VALIDATION AND APPLICATION OF EXPERIMENTAL FRAMEWORK FOR THE STUDY OF VOCAL FATIGUE

By

Mark Leslie Berardi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Communicative Sciences and Disorders—Doctor of Philosophy

2020

ABSTRACT

VALIDATION AND APPLICATION OF EXPERIMENTAL FRAMEWORK FOR THE STUDY OF VOCAL FATIGUE

By

Mark Leslie Berardi

In recent years, vocal fatigue research has been increasingly studied particularly with application to the reduction of its impact on schoolteachers and other occupational voice users. However, the concept of vocal fatigue is complex and neither well defined or well understood. Vocal fatigue seems to be highly individualized and dependent on several underlying factors or concepts. The purpose of this dissertation is to propose and support through experimentation a framework that can identify the factors contributing to vocal fatigue. The main hypothesis is that the change in vocal effort, vocal performance, and/or their interaction through a vocal demand (load) will implicate vocal fatigue. To test this hypothesis, three primary research questions and experiments were developed. For all three experiments vocal effort was rated using the Borg CR-100 scale and vocal performance was evaluated with five speech acoustic parameters (fundamental frequency mean and standard deviation, speech level mean and standard deviation, and smoothed cepstral peak prominence).

The first research question tests whether perceived vocal effort can be measured reliably and if so, how vocal performance in terms of vocal intensity changes with a vocal effort goal. Participants performed various speech tasks at cued effort levels from the Borg CR-100 scale. Speech acoustic parameters were calculated and compared across the specific vocal effort levels. Additionally, the test-retest reliability across the effort levels for speech level was measured.

Building from that experiment, the second research question was to what degree are vocal performance and vocal effort related given talker exposure to three equivalent vocal load levels. This experiment had participants performing speech tasks when presented with three different equivalent vocal load scenarios (communication distance, loudness goal, and background noise); for a given load scenario, participants rated their vocal effort associated with these tasks. Vocal effort ratings and measures of vocal performance were compared across the vocal load levels.

The last research question built on the previous two and asked to what degree do vocal performance, vocal effort, and/or their interaction change given a vocal load of excess background noise (noise load) over a prolonged speaking task (temporal load). To test this, participants described routes on maps for thirty minutes in the presence of loud (75 dBA) background noise. Vocal effort ratings and measures of vocal performance were compared throughout the vocal loading task.

The results indicate that elicited vocal effort levels from the BORG CR-100 scale are distinct in vocal performance and reliable across the participants. Additionally, a relationship between changes in vocal effort and vocal performance across the various vocal load levels was quantified. Finally, these findings support the individual nature of the complex relationship between vocal fatigue, vocal effort, and vocal performance due to vocal loads (via cluster and subgroup analysis); the theoretical framework captures this complexity and provides insights into these relationships. Future vocal fatigue research should benefit from using the framework as an underlying model of these relationships.

To Luke and Sam,
may you never lose your voice.

ACKNOWLEDGEMENTS

First, I acknowledge that this research was in part supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award Number R01DC012315. Thank you to the entire Communicative Sciences and Disorders department at Michigan State University for their incredible support throughout my PhD journey. Additionally, thank you to the many students from the Voice Biomechanics and Acoustics Laboratory for their help in testing and running the experiments.

I would like to thank my colleagues Miriam van Mersbergen from the University of Memphis and Susanna Whitling from Lund University for their valued collaborations.

I would also like to thank my committee members: Dirk Colbry, Peter Lapine, and Jeff Searl for their excellent guidance and support throughout my entire doctoral training. I would additionally like to express the deepest appreciation to my committee chair and mentor, Eric Hunter. Dr. Hunter is not only brilliant and resourceful, but also kind, compassionate, and patient. When reflecting on his impact on this work, I am reminded of the quote by Sir Issac Newton, “If I have seen further, it is by standing upon the shoulders of giants.” I was truly blessed to have the opportunity to be mentored by such an extraordinary person.

Finally, I am grateful for the support of all my family, in particular, my in-laws David and Ellen, my brother Andrew, and my dad. I am most grateful for my wife Emily who sacrificed so much for me to endure the PhD journey. She was encouraging and understanding when I needed it most and my accomplishments would be nothing without her. I am additionally grateful for her time and expertise in editing this document.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

The prevalence of occupational voice problems results in a financial burden on society. Within the United States, 25% to 35% of the working population are occupational voice users (Titze, Lemke, & Montequin, 1997). Occupational voice users are "individuals where occupational performance depends on vocal quality, health, or endurance" (Hunter & Banks, 2017). Teachers represent 4.2% of the working population, however, teachers are about 20% of the cases in voice clinics (Titze et al., 1997). A 2001 study (Verdolini & Ramig, 2001) estimates the societal cost of voice problems in teachers in the U.S. to be on the order of $2.5 billion annually. A more recent 2015 study in Colombia found the average indirect cost of a teacher with voice symptoms to be around US$492 a month (Cantor Cutiva & Burdorf, 2015).

Among the various voice problems reported in teachers, vocal fatigue is the most common. Calas and colleagues (1989) state that out of 100 teachers referred for dysphonia (difficulty in speaking), 96 reported experiencing vocal fatigue. Similarly, Gotaas and Starr (1993) report that 80% of teachers reported symptoms of vocal fatigue as compared to 5% of the general population. Hunter and Banks (2017) reported that teachers exhibited higher levels of vocal fatigue as compared to vocally-healthy adults. Other occupations that report elevated levels of vocal fatigue include but are not limited to: call-center workers (Ben-David & Icht, 2016; Lehto, Laaksonen, Vilkman, & Alku, 2008), aerobics instructors (V. I. Wolfe, Long, Youngblood, Williford, & Olson, 2002), radio broadcasters (Cantor-Cutiva, Bottalico, & Hunter, 2018; Guzmán, Malebrán, Zavala, Saldívar, & Muñoz, 2013; Timmermans, De Bodt, Wuyts, & Van de Heyning, 2003), and singers (Carroll et al., 2006; Kitch & Oates, 1994; Tepe et al., 2002; Yiu & Chan, 2003).

Despite the prevalence of vocal fatigue within occupational voice users, this condition is not well defined or well understood. Previous attempts to quantify vocal fatigue are generally inconclusive and in several cases contradictory (Laukkanen & Kankare, 2006; Welham & Maclagan, 2003). One reason for this is that vocal fatigue is a complex condition that has multiple possible underlying mechanisms. For example, in some cases vocal fatigue is used as a symptom (clinical history; Nanjundeswaran, Jacobson, Gartner-Schmidt, & Verdolini Abbott, 2015) and in other cases it describes the physiological alteration from vocal over use (i.e. fatigued tissue; Boucher & Ayad, 2010). The lack of a consensus definition may be a major contributor to the inconclusiveness in vocal fatigue results, and the lack of coherent framework prevents the interpretation of this misunderstanding. Another possible factor relating to this inconsistency and misunderstanding is the person to person variation in vocal fatigue. Several studies support the notion of vocal fatigue being characteristically different for each individual (Kitch, Oates, & Greenwood, 1996; Ternström, Bohman, & Södersten, 2006) while others showed inconsistencies within an individual (Remacle, Garnier, Gerber, David, & Petillon, 2018). While these studies are methodologically sound, they are missing a framework for vocal fatigue that allows for a complex interaction of mechanisms (e.g. physiological and perceptual) and individual variation. In the present work, a framework is developed through the connection of vocal fatigue and related concepts of vocal load, vocal loading, vocal performance, and vocal effort since these concepts are better defined. The purpose of this dissertation is to propose and support through experimentation a framework for vocal fatigue that can identify the factors contributing to vocal fatigue.

The remainder of the chapters are outlined as follows. Chapter 2 provides the relevant background information on vocal fatigue and related concepts including vocal load, vocal loading, vocal performance, and vocal effort. From this background, a theoretical framework of the relationships between these concepts to guide the study vocal fatigue is presented. The main hypotheses and research questions are presented based on the gaps in the literature and the framework. Chapter 3 provides the methodological approach to the experiments. Chapter 4 is a presentation of the results of the experiments. Chapter 5 provides a discussion of the results. Finally, Chapter 6 contains a summary of the conclusions from the discussion chapter. Additionally, the references and appendices are included at the end of this document.

CHAPTER II: BACKGROUND

2.1 Historical Background of Vocal Fatigue

In recent years, vocal fatigue research has been increasingly studied (Cantor-Cutiva, Banks, et al., 2018). This research demonstrates vocal fatigue as more than a singular mechanism but a dynamic system of multiple possible underlying and contributing mechanisms. To better delineate these uses and mechanisms, this section reviews the definitions and use of vocal fatigue, the previously proposed mechanisms, and attempts at measuring these mechanisms and their associated experiments. Additionally, this section highlights the complexity of vocal fatigue that has resulted in inconclusive attempts to quantify vocal fatigue.

2.1.a Review of definitions and use of vocal fatigue

Vocal fatigue is a common term that carries intrinsic meaning. However, within the literature there is not a common definition of vocal fatigue (Hunter & Titze, 2009). In a review of vocal fatigue, Welham and Maclagan (2003) define vocal fatigue as "negative vocal adaptation that occurs as a consequence of prolonged voice use." This negative vocal adaptation is described as "a perceptual, acoustic, or physiologic concept, indicating undesirable or unexpected changes in the functional status of the laryngeal mechanism." Focusing on the self-perceptual characteristic of fatigue, Vilkman (2004) defines vocal fatigue as "a subjective term, which refers to negative sensations related to voicing." Emphasizing the physiological nature of vocal fatigue, McCabe and Titze (2002) define vocal fatigue as "a progressive increase in phonatory effort accompanied by a progressive decrease in phonatory capabilities" and contributing factors may include "the central and/or peripheral fatigue of the respiratory subsystem, the phonatory subsystem, and the resonance/articulatory subsystems."

