative balance or imbalance of multiple forces occurring at the level of the glottis. This reiterates the notion that jitter is considered to be an acoustic index of the physiologic stability of vocal phonation. Orlikoff and Kahane (1991) offer two possible explanations for the increased jitter at the soft phonations. First, for the "soft" (low SPL) condition, they indicated that subjects had more difficulty sustaining the phonation steadily, similar to Jacob's (1968) observations. Second, they indicate that some research has shown that soft phonation is associated with a longer glottal open quotient and shorter vocal fold contact duration, which could cause increased vocal turbulence and greater randomness in the vocal fold vibration.

Gelfer (1995) studied 29 adult females with normal voices. Her study examined jitter, shimmer and signal-to-noise ratio (SNR) as functions of vowel, fundamental frequency and intensity. Each subject phonated both /i/ and /ɑ/ at three different intensities (60, 70 and 80 dB, ± 5 dB) and at two frequencies: speaking fundamental frequency (SFF) (determined from a reading passage), and one octave above SFF. All vowels were tape-recorded

in a quiet room using a head-worn microphone 1/2" from the subject's mouth. Subjects matched the target frequencies as presented with a pitch pipe and monitored intensity with a sound-level meter positioned 12" from the subject's mouth. A three-second sample of each phonation was recorded and a 1500 msec central sample was extracted for computer analysis. Data for jitter were reported in percent, as in Orlikoff and Kahane (1991).

Gelfer (1995) also reported that the highest jitter levels were associated with the lowest intensity condition, with jitter generally decreasing with incresing intensity. There was a significant frequency-intensity interaction, with the jitter at the low-frequency, low-intensity condition significantly greater than the jitter at any other frequency-intensity condition. Her results were generally consistent with both Jacob (1968) and Orlikoff and Kahane (1991), and the jitter data for the three intensities at the subjects' SFF are comparable in magnitude with Orlikoff and Kahane's (1991) jitter data for male voices at an unspecified, comfortable pitch.

Gelfer (1995) hypothesized that as both frequency and intensity increase, the muscular forces in the vocalis and cricothyroid muscles of the larynx become more balanced, making the phonatory system more robust to minute changes that could lead to increased fundamental frequency perturbations. In contrast, at the low-frequency, low-intensity condition, the activity of the cricothyroid muscle is low, so any normal fluctuations in the activity of the vocalis muscle could lead to more irregularities in vocal fold vibration, e.g. jitter. She also noted that it is likely that the SFF's for her subjects were quite low within their overall frequency ranges, which, given the nature of her reported data, agrees with the notion that jitter is related to relative pitch within a subject's range (Orlikoff & Baken, 1990).

Differences Between Trained and Untrained Voices

Studies by Bartholomew (1934) and Wolf, Stanley and Sette (1935) are examples of early attempts to quantifiably describe the voices of trained singers. Rubin and LeCover (1967) measured vocal intensity, subglottic pressure and airflow relationships in singers to enhance the field of vocal pedagogy. They reported significant differences in the efficiency of vocal production between two highly trained and one untrained singer. They were reluctant to draw firm conclusions, and advocated for the scientific study of voice students before and after training. To date, this type of valuable longitudinal research has not been performed.

Scientific understanding of the singing voice has advanced tremendously in the twentieth century with the work of individuals who have sought to combine the disciplines of voice science and vocal pedagogy (Appelman, 1967; Miller, 1986; Sundberg , 1987; Titze, 1994; Vennard, 1967). Yet there remains a need for improved understanding of the human voice and its ability to achieve the artistic demands placed on trained singers, who some have referred to as "the vocal equivalents of Olympic athletes" (Carrol et al., 1996).

Evidence which supports fundamental and quantifiable differences between trained and untrained voices has been reported by a number of investigators (Awan, 1991; Carroll et al., 1996; Griffin et al., 1995; Howard, 1995; Howard et al., 1990; Lindqvist, 1970; Gelfer et al., 1991; Rubin et al., 1967; Sulter et al., 1995, 1996; Sundberg, 1973; Titze, 1995; Titze & Sundberg, 1992). These investigators have examined vocal features such as the efficiency of phonation, various glottal dynamics such as glottal closure time, and phonetogram profiles. Some of their findings are summarized here.

Differences in vocal efficiency, or the economy of the transfer at the glottis from aerodynamic energy to acoustic energy, have been reported for trained and untrained voices. Lindqvist (1970) and Sundberg (1973) reported significantly increased glottal efficiency for trained singers as evidenced by a more robust source spectrum across intensity and frequency. That is, trained singers produced an acoustic signal that was richer in harmonics across the frequency and intensity range than that of untrained singers. A study by Carroll and others (1996) measured various parameters of glottal dynamics for a group of trained singers. Significant differences in mean flow rate and mean phonation time were reported for this group in comparison to published data for untrained voices. The authors concluded that trained voices were different enough from untrained voices to warrant the collection of specific clinical norms for phonatory function testing.

Titze & Sundberg (1992) also reported data that suggest that a greater degree of vocal efficiency is found in trained singers. Trained singers were reported to have greater amplitudes of vocal fold vibration for the same amount of subglottal driving pressure than did the untrained group, suggesting less energy loss in the vocal fold tissue and consequently more efficient phonation by the trained singers. Titze and Sundberg (1992) hypothesized that the trained singers were somehow able to lower their effective glottal impedance during singing.

Several recent studies have examined glottal closure times during a vibratory cycle and made comparisons between trained and untrained subjects (Howard, 1990; Howard et al., 1995; Sulter et al., 1996). Each of these investigators reported consistent and significantly longer glottal closed phases during phonation for trained singers compared to untrained singers.

Evidence was obtained from acoustic and videostroboscopic data. However, the subject selection criteria used in two of these studies were not stringent. Howard's (1995) study included some subjects with more than casual singing experience in the untrained group, and Sulter et al. (1996) defined the trained subjects as those having merely two years of choral singing experience as their vocal training.

Griffin and others (1995) have also reported an increase in glottal closure in trained singers' voices during well-supported singing when compared to the same singers' voices when asked to sing incorrectly and without proper support. This finding suggests that training in singing gives singers the ability to exert conscious control over vocal output in order to increase efficiency of the voice.

With regards to the purposes of the present study, an increased glottal closure quotient (CQ) also typically occurs as intensity is increased, and could be an explanation for the decrease in jitter values with increasing intensity reported in the literature (Gelfer, 1995). If trained singers do indeed exhibit an increased glottal CQ, then perhaps this change in glottal dynamics will also be manifested as increased vocal stability, measurable as decreased jitter magnitude and a more stable jitter profile as a function of intensity.

Voice researchers have also examined differences in phonetograms between trained and untrained singers. The phonetogram is a graphic representation of the intensity range of the voice as a function of the fundamental frequency range. Sulter, Schutte and Miller (1995) analyzed the phonetograms of 224 male and female subjects, both trained and untrained, for overall area (i.e., total frequency range plotted as a function of intensity range) and size of individual ranges of frequency and intensity. The trained

subjects had significantly larger phonetogram areas, mainly due to an extended frequency range at softer intensities. The trained group was not stringently defined in this study (minimum of two years weekly vocal lessons and/or choir rehearsals), making it difficult to generalize these results.

Awan (1991) reported similar phonetogram differences for twenty trained and twenty untrained singers. The two groups were clearly distinct: the trained singers had at least two years of formal classical training and were currently receiving training at the time of the experiment, while the untrained singers had no formal singing training or experience. Awan (1991) also reported that the trained singers demonstrated greater dynamic variations in intensity for each 10% increase in frequency level, which he interpreted as evidence that vocal training improved laryngeal dynamics and efficiency.

Simon (1927) examined the variability of pitch in vocal and instrumental sounds and reported that trained voices exhibited less cycle-to-cycle frequency variation than untrained voices. Some investigators have reported jitter data in relation to vocal training (Brown et al., 1990; Ferrand, 1995; Gelfer et al., 1991; Gelfer et al., 1991; Murry et al., 1979; Sabol, 1992; Sabol et al., 1995; Teachey et al., 1991). However, the differences between trained and untrained voices in terms of fundamental frequency variability as suggested by Simon (1927) have yet to be demonstrated.

Brown, Morris and Michel (1990) studied the voices of twenty-five aged (mean 72 years) female singers with long histories of vocal training and experience and reported jitter ratios for sustained /ɑ/. When compared to their data for aged and young nonsingers (Brown et al., 1989), the aged singers showed the lowest levels of jitter, followed by the young nonsingers and then

the aged nonsingers. However, none of these differences reached statistical significance.

In a study of the effects of prolonged reading on both highly trained singers and untrained subjects (Gelfer et al., 1991), acoustic analysis was perfomed during sustained vowel and singing tasks both before and after one hour of continuous reading aloud at 80% of the subject's maximum intensity level. The jitter values reported for sustained vowels were in general lower for the trained group in the pre-test condition, and the loud reading task had a much more detrimental effect on the untrained subjects' voices than on the trained singers' voices, suggesting enhanced vocal production efficiency for the trained singers.

The value of vocal training for improved vocal performance was investigated by Teachey, Kahane and Beckford (1991). The vocal technique of thirty professional singers with little or no formal training was evaluated and acoustic measures of voice were obtained. As a group, the subjects showed higher percent jitter than the published norms, which the authors attribute to the high incidence of various vocal pathologies discovered in this untrained group. In addition, the singers' jitter levels were inversely related to the judged efficiency their technique; that is, subjects with the poorest vocal technique showed the highest jitter magnitudes. These results must be interpreted with caution given the high incidence of vocal pathologies in the subject group. The authors noted that the poor technique itself may have been the biggest factor leading to the vocal pathologies observed, since good vocal technique by its nature produces an efficient voice that is less susceptible to pathologic changes resulting from faulty use.

Ferrand (1995) examined the effects of practice both with and without knowledge of the results on levels of jitter on subjects without any training in singing. The subjects showed significant decreases in jitter magnitude with practice in a sustained vowel task, with the group without knowledge of results showing stronger carryover after one week. Ferrand (1995) suggests that the practice may have resulted in changes in motor unit recruitment in the laryngeal musculature, causing a greater smoothing of the overall neurologic firing for muscle contraction, as described by Titze (1991). This study does seem to suggest that the human voice is capable of reducing the inherent level of jitter present. The values for baseline in this study (0.682% and 0.674%) are comparable with reported values of jitter in normal adults (cf. Horii's (1982) reported 0.660%) while the levels after two practice sessions are considerably lower (0.407% and 0.538%).

Sabol (1992, 1995) specifically addressed the issue of singing training and its effect on objective measures of vocal output. Twenty graduate voice majors with a minimum of four years of formal vocal training served as subjects and were divided into experimental and control groups with equal numbers of men and women. Objective voice measurements, both aerodynamic and acoustic, were made initially, and then the experimental group incorporated a specific set of vocal function exercises as described by Stemple (1993) into their regular voice practice every day, twice a day, for four weeks. The control group made no changes in their practice regimen during this time. All voice measurements were made again in four weeks, and the experimental group showed significant improvements in aerodynamic voice measures such as flow rate and maximum phonation time, suggesting an increase in glottal efficiency. However, the experimental group's average

jitter values actually increased after four weeks, while the control group's jitter level decreased. This finding is puzzling given the reported improvements in other areas of glottal efficiency. Sabol (1992) only discusses these data briefly, saying that since no norms for jitter in the singing voice exist and the changes observed in jitter for both groups were slight, the changes observed may have been due to daily variation in frequency perturbation. More research would be needed to draw any firm conclusions regarding the effects of these specific vocal exercises (Stemple, 1993) on vocal stability and jitter.

