A COMPARISON OF AUDITORY-PERCEPTUAL FEATURES OF VOICE QUALITY IN INDIVIDUALS WITH AND WITHOUT WILLIAMS SYNDROME

by

Sara K. Albrecht

Bachelor of Science in Education, 2007
University of Arkansas
Fayetteville, Arkansas

Submitted to the Graduate Faculty of
The Harris College of Nursing & Health Sciences
Texas Christian University
in partial fulfillment of the requirements
for the degree of

Masters of Science in Communication Sciences and Disorders

May 2009

A COMPARISON OF AUDITORY-PERCEPTUAL FEATURES OF VOICE QUALITY IN INDIVIDUALS WITH AND WITHOUT WILLIAMS SYNDROME

Thesis Approved:

____

Major Professor

____

Committee Member

____

Committee Member

____

College of Nursing & Health Sciences

ACKNOWLEDGEMENTS

I would like to give special thanks to the chair of my thesis committee, Dr. Christopher Watts. Without your constant support and great confidence in me, I would not have been able to realize this accomplishment. From the very first time I contacted you at James Madison University (before you had even formally met me in person), you had such enthusiasm about my potential, and you have maintained that ever since. I am truly grateful for all you have done and helped me do, so Thank You.

I would also like to thank my other committee members, Dr. Maria Muñoz and Dr. Raúl Prezas. Your commitment to my project is greatly appreciated, especially considering how late I got started. I am very thankful for the valuable comments you gave and the obvious improvement made to this project through your participation.

Special thanks to my roommate and friend, Sara Kutac, who truly motivated me to do my very best and was a great companion through the stressors (and celebrations!) of graduate school.

Finally, I would like to extend a special word of appreciation to my family and my best friend, Michaela Brady, who have supported and encouraged me throughout my education. Thank you mom and dad for allowing me to be independent, but still always being there for me, especially when I wanted to give up. And thank you to my brothers, Paul, Brian, and Alan, for constantly raising the bar and challenging me to never settle and to always succeed. Thank you Michaela for being a witness to my faith and a true friend in every aspect of my life – you have never given up on me.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Table Page

Figure Page

CHAPTER I

INTRODUCTION

Williams syndrome (WS) is a genetic disorder caused by a hemizygous deletion of a range of genes on chromosome 7, including the elastin gene (ELN) (Shprintzen, 1997; Ewart, et al., 1993). Individuals with WS have many unique characteristics as a result of this deletion. Since elastin is found in the vocal folds, there is a possibility that vocal quality may be affected in individuals with WS, and such changes have been documented in previous studies (Vaux, Wojtczak, Benirschke, & Jones, 2003; Watts, Awan, & Marler, 2008). More specifically, disorders of voice quality have been found when comparing individuals with WS and supravalvular aortic stenosis (SVAS) to normal controls (Watts, et. al, 2008). SVAS is another disorder characterized by ELN abnormalities, and sometimes coexists with WS (Pober & Dykens, 1996). This particular study aimed at comparing perceptual characteristics of vocal quality in individuals with WS (as a group by itself) to normal controls.

CHAPTER II

REVIEW OF LITERATURE

The deletion of the elastin gene, along with deletions of more than 20 other genes on chromosome 7, results in a unique phenotype characterized by a multisystem array of abnormalities in individuals with WS. In a hemizygous deletion, only one allele of a gene is affected while the other remains functional. The resulting phenotype occurs because the unaffected allele cannot produce enough of the gene product to make up for the deleted allele. The phenotype of WS has been well described and includes dysfunction or abnormalities in development, cognition, behavior, language, and hearing, among other systems (Pober & Dykens, 1996; Hepburn, Philofsky, John, & Fidler, 2005; Sugayama, Leone, Chauffaille, Okay, & Kim, 2007).

Williams Syndrome Phenotype

The most commonly reported abnormalities in this population include mental retardation, cardiovascular disorders, microcephaly, hyperacusis, developmental language delays and disorders, hypersociability, and a high tendency towards certain fears (Hepburn, et al., 2005; Shprintzen, 1997; Martens, Wilson, & Reutens, 2008). Mental retardation in this population has been consistently measured at a Full-Scale IQ score range between 50 and 70, or a category of mildly retarded (Howlin, Davies, & Udwin, 1998). Supravalvular aortic stenosis (SVAS) is the most common cardiovascular disorder to co-occur with WS, as both are associated with ELN abnormalities (Pober & Dykens, 1996). SVAS is a condition which causes impeded blood flow from the heart due to a narrowing of the aortic valve. The phenotypic characteristics of WS also include distinct facial features, which have commonly been described as "elfin" in nature and have specifically included descriptions of a wide mouth with full lips, prominent ear lobes, and a short, upturned nose (Bellugi, Lichtenberger, Jones, Lai, & George, 2000).

In the area of language skills, Martens, et al. (2008) reviewed the results of over 30 studies which showed that people with WS show typical, but delayed, development in the areas of complex syntax, semantics, vocabulary, plurals, and irregular past tense. Participants in this study with WS also showed atypical development in areas such as grammatical comprehension, gender agreement, morphosyntax, and pragmatics (Martens, et al., 2008). Jones, et al. (2000) state that the profile of WS "consists of a strong drive toward social interaction with other people" that is consistently rated higher than individuals with other disorders and normal individuals. Individuals with WS have also been shown to have more anxieties and fears, especially "extreme fears," than control groups (Dykens, 2003). Some of these fears include being burned, being hit by a car, a fear of nuclear war, and the fear of electric shock (Dykens, 2003). It has also been reported that voice abnormalities are a characteristic of the WS phenotype. Vaux, Wojtczak, Benirschke, and Jones (2003) state that "one of the most characteristic features of this disorder is a harsh, brassy or hoarse voice". It has also been found that the voices of individuals with WS are perceptually rated as being hoarser, rougher, and lower in pitch than matched controls (Watts, Awan, & Marler, 2008). In addition to perceptual ratings, Watts, et al. (2008) found acoustic measurements of pitch sigma and jitter to be significantly increased in WS.