Clinically, vocal fatigue is defined by its symptoms (Sapienza, Crandell, & Curtis, 1999; Solomon, 2008). Various symptoms of vocal fatigue have been reported. Kostyk and Rochet (1998) summarize 18 primary symptoms of vocal fatigue:

> hoarse vocal quality, breathy vocal quality, loss of voice, pitch breaks, inability to maintain typical pitch, reduced pitch range, lack of vocal carrying power, reduced loudness range, increased vocal effort, running out of breath while talking, unsteady voice, tension in neck or shoulder, throat/neck pain, throat fatigue, throat tightness or constriction, pain on swallowing, increased need to cough or throat clear, and discomfort in chest, ears or back of neck (Kostyk & Rochet, 1998).

Other important symptoms include an increase in fatigue throughout the day and improvement following rest (Colton, Casper, & Leonard, 2011; Gotaas & Starr, 1993; Kitch & Oates, 1994; Solomon, 2008). These definitions and symptoms provide a foundation to understand the underlying mechanisms, quantifications, and assessments of vocal fatigue.

2.1.b Review of potential mechanisms of vocal fatigue

In exercise science, fatigue is defined in terms of central and peripheral; central fatigue includes “factors that reside in the brain,” while peripheral fatigue refers to the muscles themselves (Davis, 1995). As stated above, this approach has been reflected in discussions of vocal fatigue where central vocal fatigue includes the “compensatory functional changes” that could manifest as an increase in effort or feeling of increased muscular tension, and the peripheral vocal fatigue consists of the neuromuscular and biomechanical factors within the respiratory, phonatory, and articulatory subsystems (Mccabe & Titze, 2002). Additionally, the etiology of vocal fatigue is proposed to be either organic or functional. An organic factor of fatigue is when “an altered physiology affects the function of the voice, and therefore, phonation.” A functional factor of fatigue is when “an inefficient use of the voice affects phonation, and ultimately, the physiology” (Mccabe & Titze, 2002).

One of the subsystems of voice production that potentially influences vocal fatigue is the respiratory system. Previous work has shown a relation between pulmonary function and reports of vocal fatigue (Hunter, Maxfield, & Graetzer, 2019). The respiratory system can be broken down into the airway (bronchi, trachea, larynx, pharynx, mouth, and nose), the lungs, and the muscles of respiration. The major muscles in respiration are the diaphragm and the intercostal muscles. There are also accessory muscles used (e.g. sternocleidomastoid and scalene muscles). The respiratory system is one of the main subsystems of the voice; it provides the air supply and pressure needed for phonation. Therefore, in theory, fatigue in the respiratory system may result in vocal fatigue. However, this effect has not been well observed (Welham & Maclagan, 2003). The respiration physiology literature states that “in general, the capacity of the pulmonary system far exceeds the demands required for ventilation and gas exchange during exercise” and that the respiratory system is only limiting in elite athletes (McKenzie, 2012). Leanderson and Sundberg (1988) note that only 50% of vital capacity is used in initiating speech as opposed to up to 100% of vital capacity used in initiating singing. Another area of potential respiratory fatigue is in individuals with pulmonary disorders. Dysphonia was noted for patients with obstructive (e.g. asthma (Abdul Latif Hamdan et al., 2017)) and restrictive (e.g. cystic fibrosis; Lourenço, Costa, & Da Silva Filho, 2014; Willis, Michael, Boyer, & Misono, 2015) pulmonary disorders. While speech within healthy adults would not fatigue the respiratory system, task-specific events and disordered conditions may contribute to respiratory fatigue.

The phonatory system could potentially have muscle or tissue fatigue. Muscle fatigue could exist within the intrinsic or extrinsic muscles. The intrinsic muscles are a set of paired adductor muscles (closing of the glottis; e.g. cricothyroid, lateral cricoarytenoid, transverse arytenoid, oblique arytenoid, and thyroarytenoid muscles). The only abductor muscles (opening of the glottis) are the posterior cricoarytenoid muscles (Rosen & Simpson, 2008). As air flows through the larynx, adducted thyroarytenoid muscles (vocal folds or cords) vibrate, providing a harmonic acoustic source for voiced speech sounds.

Peripheral fatigue in this system can be categorized as either neuromuscular or biomechanical, in other words, laryngeal muscle fatigue or laryngeal tissue fatigue respectively (Mccabe & Titze, 2002; I. R. Titze, 1999). Neuromuscular fatigue could be defined as “any exercise-induced decrease in a muscle’s ability to develop force or power” (Boyas & Guével, 2011). Prolonged or repeated muscle contractions lead to changes in the chemical state of the muscle as the body tries to maintain the level of force being produced and resist the fatigue (Boyas & Guével, 2011). The chemical changes include depletion of energy compounds (e.g. glycogen and adenotriphosphate [ATP]) and the accumulation of lactic acid (Mccabe & Titze, 2002; Welham & Maclagan, 2003). The glycogen depletion is related to long-term submaximal muscle contractions, while lactic acid accumulation is related to short-term maximal muscle contractions (Katch, 2009; Welham & Maclagan, 2003). It is suggested that in occupational voice users, the laryngeal muscles activate more than 1,800 times per hour (Titze, Hunter, & Švec, 2007). This implies that the laryngeal muscles are contracting in a prolonged and repeated manner.

The sustainability of muscle contraction and its inherent fatigue properties can be predicted by the motor unit of the muscle (Potvin & Fuglevand, 2017). Boucher and Ayad (2010) summarize the motor unit muscle fibers from Pette and Staron (1990) saying, “Type I fibers are the slowest, most fatigue-resistant, and generate less force compared with Type IIa fibers, which are partly fatigable and fast, and Type ‘IIb’ fibers (or IIx), which are the fastest and most fatigable.” Studies using artificial stimulation of feline and canine thyroarytenoid muscles describe high levels of fatigue resistance and that human thyroarytenoid muscles contained more than twice the proportion of Type I to Type II fibers (implicated greater fatigue resistance) than in canine and four times the proportion than in feline specimens (D. S. Cooper & Rice, 1990; Edstrom, Lindquist, & Martensson, 1974; Mascarello & Veggetti, 1979; Zealer, 1983). One histochemical study of human laryngeal muscles by Claassen and Werner (1992) indicates that on average, the thyroarytenoid muscles consist of 53% of Type I, 36% of Type IIa, and 5% of Type IIb. The lateral cricoarytenoid muscles consist of 51% of Type I, 35% of Type IIa, and 14% of Type IIb. This fiber distribution suggests an overall slow, fatigue-resistant property of the intrinsic laryngeal muscles (Solomon, 2008; Welham & Maclagan, 2003). However, the adductor muscles in humans are actually quite fast (Hast, 1969; Sahgal & Hast, 1974). Boucher and Ayad (2010) explain this discrepancy to be the result of previously misidentified “hybrid” fibers in facial and laryngeal muscles.

Hoh (2005) states that previous histochemical analyses are invalid because they ignore the presence of these hybrid fibers and that single-fiber protocols must be used. Wu, Crumley, Armstrong, and Caiozzo (2000) applied a single-fiber protocol of the thyroarytenoid muscle and report 30% Type I fibers, 49% Type II fibers, and 21% Type IIx fibers. Shiotani, Westra, and Flint (1999) report that the thyroarytenoid muscle has an average of 13.5% Type I fibers, 49.2% Type IIa fibers, and 37.3% Type IIx fibers, while the lateral cricoarytenoid muscle contained 18.8% Type I fibers, but 57.1% of Type IIa and 24.1% of Type IIx fibers. Hoh (2005) explains that this variance is likely a result of plasticity effects. These studies showing a higher proportion of Type II fibers support the reports of the high contraction speeds of the laryngeal adductor muscles (D’antona et al., 2002; Hast, 1969; Sahgal & Hast, 1974). The implications are that the intrinsic laryngeal muscles may be affected by neuromuscular fatigue and that there may be a significant amount of individual variation in the fatigue of laryngeal muscles.

In addition to the intrinsic laryngeal muscles, Solomon (2008) suggests that neuromuscular adjustments in the extrinsic laryngeal muscles (e.g. suprahyoid and infrahyoid [or strap muscles] groups) can contribute to the overall stiffness of the larynx and tension in the laryngeal muscles. This excessive tension can result in co-contraction of agonistic muscles and other inefficient muscle contractions that contribute to vocal fatigue (Solomon, 2008).

Prior to a discussion on biomechanical fatigue, consider the composition of the vocal folds. A vocal fold consists of five main layers. The superficial layer (epithelium) consists of squamous epithelium (which is continuous throughout the trachea, pharynx, and mouth). The next three layers have differing consistencies but are often referred to together as the lamina propria. The superficial layer of the lamina propria can be compared to the consistency of soft gelatin and is also referred to as Reinke’s space. This layer is associated with Reinke’s edema (swelling of this layer due to fluid; often associated with smokers). The next two layers are sometimes referred to together as the vocal ligament. The intermediate lamina propria is characterized by elastic fibers (a rubber band-like consistency). The deep lamina propria primarily contains collagenous fibers (thread-like consistency). The final layer is the vocalis muscle of the thyroarytenoid which is the main body of the structure and is very stiff (Rosen & Simpson, 2008).

In the case of biomechanical fatigue, parallels are often drawn to the use of fatigue in the material sciences. Here fatigue is “progressive structural damage that results from stress (force per unit area) imposed by strain [a measure of elongation or deformation] on the material” (Solomon, 2008). During regular phonation, the lamina propria is subject to continuous stress and strain. Stress from the numerous collisions of the two folds and strain from anterior-posterior posturing (elongation or shortening) as well as other deformations from the various laryngeal muscles. Titze (1994) discusses mechanical stresses on the vocal folds and concludes that "excessive collision and acceleration may be responsible for the greatest tissue damage, even though they do not account for the greatest stresses." Female teachers' vocal folds are estimated to vibrate on the order of 1.4 million times during their work day (I. R. Titze et al., 2007; E Vilkman, 2004). The magnitude of the number of collisions for a population with high prevalence of vocal fatigue is an intriguing correlation. However, two reviews of vocal fatigue concluded that while the mechanical stresses exist, their contributions to lamina propria fatigue (and vocal fatigue in general) are not well understood (Solomon, 2008; Welham & Maclagan, 2003).