The difference in jitter between trained and untrained singers has been studied by Murry, Large and Dalgaard (1979), but the results of the study are difficult to interpret due to insufficient control of subject and task variables. Four adult females (ages 26-37) with extensive singing training were compared with five males (ages 55 to 71) on sustained vowels. The trained group produced four sung and one spoken tokens of /ɑ/, varying both pitch and singing mode (straight tone versus vibrato), while the untrained group produced a single spoken token of /ɑ/. Two different frequency perturbation measurements were obtained, magnitude perturbation factor (MPF) (Lieberman, 1961, 1963) and directional perturbation factor (DPF) (Hecker & Kreul, 1971). Perceptual judgments of vowel roughness of each token were also made. The results reported suggest possible differences between the two groups for MPF, but the experimental design does not lend itself to adequate interpretation of the results.

II. PURPOSE

The purpose of this study was to determine if trained singers have smaller values of jitter than subjects with untrained voices, and if jitter is affected differently for trained and untrained voices as intensity is varied. The previously reported frequency-intensity relationship (highest jitter values associated with the low-frequency, low-intensity condition) for vocal jitter was also compared for the two experimental groups. Specifically, the following questions were asked:

- Is there a difference in mean jitter between trained and untrained voices?
- Is there a difference in mean jitter between trained and untrained voices as intensity is varied at 60, 70 and 80 dB SPL?
- Is there a difference in mean jitter between trained and untrained voices as frequency levels for the two groups are increased from the subject's speaking fundamental frequency to one octave above that level for each of three different intensity levels?

CHAPTER TWO: METHODS

I. SUBJECTS

Twenty adult females were selected as paid volunteers for this study (age range 20-33 years, mean age = 25.6 years). The subjects were assigned to two different groups, trained singers (Group 1) and subjects with untrained voices (Group 2). Group 1 consisted of 10 subjects who had a minimum of six years of formal training (range 6-14 years, mean = 8.5 years) in singing in the style of Western classical music and were enrolled as vocal performance majors (all but one subject were studying at the graduate level). Group 2 consisted of 10 subjects who had no vocal training and no solo or choral singing experience beyond casual participation in school or church activities.

All subjects met the following criteria for inclusion in the study. Each subject:

(1) reported negative histories for chronic or acute vocal pathology, including polyps, nodules, persistent hoarseness or other voice disorders.

(2) was free of upper respiratory illness or allergic symptoms at the time of the voice measurements.

(3) was a non-smoker (Sorensen & Horii, 1983), operationally defined as either having never smoked or not smoked within the last five years (Stoicheff, 1981).

(4) passed a bilateral pure tone hearing screening at 500, 1000, 2000 and 4000 Hz at 25 dB HL.

II. PROCEDURES

Each subject stood in an Acoustic Systems sound-treated booth (Model RE-144), positioned 6" from a boom-mounted Shure SM48 dynamic cardioid microphone to insure consistent microphone-to-mouth distance and a high signal-to-noise ratio for the recording of all voice samples. The microphone output was input directly into the Computerized Speech Laboratory (CSL), Model 4300B using the Real Time Pitch software option (Kay Elemetrics, Pine Brook, NJ) for digitization and analysis of all voice samples.

Before the main experimental task was initiated, each subject's mean speaking fundamental frequency (SFF) was determined by having them read a portion of the Rainbow Passage (Fairbanks, 1960). The CSL was used to analyze 10 seconds of the reading sample and calculate an average value for SFF in Hertz. The corresponding musical pitch in semitones nearest to the SFF was determined (Baken, 1987, p. 487) so that the subject could be easily cued with an electronic piano keyboard (tuned to A4 = 440 Hz) for consistent frequency output during the experiment. This pitch and the one 12 semitones (one octave) above it were the targets for production by each subject. SFF was selected as a frequency relatively low in the subject's frequency range and the frequency one octave above SFF was sufficiently higher in the subject's range for determination of any significant frequency-intensity interaction for jitter (Gelfer, 1995).

Subjects were instructed to sustain the vowel /ɑ/ as steadily as possible on each of the two target frequencies and at three different intensities 60, 70 and 80 dB SPL (Weighting Network C) for at least 5-6 seconds. The subjects were instructed to cue themselves with the keyboard for the correct target pitches and were told to maintain consistent a pitch for each experimental

intensity. The subjects were instructed to monitor the intensity of their vocalizations using a Realistic Sound Level Meter (Model 33-2050) set on the C weighting network and positioned 12" from the mouth. Each of the 6 frequency and intensity combinations were presented to each subject in a random order. A total of three tokens of each frequency and intensity combination was elicited, and the median jitter value (or modal value, in the instance of a repeated value) was used for data analysis.

Subjects were given a sufficient number of trials to accurately produce each frequency and intensity combination before capturing the signal. The Real Time Pitch software was configured to capture exactly 4 seconds of phonation. Data collection was initiated following voice onset and the subject continued to phonate after data collection terminated. This procedure was employed since voice onset and offset are marked by extreme values of vocal jitter (Baken, 1987; Lieberman, 1961; Karnell, 1991b). Mean fundamental frequency and jitter analysis were performed for each sample immediately after collection, and following application of the pitch smoothing option with the software to eliminate spurious pitch values. Only frequencies that fell within ± 1 semitone of the target frequency were accepted for analysis (Acoustical Society of America, 1960). In addition, the experimenter visually monitored the dial of the sound level meter, and productions were accepted only if the needle remained within ± 2 dB of the zeroed position during the entire phonation.

The CSL equipment calculates jitter as relative average perturbation (RAP) (Koike, 1973). This jitter index is unique among frequency-related measures of jitter (mean jitter divided by mean period) in that it determines a continuous three-point average over immediately adjacent cycles of the pitch

period for calculation of average absolute jitter (Baken, 1987). This has the effect of averaging out any longer-term pitch variations from the jitter measurement, and providing a jitter value that was more representative of the small cycle-to-cycle frequency perturbation of interest in this study.

CHAPTER THREE: RESULTS

I. RELIABILITY

Reliability was calculated for randomly selected samples of intrasubject data, repeated jitter measurements by the CSL software, and repeated digitizations of audiotaped vowel tokens. For each set of reliability data, both the standard error of the difference and a Pearson correlation coefficient were determined. The standard error calculation gives the standard deviation of the sampling distribution and provides a good estimation of measurement reliability (Blommers & Forsyth, 1977). The Pearson coefficient was also calculated since this is the reliability procedure often used by other researchers (e.g., Gelfer, 1995). All of the data used for reliability calculations are presented in Appendix 1.

Intrasubject Reliability

The entire sequence of 18 vowel tokens (3 intensities X 2 frequencies X 3 trials) was repeated in a different random order for two trained (T3 and T4) and two untrained (U8 and U5) subjects and the median jitter values for time 1 and time 2 were compared to provide intrasubject reliability data for jitter. Standard error values and Pearson correlation coefficients were as follows: for subject T3, SE = 0.03 and r = 0.95; for subject T4, SE = 0.14 and r = 0.93; for subject U8, SE= 0.15 and r = 0.98; for subject U5, SE = 0.22 and r = 0.98. These values indicated an acceptable degree of consistency in jitter performance for this subject sample, and therefore acceptable intrasubject reliability for this study.

Jitter Measurement Reliability

All vocal samples were saved to magnetic disk, and a random portion (10%) of the experimental samples were re-analyzed by CSL to determine the

reliability of the jitter measurement. The same analyses were performed on these data. Standard error was 0.004, and the Pearson correlation coefficient was 0.99, both indicating good reliability for the analysis procedure, which was expected by the investigator.

Signal Digitization Reliability

Two of the subjects' vocalizations were also audiotaped during production, input into CSL, analyzed, and then re-input and re-analyzed. Jitter values for the first and second digitization trial were compared to determine the reliability of the digitization process. Due to the addition of noise to the signal inherent in the audiotaping of the samples, one of the samples could not be analyzed from the tape using the CSL program. For the remaining 11 samples, standard error (SE = 0.04) and a Pearson correlation coefficient (r = 0.99) were also calculated, both indicating good reliability of the digitization process, which was also as expected.

II. RESULTS

The median jitter values for three trials for each of the twenty subjects at each of the six experimental conditions are listed in Table 1.* The mean, range, and standard deviation for each condition are also presented in Table 1 and Figure 1 for both trained and untrained subjects. The jitter value at each of the three trials is presented for trained subjects in Appendix 2 and for untrained subjects in Appendix 3. The first value listed for each condition is median value, which was used for statistical analysis.

---

* The six experimental conditions are labeled descriptively in Table 1 and in the remainder of the text as follows: "low" and "high" refer to the frequency levels SFF and one octave above SFF, respectively; "soft," "medium," and "loud" refer to the intensity levels 60, 70 and 80 dB, respectively (e.g. low-medium refers to the condition of SFF at 70 dB).

TABLE 1. Median jitter values for each subject with the mean, range and standard deviation (SD) listed for each condition. Jitter values are relative average perturbation (RAP) in percent and are presented for each of 10 trained and 10 untrained subjects for sustained phonations of /ɑ/ at 2 different frequencies (SFF and one octave above SFF) and 3 different intensities (60, 70 and 80 dB SPL).

	SFF			Octave above SFF		
	60 dB low-soft	70 dB low-med	80 dB low-loud	60 dB high-soft	70 dB high-med	80 dB high-loud
TRAINED						
1	0.67	0.35	0.24	1.00	0.29	0.39
2	0.68	0.39	0.29	0.94	0.46	0.40
3	0.52	0.26	0.23	0.73	0.41	0.35
4	1.30	0.33	0.30	0.50	0.48	0.36
5	1.14	0.41	0.41	0.60	0.48	0.31
6	1.27	0.27	0.37	0.56	0.36	0.39
7	2.19	0.51	0.38	1.15	0.39	0.36
8	0.38	0.24	0.67	0.55	0.43	0.29
9	0.93	0.26	0.22	0.60	0.41	0.31
10	0.68	0.26	0.34	0.46	0.44	0.38
Mean	**1.06**	**0.33**	**0.35**	**0.71**	**0.42**	**0.35**
Range	0.38 - 3.01	0.24 - 0.51	0.22 - 0.67	0.46 -1.15	0.29 - 0.48	0.29 - 0.40
SD	0.75	0.09	0.13	0.24	0.06	0.04
UNTRAINED						
11	0.76	0.37	0.22	0.63	0.45	0.48
12	0.66	0.32	0.33	0.70	0.47	0.28
13	2.57	0.29	0.19	0.43	0.71	0.39
14	1.68	0.45	0.26	0.76	0.81	0.34
15	0.74	0.32	0.35	0.62	0.54	0.43
16	2.02	0.42	0.35	0.92	1.09	0.59
17	0.65	0.38	0.27	1.08	0.48	0.37
18	2.00	0.30	0.20	0.67	0.53	0.42
19	1.09	0.30	0.19	0.72	0.50	0.35
20	0.74	0.25	0.22	1.01	0.60	0.37
Mean	**1.29**	**0.34**	**0.26**	**0.75**	**0.62**	**0.40**
Range	0.65 - 2.57	0.25 - 0.45	0.19 - 0.35	0.43 - 1.08	0.45 - 1.09	0.28 - 0.59
SD	0.71	0.06	0.07	0.20	0.20	0.09