Voice Abnormalities in Genetic Disorders

Many other genetic disorders have also been associated with vocal abnormalities. For example, several studies have reported hypernasality and increased pitch in Prader-Willi Syndrome, which is a genetic disorder linked with chromosome 15, characterized by hypotonia, developmental delays, and mental retardation (Lewis, Freebairn, Heeger, & Cassidy, 2002; Akefeldt, Akefeldt, & Gillberg, 1997). Additionally, Lewis, et al. reported that individuals with Prader-Willi Syndrome are often judged as having harsh or hoarse voices (2002). The voices of individuals with Down's syndrome have also been consistently perceived as hoarse (Moran, 1986; Wold & Montague, 1979). It has been found that slight roughness is perceived in the voices of individuals with Goldenhar syndrome and they also have measures of increased jitter (Van Lierde, Van Cauwenberge, Stevens, & Dhooge, 2004). John Van Borsel (2004) completed a meta-analysis, based on a table found in Shprintzen's book (Genetics, Syndromes, and Communication Disorders), of 299 genetic syndromes to identify how many are associated with disorder of voice and/or resonance. Of the 299 syndromes he examined, voice disorders occurred in 118 of them, and was most often noted to be a problem with vocal quality, usually hoarseness (Van Borsel, 2004). Some genetic disorders that are associated with vocal hoarseness include Apert syndrome, Brachial plexus neuropathy, CHARGE association, Cowden syndrome, Down syndrome, Fetal alcohol syndrome, Hay-Wells syndrome, Hunter syndrome, Rapp-Hodgkin syndrome, Urbach-Wiethe syndrome, and many others (Shprintzen, 1997). Pitch problems, specifically increased pitch, are also reported in about 25% of genetic disorders, including Bloom syndrome, Cartilage-hair hypoplasia, Cri-du-chat syndrome, Hypochondroplasia, Pyknodysostosis, and many others (Van Borsel, 2004; Shprintzen, 1997). All of these genetic disorders have been identified as being associated with a voice disorder, but there is not enough information available to elucidate the cause. Sataloff (1995) recommends that more research in this area is needed and should focus on the relationship between genetics and voice. The current evidence does suggest that gene expression is important for vocal fold development and vocal function. In addition, because the genotypes have been mapped, many of these genetic disorders open up the opportunity to study specific genetic influences on vocal fold physiology.

Human Vocal Fold

Altered vocal fold physiology is at the root of many voice production abnormalities, including those associated with specific genetic disorders (Shprintzen, 1997). The normal human vocal fold is a complex structure that is responsible for the production of voiced sounds in which subtle changes in physiology can produce rather significant changes in perception. The foundation of the vocal fold is the vocalis muscle, which is deep to the lamina propria and the outer epithelium. The vocalis muscle gives the vocal fold its mass and tone (Stemple, Glaze, & Klaben, 2000). The anatomy of the lamina propria was first described by Minoru Hirano (1981) and consists of three layers, the superficial layer, the intermediate layer, and the deep layer. These three layers were classified based on their fibrous make-up, that is, their relative proportion of elastic and collagenous fibers. The superficial, or outermost layer of the lamina propria, is made up of loose tissue (e.g., ground substance) and very few collagen or elastic fibers (Gray, Titze, Alipour, & Hammond, 2000). Elastic fibers are most predominant in the intermediate layer, while the deep layer consists mostly of collagenous fibers (Stemple, et al., 2000). Within the lamina propria, and the vocal fold structure as a whole, the deepest layers are the most dense and each layer gets progressively more flexible as it gets closer to the surface. The epithelium is the thinnest and most pliable layer of the vocal fold, allowing it to vibrate easily (Stemple, et al., 2000). The movement of the vocal folds is commonly described as a mucosal wave, which is initiated in the deepest portion of the folds and continues to the outermost portion (Gray, et al., 2000).

Differences in the physiology of the vocal folds will cause perceptual differences, as well. The most basic illustration of this is a difference in mass; as a general rule, the larger mass of a man's vocal folds produces a lower pitch than the smaller mass of a woman's vocal folds (Stemple, et al., 2000). At a disordered level, pathologies such as vocal hemorrhage, Reinke's edema, and polyps may result in a decrease in pitch, most likely because of how these pathologies increase the mass of the folds (Stemple, et al., 2000). Many vocal fold pathologies result in the vocal folds being unable to vibrate periodically, which causes the perception of roughness. This is the case in pathologies such as vocal nodules, polyps, and vocal fold paralysis. Any disorder that does not allow the vocal folds to completely adduct (vocal nodules, bowed vocal folds, contact granulomas) may give the voice a breathy quality, because it allows more air to escape than normal.

Perceptual Qualities of Voice

There are a number of perceptual qualities that can be measured within the human voice. Among others, these include perceptions of overall vocal quality, such as roughness, breathiness, and hoarseness. Roughness is often a result of irregular vocal fold vibration which is perceived as an unpleasant voice that may be coarse or raspy (Awan & Lawson, in press; Hirano, 1981). The quality of breathiness can be a result of hypoadduction and is usually perceived as excess air escaping through the glottis (Awan & Lawson, in press; Hirano, 1981). Hoarseness can be perceived as a voice producing “both breathy and rough qualities simultaneously” (Awan & Lawson, in press). While these terms are defined similarly across publications, measures of perception have inherent issues of reliability. Two different people may listen to the same voice and one may rate the voice as rough and one may rate the voice as hoarse. Issues related to the accuracy of ratings, therefore, are a concern. When measuring the perceptual qualities of voice, then, it can be helpful to have an anchor voice available, as well as a definition of the parameter being measured (Awan & Lawson, in press). In fact, Awan and Lawson (in press) found that when these two types of anchors were used independently, reliability was greater than without an anchor, but when both anchors were used by perceptual judges, reliability was significantly stronger.

CHAPTER III

STATEMENT OF PURPOSE

Previous research has demonstrated that the elastic fibers within the vocal folds of an individual with WS were altered in structure and number (Vaux, et al., 2003). In addition, when individuals with SVAS and WS were studied together as one group, voice quality and acoustic measures of vocal fold physiology were found to be different than normal controls (Watts, et al., 2008; Watts, Marler, & Urban, 2007). This evidence suggests that the vocal fold structure, as described previously, is likely different in the WS population compared to a normal population. However, to date, the vocal function of individuals with WS has not been studied separately from that of SVAS. The purpose of the present study was to investigate the auditory-perceptual characteristics of voice quality in a group of individuals with WS compared to normal controls, in an attempt to replicate previous findings and further determine if the perceptual characteristics previously reported in SVAS and WS voices were maintained when WS was studied as a separate group. The specific research questions of this study were: (1) Do significant differences exist in the ratings of mean frequency of voice quality category in individuals with Williams syndrome compared to normal controls? (2) Was the severity of voice quality deviation from normal rated differently in individuals with Williams syndrome than in normal controls? (3) Was the vocal pitch of individuals with Williams syndrome rated differently than that of normal controls?

CHAPTER IV

METHODOLOGY

Participants

Participants in this study included speakers and perceptual judges. Two groups of speaking participants were recruited for this study. An experimental group consisting of 16 individuals with WS (8 male, 8 female) were recruited and recorded previously. Inclusion criteria for these participants consisted of (1) genetically confirmed WS genotype via fluorescence in situ hybridization (FISH) testing and (2) confirmation of the WS behavioral phenotype by a medical professional. A convenience sample of 16 normal controls was recruited through class announcements, word of mouth, and phone calls from the investigator's acquaintances from the Dallas/Fort Worth metroplex and the Greater Houston area. The control group was matched to WS participants on two variables: (1) gender and (2) chronological age. Participants were gender-matched equally (8 normal males and 8 normal females). For age, the normal group as a whole fell within the range (+/- 2 years) of ages recruited for the experimental group (range = 16 years to 56 years for males; 18 years to 34 years for females), and were within 6 months of the group's mean age (mean = 28.6 years for males; 24.6 years for females). In addition, each participant was matched within 6 months for children below 18 years of age and within 3 years for adults. Inclusion criteria for the normal controls consisted of the following: self-reported ability to read, negative history for voice disorder or neurological disorder, non-smoker, and no current voice complaints. Table 1 includes age and gender information for all speaking participants. No participants were under 16 years of age.