Other biomechanical factors of vocal fatigue include increased viscosity, stiffness, and lesions (I. R. Titze, 1994). The vocal folds' viscous properties allow for lubrication and shock absorption (Solomon, 2008). McCabe and Titze (2002) suggest that a change in tissue viscosity could be a result of vibrations of the vocal folds that "may lead to an inability of a stable blood circulation throughout the lamina propria to remove inhibitory elements such as lactic acid, or an accumulation of heat." In the case of prolonged, high-pitched phonation, increased tissue viscosity was measured, as well as increased frictional energy loss and an increase in heat dissipation (Donald S. Cooper & Titze, 1985). Solomon (2008) concludes that an interaction between viscosity changes and exposure to stress and strain on the tissues is "likely to exacerbate tissue" and that the "importance of non-muscular biomechanical properties on vocal function makes the study of vocal fatigue more complex than the study of fatigue involving most other skeletal muscles."

The third subsystem relating to vocal fatigue is the articulatory and resonance system. There are three main resonating cavities in this system: the pharynx (comprising the laryngopharynx, oropharynx, and nasopharynx cavities), oral cavity, and nasal cavity. The major articulators used in speech are the velum (or soft palate), hard palate, teeth, tongue, lips, and cheeks. The muscles used in the articulator and resonance systems primarily consist of fatigue-resistant Type I and IIa muscle fibers (this is the case for the tongue, orbicularis oris [lips], and buccinator [cheek] muscles (Solomon, 2006)). Like the respiratory system, normal speech in healthy individuals is not fatiguing. However, individuals with varied disorders may experience fatigue in the articulatory and resonance muscles (e.g. lingual fatigue in myasthenia gravis (Wenke, Goozee, Murdoch, & LaPointe, 2006) or lingual and lip fatigue in laryngectomees (Searl & Knollhoff, 2018).

2.1.c Review of attempts to measure vocal fatigue

Many methods are used to assess and quantify vocal fatigue. In general, these methods attempt to measure either the physiological mechanisms of vocal fatigue or the perception of vocal fatigue experienced by the individual. These measurement methods are categorized as physiological, aerodynamic, acoustic, or self-report assessments. Physiological methods include assessments where the muscles (particularly the intrinsic laryngeal muscles) are measured directly. Aerodynamic methods are assessments that measure flow or pressure of the aerodynamic source of phonation. Acoustic methods include both objective and perceptual measures where objective measures are calculated from acoustic samples and perceptual measures are judgements of the acoustic samples by listeners. Finally, self-report assessments are subjective ratings of the subjects’ own experience of vocal fatigue. Most vocal fatigue studies include multiple measurement methods.

Physiological assessment of vocal fatigue attempts to directly measure the mechanism of the larynx. This is done through imaging, electroglottography (EGG), electromyography (EMG), or through physical and computational models. The types of laryngeal imaging used to assess vocal fatigue include rigid endoscopy (Eustace, Stemple, & Lee, 1996; A L Hamdan, Sibai, & Rameh, 2006; Kelchner, Toner, & Lee, 2006; Linville, 1995; Niebudek-Bogusz, Kotyło, & Śliwińska-Kowalska, 2007; Pearl Solomon & Stemmle DiMattia, 2000), flexible nasal endoscopy (D’haeseleer et al., 2016), and high speed endoscopy (Doellinger, Lohscheller, McWhorter, & Kunduk, 2009; Whitling, Lyberg-Åhlander, & Rydell, 2017). Various measures are observed and calculated from laryngeal imaging. These include vocal fold edema (Scherer et al., 1987), glottal closure patterns (subjective evaluation of the open and closure of the vocal folds (Eustace et al., 1996; Kelchner et al., 2006; Linville, 1995; Niebudek-Bogusz et al., 2007; Stemple, Stanley, & Lee, 1995)), laryngeal length-to-width ratio of the glottis (Yiu et al., 2013), amplitude of vocal fold vibration (Eustace et al., 1996; Niebudek-Bogusz et al., 2007; Solomon, Glaze, Arnold, & van Mersbergen, 2003; Vintturi et al., 2001), phase symmetry of vibration (Eustace et al., 1996), and quality of mucosal wave (Eustace et al., 1996; Niebudek-Bogusz et al., 2007). Other studies use laryngeal imaging as a screening to confirm the presence or absence of laryngeal pathology (D’haeseleer et al., 2016). Some of the studies report no laryngeal imaging changes associated with vocal fatigue (Eustace et al., 1996; Kelchner et al., 2006; Niebudek-Bogusz et al., 2007; Whitling et al., 2017), while others reported contradicting changes (Solomon, Glaze, Arnold, and van Mersbergen, 2003) report a decrease in amplitude of vocal fold vibration associated with vocal fatigue while Vinturri et al.(2001) report an increase). Observations of incomplete glottal closure, such as an anterior glottal chink or a spindle-shaped glottis (Linville, 1995; Pearl Solomon & Stemmle DiMattia, 2000; Solomon et al., 2003; Stemple et al., 1995), suggest effects of neuromuscular fatigue on the adductor muscles. Observations of vocal fold edema and changes in quality of the mucosal wave support the theories of tissue changes in vocal fatigue. The major drawback to laryngeal imaging is that it is a relatively invasive measurement requiring particular equipment. Additionally, the quantitative measurements of imaging are computationally expensive and partly subjective (Doellinger et al., 2009), although recent research aims to address those concerns (Naghibolhosseini, Deliyski, Zacharias, de Alarcon, & Orlikoff, 2018; Poburka, Patel, & Bless, 2017; Zacharias, Deliyski, & Gerlach, 2018).

Electroglottography (EGG) measures the level of contact between the vocal folds during voicing. It does this with a pair of electrodes placed on the surface of the neck such that the larynx is between them. The electrodes measure the electrical impedance of the larynx which changes with the contact area of the vocal folds. This can be useful in estimating the time the vocal folds spend open and closed (Childers, Hicks, Moore, Eskenazi, & Lalwani, 1990). Due to obvious variations in biology and therefore neck impedance, EGG cannot be used to absolutely measure glottal contact area. The most commonly used EGG measure is the closed quotient (the ratio of closed time and combined closed and open time). Electroglottography presents a less invasive physiological measurement of vocal function. However, the usefulness of EGG in studies of vocal fatigue is inconclusive (Buekers, 1998; Laukkanen, Mäki, & Leppänen, 2009; V. I. Wolfe et al., 2002).

Electromyography (EMG) is a more direct measurement of the electrical activity in muscles and has been used to show changes in the laryngeal muscles (the lateral cricoarytenoid in particular) as a result of vocal fatigue (Boucher, Ahmarani, & Ayad, 2006; Boucher & Ayad, 2010; Rubin et al., 2005). These results strongly support the proposition that the laryngeal muscles are not fatigue resistant. Although EMG is useful in investigating neuromuscular fatigue, it is highly invasive and not practical in most vocal fatigue research. Some recent work has used surface EMG on the neck to estimate laryngeal muscle activations with the intent to apply the methods to vocal fatigue research (N. R. Smith et al., 2016).

As stated above, the stress and strain of vibration on vocal fold tissues may lead to tissue fatigue and, thus, vocal fatigue. Directly measuring these stresses and strains in vivo is problematic, therefore modeling becomes a useful tool.

Researchers use physical or synthetic (Drechsel & Thomson, 2008; Spencer, Siegmund, & Mongeau, 2008; I. R. Titze, 1994), computational (Horáček, Laukkanen, Šidlof, Murphy, & Švec, 2009; Tao & Jiang, 2007), and ex vivo human and animal (Doellinger & Berry, 2006; Matsushita, 1975; Katherine Verdolini, Chan, Titze, Hess, & Bierhals, 1998) models to study the vibratory properties of the vocal folds. Models are continually being used to study biomechanical changes to vocal folds.

Aerodynamic measurements serve two primary assessments of the vocal fatigue. First, to investigate biomechanical changes in the vocal folds and how that might impede with the aerodynamic output. Second, to assess the pulmonary capacity of the research subjects. The most common and consistent aerodynamic measure used in vocal fatigue research is phonation threshold pressure (PTP). This is the minimal level of lung pressure below the vocal folds required for sustained phonation. True PTP is measured with a tracheal puncture, but it can be well estimated noninvasively by measuring the oral pressure during a bilabial stop consonant (Fisher & Swank, 1997). Many (but not all) have found significant increases of PTP with vocal fatigue (Chang & Karnell, 2004; Enflo, Sundberg, & McAllister, 2013; Erickson-Levendoski & Sivasankar, 2011; Kagan & Heaton, 2017; Pearl Solomon & Stemmle DiMattia, 2000; Erkki Vilkman, Lauri, Alku, Sala, & Sihvo, 1999; Whitling et al., 2017). This finding suggests biomechanical changes in the vocal folds resulting in increased viscosity and therefore more pressure is required for them to begin sustained oscillation. This is further supported by the relationship between PTP and hydration (Pearl Solomon & Stemmle DiMattia, 2000; Mahalakshmi Sivasankar, Erickson, Schneider, & Hawes, 2008). Spirometry is used to measure pulmonary capacity in studies of vocal fatigue (Koufman & Blalock, 1988). Maxfield, Hunter, and Graetzer (2016) used spirometry to show a relationship between vocal fatigue and pulmonary function.

Acoustic assessment is the most common type of assessment of vocal fatigue. This is likely because it is a noninvasive measurement using relatively common equipment (i.e. microphone and signal recorder). The acoustic properties of speech are directly related to physiological mechanisms that produce speech (air supply, vocal fold vibration, and pharyngeal, oral, and nasal articulations). The most obvious relationship is that the measured speaking fundamental frequency (F0; usually reported as cycles per second or Hertz but sometimes as semitones which are $1/12^{th}$ of a doubling of frequency) relates to the number of vocal fold vibratory cycles per second. If there are neuromuscular or biomechanical changes, then the acoustic signal should change accordingly.