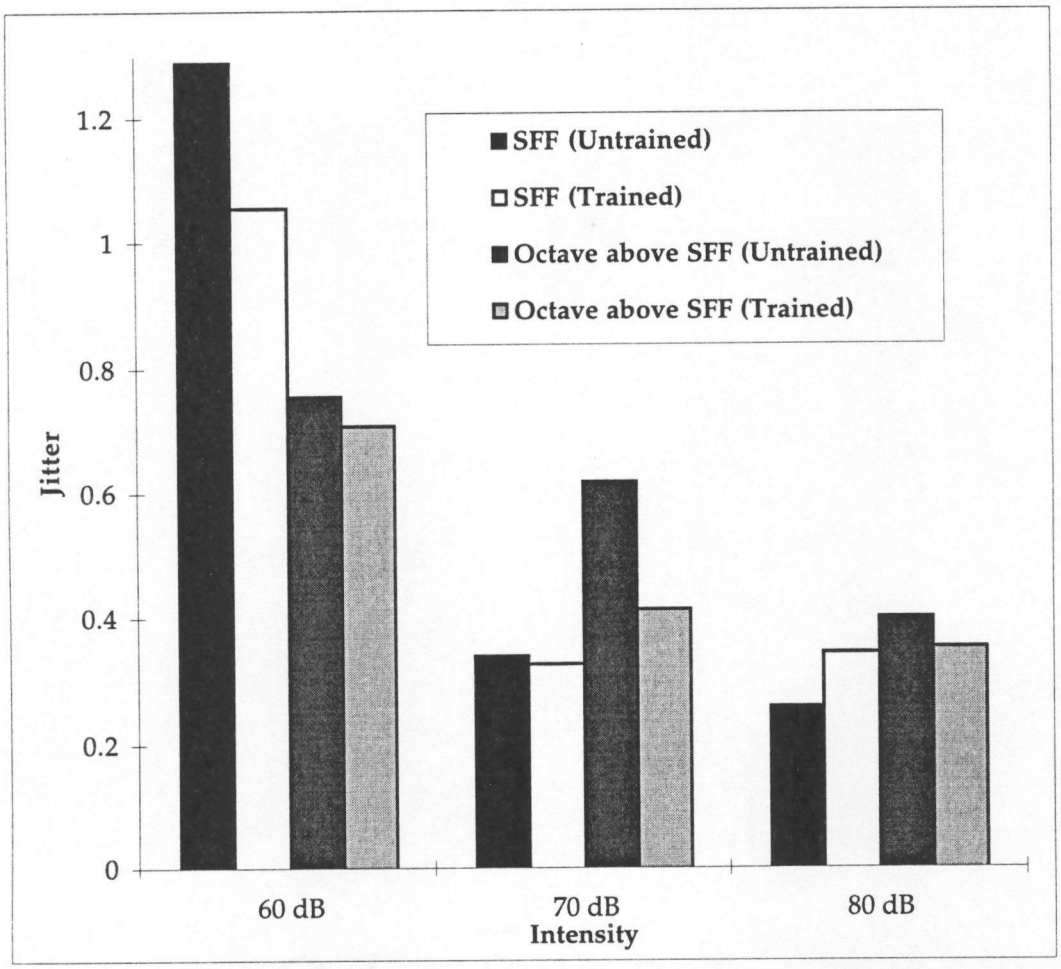

Figure 1: Mean jitter for each experimental condition. The data are from each of 10 trained and 10 untrained subject's median jitter values at two frequencies (SFF and one octave above SFF) and three intensities (60, 70 and 80 dB SPL).

From initial examination of Table 1, the low-soft (SFF at 60 dB) condition was associated with the highest mean jitter and standard deviation for both the trained (M = 1.06; SD = 0.75) and the untrained (M = 1.29; SD = 0.71) groups. The high-soft condition (one octave above SFF at 60 dB) also had a significantly higher mean jitter and standard deviation than the remaining conditions for the trained subjects (M = 0.71; SD = 0.24). The untrained subjects also had higher mean jitter at the high-soft condition in comparison to the remaining conditions. Mean jitter values, standard deviations and ranges at the two soft conditions were much larger than the respective magnitudes at all other conditions. The only exception is that the range and standard deviation for the high-medium (one octave above SFF at 70 dB) condition for the untrained voices was essentially identical to that for the high-soft condition.

Figure 1 demonstrates that the spread of mean jitter values for each frequency for the two groups decreased with increasing intensity, with very little difference in the jitter values at 80 dB. Mean jitter tended to decrease with increasing intensity for both groups of subjects at both SFF and one octave above SFF, with the exception of the trained voices, where the mean jitter increased from the low-medium to the low-loud condition. From observation, the mean jitter values were higher in magnitude for the untrained voices at each condition except for low-loud, where the untrained voices had a lower mean jitter value than the trained voices. There were several untrained voices with very low jitter values at this condition (around 0.19 or 0.20), and one trained voice with a jitter value of 0.67. These extreme scores at this condition may account for the reversal in the trend for untrained voices to have higher mean jitter values than the trained singers.

First Analysis of Variance

A multivariate repeated measures analysis of variance (ANOVA) was performed to determine the effects of group, intensity and frequency on vocal jitter. The results of this ANOVA are summarized in Table 2. There was a significant main effect for intensity on mean jitter ($F = 62.78$, df = 2, $p < 0.001$). There was also a significant two-way interaction between frequency and intensity for mean jitter ($F = 10.76$, df = 2, $p < 0.001$). The low-soft condition had the highest level of mean jitter for both groups. The interaction between these factors is clearly illustrated in Figure 2, with mean jitter decreasing dramatically from SFF to one octave above it for the 60 dB condition, and increasing slightly with frequency for both the 70 dB and 80 dB conditions.

Group Difference in Mean Jitter

The first experimental question involved the difference in mean jitter between trained and untrained voices. While the overall mean jitter for the trained voices (0.53) was lower than for the untrained voices (0.61), this was not a significant difference ($F = 1.76$, df = 2, $p = 0.19$).

Group Difference in Mean Jitter with Changes in Intensity

The second experimental question involved the difference in mean jitter between trained and untrained voices as intensity was varied at 60, 70 and 80 dB SPL. This two-way interaction between intensity and group was also not significant ($F = 0.10$, df = 2, $p = 0.22$). Figure 3 illustrates a decrease in jitter with higher intensity, with a similar performance for both trained and untrained voices.

Group Difference in Mean Jitter with Changes in Frequency at Each Intensity

The third experimental question involved the difference in mean jitter between trained and untrained voices as frequency was varied from SFF to

TABLE 2. Analysis of variance computed for the main effects, 2- and 3-way interactions among the experimental factors frequency (freq), intensity (int), and subject group (group) on the dependent variable jitter. df = degrees of freedom, *F* = the calculated *F* value, and p = probability (N = 20).

Souse of variation	Sum of squares	df	*F*	p
Main effects				
freq	0.07	1	1.05	0.32
int	8.21	2	62.78	< 0.001
group	0.32	2	1.76	0.19
2-way interactions				
freq x int	1.94	2	10.76	< 0.001
freq x group	0.003	1	0.04	0.84
int x group	0.20	2	0.10	0.22
3-way interaction				
freq x int x group	0.32	2	1.76	0.19

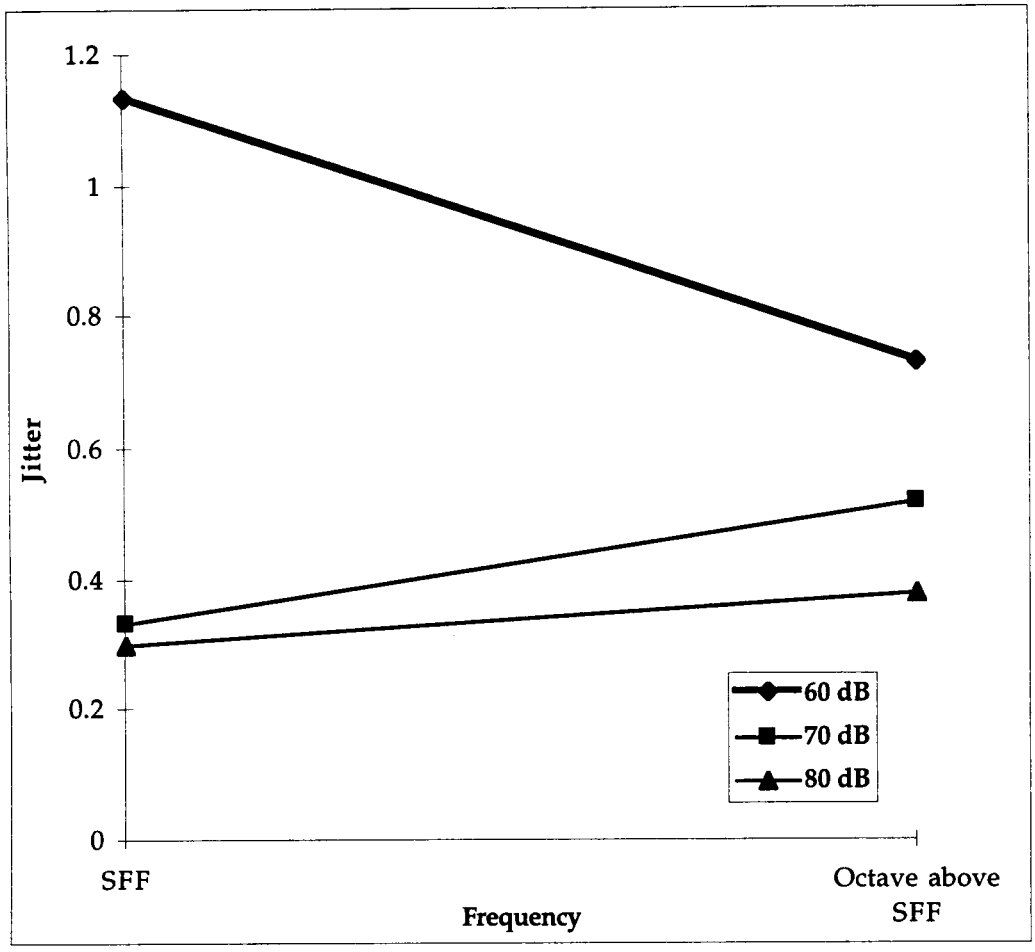

Figure 2: Mean jitter as a function of frequency comparing the three different experimental intensity levels of 60, 70 and 80 dB SPL.