Table 1: Speaking Participants

Group	**Gender**	**Age (years)**	**Group Mean/Range**
WS			
L	M	47	
B	M	22	
S	M	31	
D	M	16	
G	M	17	
J	M	56	
J	M	18	*WS Male Mean: 28.6*
B	M	22	*WS Male Range: 16-56*
B	F	19	
K	F	19	
K	F	18	
H	F	20	
D	F	28	
M	F	31	
T	F	28	*WS Female Mean: 24.6*
K	F	34	*WS Female Range: 18-34*
Normal			
RA	M	55	
BA	M	23	
PM	M	50	
AA	M	20	
AA	M	16	
SA	M	27	
CP	M	16	*Normal Male Mean: 28.6*
JF	M	22	*Normal Male Range: 16-55*
KW	F	19	
KG	F	34	
JB	F	27	
JS	F	19	
BG	F	28	
CC	F	27	
DG	F	19	*Normal Female Mean: 24.1*
AK	F	20	*Normal Female Range: 19-34*

Twenty students enrolled in the Communication Sciences and Disorders program at TCU were recruited to serve as perceptual judges. For inclusion in the study, judges were 18 years of age or older and had self-reported hearing within normal limits.

Instrumentation

A portable digital audio recorder (Marantz PMD 670) was used to record normal voices (a digital recorder was also previously used to record the WS voices). These were then digitally transferred to the Computerized Speech Lab (CSL - KayPENTAX, Lincoln Park, NJ) and saved as .wav files for later auditory manipulation. An AKG Acoustics head-mounted microphone was connected to the digital recorder for direct analog-to-digital conversion of vocal productions at 44.1KHz sampling rate. The microphone was placed approximately 3cm off the left corner of the mouth for all speaking participants.

The Anchors software program, developed by Dr. Shaheen Awan at Bloomsburg University, was used to collect auditory-perceptual ratings of voice quality type, severity, and pitch for the recorded voice stimuli. The validity of the software for use in auditory-perceptual ratings has been established in previous studies (Awan & Roy, 2005; Awan & Roy, 2006; Awan & Lawson, in press, Watts, et al., 2008). The auditory-perceptual rating task is described in more detail in the Procedures section.

For perceptual training sessions and experimental data collection sessions, a standard desktop computer wired to circumaural headphones was used to present auditory stimuli to perceptual judges at a comfortable listening level, as established by individual judges.

Procedures

Speaker Recordings: All speaking participants (WS & control) were recorded with similar equipment in a similar low-background noise environment (<45dbSPL), confirmed by a portable sound level meter. Speaker participation took approximately 20 to 30 minutes. Recording began after an explanation of the study procedures and associated risks were given to the participants and caregivers (in the case that the participant was under 18) and the consent form was signed (individuals younger than 18 were not allowed to participate without the signed consent of a caregiver) (see Appendix A). Demographic and history information was then collected (see Appendix C).

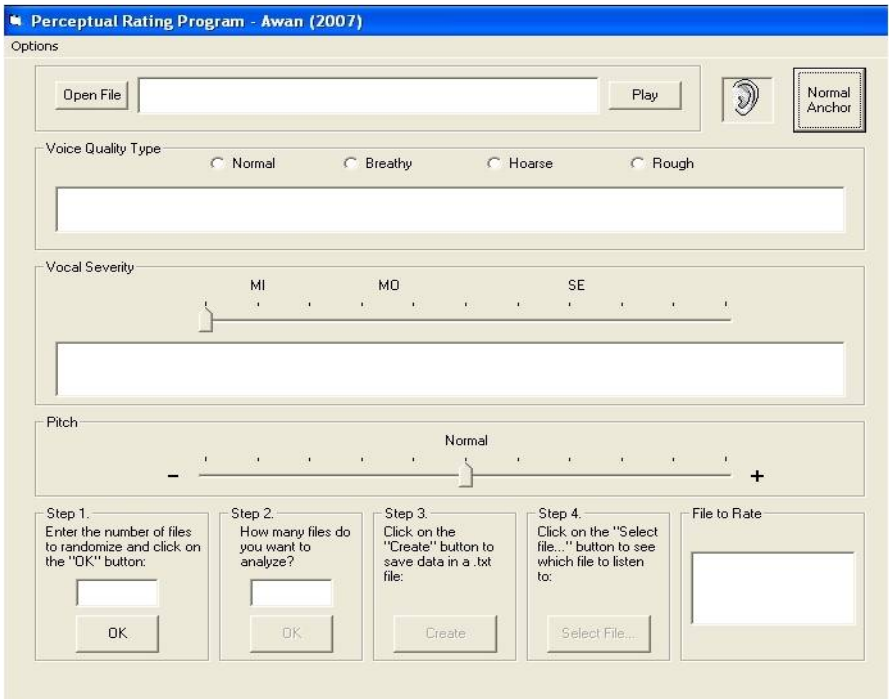
All speaking participants were instructed to take a sufficient breath and sustain the vowel /a/ at a comfortable pitch and loudness until asked to stop (approximately 5 seconds) and repeated this task three times. The same participants were also asked to sustain other vowels and participate in oral reading tasks (see Appendix D). This data was collected for a separate study, which will not be described here.

The three digitized, sustained /a/ vowels produced by the WS and control participants were copied onto the hard-drive of a standard desktop computer. Each vowel was then cropped to a one-second sample. The first /a/ recorded by each participant was used as the experimental stimuli for auditory-perceptual tasks. These stimuli were duplicated, randomly assigned a file number (1-64), and saved to the computer.. The same computer was used to control the experimental software (Anchors), playback of stimuli, and recording of participant responses. The stimuli were duplicated to allow each judge to listen to and rate each voice twice, while maintaining the random order of presentation.

Perceptual Judgments: Participation by the judges in this study took approximately one hour. Any student recruited for participation was informed of her/his choice to refuse to participate or withdrawal without impacting their academic standing. An explanation of the study procedures and associated risks was given to the participants and the consent form was signed (see Appendix B). Prior to the judgment task, judges participated in a 20-minute training session. During the training session, instructions regarding use of the computerized response form and a review of the operational definitions for the voice quality categories (normal, breathy, hoarse, rough), pitch (low pitch, high pitch) and severity (normal, mild, moderate, severe) were provided. Voice quality categories were defined for each judge as they appeared in the Anchors software: Normal - "The voice does not differ substantially from your expectations in terms of parameters such as quality", Breathy – "Commonly perceived as a whispery or airy voice; associated with hypoadduction", Rough – "Commonly perceived as an unpleasant voice and associated with terms such as coarse, strident, low-pitched noise, and rasping; associated with hyperadduction and irregular vocal fold vibration", Hoarse – "A hybrid descriptor that denotes a voice with both breathy and rough qualities simultaneously" (Awan & Roy, 2005). Each judge then listened to pre-recorded voice samples which are representative for the range of voice categories (breathy, rough, hoarse) and severities (mild and severe) possible for individuals sustaining the vowel /a/. These representative stimuli have been validated in three published studies (Awan & Roy, 2005; Awan & Roy, 2006, Watts, et al., 2008). After listening to these representative stimuli, judges compared them to the normal anchor and were guided through the three steps of rating each voice (quality, severity, and pitch). Before completing the training session, judges were given an opportunity to ask questions and make sure the task was clear. After the training session ended, judges were not able to listen to the representative stimuli again for purposes of comparing them to the experimental stimuli.