Many different objective measures are calculated from the acoustic signal. The most common measures used in the assessment of vocal fatigue relate to F0’s minimum (Cho, Yin, Park, & Park, 2011; Stemple et al., 1995), mean (Cho et al., 2011; D’haeseleer et al., 2016; Jonsdottir, Laukkenen, & Siiki, 2003; Laukkanen, Ilomäki, Leppänen, & Vilkman, 2008; Laukkanen & Kankare, 2006; Laukkenen et al., 2004; Lehto, Laaksonen, Vilkman, & Alku, 2006; V. I. Wolfe et al., 2002), maximum (Cho et al., 2011; D’haeseleer et al., 2016; M. Sivasankar, 2002), and variance (Ben-David & Icht, 2016; Cho et al., 2011; V. I. Wolfe et al., 2002). The voicing intensity or speech level (reported as a sound pressure level in decibels) is also commonly measured in terms of its minimum (D'haeseleer et al., 2016; Erkki Vilkman et al., 1999), mean (Ben-David & Icht, 2016; Jonsdottir et al., 2003; Laukkanen et al., 2008; Laukkanen & Kankare, 2006; Laukkenen et al., 2004; Lehto et al., 2006; V. I. Wolfe et al., 2002), and variance (Ben-David & Icht, 2016; Bottalico, Cantor Cutiva, & Hunter, 2017; Cho et al., 2011). Sometimes the range of F0 and the range of voicing intensity are combined to calculate a voice range profile (Damsté, 1970; I. R. Titze, 1992) which is used to investigate changes with vocal fatigue (E. Holmberg, Ihre, & Södersten, 2007; Wingate, Brown, Shrivastav, Davenport, & Sapienza, 2007). Measures of perturbation, which include jitter (cycle-to-cycle frequency instability), shimmer (cycle-to-cycle amplitude instability), and harmonic-to-noise ratio (HNR; ratio of the harmonic energy to noise energy in the acoustic signal), are also commonly used in vocal fatigue assessment (Cho et al., 2011; D'haeseleer et al., 2016; Gelfer, Andrews, & Schmidt, 1991; Laukkanen et al., 2008; Laukkanen & Kankare, 2006; Scherer et al., 1987; Verstraete, Forrez, Mertens, & Debruyne, 1993; V. I. Wolfe et al., 2002). Less common but important acoustic measures include relative fundamental frequency (RFF; measure of the stability of the offset and onset of vocal fold vibration in speech) and cepstral peak prominence (CPP; a robust spectral-cepstral measure of vibratory periodicity). Vocal fatigue studies have only recently started to use these two measures, but they present more promising results than the previously mentioned measures (Fujiki, Chapleau, Sundarrajan, McKenna, & Sivasankar, 2017; Gorham-Rowan, Berndt, Carter, & Morris, 2016; Kagan & Heaton, 2017). A recent study measures changes in the formants (acoustic resonances of the vocal tract) associated with vocal fatigue (Pellicani, Fontes, Santos, Pellicani, & Aguiar-Ricz, 2018). Several studies utilized the commercial multidimensional voice program (Deliyski, 1993) to calculate F0, perturbation measures, and a dozen other acoustic parameters (Boucher, 2008; Boucher & Ayad, 2010; Buekers, 1998; D'haeseleer et al., 2016; Pellicani, Ricz, & Ricz, 2015). Although speech acoustic measures are commonly used, their results vary. In almost every common acoustic measure mentioned above, studies report increases, decreases, or no change associated with vocal fatigue. Kitch, Oates, and Greenwood (1996) report the acoustic parameters changing in opposite directions with different participants within the same study. Another vocal fatigue study reports considerable inter- and intra-subject variability with these acoustic measures (Remacle et al., 2018).

Another way the acoustic signal is used is in perceptual ratings of voice quality. In these cases, listeners judge the quality of voice based on predetermined scales. The two most common instruments are the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale (Cho et al., 2011; D'haeseleer et al., 2016; Faham et al., 2017) and the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) scale (Kagan & Heaton, 2017; Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009). In studies of vocal fatigue, these instruments are used as either measurement variables or screenings for vocal health. Another instrument, inability to produce soft voice (IPSV) is a self-evaluation from a subject after they try to produce a high-pitch, low-amplitude vocal sound (Bastian, Keidar, & Verdolini-Marston, 1990; Halpern, Spielman, Hunter, & Titze, 2009; Hunter & Titze, 2009; Kagan & Heaton, 2017). This is different from other self-ratings (described below) as it is based on the performance of a task that would be more difficult for someone experiencing vocal fatigue, rather than a rating of internal feelings of fatigue or effort.

The previous measurements aim to assess peripheral vocal fatigue. Central vocal fatigue is assessed through self-reporting of the subjects, typically in the form of questionnaires. A variety of self-reporting instruments have been used to study vocal fatigue. One of the most common instruments is self-ratings of vocal or phonatory effort during the vocal fatigue experiment (Carroll et al., 2006; Chang & Karnell, 2004; A L Hamdan et al., 2006; Mccabe & Titze, 2002). Other commonly asked self-rating questions include laryngeal pain, laryngeal fatigue, voice quality, or prior vocal complaints (Buekers, 1998; Fabron et al., 2015; Kelchner et al., 2006; Laukkanen et al., 2008). Most of these instruments ask subjects to provide a rating from 1 to 10 (or something similar). These instruments are continually developed using mechanisms from the survey literature including visual analog scales (VAS; Lehto et al., 2008; Pellicani et al., 2015), Likert scales (Whitling et al., 2017), and the Borg CR-10 scale (Ford Baldner, Doll, & van Mersbergen, 2015; van Leer & van Mersbergen, 2017). Within the general field of voice assessment, indexes have been developed and a common one is the Voice Handicap Index (VHI or VHI-1; Jacobson et al., 1997). This has been applied to vocal fatigue research (Cho et al., 2011; Niebudek-Bogusz et al., 2007; Wingate et al., 2007) but also led to the creation of an index specifically for vocal fatigue. Nanjundeswaran and colleagues developed the Vocal Fatigue Index (VFI) to be used in self-reporting of trait vocal fatigue (Nanjundeswaran et al., 2015; Nanjundeswaran, van Mersbergen, & Morgan, 2017). Although it is a recently developed instrument, it has already been used to show differences in vocal fatigue in occupational voice users (dos Santos, Silverio, Dassie-Leite, Costa, & Siqueira, 2018; Hunter & Banks, 2017). Self-ratings are important because they aim to measure the perceived experience of vocal fatigue. This contributes to understanding a different facet of fatigue not detected using other assessments. However, the measurement of peripheral vocal fatigue may only be quantifying the trait or systemic factor of fatigue. As a result, these instruments cannot be used to measure changes in vocal fatigue throughout a communication event thereby limiting its application.

2.1.d Review of experiments in vocal fatigue

Quantification of vocal fatigue requires a subject who is vocally fatigued. There are different methods of experimentation that attempt to induce vocal fatigue. The most common type of vocal fatigue experiment is a vocal loading task (VLT; the term vocal loading is described in more detail in the following section). Here the subjects undergo some variation of a prolonged phonation task and it is assumed that due to the VLT the subject will experience some amount of vocal fatigue. Vocal loading tasks have taken many different forms. In a review of VLTs by Fujiki and Sivasankar (2017), it was found that the most common duration of VLT was two hours with some being as short as 15 minutes and as long as 3.75 hours. Some of the VLT involved multiple experimental sessions which ranged from two to seven sessions. A few of the studies allowed for the subjects to terminate the experiment when they felt fatigued. The most common type of loading task was prolonged, loud reading. When a loud voice was elicited, several used ambient noise in the VLT (De Bodt, Wuyts, Van De Heyning, Lambrechts, & Abeele, 1998; Erickson-Levendoski & Sivasankar, 2011; Whitling et al., 2017), while others required a vocal loudness target (Gelfer et al., 1991; Linville, 1995; Neils & Yairi, 1987). Others have investigated the relationship between acoustic environment (in particular noise and reverberation) and vocal fatigue (Bottalico, Cantor Cutiva, et al., 2017; Bottalico, Graetzer, & Hunter, 2016; Kristiansen et al., 2014). One limitation to many of these studies is that measurements are only taken at the beginning and the end of the VLT. A few studies have also taken measurements at periodic intervals during a VLT (Boucher, 2008; Buekers, 1998; Laukkenen et al., 2004; Xue, Kang, Hedberg, Zhang, & Jiang, 2019) which may provide more information of how vocal fatigue develops through~~out a VLT~~ way to investigate. The other ~~major~~ limitation is that many of the studies assume that vocal fatiguing is occurring as a result of the vocal loading—which may not be the case.

Vocal loading tasks are likely the most common way to investigate vocal fatigue because the researchers can control many of the factors in the experiment. However, others have taken a more ecologically valid approach to vocal loading by assessing vocal fatigue throughout a workday in real world environments. This has been done for teachers (Halpern et al., 2009; Jonsdottir et al., 2003; Laukkanen & Kankare, 2006; Rantala, Paavola, Körkkö, & Vilkman, 1998; Remacle et al., 2018), call-center workers (Ben-David & Icht, 2016; Lehto et al., 2008), and radio broadcasters (Cantor-Cutiva, Bottalico, et al., 2018). Other methods use surveys in populations where vocal fatigue is a common complaint (e.g. teachers and individuals with voice disorders) to study the prevalence of vocal fatigue and possible associated factors (Bastian & Thomas, 2016; dos Santos et al., 2018; Hunter & Banks, 2017; Munier & Kinsella, 2008). Another methodology divides the subject pool into groups based on self-reported vocal symptoms (D'haeseleer et al., 2016; Faham et al., 2017; Ilomäki, Kankare, Tyrmi, Kleemola, & Geneid, 2017; Laukkanen et al., 2008) or measured level of overall fatigue (Cho et al., 2011) and studies the differences between the defined groups. Throughout all of these studies, and in particular VLT, the only consistent measure associated with vocal fatigue has been perceived vocal effort.

2.2 Historical Background of Vocal Effort

Perceived effort or exertion given a task or from the result of a task has been the focus of research in a broad range of fields such as exercise science, cognitive science and psychology, and audiology (listening effort). Within the context of vocal fatigue, perceived vocal effort was the only consistent measurement of vocal fatigue in the vocal loading tasks. To better understand vocal effort and its relationship to vocal fatigue, vocal load, and vocal loading, this section reviews the definitions, uses and measurements of vocal effort.