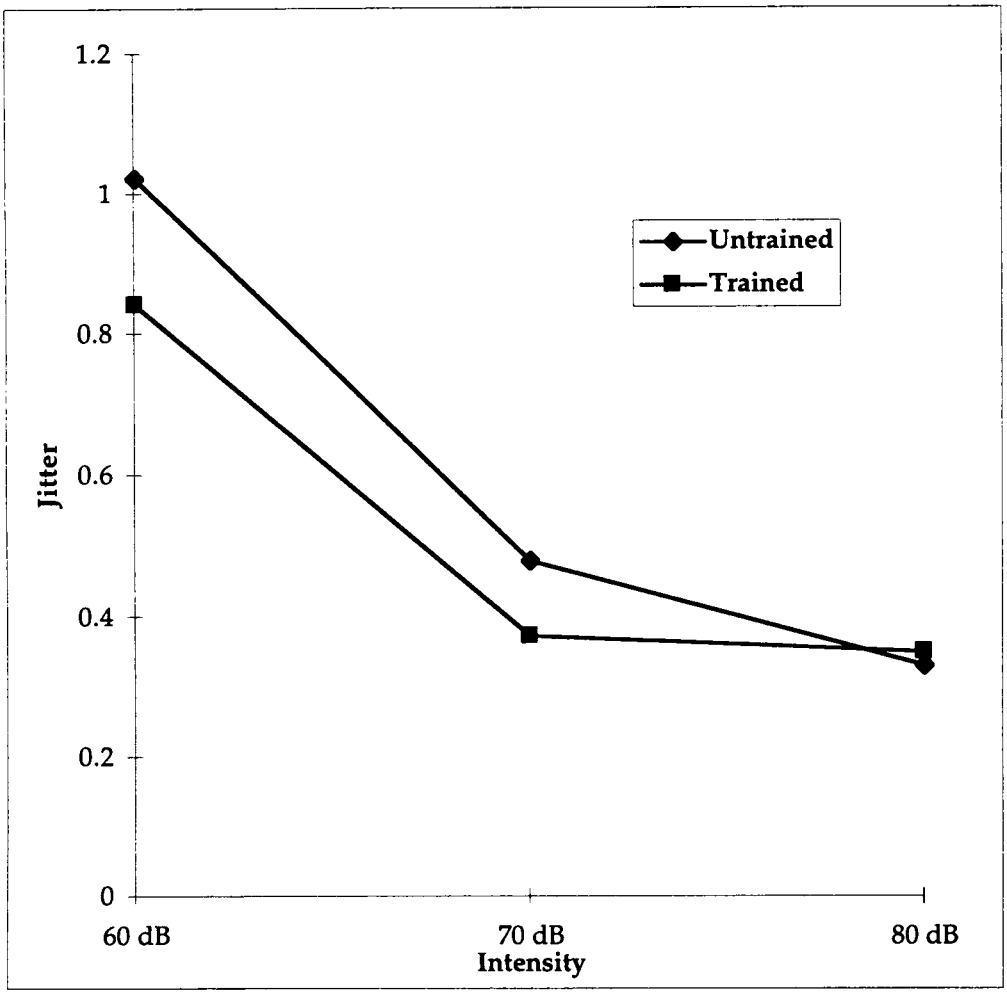

Figure 3: Mean jitter as a function of the three experimental intensities (60, 70 and 80 dB SPL) comparing the 10 trained versus 10 untrained voices.

one octave above for each of the three intensity levels. The analysis of variance indicated that this three-way interaction between frequency, intensity and group was not significant ($F = 1.76$, df = 2, p = 0.19).

Second Analysis of Variance

Mean jitter for both groups of subjects at the 60 dB conditions (especially at SFF) was much greater in comparison to mean jitter at the remaining four conditions. Mean jitter values at the low-soft and high-soft conditions for both groups were compared using a post-hoc t-test for independent samples. These t-tests did not reveal a significant difference between the trained and untrained voices for either the low-soft ($t = -0.71$, p = 0.47) or the high-soft ($t = -0.46$, p = 0.65) conditions. Given this lack of difference in group performance for these conditions, and the relative extremes in jitter magnitude and variability for the jitter values at these two conditions compared with all other conditions, a second multivariate repeated measures ANOVA was performed, with the data for the two 60 dB conditions excluded from analysis. The results of this second ANOVA are summarized in Table 3.

There was a significant interaction between frequency and intensity ($F = 6.21$, df = 1, p = 0.20) with this second analysis. This significant interaction was further analyzed using post hoc t-tests and is illustrated in Figure 4. A two-tailed independent samples t-test indicated that the mean difference in jitter between the low and high fundamental frequencies was significantly different between the 70 dB and 80 dB conditions ($t = 2.12$, p = 0.04). A two-tailed one-sample t-test revealed that for both the 70 dB ($t = 2.40$, $p < .001$) and 80 dB (t = 4.74, p = 0.03) conditions, jitter increased significantly with increasing frequency.

TABLE 3. Analysis of variance computed with data for only the 70 dB and 80 dB conditions for the main effects, 2- and 3-way interactions among the factors frequency (freq), intensity (int), and subject group (group) on the dependent variable jitter. df = degrees of freedom, *F* = the calculated *F* value, and p = probability (N = 20).

Source of variation	Sum of squares	df	*F*	p
Main effects				
freq	0.34	1	34.35	< .001
int	0.15	1	19.42	< .001
group	0.04	1	2.32	0.15
2-way interactions				
freq x int	0.06	1	6.21	0.020
freq x group	0.13	1	13.61	0.002
int x group	0.08	1	10.71	0.004
3-way interaction				
freq x int x group	0.004	1	0.43	0.52

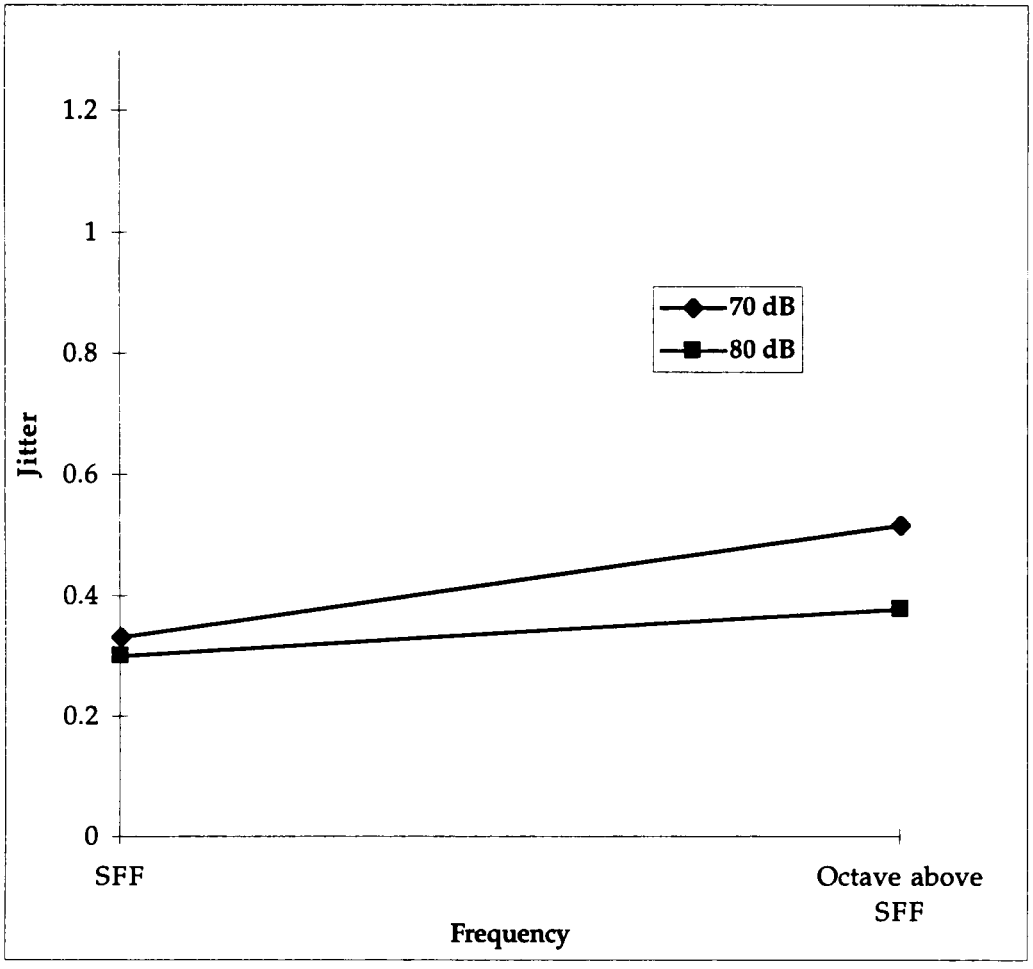

Figure 4: Mean jitter as a function of frequency comparing the medium (70 dB) and loud (80 dB) conditions at SFF and one octave above SFF for all subjects in both groups.

With the second ANOVA, a significant frequency by group interaction for mean jitter was revealed ($F = 13.61$, df = 1, $p = 0.002$). This was a difference not considered in the original experimental questions. This frequency-group interaction was also analyzed using post hoc t-tests and is illustrated in Figure 5. The mean difference between the low-frequency (SFF) and high-frequency (octave above SFF) conditions was compared for trained and untrained voices using a two-tailed independent samples t-test. This revealed a significant difference between the two groups ($t = 3.56$, $p = 0.001$). Next a two-tailed one-sample t-test was calculated for each group to determine if the mean change in jitter from low to high frequency was significant. For the untrained group, the mean change in jitter from low to high frequency was significant ($t = 6.10$, $p < .001$), while for the trained group, the mean change was not significant ($t = 1.60$, $p = 0.13$). That is, mean jitter for the trained voices was not significantly different for the two experimental frequencies, while for the untrained voices, mean jitter significantly increased with the change from SFF to one octave above.

Group Difference in Mean Jitter

As with the first ANOVA, there was no significant difference in the overall mean jitter between trained and untrained voices ($F = 2.32$, df = 1, $p = 0.15$).

Group Difference in Mean Jitter with Changes in Intensity

Using the data for only the 70 dB and 80 dB conditions, there was a significant interaction between intensity and group ($F = 10.71$, df = 1, $p = 0.004$). The same series of post-hoc t-tests was performed to further evaluate the intensity-group interaction, which is illustrated in Figure 6. A two-tailed

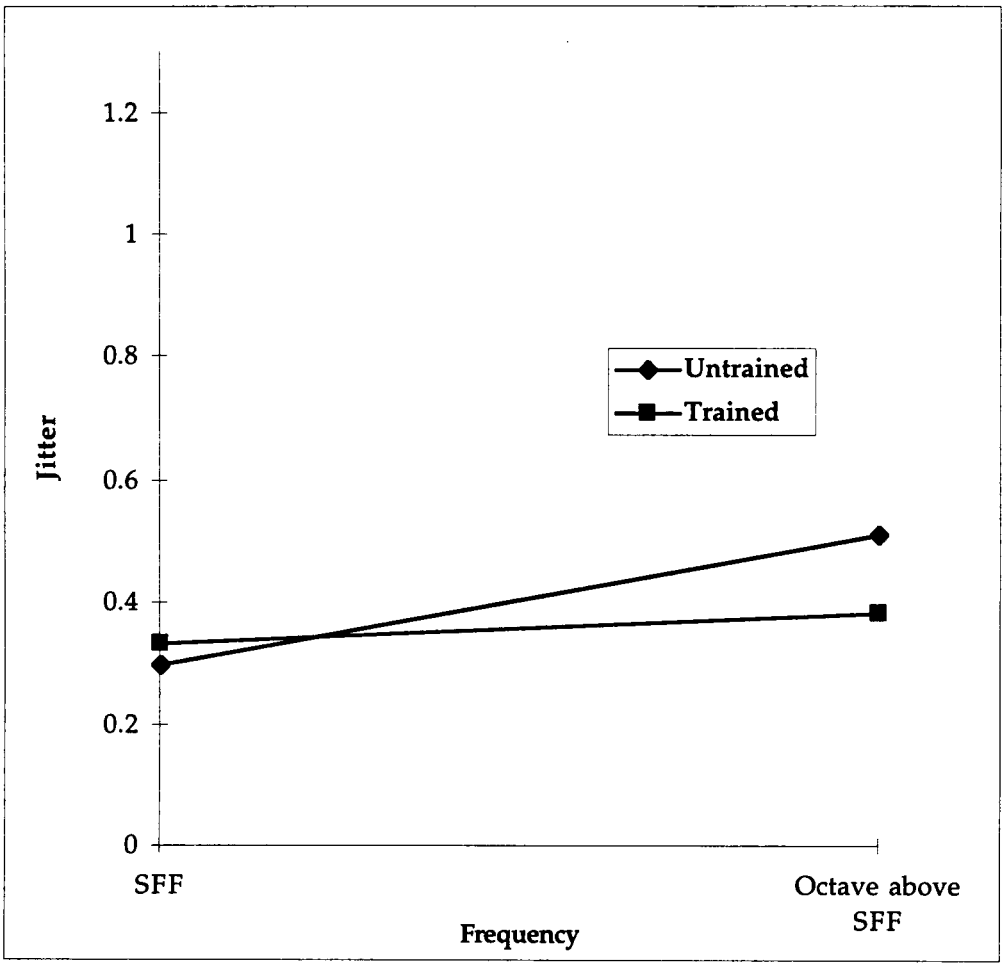

Figure 5. Mean jitter as a function of the two experimental frequencies (SFF and one octave above SFF) comparing trained versus untrained voices, without the data for the soft condition (60 dB) used for analysis.