Once the training session was complete, judges listened to one stimulus at a time while seated and wearing the circumaural headphones. The Anchors software, depicted in Figure 1, allowed the user to select samples for playback from the CD-ROM, in an order that was randomized for each judge. After listening to each stimulus, the judges were asked to make ratings regarding the quality, severity, and pitch of that voice. For voice quality ratings, judges could choose “normal”, “breathy”, “hoarse”, or “rough”. For severity ratings, listeners used a visual analog (VA) scale, and designated the severity of each voice’s quality. The far left of the scale indicated normal (value=0), and the far right of the scale indicated greatest severity (value=100). The scale did not have corresponding numbers, but was marked from left to right with anchors to indicate mild (MI), moderate (MO), and severe (SE), and a definition for them would appear when the mouse scrolled over each one. The “MI” label appeared above the second tick mark (value = 10). The “MO” label appeared between the fourth and fifth tick marks (value = 35). The “SE” label appeared above eighth tick mark (value = 70). The listener could place a mark at any point along the line. Pitch was also rated by the listeners on a 0-100 point visual analog scale, marked off in increments of 10 with “-” on the far left (low pitch) and “+” on the far right (high pitch). The Anchors Program set the pitch marker by default at 50 (or in the exact middle of the scale) to indicate a “normal” pitch. The listener moved the marker to the left (low) or right (high) of normal to indicate their perception of pitch for each voice. If the pitch was perceived to be normal, the listener was instructed to leave the marker at the mid-point. The Anchors program only records the ratings that the judges made; the program itself does not assess the voices. Therefore, there is no “right” or “wrong” rating on any parameter.

Figure 1. Screen shot of Anchors Software Program.



Judges were allowed to replay each sample as many times as necessary before rating it. In addition, judges were allowed to compare each voice sample with the two pre-selected “normal” voice samples (i.e., normal male and normal female) used in the training session. These samples served as an external auditory standard (e.g., a perceptual anchor). The use of perceptual anchors in this way have been shown to be beneficial to auditory-perceptual judgments of voice quality and severity as they help to reduce listener-related variability (Awan & Lawson, in press; Awan & Roy, 2006; Eadie & Baylor, 2006). The same two normal perceptual anchors were used for all judgments, for every judge. In addition to the perceptual anchor, judges also had access to the operational definitions for each rating, as reviewed in the training session. For purposes of measuring intra-judge reliability, every judge listened to and rated each voice sample twice (first rating and second rating) during the study.

Analyses

SPSS was utilized for all statistical analyses. In this study, data were collected for 3 separate measures. The first measure was the mean frequency of rating in each voice quality category (“normal”, “hoarse”, “breathy”, or “rough”) for WS & normal speakers. For this measure, the analysis utilized two independent variables: (1) group, a between-subject variable, and (2) voice quality category, a within-subject or repeated measures variable. When making ratings, the twenty judges assigned each WS and normal speaker to one of the voice quality categories. The mean frequency of rating in each voice quality category was then computed for each speaker, pooling across the 20 judges. This created four separate levels of the voice quality category variable (“normal”, “hoarse”, “breathy”, “rough”). The voice quality category variable was repeated measures because each speaker had the opportunity to be rated in every level of the variable. The design of this analysis is illustrated in Table 2, which shows group (WS, Normal) as one independent variable and voice quality category (“normal”, “hoarse”, “breathy”, “rough”) as a second independent variable. The dependent variable is shown as the mean frequency of rating.

Table 2. Organization of the data for ratings of voice quality, with group and voice quality category as the two independent variables.

	Voice Quality Category			
Group	“Normal”	“Hoarse”	“Breathy”	“Rough
WS	*# Mean frequency*	*# Mean frequency*	*# Mean frequency*	*# Mean frequency*
Normal	*# Mean frequency*	*# Mean frequency*	*# Mean frequency*	*# Mean frequency*

A two-way mixed ANOVA (group x voice quality category, with repeated measures on quality) was used to assess the effects of group and voice quality category.

The second measure was perceived vocal severity. For this measure, the analysis utilized only one independent variable, group (WS, Normal), with mean severity rating (ranging from 0 – 100) as the dependent variable. An independent samples t-test was applied to this data to assess group differences.

The third measure was perceived vocal pitch. For this measure, the analysis again only utilized one independent variable, group (WS, Normal), with mean pitch rating (ranging from 0 – 100) as the dependent variable. An independent samples t-test was applied to this data to assess group differences.

Intra-judge reliability was assessed for the ratings of voice quality category, severity and pitch ratings using a Pearson product-moment correlation, which assessed the degree of relationship between the first and second ratings of each stimulus by the judges. All reliability correlations were high, suggesting a large degree of reliability within each judge across all ratings, and all comparisons were statistically significant at the .01 level. Correlation coefficients for the specific intra-judge reliability measures were as follows: Voice quality – ratings of “normal” (r = .90), ratings of “breathy” (r = .80), ratings of “rough” (r = .63), ratings of “hoarse” (r = .80); Pitch (r = .97); Severity (r = .95).