2.2.b Review of measurements of vocal effort

The previous section concludes that vocal effort is more nuanced than vocal intensity and therefore requires other measurements. There are two main types of measurements of vocal effort: perceptual and physiological. The perceptual measurements include either the perception of vocal effort from the talker or from a listener. Physiological estimates of vocal effort generally include aerodynamic and acoustic. Understanding the types of measurements that are effective in quantifying vocal effort will contribute to the understanding of the phenomenon.

The earliest type of measurements of vocal effort are from listeners. Many of these studies used scales developed for that particular study and were not used in any standard way. One example of such a scale is a nine-point vocal effort rating of prerecorded harsh voices (Thomas-kersting & Casteel, 1989). Eventually two main listener rating scales were developed, validated, and standardized for voice research. One is the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale (Hirano, 1981). Here strain is defined as the "perception of excessive vocal effort." Another perceptual scale is the Consensus Auditory Perceptual Evaluation–Voice (CAPE-V) scale (Kempster et al., 2009) which includes the same definition and application of the word strain. Beyond these two scales and others that were singularly developed, visual-analog scales (VAS) have been used (Eadie & Stepp, 2013). This perspective of vocal effort is important clinically as clinicans are the listener-judges for the evaluation of vocal effort.

For the quantification of self perception of effort, VAS have traditionally been the most common (Paes, 2017; Shewmaker, Hapner, Gilman, Klein, & Johns, 2010; Södersten, Granqvist, Hammarberg, & Szabo, 2002; Tanner et al., 2010; Warrick et al., 2000). There has been recent research that report vocal effort scales to match the development of standardized physical exertion scales, namely the Borg CR-10 (van Leer & van Mersbergen, 2017). The Borg CR-10 uses a logarithmic distribution of anchored statements about vocal effort (such as “slight vocal effort” or “severe vocal effort”) that correlates more accurately to how physical exertion is perceived. This same design has been implemented into a single-question pictorial scale on perceived vocal effort/exertion called the OMNI Vocal Effort Scale (OMNI-VES; Shoffel-havakuk et al., 2019). One problem with the Borg CR-10 and OMNI-VES is that the majority of the scale is in the difficult exertion portion of the scale and the healthy voices tend to not be in that section of the scale resulting in a resolution problem. Visual analog scales do not have this problem because they are continuous. However, using VAS loses the benefits of using the anchors in the Borg CR-10 or OMNI-VES. One possible solution is to convert the Borg CR-100 to a vocal effort scale. This scale would have the benefits of the anchors of the Borg CR-10 with the resolution closer to the VAS. Additionally, the VAS is a linear scale while the Borg CR-10 has logartihmic spacing of the anchors that correlate best with human perception of exertion (Borg & Löllgen, 2001). Here the vocal-effort adapted Borg CR-100 scale is used to quantify percieved vocal effort.

Self-rating surveys have been developed, standardized and widely used (similar to GRBAS or CAPE-V) that contain components related to vocal effort. The Voice-Handicap Index (VHI; Jacobson et al., 1997) has been used to measure self ratings of vocal effort (Sampaio & Jos, 2012). The VHI asks “I use a great deal of effort to speak” which relates to the measurement of vocal effort. A similar scale to the VHI was developed for the use of vocal fatigue research, the Vocal Fatigue Index (VFI; Nanjundeswaran et al., 2015). This survey asks “I experience increased sense of effort with talking,” “it is effortful to produce my voice after a period of voice use," and "the effort to produce my voice decreases with rest" which all relate how vocal effort and vocal fatigue interact. These scales are very useful in collecting state levels of vocal effort. They reflect the person's overall state of vocal effort and cannot be used to look at changes in vocal effort across an experiment.

In many cases, vocal effort is regarded as a perceptual condition, however attempts have been made to quantify vocal effort. One of the most common physiological measurements of vocal effort is phonation threshold pressure (PTP; Pearl Solomon & Stemmle DiMattia, 2000; Rosenthal et al., 2014; Solomon, Glaze, Arnold, & van Mersbergen, 2003). Phonation threshold pressure is the minimum level of lung pressure required to produce sustained vocal fold oscillations. Typically PTP cannot be measured directly (this would require a tracheal puncture) but can be well estimated by measuring the intraoral pressure of a bilabial plosive. It is unclear as to whether high PTP results in higher vocal effort or whether excessive vocal effort results in higher PTP. On one hand a lower PTP for a given task could represent improved vocal ability which would likely mean lower vocal effort on certain tasks. It could also be the case that elevated vocal effort could cause extra tension or miscoordination of the vocal mechanism which would result in increases in PTP. Either way, PTP seems to be a measure strongly linked to vocal effort. The major drawback is that the measurement is not as easy as acoustic measurements and its variability is affected by many other factors such as hydration (Verdolini et al., 1994) and fatigue (Chang & Karnell, 2004).

Many different acoustic measurements have been used to quantify change in vocal effort. Some have already been mentioned (fundamental frequency, formant amplitude and frequency, speech level). Some of the acoustic measures are time based while others are spectral based. A promising time-based measure is relative fundamental frequency (RFF). Relative fundamental frequency estimates differences in vocal fold vibration during voiced onsets. During regular speech, voicing starts (abducts) and stops (adducts) for different speech sounds. The offsets and onsets from the abduction and adduction result in changes to the rate of vibration of the vocal folds until the vibration is steady. The rate of vocal fold vibration is the fundamental frequency. Relative fundamental frequency compares the fundamental frequencies of the individual glottal pulses after a voiceless consonant to the fundamental frequency of the steady-state vowel. Lien and Stepp (2015) used RFF to track changes in healthy individuals as they changed their vocal effort. The authors also showed strong relationships between RFF and perceptual and aerodynamic measures of vocal effort. Relative fundamental frequency has been used to show differences in populations with characteristically high vocal effort: vocal hyperfunction (Stepp, Sawin, & Eadie, 2012) and ADSD (Eadie & Stepp, 2013). Despite the promising nature of RFF, in practice it is subjective to estimate the glottal pulses relating to vocal offset or onset (e.g. it is not always clear as to when the vowel begins or what exactly is the steady state). The analysis can also be computationally expensive. Despite these shortcomings, it offers a physiologically-minded acoustic measure of vocal effort that has been well validated.

Other acoustic measures which relate to vocal effort include cepstral-peak prominence (CPP) and mel-frequency cepstral coefficients (MFCC). Both of these measures use the cepstrum which is the Fourier transform of the power spectrum of speech. Cepstral peak prominence is a measure of the decibel difference between the magnitude of the strongest cepstral peak and a regression of the average cepstral energy. The CPP often relates a measurement of the periodicity of the voice and has been widely used in voice disorder research (Leong et al., 2013). Rosenthal, Lowell, and Colton (2014) showed higher values of CPP in maximal effort speech and significantly lower CPP during minimal effort speech. Although this seems promising, it was noted that CPP often shares a linear relationship with vocal intensity and in this case, the maximal effort and minimal effort speech had a significant difference in vocal intensity. It is hard to discern whether the difference in CPP is a result of the change in intensity or the change in effort. Mel-frequency cepstral coefficients are bands of energy with the cepstrum. They are commonly used in speech signal processing, particularly in the area of speech recognition. Zelinka, Sigmund, and Schimmel (2012) used MFCC to train a computer to automatically classify categorical vocal effort levels from a database. The main drawback of MFCC is the lack of interpretation. Although the computer could correctly identify levels of vocal effort, the process it used to do that is not interpretable.

2.3 Theoretical Framework

This section develops a theoretical framework to address vocal fatigue from the related concept of vocal effort. This is because the perceived effort, compared to fatigue, is more universally studied and defined in a wide range of fields. This framework is partially modeled after the listening effort framework (Pichora-Fuller et al., 2016). The proposed theoretical framework of vocal effort includes vocal fatigue and other related concepts from the voice literature such as vocal load, vocal loading, vocal ability, and vocal performance. Here they are provided with specific definitions to establish the framework. Several of these concepts are discussed in recent reviews (Cantor-Cutiva et al., 2018; Hunter et al., 2020).

2.2.a Review of definitions and use of vocal effort

One of the early definitions of vocal effort focuses on communicative distance as the catalyst for vocal effort. Traunmüller and Eriksson (2000) state that vocal effort is the "quantity that ordinary speakers vary when they adapt their speech to the demands of an increased or decreased communication distance." Here vocal effort is quantified by having listeners rate the perceived talker-listener distance from a talker's speech sample. An instinctual criticism of this approach is whether the listeners are actually rating the vocal effort and not just the loudness of the voice. If while using a loudspeaker the desire is to increase the acoustic radiation (in other words the communication distance), one needs to only increase the amplitude of the loudspeaker's output. In the case of the voice, if the only goal is the increase of speech acoustic radiation, then, in theory, one only needs to be louder. However, this study accounted for that. The authors investigated a wide range (0.3-187.5 m) of communication distance but also randomly modulated the amplitude of the speech samples. The listeners were still reliable in rating differences in the perceived talker-listener distance, which the authors use to quantify vocal effort. They conclude that vocal effort is a physiological exertion different from vocal intensity that accounts for changes in communication distance. This definition of vocal effort is widely used (being cited over 250 times) but it depends on subjective observations of listeners and vocal effort is arguably a sensation experienced and measurable only from the talker (Hunter et al., 2020).

The notion of vocal effort being related to communication distance extends to other studies. However, instead of listeners rating a perceived talker-listener distance, the actual talker-listener distance was used as vocal effort. In these cases, the greater the distance, the greater the vocal effort. Liénard and Di Benedetto (1999) used three different distances (close—0.4 m, normal—1.5 m, and far—6 m) and Pelegrín-García, Smits, Brunskog, and Jeong (2011) studied vocal effort at four distances (1.5, 3, 6, and 12 m). These studies found changes in fundamental frequency, formant frequencies and amplitudes, and sound pressure level of the speech to be related to this distance based vocal effort, or rather increases in communication distance. Again, the pattern of speech characteristics, other than the loudness, changed, with increasing talker-listener distance which is consistent with Traunmüller and Eriksson. These studies illustrate one possible component of vocal effort—vocal adaptations to communication distance. It is straightforward to illustrate how vocal effort would increase with communicative distance. Recall any experience of trying to communicate with someone across a large room—this requires greater vocal effort than talking to someone within close proximity. However, an observable example of trying to converse with a close neighboor in a noisey environment refutes that this is the only component of vocal effort.