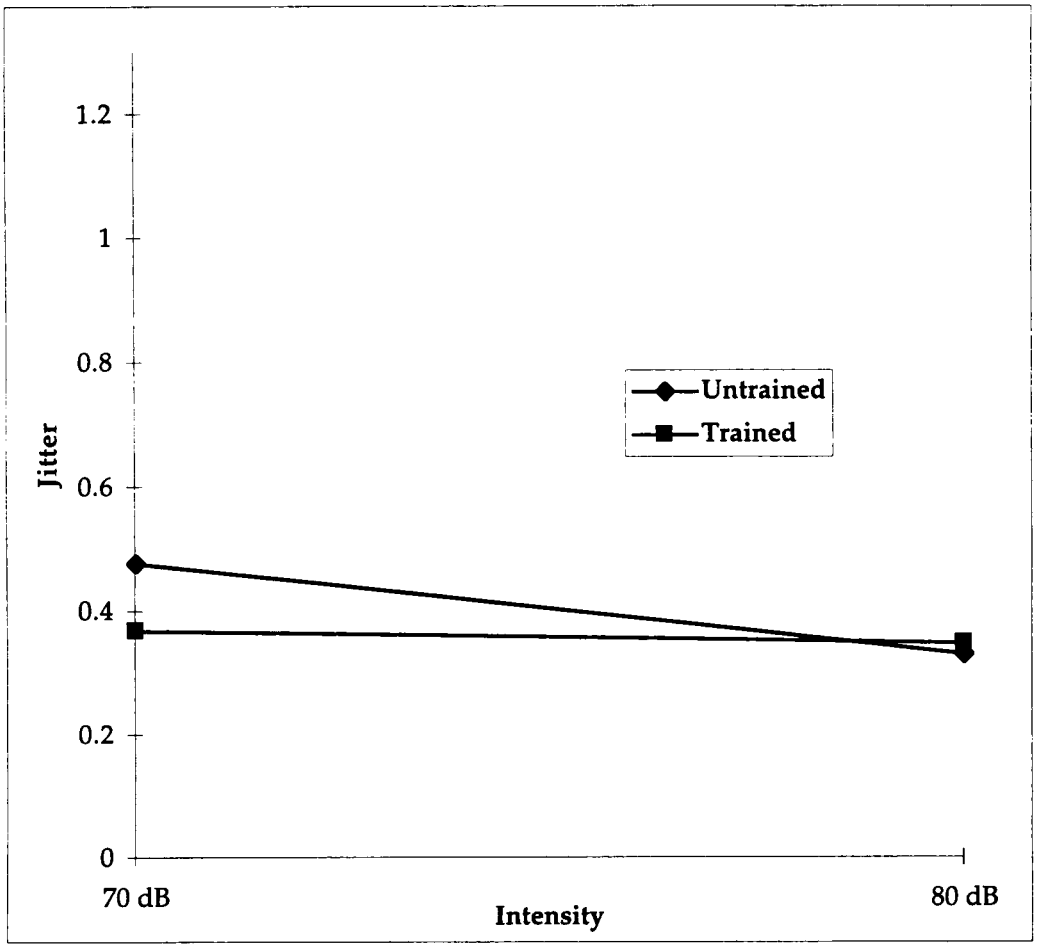

Figure 6: Mean jitter as a function of intensity comparing 10 trained versus 10 untrained voices for the medium (70 dB) and loud (80 dB) conditions.

independent sample t-test indicated that there was a significant difference in the change in mean jitter between the trained and untrained voices from 70 dB to 80 dB (t = -2.92, p = 0.01). A two-tailed one-sample t-test indicated that for the trained voices the 70 dB and 80 dB conditions did not have a significant difference in mean jitter (t = -0.75, p = 0.46), while the untrained voices had a significant decrease in mean jitter from 70 to 80 dB (t = -4.67, p < .001).

Group Difference in Mean Jitter with Changes in Frequency at each Intensity

With this second ANOVA, the 3-way interaction between frequency, intensity and group was once again not significant (F = 0.43, df = 1, p = 0.52).

CHAPTER FOUR: DISCUSSION

This study was designed to determine if any differences existed between trained voices and untrained voices for mean vocal jitter of a sustained vowel. Jitter was measured in six specific conditions with varying levels of intensity and fundamental frequency to determine the relationship between jitter and these two vocal parameters for these two groups. Two separate analyses of the data were conducted, one with the entire data set, and one without the mean jitter values at the two soft (60 dB) conditions. The rationale for this second analysis of the data came predominantly from the extreme variability found in the jitter values for both groups of subjects at these conditions, especially at the low-soft condition. Previous researchers have reported that the low-soft condition is typically associated with the highest jitter values, the most variability in jitter values, and the most reported subject difficulty with sustaining a steady-state vowel at this vocal condition (Gelfer, 1995; Jacob, 1968; Orlikoff & Kahane, 1991).

In the present study, the low-soft condition was also associated with the highest overall levels of mean jitter and variability in jitter values. The high-soft condition also resulted in high values of mean jitter and large variability. Both of these conditions were the ones that the majority of subjects found to be the most difficult to execute. Perceptually, the low-soft condition was associated with a breathy quality, which Jacob (1968) also reported for his subjects phonating at the softest possible level that each could produce with a steady pitch. The high-soft condition was identified by most subjects in the present study as the single-most difficult condition to execute, and subjects typically required more practice trials for the high-soft vowel than for any other condition to meet the intensity criteria. This condition was associated

perceptually with constant fluctuations in intensity and occasional voice breaks.

The trained singers in particular reported great difficulty sustaining the vowel at the 60 dB criterion level, especially at the higher frequency. All of them commented that the sound they were required to produce was much softer than anything they would ever produce in a singing context. In fact, the "soft," "medium," and "loud" descriptive labels for the 60, 70 and 80 dB SPL productions were considered by the trained singers to be quite contrary to their typical perceptual loudness designations. Most of these subjects judged the "loud" phonation to be equivalent to a medium-soft production in a typical performing context. In contrast, the untrained subjects tended to find the "loud" production rather difficult, and required much more vocal effort than they were accustomed to exerting. All subjects seemed to consider the "medium" 70 dB level to be relatively easy to produce. This level was determined to be the "most comfortable" loudness level for two untrained pilot subjects.

The extremely high jitter values found at the soft intensities, especially at the lower frequency, can be attributed to the highly unstable productions that most subjects demonstrated at this intensity level. The breathy quality at the low-soft condition suggests that the subjects glottal closure time was reduced, resulting in less stable phonation and elevated jitter values. The high-soft condition also tended to be perceptually quite different from the other conditions, with occasional evidence of small pitch or voice breaks in the voice signal for some subjects, suggesting increased instability.

It is clear that the 60 dB phonation for these subjects was an extremely difficult task, and therefore not representative of typical vocal dynamics.

Both groups performed similarly on this task, perhaps confirming that this mode of production is difficult for anyone, regardless of vocal use history. Jacob (1968) concluded that his low-soft condition was associated with too much vocal instability to be useful for experimental or clinical purposes. While the exact intensity or frequency values used by his subjects were not reported, the subjects' basal frequency and intensity levels were determined, and the productions were elicited at a basal intensity and at 25% above the basal modal frequency. The resulting low-soft condition was most likely comparable to the 60 dB level at SFF elicited in the present study.

Gelfer (1995) and Orlikoff and Kahane (1991) reported mean jitter values that were highest for soft phonations, but did not report the same extremes in mean values or variability reported here. This is most likely due to the less stringent intensity criteria that they used compared to the present study. Gelfer (1995) allowed a 5 dB variation above and below the target, with a minimum of 5 dB between conditions, while Orlikoff and Kahane (1991) allowed intensity to vary from 60-68 dB at the soft condition. Both studies measured dB SPL at a microphone distance of 12", as in the present study. In the present study, subjects could vary only 2 dB around the target intensity. Therefore, the subjects were likely phonating consistently softer in this study than previous studies, making the task considerably more difficult.

The decision to analyze the partial data set was based on these observations made during data collection regarding the subjects' difficulty with the soft condition at both frequencies, compared with the relative ease in production for the other intensities, and also the resulting extremes in mean jitter and variability at the soft conditions. Since the subject groups' mean jitter values were not statistically different at these two conditions, the

rationale for the second analysis was further supported. The different results obtained with the two analyses reveal some valuable information regarding the dissimilar performance of the two groups at the medium and loud intensities, which was masked by the extreme variability at the two soft conditions.

The first experimental question for this study asked if there was a difference in mean jitter between trained and untrained voices. While there was no significant difference in overall mean jitter between the two subject groups with either analysis, excluding the data from the soft conditions revealed significant differences between the groups in the form of interactions between both intensity and frequency. The group difference in mean jitter with changes in intensity provided an affirmative answer to the second experimental question posed in this study. The significant difference in mean jitter between groups with changes in frequency was not originally anticipated.

The remaining experimental question was if the frequency-intensity relationship for vocal jitter found in previous studies was different for trained than for untrained voices. This three-way interaction between frequency, intensity and group was not statistically significant with either the full data set or the reduced data set. Both groups demonstrated similar relationships between frequency, intensity and jitter, with jitter actually increasing slightly with increasing frequency at both the medium and loud intensities, with a larger significant increase for the medium condition (Fig. 4). Both groups demonstrated a significant decrease in jitter with increasing frequency for the soft condition (Fig. 2).

A possible explanation for this relationship may involve the relationship between fundamental frequency and intensity control at the level of the vocal folds. With an increase in frequency and an accompanying increase in vocalis muscle tension, the instability of the soft phonation (due to a shorter glottal closure phase) was significantly compensated for, while the overall instability remained relatively high compared to the medium and loud conditions, where the glottal closure time was longer. The unexpected slight increase in jitter from the low to the high frequency at 70 and 80 dB will be discussed further below.

The greater stability of the trained voices suggested by the essentially constant level of vocal jitter with changes in frequency and intensity (from 70 dB to 80 dB) is the most significant finding of this study. From the results of the post-hoc analyses and in examining Figures 5 and 6, it is apparent that the stability of the untrained voices varied significantly as frequency and intensity changed, while the trained voices remained essentially constant in stability with these parameter changes, as indicated by the mean jitter levels. For the untrained voices, jitter decreased significantly with increasing intensity, and increased significantly with increasing frequency. While the changes in jitter between conditions were in these same directions for the trained singers, the differences were not statistically significant.

The decrease in jitter with intensity is consistent with what has been reported by previous researchers, while the increase in jitter with frequency (Fig. 5) was unexpected, and was another significant finding of the present study. Both Gelfer (1995) and Jacob (1968), reported a significant decrease in jitter with increasing frequency. Horii (1979) and Orlikoff and Baken (1990) also reported a general trend for jitter to decrease with increasing frequency.