CHAPTER V

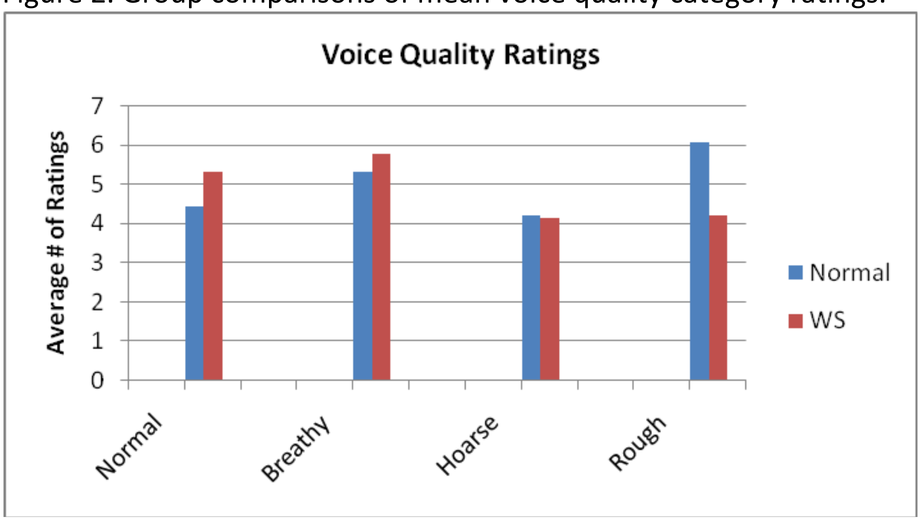
RESULTS

Quality Category: The judges rated each voice for voice quality as either “normal”, “breathy”, “rough”, or “hoarse”. The mean frequency of rating for each category was computed, pooling the data across the 20 judges. The average data for group and voice quality category is presented in Table 3. Normal voices were rated as “normal” an average of 4.44 times (or 11.1% of the time), with a standard deviation (SD) of 3.076. WS voices were rated as “normal” an average of 5.31 times (or 13.3% of the time), with a standard deviation of 4.045. Ratings of “breathy” were assigned to normal and WS voices an average of 5.31 times (SD=3.754) and 5.75 times (SD=3.786), respectively. In other words, normal voices were rated “breathy” 13.3% of the time, and WS voices were rated “breathy” 14.4% of the time. A “hoarse” rating was given an average of 4.19 times (or 10.5% of the time) with a standard deviation of 3.016 for normal voices, and an average of 4.12 times (10.3% of the time) with a standard deviation of 2.872 for WS voices. Normal voices were rated as “rough” an average of 6.06 times (15.2%), with a standard deviation of 2.720. WS voices were rated as “rough” an average of 4.19 times (10.5%), with a standard deviation of 3.167. The descriptive statistics for quality ratings are presented in Table 2, as well as illustrated in Figure 2. Based on the descriptive statistics, there did not appear to be any large difference between the two groups (normal & WS) in the way they were rated on any level of voice quality category, though the WS voices were rated as “normal” slightly more often than normal voices and normal voices were rated as “rough” slightly more often than WS voices.

Table 3: Descriptive Statistics for voice quality category in each group.

	Group	Mean	Std. Deviation	N
Normal	Normal	4.44	3.076	16
	WS	5.31	4.045	16
	Total	4.87	3.563	32
Breathy	Normal	5.31	3.754	16
	WS	5.75	3.786	16
	Total	5.53	3.716	32
Hoarse	Normal	4.19	3.016	16
	WS	4.12	2.872	16
	Total	4.16	2.897	32
Rough	Normal	6.06	2.720	16
	WS	4.19	3.167	16
	Total	5.13	3.056	32

Figure 2. Group comparisons of mean voice quality category ratings.



Statistical Analyses for Quality Category: Different statistics were used to analyze the Quality data than the Severity and Pitch data. For Quality, a two-way (group x voice quality category) mixed Analysis of Variance (ANOVA) with repeated measures on the voice quality variable was completed in order to determine whether each independent variable, group (WS or normal) and quality (normal, breathy, rough, hoarse), or the interaction of these variables had an effect on the dependent variable (mean frequency of rating). For this analysis, an alpha level of .05 was used as the criterion for statistical significance. The results of the ANOVA did not show any significant effects for quality (F[3,90]=.732, p=.535) or group (F[1,30]=1.000, p=.325). The interaction of group and quality also did not prove statistically significant (F[3,90]=.801, p=.496). Table 4 presents the results of the ANOVA for the within-subject variable (voice quality) and the interaction effects. Table 5 presents the results of the ANOVA for the between group (WS vs. Normal) effects.

Table 4: Tests of within-subjects effects for voice quality and interaction effect for group x voice quality category.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Category	Sphericity Assumed	32.031	3	10.677	.732	.535
	Greenhouse-Geisser	32.031	2.255	14.206	.732	.500
	Huynh-Feldt	32.031	2.529	12.668	.732	.514
	Lower-bound	32.031	1.000	32.031	.732	.399
Category * Group	Sphericity Assumed	35.031	3	11.677	.801	.496
	Greenhouse-Geisser	35.031	2.255	15.536	.801	.466
	Huynh-Feldt	35.031	2.529	13.854	.801	.478
	Lower-bound	35.031	1.000	35.031	.801	.378
Error(Category)	Sphericity Assumed	1311.938	90	14.577		
	Greenhouse-Geisser	1311.938	67.645	19.394		
	Huynh-Feldt	1311.938	75.856	17.295		
	Lower-bound	1311.938	30.000	43.731		

a. Computed using alpha = .05

Table 5: Tests of between-subjects effects for voice quality

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	3100.781	1	3100.781	3969.000	.000
Group	.781	1	.781	1.000	.325
Error	23.438	30	.781		

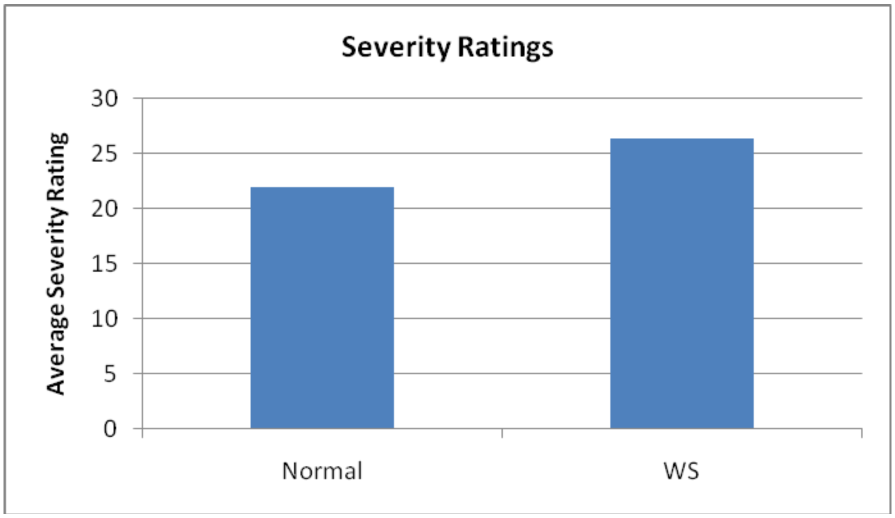
a. Computed using alpha = .05

<u>Severity</u>: On the visual analog scale from 0-100, normal voices were assigned an average severity rating of 21.88 with a SD of 10.492. WS voices were rated at an average severity of 26.28 with a SD of 14.332. Overall, WS voices were rated only slightly more severe than normal voices. However, the higher standard deviation for the WS voices shows that there was greater variability in the ratings of severity for this group as compared to the normal control group. Statistical data for these severity ratings is presented in Table 6. The average ratings assigned are also presented visually in Figure 3.

Table 6: Group Statistics for measures of voice severity and pitch.