In addition to communication distance, vocal effort may be needed to accommodate within an environment with excessive noise. It has been well understood that speech changes in response to background noise, this is often called the Lombard effect (Lombard, 1911). Simply put the Lombard effect is a reflex to increase one's vocal level in response to exposure of increased noise levels. Junqua (1993) explains the Lombard effect within the context of vocal effort as the "reflex that takes place when a speaker modifies his vocal effort while speaking in the presence of noise." It has been previously shown that vocal effort is more than an increase in vocal level. If the Lombard effect is truly a vocal effort modulation, then other speech characteristics should change. Similar changes in speech patterns (such as fundamental frequency and formant frequencies) that were found in increased communicative distances relate to vocal effort in response to background noise (Lu & Cooke, 2009; Vogel, Fletcher, Snyder, Fredrickson, & Maruff, 2011). Another analogy can be drawn between Traunmüller and Eriksson (2000) and Södersten, Granqvist, Hammarberg, and Szabo (2002) who showed listener judgements of vocal effort in increasing levels of background noise. Södersten, Granqvist, Hammarberg, and Szabo also measured the talker's perception of vocal effort and found that this also increased with the noise. These studies show similarities between vocal effort in the context of communicative distance and in the context of excessive background noise. Both of these conditions are environmental factors that affect the successful transmission of a spoken message. It can then be assumed that other environmental factors may similarly affect vocal effort.

Another communication environment factor that could relate to increases in vocal effort is room reverberation. Reverberation presents a slightly different problem than static background noise. Here the speech of the individual contributes to the noise in the form of reverberation. As the talker increases their vocal energy, the energy of the “noise” also increases. Recent research has shown that room reverberation can increase vocal effort (Berardi, Hunter, & Leishman, 2015; Bottalico, Graetzer, et al., 2017; Hunter et al., 2015). However, Bottalico (2017) demonstrates that the effect of reverberation on vocal effort is not linear and that some reverberation can be benificial to the talker and reduce vocal effort. In this study, the participants spoke and then rated their vocal effort in three different reverberation conditions (anechoic, reverberant, and semi-reverberant). Additionally, the participants were recorded with and without acrylic glass panels 0.5 m from the talker to provide strong early reflections. If vocal effort in reverberation were similar to the Lombard effect then it would be the case that vocal effort would be the smallest in the anechoic condition and the most in the reverberant condition with the panels adding more of a response to vocal effort. Bottalico reported that vocal effort was the highest in the anechoic condition and that the presence of the panels reduced vocal effort. This implies that a certain amount of reverberation is actually preferred for support in the auditory feedback. When comparing the combination of noise and reverberation on vocal effort, Cipriano, Astolfi, and Pelegrín-García (2017) reported that vocal comfort related more with the noise annoyance than the room’s reverberation. Although the vocal effort response to room reverberation is not as predictable as communication distance or background noise, it still is an environmental factor in vocal effort.

In each of these cases, talker-listener distance, background noise, and room reverberation, have been shown to relate to the notion that vocal effort is an accommodation to the limitations of the communication environment. These studies also illustrate that vocal effort is more than increasing vocal intensity. However, in many cases vocal effort is used analogously with vocal intensity and is measured as the sound pressure level of speech in decibels (dB). In fact the international standard “Ergonomics—Assessment of speech communication” (ISO 9921:2003) defines vocal effort to be the “exertion of the speaker, quantified objectively by the A-weighted speech level at 1 m distance in front of the mouth.” This measure of vocal effort is widely used (Cushing, Li, Cox, Worrall, & Jackson, 2011; Rosenthal et al., 2014) despite vocal effort being more nuanced than vocal intensity. Another example of the difference between vocal effort and vocal intensity is in the speech-loudness-effort hypothesis by Rosenblum et al (1991). This hypothesis states that the perception of loudness is based on a listener’s perception of the vocal effort of the talker and not the intensity of speech signal. The authors summarized research which showed perceived loudness to be more related to listener-estimated vocal effort than the SPL of the signal. Whether or not this hypothesis is completely accurate, it is the case that loudness perception and vocal effort are more than the amplitude of a speech signal.

Vocal intensity alone is a limiting measure for vocal effort, but it has been used to differentiate a categorical definition of vocal effort. For example, Skinner et al (1997) defined specific dB values for different vocal effort levels: causal (56 dB), normal (60 dB), loud (74 dB), and shout (83 dB). Others have used this similar pattern to operationalize vocal effort (although not always with specific decibel values). Cushing, Li, Cox, Worrall, and Jackson (2011) used “hushed, normal, raised, loud, and shout”; Zelinka, Sigmund, and Schimmel (2012) used “whisper, soft, normal, loud, and shouting”; while Holmberg, Hillman, Perkell, Guiod, and

Goldman (1995) simply used "comfortable and loud" as a binary vocal effort. These categorical definitions of vocal effort are straightforward to implement but can be problematic. With the exception of Skinner et al, there is no regulation for the variation of individual perception of these categories. Individuals have different speaking styles so one person's normal could be another person's loud. And even in the case of the set dB values, two individuals could have similar vocal levels with different vocal efforts. Another body of literature suggests that speaking style does have an impact on vocal effort. There is evidence that word stress (Mooshammer & Mooshammer, 2010; Sluijter & Van Heuven, 1996), hyper-articulation of fricatives (Meynadier, El Hajj, Pitermann, Legou, & Giovanni, 2018), and even posture (Lagier et al., 2010) can affect vocal effort. These are not the only possible individual factors that could influence vocal effort.

Vocal effort has also been defined as a symptom of other vocal conditions. Voice disorders have a higher prevalence and magnitude of reported vocal effort (Altman, Atkinson, & Lazarus, 2005; Roy, Merrill, Gray, & Smith, 2005; Smith et al., 1998). Some of the specific voice disorders associated with vocal effort include essential tremor (Warrick et al., 2000), muscle-tension dysphonia (Roy, Smith, Allen, & Merrill, 2007), and vocal nodules (Hillman, Holmberg, & Perkell, 1989). Adductor spasmodic dysphonia (ADSD) has been the most studied in relation to excessive vocal effort (Cannito, Doiuchi, Murry, Woodson, & York, 2012; Eadie & Stepp, 2013; Roy et al., 2007; Shoffel-havakuk et al., 2019; M. E. Smith, Roy, Wilson, & Hypothesis, 2006). In these cases, there are structural (vocal nodules) or neurological (ADSD) differences in these individuals that result in increased vocal effort. Other physiological conditions have been shown to result in excessive vocal effort including vocal fold stiffness (Katherine Verdolini et al., 1994), internal temperature (Sandage, Connor, & Pascoe, 2013), and dehydration (Pearl

Solomon & Stemmle DiMattia, 2000). A comprehensive definition of vocal effort would need to contain components of communication messages or objectives, environments, and capabilities.

2.3.a Theoretical framework for vocal effort and related terms

The Fifth Eriksholm Workshop on "Hearing Impairment and Cognitive Energy" developed the Framework for Understanding Effortful Listening (FUEL). This framework defines effort as "the deliberate allocation of resources to overcome obstacles in goal pursuit when carrying out a task" (Pichora-Fuller et al., 2016). Although this definition was developed for listening effort, it can provide the scaffolding for a framework on vocal effort. There are three key parts to this definition. First, "the deliberate allocation of resources to overcome" implies that effort is a cognitive and active process and that the available resources play a role in effort. Second, "obstacles in goal pursuit" means that effort is a result of some limitation to clarity of communication. Lastly, "when carrying out a task" means that effort is task-based which means that some sort of communication needs to be initiated for vocal effort to occur. This definition fits well with the previously mentioned components of a vocal effort definition.

Above it is stated that a "comprehensive definition of vocal effort would need to contain components of communication messages or objectives, environments, and capabilities." Incorporating these ideas with the definition of effort, vocal effort (as defined in this document) is then the deliberate allocation or exertion of cognitive or physiological resources to adapt existing communication capability to overcome internal (i.e. voice problems such as fatigue or functional, structural, or neurological impairments) and external (i.e. communication environments including talker-listener distance, background noise, room reverberation, or communciation intent) obstacles in goal pursuit of a voice-related communication objective or task. This proposed definition of vocal effort can be used to redefine and relate common terms used in voice research.

The terms that will be described are vocal load, vocal loading, vocal ability, and vocal fatigue and they will be hereafter referred to without "vocal" in the term. Within this framework, the load(s) are the obstacles that are in opposition to the communication objective. The loading is the allocation of resources to adapt or overcome the load. In other words, loading is vocal response to the demand of the load. Ability is the existing voice-related communication capability to handle or endure the particular load. Performance is the general use of the vocal mechansim. Finally, fatigue is the loss of resources as a result of the loading. To try to put this all in context together, effort is the exertion in response to loading in order to overcome a load towards a communication objective. Additionally, fatigue is the change in one's performance and/or effort as a result of loading. This is because an indiviudal can use more effort to maintain constant performance with loading, or lose performance with the absence of increased effort. A quick validity check of the framework is that an individual with prolonged vocal effort would eventually experience fatigue—this is consistent with the previous results.

Additionally, this framework supports the previously reported use of vocal effort. For example, "quantity that ordinary speakers vary when they adapt their speech to the demands of an increased or decreased communication distance" from Traunmüller and Eriksson (2000). Here the loading is the adaptation of their speech to a load of increased or decreased communication distance. More importantly is that the framework accounts for the observations that individuals can greatly vary in their response to the same load. This framework takes into account the ability that may affect how much effort is needed to overcome a load.

2.3.b Summary of proposed defintions for voice-related terms

For these terms two definitions are given. The first definition is within the framework of vocal effort as described above and the second is a more general term.