However, the relationship was not monotonic, but in fact quasi-exponential as frequency increased. Horii (1979) reported an increase in absolute jitter past a certain fundamental frequency in normal male subjects. Orlikoff and Baken (1990) suggested that jitter was related more to relative fundamental frequency within a speaker's overall range and that frequency-related jitter indices such as jitter in percent in fact overcorrected for jitter values at higher fundamental frequencies in female voices. The present study utilized a frequency-related jitter index, RAP, which is an average perturbation value divided by mean period, expressed as percent. It is possible that with the female voices in this study, the RAP jitter index elevated the actual amount of frequency perturbation present for the higher fundamental frequency. Since only two frequencies were examined, it could also be that jitter actually would have been found to decrease with increasing frequency in smaller intervals above SFF, and then increase again at a frequency level at or near one octave above SFF, similar to the pattern reported by Horii (1979).

Yet another explanation for this jitter increase with increasing frequency may involve the nature of the specific fundamental frequency that was one octave above most of the subjects' SFF. Fifteen of the twenty subjects' higher frequency target was located between approximately 392 Hz and 494 Hz, or between G and B above middle C in musical notation. Many professional singers and teachers of singing have identified the location of a register break for the typical female voice in this same fundamental frequency interval. Since register breaks are associated with a physiologic transition in the mode of vocal fold vibration, this could explain an increase in jitter, or vocal instability, at this frequency. Trained singers are typically taught to minimize the effects of register changes and achieve a consistent vocal quality

throughout the frequency range, which may explain why the trained voices had relatively little increase in mean jitter at SFF to one octave above it (0.34 to 0.38), while the untrained voices did show a statistically significant increase in mean jitter with frequency (0.30 to 0.51). The untrained subjects may have been unable to control phonation as efficiently in the vicinity of the unstable register shift.

Referring again to Figure 6, the difference between the two groups regarding the decrease in jitter as intensity increased from 70 to 80 dB is illustrated. The direction of the change in jitter for both groups is in agreement with previous studies. However, the decrease in jitter for the trained singers was not significant (0.37 to 0.35), while the change for the untrained singers was significant (0.48 to 0.33). Apparently, as for the two frequency conditions, the trained singers used a phonation pattern with a similar amount of stability for both the medium and loud productions. This is supported by perceptual observations made during data collection. The trained singers used a consistent vocal quality for both of these conditions. In contrast, the untrained singers apparently used a different phonation pattern for the two intensities, significantly increasing vocal stability from 70 to 80 dB. These data are supported by perceptual observations. The untrained subjects used a strained vocal quality with an audible increase in vocal tension. The increased vocal tension had the effect of stabilizing the laryngeal mechanism and decreasing the amount of jitter due to muscular forces as well as turbulent forces from airflow. The untrained voices actually had a slightly lower mean jitter value (0.33) at the loud condition than the trained voices (0.35), presumably due to this marked change in mode of vocal fold vibration.

One possible explanation is that the untrained subjects used a more "pressed" style of phonation as intensity increased, while the trained singers used a "flow" phonation style regardless of intensity. Flow phonation is a vocal production mode in which there is an ideal balance between airflow and laryngeal adduction, maximizing the transfer from aerodynamic to mechanical energy at the glottis and allowing the vocal folds to vibrate with the full range of harmonic partials present. In pressed phonation, this freedom to vibrate is reduced, creating a damping effect, which acoustically flattens the harmonic spectrum, and perceptually results in a tense vocal quality. The untrained subjects increased vocal tension to achieve greater intensity instead of increasing airflow, as the trained singers likely did. This had the effect of decreasing vocal jitter considerably, but at the expense of vocal quality. The trained singers, due to their enhanced ability to coordinate airflow and vocal tension, were able to phonate with the same degree of vocal stability at 70 and 80 dB, and without changing vocal quality.

Returning to Figure 4, the interaction between frequency and intensity is illustrated for the 70 and 80 dB conditions. The possible reasons for the unexpected increase in jitter with frequency have already been addressed. It is important to note that the change was significant for both the 70 dB and 80 dB conditions, but the magnitude of the change was greater for the 70 dB condition. The reason for this difference may be due to the increased glottal closure phase at the loud condition. Phonation at this intensity level showed more resistance to a change in stability (i.e. jitter) than did phonation at the medium intensity level.

If indeed there was increased instability at the higher frequency due to the presence of a register break in the vicinity of that frequency (or due to

some unknown factor), then phonation at the louder intensity counteracted this instability by increasing glottal closure time. This phenomenon may be compared to an increase in mechanical inertia. The laryngeal system producing a greater intensity is analogous to a body with larger mass which will tend to stay in the same state of motion more than a body with a smaller mass, which has less inertia and is more easily "pushed" into a different state of motion. There are numerous potential sources of jitter in the human vocal system (Titze, 1994), and it is possible that a change in laryngeal dynamics that can increase the "resistance" of the vocal system to perturbation (such as increased intensity by means of a longer glottal closure phase) will reduce the level of vocal jitter.

This possible effect is evident in the data for the 80 dB condition. When the mean jitter values are compared for both low and high frequency and for both groups at the 80 dB (loud) condition, it is clear that the variability in mean jitter level is comparatively small. Figure 7 (cf. Figure 1) clearly illustrates the decreasing variability in mean jitter for both groups at each frequency as intensity was increased. It seems that at the loudest intensity level, all subjects' vocal mechanisms behaved similarly, regardless of frequency. It may be that this intensity level had the effect of stabilizing the vocal mechanism so that both frequency effects and effects of vocal training had minimal impact on the mean jitter values. Figure 7 also illustrates the comparatively large variation in mean jitter at the 60 dB condition, where both trained and untrained subjects had increased instability of phonation.

The finding that vocal intensity has a significant effect on vocal stability for both trained and untrained voices is supported by the data in this study. At a very soft or basal intensity as used in this study (60 dB), both

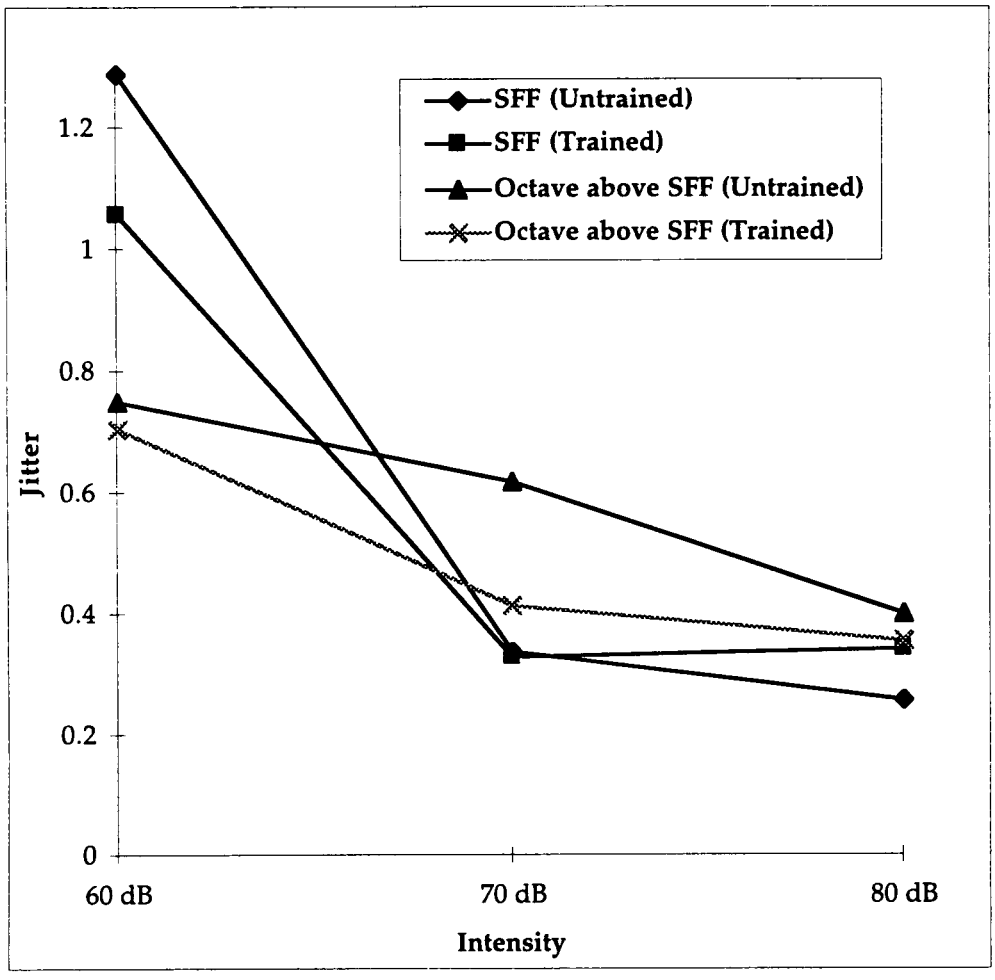

Figure 7: Mean jitter for each experimental condition, illustrating the reduction in variability in mean jitter as intensity increases. The data are from each of 10 trained and 10 untrained subject's median jitter values at two frequencies (SFF and one octave above SFF) and three intensities (60, 70 and 80 dB SPL).

trained and untrained voices showed a comparatively large degree of vocal jitter at both low and high frequencies. Jitter for both groups decreased dramatically with increasing intensity for both frequency levels at this basal intensity. For the remaining conditions, the trained singers were able to manage both changes in intensity and also in frequency to a greater degree than the untrained subjects. The trained singers demonstrated consistently stable phonation regardless of intensity or frequency level, as measured by their relatively constant amount of vocal jitter.

CHAPTER FIVE: SUMMARY AND CONCLUSIONS

Ten trained female singers and ten female speakers with no training in singing sustained an /ɑ/ vowel for four seconds under six experimental conditions. There were two levels of fundamental frequency, the subject's speaking fundamental frequency (SFF) and an octave above SFF, and three levels of intensity, 60, 70 and 80 dB SPL measured at a microphone-to-mouth distance of 12". For each vowel token, intensity and frequency were monitored to maintain experimental criteria, and the vowel samples were digitally analyzed to obtain vocal jitter expressed as relative average perturbation (RAP) in percent. Mean jitter values were compared across frequency, intensity and group factors using an analysis of variance technique.

Due to extreme variability in mean jitter and subject difficulty phonating at the soft intensity, a second ANOVA was calculated using the data from the medium and loud conditions. There were significant differences between the groups with respect to both frequency and intensity. For both factors, the trained singers demonstrated no significant change in mean jitter from SFF to one octave above SFF or from 70 to 80 dB. In contrast, the untrained singers demonstrated a significant increase in vocal jitter with frequency, and a significant decrease in jitter with intensity. There were no significant differences between the trained and untrained singers for overall mean jitter or for the jitter-frequency interaction. The implications are that both groups of subjects had difficulty producing stable phonation at 60 dB, but that the trained singers were able to produce stable phonation regardless of frequency or intensity for the remaining experimental conditions, while the untrained subjects phonation stability varied significantly with changes in frequency and intensity.

From the results obtained in this study, it may be concluded that the differences in vocal production technique used by a trained singer may indeed have the effect of increasing vocal "inertia," allowing the production of a steady-state mode of vibration that is less prone to changes in frequency perturbation as fundamental frequency and intensity are varied. The level of mean jitter was essentially unaltered by changes in frequency or intensity for trained voices, while mean jitter varied considerably with changes in frequency and intensity for the untrained voices. These subjects used a different means of vocal production for all four remaining combinations of intensity and frequency. The complex interrelationship between frequency and intensity control in the human voice has long been recognized, and the trained singers in this study demonstrated a markedly different dynamic between these parameters than did the individuals with no vocal training. This difference was predicted from *a priori* assumptions about vocal training and the musical demands required of a trained singer, and the data in this study support this notion with evidence from measurements of vocal stability.