	Group	N	Mean	Std. Deviation	Std. Error Mean
Severity	Normal	16	21.88	10.492	2.623
	WS	16	26.28	14.332	3.583
Pitch	Normal	16	42.09	10.081	2.520
	WS	16	42.78	10.751	2.688

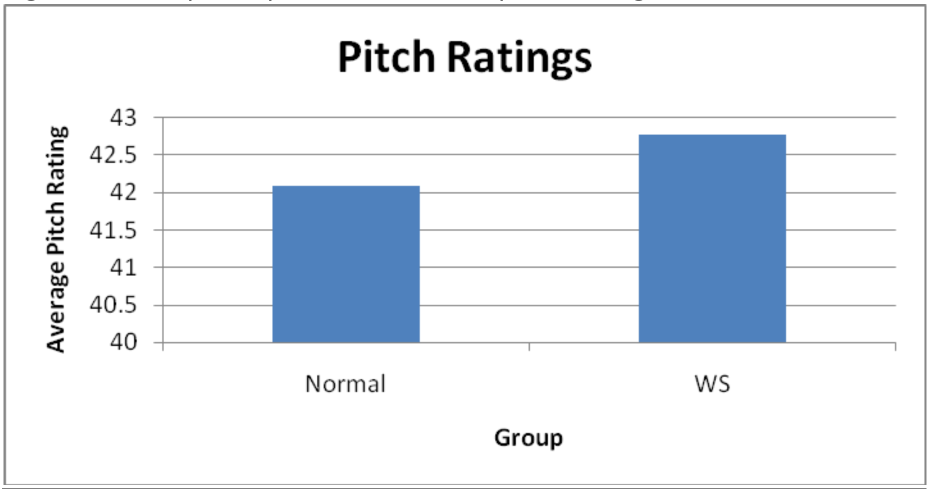
Figure 3. Group comparisons of mean severity ratings.



Statistical Analyses for Severity: To analyze Severity and Pitch data, two separate independent samples t-tests were conducted, one for severity and one for pitch. For both of these analyses, group (WS vs. Normal) was the independent variable. When applied to the severity data, no significant effects were found ($t[1,30]=-.989$, $p=.330$). These results are presented in Table 7 and show that the two groups (WS & normal) were not rated significantly different for vocal severity.

Pitch: On the visual analog scale from 0-100, normal voices were given an average pitch rating of 42.09 with a SD of 10.081. WS voices were rated at an average pitch of 42.78 with a SD of 10.751. Both groups were assigned an average pitch that was slightly lower than normal. Statistical data for these pitch ratings is presented in Table 6. The average ratings assigned are also presented visually in Figure 4.

Figure 4. Group comparisons of mean pitch ratings.



Statistical Analyses for Pitch: An independent sample t-test also was applied to the pitch data and there was not a significant difference between the groups (t[1,30]=-.186, p=.854). These results are presented in Table 7 and show that the two groups (WS & normal) were not rated significantly different for pitch.

Table 7: Independent Samples tests comparing group effects on severity and pitch ratings.

		t-test for Equality of Means		
		T	Df	Sig. (2-tailed)
Severity	Equal variances assumed	-.989	30	.330
	Equal variances not assumed	-.989	27.490	.331
Pitch	Equal variances assumed	-.186	30	.854
	Equal variances not assumed	-.186	29.877	.854

CHAPTER VI

DISCUSSION

The purpose of this study was to determine whether WS voices would be assigned a different vocal quality, severity, and pitch rating as compared to normal voices. This investigation was conducted in order to replicate previous findings of perceptual characteristics that have been linked with individuals with ELN abnormalities. The specific research questions were as follows: (1) Do significant differences exist in the ratings of mean frequency in individuals with Williams syndrome compared to normal controls? (2) Was the severity of voice quality deviation from normal rated differently in individuals with Williams syndrome than in normal controls? (3) Was the vocal pitch of individuals with Williams syndrome rated differently than that of normal controls?

The results of the study were not definitive in proving a qualitative or quantitative difference between normal voices and voices of individuals with WS. People with WS are missing the elastin gene (ELN), a protein found in the vocal folds (among other places), and this anatomical difference is believed to be the cause of disordered vocal quality in these individuals. According to previous research in this area, voices of people with WS have been perceptually categorized as hoarse, rough, and low in pitch (Watts, Awan, & Marler, 2008). Not only have WS voices been categorized this way, but disordered voice quality has been used as a phenotypic characteristic of WS (Vaux, Wojtczak, Benirschke, & Jones, 2003). Because of these findings and the simultaneous deletion of ELN, the hypothesis of this study was that differences in perceptual measurements of voice would be found between individuals with WS and normals. If this was found, a direct relationship between ELN deletion and voice quality could be supported.

Unfortunately no significant differences, on any measure between individuals with WS and normals, were found. This resulted in retaining of the null hypothesis, which stated that the perceived vocal function of individuals with WS and normals would not be different. Visual analysis of the raw data did reveal slight differences in some areas of the data. The largest differences found in the voice quality ratings were in the categories of "normal" and "rough." Ratings of "normal" were assigned to WS voices an average of 2% more than the same rating being assigned to normal voices. Ratings of "rough" were assigned to normal voices an average of 5% more than the same rating being assigned to WS voices. The vocal severity ratings assigned to WS voices had a larger amount of deviation from the mean than the severity ratings of normal voices. This means that, while the average severity rating for WS voices was 26 (on a 100-point scale), most of the ratings assigned to this group fell between 12 and 40. Severity ratings for the normal voices mostly fell between 11 and 31 (on a 100-point scale). On closer inspection of this data, several WS voices were rated at a severity of 70 or higher (10%), and only 4% of normal voices were rated at a severity of 70 or higher.

The lack of statistical significance does not support the notion that all individuals with ELN abnormality manifest vocal fold alterations which affect vocal function. In contrast to the present study, Watts et al. (2008) did find perceptual differences in a group of individuals with ELN abnormality, although this group included both individuals with WS and SVAS. Previously Watts et al. (2007) had also found acoustic differences in the voices of individuals with SVAS only. Future research will need to investigate the nature of acoustic and perceptual differences in WS and SVAS populations directly, as the results of the present study suggest that vocal function secondary to these two disorders might be affected in a different manner. A replication of this study is also recommended to confirm the divergent findings.

There are many possible reasons why the voices were rated with no significant differences. First, each recording was only one second in length which does not allow the listener much information from which to assign a rating. It is not known if voice quality ratings are different when comparing a sustained vowel to connected speech, although it is possible. Future research may need to address this issue in both normal and disordered voices, as the clinical relevance related to that question is high. Second, the equipment used to record the normal voices was different from that used to record the voices of people with WS (and these two groups were recorded in two different physical locations). The new equipment purchased to record the normal voices should have produced recordings of equal quality to the voices of WS participants. However, the investigators judged the quality of the recordings to be diminished as compared to the quality of WS recordings. Another factor which might have influenced ratings of quality was the presence of amplitude clipping in some of the WS audio files. To eliminate the influence of these possible confounding variables, future research done in this area should utilize the exact same recording device, set at an input level consistent with the dynamic range of the recording device, for all voices.

CHAPTER VII

CONCLUSION

In conclusion, this study found no difference in the auditory-perceptual features of voice quality in Williams syndrome as compared to normal controls. The findings of this study are contradictory to previously published data on the perceptual voice quality differences in individuals with ELN abnormalities. Additional research in this area is warranted to replicate the findings of previous studies and to better control for possible confounding variables in this study. It should be noted however, that this is the first study to isolate WS and compare perceptual ratings of WS voices to normal voices. Previous research has studied the ratings of individuals with WS and SVAS as a combined group compared to the ratings of normal controls. Further research will be needed, incorporating larger sample sizes, before definitive evidence can be generated to clear the picture of the nature of voice quality in the WS population.