Vocal Effort:

(1) the deliberate allocation or exertion of cognitive or physiological resources to adapt existing communication capability to overcome internal (i.e. voice limitations such as fatigue or functional, structural, or neurological impairments) and external (i.e. communication environment demands such as talker-listener distance, background noise, and room reverberation) obstacles in goal pursuit of a voice-related communication objective or task

(2) the perception of exertion as a result of vocal loading (see below) to a perceived vocal load (see below)

Vocal Load:

(1) the obstacles that are in opposition to the communication objective,

(2) the vocal constraint including both internal (e.g. hydration, fatigue, vocie disorders, etc.) and external (e.g. acoustic environment, communication intent, etc.) to a particular voice-related communication task which is independent of the individual's physiology to perform the task or their perception of the task

Vocal Loading:

(1) the allocation of resources to adapt or overcome the load

(2) the change of an individual's manner of voicing (or vocal performance) to accommodate the vocal load in a particular voice-related communication situation (e.g. Lombard effect in the presence of noise, clear speech with hearing-impared

listeners, etc.)

Vocal Ability:

(1) the existing voice-related communication capability to handle or endure the particular load

(2) an individual's intrinsic voice physiology, experience, and perception of vocal load for a particular voice-related communication situation

Vocal Performance:

(1) the general use of the vocal mechanism

(2) the way an individual accomplishes a vocal task for a particular voice-related communication situation (i.e. fundamental frequency or amplitude of vocal fold vibration, vocal fold closure or posture, resonance, etc.)

Vocal Fatigue:

(1) the change in one's vocal performance and/or effort as a result of vocal loading

(2) the physiological and/or perceptual manifestation of a change in the voice that influences an individual's intrinsic voice physiology, experience, and perception of vocal load for a particular voice-related communication situation which may be a result of vocal loading or vocal effort

2.4 Hypothesis and Research Questions

The presented framework provides a novel and useful approach towards understanding vocal fatigue and related concepts. In particular, this framework allows for a dynamic system of underlying mechanisms of vocal fatigue as well as explaining the variability in previous experiments and accounting for individual vocal ability. Viewing vocal fatigue from the perspective of this framework, vocal fatigue could be measured as the change in vocal performance or vocal effort (or a combination of the two) with and without a vocal load. This perspective allows for the study of the factors that contribute to vocal fatigue because it reduces the attempt to measure the complex subsystems of vocal fatigue to practical measurements of vocal loading (change in vocal performance from a vocal load) and vocal effort.

2*.4.a Main Hypothesis*

The primary goal of this dissertation is to experimentally validate one of the primary assumptions of the proposed framework for vocal fatigue (e.g. a talker's experience of vocal fatigue is related to the change in vocal performance and/or vocal effort when responding to a vocal load). Therefore, the overarching hypothesis of this research is:

H0: The changes in vocal performance, vocal effort, and/or their interaction through a vocal load will implicate vocal fatigue.

2.4.b Research Questions

Towards the purpose of testing H0, three sub-hypotheses and research questions were developed. Fundamental to the utility of the framework is the ability to reliably measure vocal effort and show a direct relationship between effort and vocal performance. Thus, the first research question is:

Q1: Can perceived vocal effort be measured reliably and if so, how does vocal performance in terms of vocal intensity change with vocal effort?

Associated with this research question is Hypothesis 1:

H1: Vocal performance in terms of vocal intensity will be distinct for each vocal effort level and be consistent within and across participants.

To test H1, Experiment 1 consists of participants performing various speech tasks at specific effort levels from the Borg CR-100 scale.

With a reliable vocal effort scale and a relationship between vocal performance and vocal effort quantified (results of Experiment 1), the next step (Experiment 2) is to test the effects of different levels of vocal loads. In other words, the relationship between vocal performance and vocal effort was tested for different vocal loads to examine the interaction between vocal load, vocal performance, and vocal effort. To control for the effects of different loads, three levels of three different equivalent loads were used (these are loads that theoretically have equivalent vocal loading, for more detail see 3.2). Thus, the second research question is:

Q2: To what degree are vocal performance and vocal effort related given three equivalent vocal load levels?

Associated with this research question is Hypothesis 2:

H2: The vocal performance and vocal effort will be constant within equivalent load conditions.

To test H2, Experiment 2 consists of participants performing communicative tasks in three load levels of three different load source (where each load source should be equivalent) and rate their vocal effort associated with these tasks.

Experiment 2 establishes relationships between measurements of vocal performance and vocal effort across a vocal load. Testing the main hypothesis (H0) requires investigation of changes in these measurements within a controlled vocal load over time. Therefore, the third research question is:

Q3: To what degree do vocal performance, vocal effort, and/or their interaction change given a combined vocal load of excess background noise over time?

Associated with this research question is Hypothesis 3, which is a more specific iteration of H0:

H3: The measured changes in vocal performance, vocal effort, and/or their interaction will change through a vocal load (background noise and prolonged speaking).

To test H3 and by extension H0, Experiment 3 consists of participants performing a vocal loading task in the presence of excess background noise for an extended duration while rating their vocal effort levels throughout.

CHAPTER III: METHODOLICAL APPROACH

This dissertation presents a theoretical framework for vocal fatigue and validates several aspects of it through experimentation. This validation occurs through three experiments detailed in this chapter. These experiments directly test the research questions and hypotheses found in 2.4. First, descriptions of the measurement variables used throughout all of the experiments are provided. Then the experiments are presented in sequential order. The experiments build in numerical order, therefore, methodological approaches described in earlier experiments will only be referenced in the later experiments.

3.1 Measurement

Each of the experiments use similar methodology for the measurement variables. This section provides the general details for these variables that are used through all of the experiments.

3.1.a Vocal Effort Measurement

Perceived vocal effort was measured using the Borg CR-100 (Fig. 3.1). Borg scales have been used to quantify the perception of pain, exertion and effort generally (Borg & Löllgen, 2001; Fanchini et al., 2015). The Borg scales combine the mathematical precision of a direct magnitude estimation scale and the usability of a visual analog scale. The Borg CR-10 has been successfully adapted for applications of vocal effort (Borg & Löllgen, 2001; Fanchini et al., 2015; van Leer & van Mersbergen, 2017). Here the Borg CR-100 scale is modified from the previous Borg CR-10 scale used for vocal effort but using the more granular intervals of the Borg CR-100. The only change, other than the more granular scaling is that “very very” from the

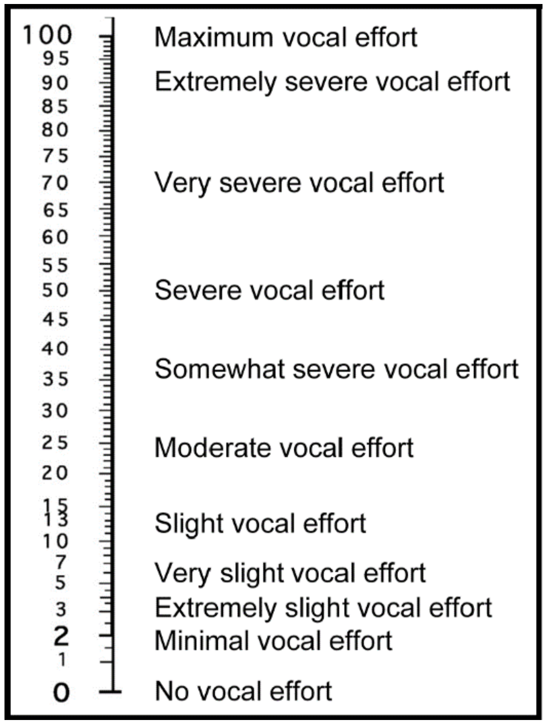

Figure 3.1 Borg CR-100 scale used to measure the self-perception of vocal effort.

vocal effort scale was replaced with "extremely" which is consistent from other applications of the Borg CR-10 scale. In Experiment 1, this scale is used as an independent variable while in Experiments 2 and 3 it is used as a dependent variable.

3.1.b Acoustic Measurement

The specific speech samples elicited varied across the experiments, but they were processed in a similar manner. After segmentation, the speech samples were processed to have all of the non-voicing segments removed (Maryn, De Bodt, & Roy, 2010; Rubin et al., 2019). While there are many acoustic speech parameters which could be used to reflect vocal performance, five were chosen here that represent basic performance qualities: pitch, pitch range, volume, dynamic range, and quality. Five acoustic parameters were derived from each voice concatenated speech segment: mean fundamental frequency (F0), standard deviation of fundamental frequency

(F0sd), speech level (SL), standard deviation of speech level (SLsd), and smoothed cepstral peak prominence (CPPS).

Fundamental Frequency

The fundamental frequency is the lowest mode of vocal fold oscillation during speech. While this measurement (cycles per second or Hertz) is a cyclic quantity, it also relates to perception of pitch. The mean and standard deviation of F0 were measured as part of the vocal performance evaluation.

The F0 was estimated using Praat's built in pitch capabilities. For the male participants estimation, an F0 range of 65 Hz to 350 Hz was used. While for the female participants, an F0 range of 150 Hz to 800 Hz was used. Table 3.1 contains the other Praat parameters used. Praat estimates the F0 in small increments over time through the speech sample resulting in a time-based array of frequency values . In the results below, F0 represents the average measured F0 from a sample, and F0sd represents the standard deviation of the measured F0 from a sample. Since the perception of pitch is nonlinear (i.e. a doubling of pitch from 100 Hz to 200 Hz is fewer Hz than a doubling of pitch from 200 Hz to 400 Hz), semitones (1/12th of an octave or doubling of pitch) are used for F0. This allows for comparable measures across all of the participants since habitual F0 significantly varies. The F0 measurements were converted to semitones as follows:

$$ST_n = 12 \log_2 \left(\frac{F_n}{F_{ref}} \right) \qquad \text{Eq. 3.1}$$

where F_n is the measured frequency in Hertz from sample n and F_{ref} is the reference frequency.

Table 3.1 Praat specifications for fundamental frequency calculation

Praat Parameter	Value
Analysis Method	Autocorrelation
Octave Cost	0.01
Max number of candidates	15
Silences threshold	0.03
Voice threshold	0.45
Octave-jump cost	0.35
Voice/unvoiced cost	0.2

The reference frequency is slightly different for each experiment, but all represent a habitual or baseline frequency measurement of each participant. A semitone is 1/12th of an octave where an octave is doubling of frequency. This is used to normalize the change in F0 across many participants with different baseline F0.