Future investigators need to examine the differences between trained and untrained voices even further. The findings of the present study need to be replicated with larger numbers of subjects. Standardizing the procedures used, including jitter index, analysis software, microphone type and placement, and intensity criteria will be necessary to make adequate comparisons among various studies. One ongoing difficulty in vocal perturbation research is the fact that there are no two studies that are completely comparable, so any conclusions drawn from comparing results have to be made cautiously at best, and are often not even possible. The data

from the present study are most comparable with Gelfer (1995), but there are still significant methodological differences between the two studies that make comparing results problematic.

One potential area to explore for future studies would be the relationship between absolute intensity and perceived loudness levels, from "as soft as possible with a good sound," through "most comfortable," to "as loud as possible with a good sound." Another experimental paradigm might involve determining the most comfortable loudness of phonation for each subject and measuring jitter at that level and at intensities ± 10 dB around this comfortable intensity. Comparing the differences between mean jitter in trained and untrained voices with these self-determined loudness parameters could be informative, given the disparity in the subjects' perception and difficulty level for the pre-determined intensities used in this study.

Future research should also include standardized perceptual judgments of vocal quality, airflow measures and videostroboscopic examinations to accompany jitter measurements. These additional measures would help confirm some of the observational conclusions drawn in this study to support the jitter findings. Finally, researchers should continue and refine longitudinal or pre- and post-training experimental designs (cf. Sabol et al., 1995), where acoustic, aerodynamic, perceptual and videostroboscopic measures of voice are made both before and after vocal training. Studies such as these would effectively document the changes made in the vocal mechanism over the course of training and allow professionals in voice science and singing training to improve training methods and effectively monitor the training process.

REFERENCES

REFERENCES

Acoustical Society of America. (1960). American standard acoustical terminology. New York: Author.

Appelman, D. R. (1967). The science of vocal pedagogy. Bloomington: Indiana University Press.

Askenfelt, A. G., & Hammarberg, B. (1986). Speech waveform perturbation analysis: A perceptual-acoustical comparison of seven measures. Journal of Speech and Hearing Research, 29, 50-64.

Awan, S. N. (1991). Phonetographic profiles and F_0-SPL characteristics of untrained versus trained vocal groups. Journal of Voice, 5, 41-50.

Baken, R. J. (1987). Clinical Measurement of Speech and Voice. Boston: College Hill Press.

Bartholomew, W. T. (1934). A physical definition of 'good voice quality' in the male voice. Journal of the Acoustical Society of America, 6, 25-33.

Blommers, P. J., & Forsyth, R. A. (1977). Elementary Statistical Methods in Psychology and Education. Boston: Houghton Mifflin Company.

Bielamowicz, S., Kreiman, J., Gerratt, B. R., Dauer, M. S., & Berke, G. S. (1996). Comparison of analysis systems for perturbation measurement. Journal of Speech and Hearing Research, 39, 126-134.

Bough, Jr., I. D., Heuer, R. J., Sataloff, R. T., Hills, J. R., & Cater, J. R. (1996). Intrasubject variability of objective voice measures. Journal of Voice, 10, 166-174.

Brown, Jr., W. S., Morris, R. J., & Michel, J. F. (1989). Vocal jitter in young adult and aged female voices. Journal of Voice, 3, 113-119.

Brown, Jr., W. S., Morris, R. J., & Michel, J. F. (1990). Vocal jitter and fundamental frequency characteristics in aged, female professional singers. Journal of Voice, 4, 135-141.

Carroll, L. M., Sataloff, R. T., Heuer, R. J., Spiegel J. R., Radionoff, S. L., & Cohn, J. R. (1996). Respiratory and glottal efficiency measures in normal classically trained singers. Journal of Voice, 10, 139-145.

Casper, J. K. (1984). Frequency perturbation in normal speakers: A descriptive and methodological study. (Doctoral dissertation, Syracuse University, 1983). Dissertation Abstracts International, 45, 525-B.

Deem, J. F., Manning, W. H., Knack, J. V., & Matesich, J. S. (1989). The automatic extraction of pitch perturbation using microcomputers: some methodological considerations. Journal of Speech and Hearing Research, 32, 689-697.

Doherty, E. T., & Shipp, T. (1988). Tape recorder effects on jitter and shimmer extraction. Journal of Speech and Hearing Research, 31, 485-490.

Fairbanks, G. (1960). Voice and Articulation Drillbook, 2nd ed. New York: Harper & Row.

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

Ferrand, C. T. (1995). Effects of practice with and without knowledge of results on jitter and shimmer levels in normally speaking women. Journal of Voice, 9, 419-423.

Fukazawa, T., Blaugrund, S. M., El-Assuootya, A., & Gould, W. J. (1988). Acoustic analysis of hoarse voice: a preliminary report. Journal of Voice, 2, 127-131.

Gelfer, M. P. (1995). Fundamental frequency, intensity, and vowel selection: effects on measures of phonatory stability. Journal of Speech and Hearing Research, 38, 1189-1198.

Gelfer, M. P., Andrews, M. L., & Schmidt, C. P. (1991). Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers. Journal of Voice, 5, 158-167.

Gelfer, M. P., & Fendel, D. M. (1995). Comparisons of jitter, shimmer and signal-to-noise ratio from directly digitized versus taped voice samples. Journal of Voice, 9, 378-382.

Glaze, L. E., Bless, D. M., & Susser, R. D. (1990). Acoustic analysis of vowel and loudness differences in children's voice. Journal of Voice, 4, 37-44.

Gould, W. J., & Korovin, G. S. (1994). Laboratory advances for voice measurements. Journal of Voice, 8, 8-17.

Griffin, B., Woo, P., Colton, R., Casper, J., & Brewer, D. (1995). Physiological characteristics of the supported singing voice. A preliminary report. Journal of Voice, 9, 45-56.

Haji, T., Horiguchi, S., Baer, T., & Gould, W. (1986). Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. Journal of the Acoustical Society of America, 80, 58-62.

Hecker, M. H. L., & Kreul, E. J. (1971). Descriptions of the speech of patients with cancer of the vocal folds. Part 1: Measures of fundamental frequency. Journal of the Acoustical Society of America, 49, 1275-1282.

Heiberger, V. L., & Horii, Y. (1982). Jitter and shimmer in sustained phonation. In N. J. Lass, (ed.) Speech and Language: Advances in Basic Research and Practice, Vol. 7 (pp. 299-232). New York: Academic Press.

Higgins, M. B., & Saxman, J. H. (1989). A comparison of intrasubject variation across sessions of three vocal frequency perturbation indices. Journal of the Acoustical Society of America, 86, 911-916.

Hollien, H., Michel, J., & Doherty, E. T. (1973). A method for analyzing vocal jitter in sustained phonation. Journal of Phonetics, 1, 85-91.

Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. Journal of Speech and Hearing Research, 22, 5-19.

Horii, Y. (1980). Vocal shimmer in sustained phonation. Journal of Speech and Hearing Research, 23, 202-209.

Horii, Y. (1982). Jitter and shimmer differences among sustained vowel phonations. Journal of Speech and Hearing Research, 25, 12-14.

Howard, D. M. (1995). Variation of electroglottographically derived closed quotient for trained and untrained adult female singers. Journal of Voice, 9, 163-172.

Howard, D. M., Lindsey, G. A., & Allen, B. (1990). Toward the quantification of vocal efficiency. Journal of Voice, 4, 205-212.

International Association of Logopedics and Phonatrics (IALP) Voice Committee Discussion of Assessment Topics. (1992). Journal of Voice, 6, 194-210.

Jacob, L. (1968). Normative study of laryngeal jitter. Unpublished Master's thesis. University of Kansas, Lawrence, KS.

Karnell, M. P. (1991a). Fundamental frequency and perturbation measurement. Seminars in Speech and Language, 12, 88-97.

Karnell, M. P. (1991b). Laryngeal perturbation analysis: minimum length of analysis window. Journal of Speech and Hearing Research, 34, 544-548.

Karnell, M. P., Hall, K. D., & Landahl, K. L. (1995). Comparison of fundamental frequency and perturbation measurements among three analysis systems. Journal of Voice, 9, 383-393.

Koike, Y. (1973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. Studia Phonologica, 7, 17-23.

Koike, Y., Takashashi, H., & Calcaterra, T. C. (1977). Acoustic measures for detecting laryngeal pathology. Acta Otolaryngologica, 84, 105-117.

Laver, J., Hiller, S., & Beck, J. M. (1992). Acoustic waveform perturbations and voice disorders. Journal of Voice, 6, 115-126.

Lieberman, P. (1961). Perturbations in vocal pitch. Journal of the Acoustical Society of America, 33, 597-603.

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. Journal of the Acoustical Society of America, 35, 344-353.

Lindqvist, J. (1970). The voice source studied by means of inverse filtering. Speech Transmission Laboratory: Quarterly Progress and Status Report, 1, 3-9.

Ludlow, C. L., Bassich, C. J., Connor, N. P., Coulter, D. C., & Lee, J. L. (1987). The validity of using phonatory jitter and shimmer to detect laryngeal pathology. In T. Baer, C. Sasaki, & K. S. Harris (Eds.), Laryngeal Function in Phonation and Respiration (pp. 492-508). Boston: Little, Brown and Company.

Ludlow, C. L., Coulter, D. C., & Gentges, F. (1983). The differential sensitivity of frequency perturbation to laryngeal neoplasms and neuropathologies. In D. Bless and J. Abbs (Eds.), Vocal Fold Physiology (pp. 381-392). San Diego: College Hill Press.

Milenkovic. P. (1987). Least mean square measures of voice perturbation. Journal of Speech and Hearing Research, 30, 529-538.

Miller, R. (1986). The structure of singing: system and art in vocal technique. New York: Macmillan.

Moore, P., & von Leden, H. (1958) Dynamic variations of the vibratory pattern in the normal larynx. Folia Phoniatrica, 10, 205-238.

Murry, T., & Doherty, E. T. (1980). Selected acoustic characteristics of pathological and normal speakers. Journal of Speech and Hearing Research, 23, 361-369.

Murry, T., Large, J., & Dalgaard, J. (1979). Vocal jitter in sung and spoken vowels. Journal of Research in Singing, 2, 28-43.

Nittrouer, S., McGowan, R. S., Milenkovic, P. H., & Beeher, O. (1990). Acoustic measurements of men's and women's voices: a study of context effects and covariation. Journal of Speech and Hearing Research, 33, 761-775.

Newman, R. A., & Emanuel, F. W. (1991). Pitch effects on vowel roughness and spectral noise for subjects in four musical voice classifications. Journal of Speech and Hearing Research, 34, 753-760.

Orlikoff, R. F. (1995). Vocal stability and vocal tract configuration: an acoustic and electroglottographic investigation. Journal of Voice, 9, 173-181.

Orlikoff, R. F., & Baken, R. J. (1990). Consideration of the relationship between the fundamental frequency of phonation and vocal jitter. Folia Phoniatrica, 42, 31-40.

Orlikoff, R. F., & Kahane, J. C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. Journal of Voice, 5, 113-119.

Pabon, J. P. H. (1991). Objective acoustic voice-quality parameters in the computer phonetogram. Journal of Voice, 5, 203-216.

Perry, C. K., Ingrisano, D. R-S., & Blair, W. B. (1996). The influence of recording systems on jitter and shimmer estimates. American Journal of Speech-Language Pathology, 5, 86-90.

Ramig, L., & Ringel, R. (1983). Effects of physiological aging on selected acoustic characteristics of voice. Journal of Speech and Hearing Research, 26, 22-30.

Rubin, H. J., LeCover, M., & Vennard, W. (1967). Vocal intensity, subglottic pressure and airflow relationships in singers. Folia Phoniatrica, 19, 393-413.

Sabol, J. W. (1992). The value of vocal function exercises in the practice regimen of singers. DMA Dissertation, The University of Cincinnati.