References

Akefeldt, A., Akefeldt, B., & Gillberg, C. (1997). Voice, speech and language characteristics of children with Prader-Willi syndrome. *Journal of Intellectual Disability Research, 41*(4), 302-311.

Awan, S.N., & Lawson, L.L. (in press). The effect of anchor modality on the reliability of vocal severity ratings. *Journal of Voice.*

Awan, S.N., & Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *Journal of Voice, 19*(2), 268-282.

Awan, S.N., & Roy, N. (2006). Toward the development of an objective index of dysphonia severity: A four-factor acoustic model. *Clinical Linguistics & Phonetics, 20*(1), 35-49.

Dykens, E.M. (2003). Anxiety, fears, and phobias in persons with Williams syndrome. *Developmental Neuropsychology, 23*(1&2), 291-316.

Eadie, T.L., & Baylor, C.R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice, 20*(4), 527-544.

Ewart, A.K., Morris, C.A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., Stock, A.D., Leppert, M., & Keating, M.T. (1993). Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Genetics, 5*, 11-16.

Gray, S.D., Titze, I.R., Alipour, F., & Hammond, T.H. (2000). Biomechanical and histologic observations of vocal fold fibrous proteins. *The Annals of Otology, Rhinology, and Laryngology, 109*(1), 77-85.

Hepburn, S., Philofsky, A., John, A., & Fidler, D.J. (2005). A case study of early development in Williams syndrome: Implications for early intervention. *Infants & Young Children, 18*(3), 234-244.

Howlin, P., Davies, M., & Udwin, O. (1998). Cognitive functioning in adults with Williams syndrome. *Journal of Child Psychology and Psychiatry, 39*(2), 183-189.

Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R. (2000). Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience, 12*(Suppl. 2), 30-46.

Lewis, B.A., Freebairn, L., Heeger, S., & Cassidy, S.B. (2002). Speech and language skills of individuals with Prader-Willi syndrome. *American Journal of Speech-Language Pathology, 11,* 285-294.

Martens, M.A., Wilson, S.J., & Reutens, D.C. (2008). Williams syndrome: A critical review of the cognitive, behavioral, and neuroanatomical phenotype. *Journal of Child Psychology and Psychiatry, 49*(6), 576-608.

Moran, M.J. (1986). Identification of Down's syndrome adults from prolonged vowel samples. *Journal of Communication Disorders, 19,* 387-394.

Pober, B.R., & Dykens, E.M. (1996). Williams syndrome: An overview of medical, cognitive, and behavioral features. *Mental Retardation, 5*(4), 929-943.

Sataloff, R.T. (1995). Genetics of the voice. *Journal of Voice, 9*(1), 16-19.

Shprintzen, R.J. (1997). *Genetics, syndromes, and communication disorders.* San Diego, CA: Singular Publishing Group, Inc.

Stemple, Glaze, & Klaben. (2000). *Clinical voice pathology: Theory and management* (3rd ed.). San Diego, CA: Singular Publishing Group.

Sugayama, S.M.M., Leone, C., Chauffaille, M. de L.L.F., Okay, T.S., & Kim, C.A. (2007). Williams syndrome: Development of a new scoring system for clinical diagnosis. *Clinics, 62*(2), 159-166.

Van Borsel, J. (2004). Voice and resonance disorder in genetic syndromes: A meta-analysis. *Folia Phoniatrica et Logopaedica, 56*(2) 83-92.

Van Lierde, K.M., Van Cauwenberge, P., Stevens, I., & Dhooge, I. (2004). Language, articulation, voice, and resonance characteristics in 4 children with Goldenhar syndrome: A pilot study. *Folia Phoniatrica et Logopaedica, 56*(3), 131-143.

Vaux, K.K., Wojtczak, H., Benirschke, K., & Jones, K.L. (2003). Vocal cord abnormalities in Williams syndrome: A further manifestation of elastin deficiency. *American Journal of Medical Genetics, 119*(A), 302-304.

Watts, C.R., Awan, S.N., & Marler, J.A. (2008). An investigation of voice quality in individuals with inherited elastin gene abnormalities. *Clinical Linguistics & Phonetics, 22*(3), 199-213.

Watts, C.R., Marler, J.A., & Urban, Z. (2007). The effects of Supravalvular aortic stenosis elastin gene mutation on voice production. *Journal of Medical Speech-Language Pathology, 15*(4), 395-405.

Wold, D.C., & Montague, J.C. (1979). Preliminary perceived voice deviations and hearing disorders of adults with Down's syndrome. *Perceptual and Motor Skills, 49*, 564.

Appendix A

CONSENT FORM – Speaking Participants
A Comparison of Auditory-Perceptual Features of Voice Quality in
Individuals With and Without Williams Syndrome
Investigator: Sara Albrecht, B.S.

Investigator's Statement

PURPOSE AND BENEFITS

You have been invited to participate in a research project which investigates vocal function in individuals with and without connective tissue disorders. Your participation, including the resulting data, represents a valuable contribution toward our understanding of how the voice is influenced by genetic makeup.

POTENTIAL RISKS

This study does not involve risks or harm any greater than those ordinarily encountered in daily life. The potential benefits from participating in this study are to provide additional theoretical and clinical information, which may enhance our understanding of how the genetic code influences voice production.

PROCEDURE

Should you decide to participate in this research study, you will be asked to sign this consent form once all your questions have been answered to your satisfaction. Before the experiment, you or your parent/guardian will be asked questions about your medical history. Prior to the experiment, you will be given specific instructions about what is required. We expect the test session to be completed within 20 to 30 minutes. You will wear a head-mounted microphone, which will fit over your ears like sunglasses, but with a microphone extending out at the corner of your mouth. You will be asked to perform a number of experimental tasks using your voice, including: saying "ahhh" for five to six seconds, reading a paragraph, and saying "eeeee" at a high pitch and a low pitch. When you are done, at a later date, approximately 20 students from TCU will listen to the recordings we make.

CONFIDENTIALITY

The results of this research will be presented at conferences. The results of this project will be coded in such a way that your identity will not be attached to the final form of this study. The researcher retains the right to use and publish non-identifiable data. While individual responses are confidential, aggregate data will be presented representing averages or generalizations about the responses as a whole. The principal investigator, co-investigators, and assistants will have access to data, but only the principal investigator will have access to your name. All data will be stored in a secure, locked location only accessible to the researcher. **Final aggregate results will be made available to participants upon request.**

PARTICIPATION & WITHDRAWAL

Your participation is entirely voluntary. You are free to choose not to participate. Should you choose to participate, you can withdraw at any time without consequences of any kind.

QUESTIONS

You may have questions or concerns during the time of your participation in this study, or after its completion. If you have any questions about the study, contact Chris Watts, Ph.D. at 817-257-6878 or c.watts@tcu.edu.

____
Name of Investigator (Printed) Date

____
Investigator’s Signature Date

Subject’s Statement

The study described above has been explained to me. I voluntarily consent to participate in this activity. I have had an opportunity to ask questions. I understand that immediate questions I may have about the research or about my rights as a participant will be answered by one of the investigators. I certify that I am at least 18 years of age.

____
Name of Participant (Printed) Date

____
Participant’s Signature Date

____
Name of Parent/Caregiver (if <18 yrs of age) Date

____
Signature of Parent/Caregiver Date

OPTIONAL SUBJECT INFORMATION: Federal guidelines require groups of subjects that participate in research experiments be representative of the general population of this region. In order to achieve this goal, it would be helpful to know the following information.

GENDER: Female_____ Male_____

ETHNIC/RACIAL ORIGIN: African-American_____ Asian/Pacific Islander_____
Caucasian_____ Hispanic_____ Native American_____ Mixed Race_____

For questions about your rights as a research subject, you may contact the chair of TCUs Institutional Review Board (IRB): Dr. Meena Shah, (817) 257-6871, m.shah@tcu.edu.
Copies to: Investigator's file, Participant

Appendix B

CONSENT FORM – Perceptual Judges (Listeners)
A Comparison of Auditory-Perceptual Features of Voice Quality in
Individuals With and Without Williams Syndrome
Investigator: Sara Albrecht, B.S.

Investigator's Statement

PURPOSE AND BENEFITS

You have been invited to participate in a research project which investigates vocal function in individuals with an inherited genetic condition. Your participation, including the resulting data, represents a valuable contribution toward our understanding of how genetic makeup influences voice production.

POTENTIAL RISKS

This study does not involve risks or harm any greater than those ordinarily encountered in daily life. The potential benefits from participating in this study are to provide additional theoretical and clinical information, which may enhance our understanding of how genes influence vocal function.

PROCEDURE

Should you decide to participate in this research study, you will be asked to sign this consent form once all your questions have been answered to your satisfaction. Before the experiment, your hearing will be screened. This will involve playing tunes to you through headphones and being asked whether or not you hear them. Prior to the experiment, you will be given specific instructions about what is required, and go through a practice session that demonstrates the experimental procedure. We expect the practice and test session to be completed within one hour. During the test session, you will be asked to listen to recorded voices and make judgments about the quality of each voice. All recordings will be played at a comfortable listening level, which will be set by you.

CONFIDENTIALITY

The results of this research will be presented at conferences. The results of this project will be coded in such a way that your identity will not be attached to the final form of this study. The researcher retains the right to use and publish non-identifiable data. While individual responses are confidential, aggregate data will be presented representing averages or generalizations about the responses as a whole. The principal investigator, co-investigators, and graduate assistants will have access to data, but only the principal investigator will have access to your name. All data will be stored in a secure, locked location only accessible to the researcher. **Final aggregate results will be made available to participants upon request.**

PARTICIPATION & WITHDRAWAL

Your participation is entirely voluntary. You are free to choose not to participate. Should you choose to participate, you can withdraw at any time without consequences of any kind. Your decision not to participate will not impact your current or future academic standing.

QUESTIONS

You may have questions or concerns during the time of your participation in this study, or after its completion. If you have any questions about the study, contact Chris Watts, Ph.D. at 817-257-6878 or c.watts@tcu.edu.

Name of Investigator (Printed) Date

Investigator’s Signature Date

Subject’s Statement

The study described above has been explained to me. I voluntarily consent to participate in this activity. I have had an opportunity to ask questions. I understand that immediate questions I may have about the research or about my rights as a participant will be answered by one of the investigators. I certify that I am at least 18 years of age.

Name of Participant (Printed) Date

Participant’s Signature Date

Investigator’s Signature Date

OPTIONAL SUBJECT INFORMATION: Federal guidelines require groups of subjects that participate in research experiments be representative of the general population of this region. In order to achieve this goal, it would be helpful to know the following information.

GENDER: Female_____ Male_____

ETHNIC/RACIAL ORIGIN: African-American_____ Asian/Pacific Islander_____
Caucasian_____ Hispanic_____ Native American_____ Mixed Race_____

For questions about your rights as a research subject, you may contact the chair of TCUs Institutional Review Board (IRB): Dr. Meena Shah, (817) 257-6871, m.shah@tcu.edu.

Copies to: Investigator's file, Participant

Appendix C

Participant Questionnaire

Name:__________________________ Date of Birth (MM/DD/YY): ____________________________

Please check the box that applies to you:

Gender: ☐ MALE ☐ FEMALE

Do you have a problem with your voice?	☐ YES	☐ NO
Have you ever smoked cigarettes on a daily basis?	☐ YES	☐ NO
Do you currently smoke cigarettes?	☐ YES	☐ NO
Have you ever been diagnosed with a voice disorder?	☐ YES	☐ NO
Have you ever been diagnosed with a neurological disorder?	☐ YES	☐ NO
Have you ever been diagnosed with a genetic disorder?	☐ YES	☐ NO

Appendix D

Voice Production Tasks

1. Take a good breath, and say the vowel /a/ at a comfortable pitch and loudness, a steady as you can until I say stop. (Have participant hold for approximately 5 seconds. Repeat 3 times.)
2. Read Grandfather Passage
3. Take a good breath, and starting at a comfortable pitch glide up on the vowel /i/ until you get to your highest pitch. (Repeat 3 times.)
4. Take a good breath, and starting at a comfortable pitch glide down on /i/ until you get to your lowest pitch (Repeat 3 times.)
5. Take a good breath, and sustain your highest pitch on the /i/ vowel, until I tell you to stop. (Repeat 3 times.)

ABSTRACT

A Comparison of Auditory-Perceptual Features of Voice Quality in
Individuals With and Without Williams Syndrome

by Sara K. Albrecht, M.S., 2009
Department of Communication Sciences and Disorders
Texas Christian University

Thesis Advisor: Christopher R. Watts, Ph.D., CCC-SLP

The goal of this research study was to obtain empirical evidence for the role of elastic fibers in the physiology of voice production, using perceptual measures of voice quality. To achieve this goal, the researcher studied individuals with Williams syndrome (WS), a genetic condition which has the hallmark characteristic of heterozygous deletions to the elastin (ELN) gene on chromosome 7. Due to alterations in normal functioning of the ELN gene, which programs elastogenesis in the connective tissues of the body, it is believed that this population can help determine how altered vocal fold histological structure, characterized by reduced and/or disorganized elastic fibers, influences vocal fold movement and the resulting sound production (voice quality). The specific aim of this project was to characterize the perceptual clinical presentation of voice quality in populations with WS via auditory-perceptual measures. Auditory-perceptual voice quality ratings were studied in 16 individuals with WS and 16 normal matched controls. Results revealed no significant differences between individuals with WS and normal controls on ratings of voice quality category, severity, or pitch. These findings diverge from those of previously published research, which will be addressed in the discussion.