Sabol, J. W., Lee, L., & Stemple, J. C. (1995). The value of vocal function exercises in the practice regimen of singers. Journal of Voice, 9, 27-36.

Scherer, R., Gould, W., Titze, I., Myers, A., & Sataloff, R. (1988). Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis. Journal of Voice, 2, 230-244.

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

Simon, C. (1927). The variability of consecutive wavelengths in vocal and instrumental sounds. Psychological Monographs, 36, 41-83.

Sorensen, D., & Horii, Y. (1983). Frequency and amplitude perturbation in the voices of female speakers. Journal of Communication Disorders, 16, 57-61.

Sorensen, D., & Horii, Y. (1984). Directional perturbation factors for jitter and shimmer. Journal of Communication Disorders, 17, 143-151.

Stemple, J. (1993). Voice therapy: clinical studies. St. Louis: Mosby.

Stoicheff, M. L. (1981). Speaking fundamental frequency characteristics of nonsmoking female adults. Journal of Speech and Hearing Research, 24, 437-441.

Sulter, A. M., Schutte, H. K., & Miller, D. G. (1995). Differences in phonetogram features between male and female subjects with and without vocal training. Journal of Voice, 9, 363-377.

Sulter, A. M., Schutte, H. K., & Miller, D. G. (1996). Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and

female subjects, and effects of varying sound intensity, fundamental frequency, and age. Journal of Voice, 10, 175-189.

Sundberg, J. (1973). The source spectrum in professional singing. Folia Phoniatrica, 25, 71-90.

Sundberg, J. (1987). The Science of the Singing Voice. Dekalb, IL: Northern Illinois University Press.

Sussman, J. E., & Sapienza, C. (1994). Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. Journal of Voice, 8, 145-156.

Teachey, J. C., Kahane, J. C., & Beckford, N. S. (1991). Vocal mechanics in untrained professional singers. Journal of Voice, 5, 51-56.

Titze, I. R. (1994). Principles of Voice Production. Engelwood Cliffs, NJ: Prentice Hall.

Titze, I. R. (1995). Singing: A story of training entrained oscillators. Journal of the Acoustical Society of America, 97, 704.

Titze, I. R., & Sundberg, J. (1992) Vocal intensity in speakers and singers. Journal of the Acoustical Society of America, 91, 2936-2946.

Vennard, W. (1967). Singing: the mechanism and the technic. New York: Fisher.

Verstraete, J., Forrez, G., Mertens, P., & Debruyne, F. (1993). The effect of sustained phonation at high and low pitch on vocal jitter and shimmer. Folia Phoniatrica, 45, 223-228.

Wilcox, K. A., & Horii, Y. (1980). Age and changes in vocal jitter. Journal of Gerontology, 35, 194-198.

Wolf, S. K., Stanley, D., & Sette, W. J. (1935). Quantitative studies on the singing voice. Journal of the Acoustical Society of America, 6, 255-266.

Wolfe, V., Cornell, R., & Palmer, C. (1991). Acoustic correlates of pathologic voice types. Journal of Speech and Hearing Research, 34, 509-516.

APPENDICES

APPENDIX 1: RELIABILITY DATA

Intrasubject Reliability

Condition	Subject T3 Time 1 (t_1)	Subject T3 Time 2 (t_2)	Subject T4 Time 1 (t_1)	Subject T4 Time 2 (t_2)
low-soft	0.52	0.37	1.30	0.86
low-med	0.26	0.26	0.33	0.39
low-loud	0.23	0.23	0.30	0.31
high-soft	0.73	0.72	0.50	0.60
high-med	0.41	0.39	0.48	0.38
high-loud	0.35	0.33	0.36	0.32
	***SE = 0.03**		**SE= 0.14**	
	Pearson r = 0.95		**Pearson r = 0.93**	

Condition	Subject U8 Time 1 (t_1)	Subject U8 Time 2 (t_2)	Subject U5 Time 1 (t_1)	Subject U5 Time 2 (t_2)
low-soft	2.02	2.58	2.57	1.42
low-med	0.42	0.35	0.29	0.29
low-loud	0.35	0.50	0.19	0.23
high-soft	0.92	0.88	0.43	0.36
high-med	1.09	1.64	0.71	0.58
high-loud	0.59	0.61	0.39	0.54
	SE = 0.15		**SE = 0.22**	
	Pearson r = 0.98		**Pearson r = 0.98**	

Jitter Measurement Reliability

Sample	Time 1 (t_1)	Time 2 (t_2)
1	0.39	.38
2	0.74	0.74
3	0.36	0.36
4	2.02	2.05
5	0.59	0.57
6	0.32	0.32
7	0.67	0.65
8	0.72	0.72
9	0.46	0.45
10	1.30	1.30
11	0.48	0.47
12	0.37	0.37
	SE = 0.004	
	Pearson r = 0.99	

Signal Digitization Reliability

Subject	Condition	Time 1 (t_1)	Time 2 (t_2)
T3	low-soft	3.42	3.24
	low-med	1.30	0.29
	low-loud	0.27	0.34
	high-soft	2.44	2.56
	high-med	0.85	0.85
	high-loud	0.36	0.37
U6	low-soft	3.48	3.60
	low-med	1.20	1.14
	low-loud	*distorted*	*distorted*
	high-soft	1.81	1.76
	high-med	3.05	2.86
	high-loud	1.12	0.90
		SE = 0.04	
		Pearson r = 0.99	

***Standard Error formula:** $SE = \frac{\sqrt{\Sigma(t_1 - t_2)^2}}{\sqrt{N(N-1)}}$ (Blommers & Forsyth, 1977)

APPENDIX 2: RAW DATA FOR TRAINED SUBJECTS

Measured jitter values for each of three tokens at each experimental intensity and frequency are listed for each of the 10 trained subjects. Median (or modal) jitter values (used for analysis) are listed first in bold.

	SFF			Octave above SFF		
	60 dB	70 dB	80 dB	60 dB	70 dB	80 dB
TRAINED	low-soft	low-med	low-loud	high-soft	high-med	high-loud
T1	**0.67**	**0.35**	**0.24**	**1.00**	**0.29**	**0.39**
	0.67	0.34	0.34	0.73	0.28	0.29
	0.77	0.38	0.20	1.15	0.43	0.50
T2	**0.68**	**0.39**	**0.29**	**0.94**	**0.46**	**0.40**
	0.45	0.40	0.35	1.00	0.43	0.38
	1.20	0.38	0.27	0.69	0.49	0.41
T3	**0.52**	**0.26**	**0.23**	**0.73**	**0.41**	**0.35**
	0.40	0.22	0.22	0.73	0.36	0.33
	0.60	0.38	0.25	0.71	0.44	0.36
T4	**1.30**	**0.33**	**0.30**	**0.50**	**0.48**	**0.36**
	1.58	0.36	0.32	0.91	0.55	0.47
	1.23	0.32	0.24	0.42	0.32	0.33
T5	**1.14**	**0.41**	**0.41**	**0.60**	**0.48**	**0.31**
	0.76	0.29	0.43	0.63	0.35	0.30
	1.34	0.56	0.36	0.54	1.14	0.35
T6	**1.27**	**0.27**	**0.37**	**0.56**	**0.36**	**0.39**
	1.65	0.26	0.38	0.42	0.35	0.38
	0.79	0.33	0.30	0.60	0.39	0.39
T7	**3.01**	**0.51**	**0.38**	**1.15**	**0.39**	**0.36**
	2.19	0.39	0.40	1.70	0.41	0.38
	4.58	0.51	0.29	0.90	0.39	0.36
T8	**0.38**	**0.24**	**0.67**	**0.55**	**0.43**	**0.29**
	0.48	0.35	0.98	0.50	0.33	0.28
	0.32	0.23	0.58	0.73	0.45	0.30
T9	**0.93**	**0.26**	**0.22**	**0.60**	**0.41**	**0.31**
	1.80	0.24	0.20	0.54	0.96	0.33
	0.73	0.34	0.24	0.67	0.34	0.30
T10	**0.68**	**0.26**	**0.34**	**0.46**	**0.44**	**0.38**
	0.64	0.25	0.36	0.51	0.50	0.36
	0.84	0.28	0.19	0.46	0.41	0.43

APPENDIX 3: RAW DATA FOR UNTRAINED SUBJECTS

Measured jitter values for each of three tokens at each experimental intensity and frequency are listed for each of the 10 untrained subjects. Median (or modal) jitter values (used for analysis) are listed first in bold.

	SFF			Octave above SFF		
	60 dB	70 dB	80 dB	60 dB	70 dB	80 dB
UNTRAINED	low-soft	low-med	low-loud	high-soft	high-med	high-loud
U1	**0.76**	**0.37**	**0.22**	**0.63**	**0.45**	**0.48**
	0.71	0.33	0.31	0.54	0.45	0.61
	0.77	0.41	0.21	0.66	0.42	0.42
U2	**0.66**	**0.32**	**0.33**	**0.70**	**0.47**	**0.28**
	1.05	0.41	0.33	0.64	0.50	0.27
	0.41	0.29	0.30	0.97	0.43	0.28
U3	**2.57**	**0.29**	**0.19**	**0.43**	**0.71**	**0.39**
	1.47	0.30	0.24	0.39	0.42	0.37
	3.66	0.29	0.19	0.71	0.74	0.45
U4	**1.68**	**0.45**	**0.26**	**0.76**	**0.81**	**0.34**
	4.89	0.43	0.25	1.00	0.59	0.35
	0.98	0.96	0.27	0.65	1.98	0.32
U5	**0.74**	**0.32**	**0.35**	**0.62**	**0.54**	**0.43**
	0.70	0.22	0.41	0.76	0.52	0.36
	0.75	0.32	0.24	0.43	0.61	0.54
U6	**2.02**	**0.42**	**0.35**	**0.92**	**1.09**	**0.59**
	4.52	0.34	0.28	0.93	0.97	0.85
	1.54	0.60	0.53	0.66	1.56	0.55
U7	**0.65**	**0.38**	**0.27**	**1.08**	**0.48**	**0.37**
	1.34	0.32	0.28	1.50	0.59	0.39
	0.37	0.39	0.23	0.61	0.45	0.32
U8	**2.00**	**0.30**	**0.20**	**0.67**	**0.53**	**0.42**
	2.06	0.31	0.25	0.55	0.58	0.43
	1.18	0.29	0.19	0.84	0.35	0.38
U9	**1.09**	**0.30**	**0.19**	**0.72**	**0.50**	**0.35**
	1.31	0.29	0.29	0.85	0.78	0.38
	0.84	0.32	0.19	0.67	0.44	0.35
U10	**0.74**	**0.25**	**0.22**	**1.01**	**0.60**	**0.37**
	2.52	0.26	0.24	0.82	0.55	0.32
	0.52	0.24	0.19	1.18	0.80	0.37

VITA

Christopher Somers Gaskill was born on May 8, 1969 in Nashville, TN. He graduated as valedictorian of Glencliff High School in 1987, and enrolled at Rhodes College in Memphis, TN. He graduated *magna cum laude* with a B.A. from Rhodes in 1991. He pursued an interdisciplinary degree with a concentration in physics and completed a senior project which examined the singing voice from the perspective of acoustics. Chris received an M.M. in choral conducting from Emory University in Atlanta, GA in 1994.

Wishing to further pursue his interests in vocal physiology, he began his master's degree program in speech-language pathology at the University of Tennessee, Knoxville in the fall of 1994. He graduated in August 1997 with a Distinguished Academic Award. He is currently employed as a speech-language pathologist at Baptist Hospital of East Tennessee in Knoxville, TN.