Speech Level

In general, the speech level or sound pressure level (SPL) is a measurement of vocal intensity and relates to the perceived loudness of the talker and is calculated as follows:

$$SPL = 20\log_{10}\left(\frac{p_{RMS}}{p_{ref}}\right) \qquad \text{Eq. 3.2}$$

where p_{RMS} is the root-mean-square (RMS) sound pressure and p_{ref} is the reference pressure of 20 micropascals (μPA). Here a reference voltage was provided from a reference microphone and calibrator to convert the microphone voltage to sound pressure. The mean and standard deviation of the SPL were measured as part of the vocal performance evaluation.

Here the mean speech level (SL) is the average SPL of multiple windowed segments of the voice concatenated voice signal as follows:

$$SL = \frac{1}{N}\sum_{i=1}^{N} SPL_i \qquad \text{Eq. 3.3}$$

where N is the number of segments in the sample. The segments were 20 msec windows with 50% overlap. The standard deviation of the speech level (SLsd) is the sample standard deviation of the SPL segments as follows:

$$SLsd = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} |SPL_i - SL|^2} \qquad \text{Eq. 3.4}$$

Smoothed Cepstral Peak Prominence

The smoothed cepstral peak prominence (CPPS) is an acoustic measure that represents the relative strength or prominence of the periodicity of the voice signal. More specifically it is the distance (measure in decibels) between the peak of the first rahmonic (dominant quefrency) and a linear regression of the smooth cepstrum. The cepstrum is defined as the fast Fourier transform (FFT) the of log magnitude power spectrum as follows:

$$C(q) = \mathcal{F}\{log|\mathcal{F}(f(t)|^2\} \qquad \text{Eq. 3.5}$$

where $\mathcal{F}$ is the Fourier operator and $f(t)$ is the time-domain speech signal. Here the CPPS was calculated with Praat and used the voice-concatenated segments. This measure has been widely used in previous voice science work as an acoustic correlate to voice quality and therefore is included as one of the vocal performance measures (Hillenbrand, Getty, Clark, & Wheeler, 2005; Maryn et al., 2010).

3.1.c Analysis Software

The acoustic signal processing was completed in Matlab with custom scripts that used Praat's functions for the voice concatenation, fundamental frequency estimation, and smoothed cepstral peak prominence measurement. Statistical analysis was completed using SPSS.

3.2 Experiment 1

The purpose of this experiment was to test hypothesis 1 (H1).

Q1: Can perceived vocal effort be measured reliably and if so, how does vocal performance in terms of vocal intensity change with vocal effort?

H1: Vocal performance in terms of vocal intensity will be distinct for each vocal effort level and be consistent within and across participants.

The independent variable for this experiment was the vocal effort goal.

The dependent variable for this experiment was the vocal performance.

In Experiment 1, the vocal effort level goal was determined using the Borg CR-100 scale (Fig 3.1). This scale was also used in the other two experiments to measure perceived vocal effort. Additionally, this experiment measured vocal performance in terms of the five voice acoustic parameters outlined above (fundamental frequency mean and standard deviation, speech level mean and standard deviation, and smoothed cepstral peak prominence). In this experiment participants produced different speech samples at specific cued vocal effort levels.

3.1.a Participants

With protocol approval of the Michigan State University's Human Research Protection Programs Human Subject's Review Board, this experiment consisted of twenty college-age participants (10 male and 10 females). The participants were recruited through an online recruiting and scheduling system at Michigan State University in the College of Communication

Arts and Sciences. The participants' time was compensated with course credit. The participants were screened for hearing limitations. The hearing screening required for inclusion consisted of pure-tone stimulation of at least 20 dB HL in both ears at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. Following informed consent (included in Appendix D), the participants proceeded with the experiment.

3.1.b Instrumentation

Acoustic Measurement

The participants were recorded with a head-mounted microphone (HMM; omni-directional; Countryman B3) placed approximately 5 cm from the participant's mouth using a digital hand-held recorder (ZOOM H5 Handy Recorder) in a sound-isolation booth ($2.5 \times 2.1 \times 2.0$ m single-walled). A reference microphone (Behringer ECM8000) was placed 50 cm from the speaker's mouth. The reference microphone was absolutely calibrated to 94 dB SPL (relative to 20 μPa) using a commercial calibrator that fit onto the microphone. This calibration was used to referentially calibrate the HMM. This was done through the participant producing a steady vowel 50 cm from the reference microphone. The voltage ratios between the calibration signal of 94 dB and vowel productions were used to calibrate the HMM. The speech was sampled at 44.1 kHz with 16-bit resolution.

Stimulus Presentation Instrument

The stimulus was present on a laptop using the open-source python-based program, PsychoPy (Peirce et al., 2019). Additionally, all of the user's non-acoustic and non-survey responses were recorded through this program. A schematic of the instrumentation used is contained in Fig. 3.2

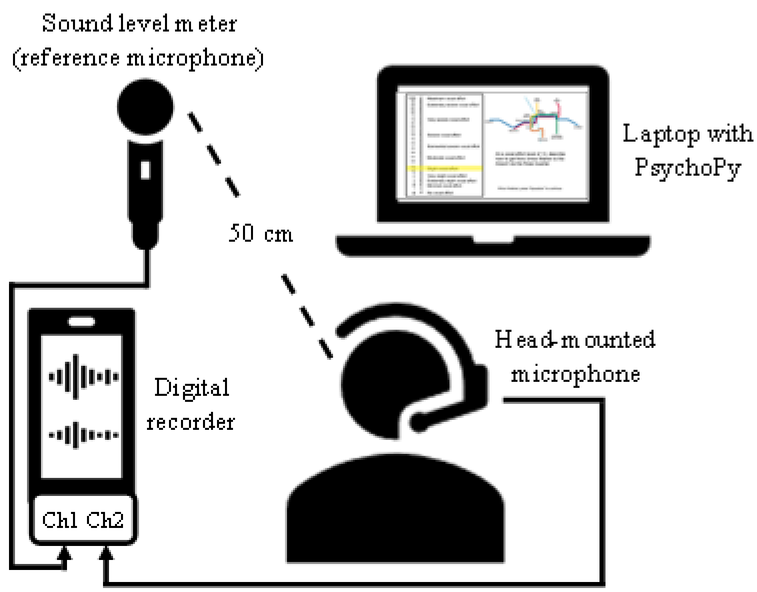

Figure 3.2 Schematic for the instrumentation used in experiment 1

3.1.c Stimuli

Throughout the experiment the participants were instructed to read a particular stimulus at a specific vocal effort level from the Borg CR-100 scale (Fig. 3.1). This study used three types of speech stimuli with three variations of each type. For a complete list of the speech stimuli used, see Appendix A: Experiment 1 Stimuli. The speech stimuli were chosen such that it would take a typical talker 12-16 seconds to produce the utterance.

The first type of speech stimuli was automatic speech. This is speech that it is generally accepted that any native English speaker could produce without a script. Here participants were asked to either state the alphabet (English), count to twenty-five, or say the names for days of the week and months of the year.

The second type of speech stimuli was reading aloud standard sentences. These sentences were excerpts from standard speech acoustic reading passages, the Marvin Williams Passage

(Švec, Titze, & Popolo, 2005), the Rainbow Passage (Fairbanks, 1960), and the Stella Passage (Weinberger & Kunath, 2011). (See Appendix A for the exact sentences used.)

The third type of speech stimuli used was a map description task. Here participants were asked to describe a specific route on a Portland, Oregon subway map (Fig. 3.3). The image and a description of the desired path were shown to the participants. The routes were all in the form of "describe how to get from A to B via C" where "A", "B", and "C" were specific points on the map. There were three different routes used during the experiment (see Appendix A).

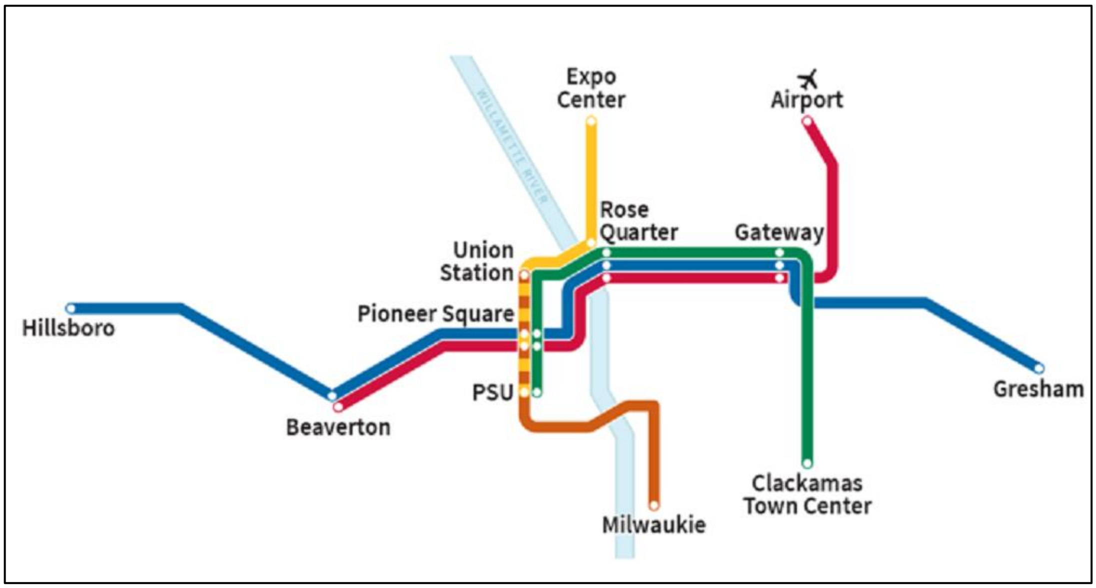

Figure 3.3 Subway map used for the map description speech task.

3.1.d Procedure

Prior to the main experiment, the participants first went through a tutorial. The tutorial and experiment were narrated using artificial speech from a text script via WaveNet (Van Den Oord et al., 2016). An artificial narrator was used to remove any possible experiment-administration bias and to keep all narration samples acoustically similar (e.g. later narration samples could be generated to account for changes in the experiment without any perceptual change in when the sample was created).

The tutorial consisted of having the participant first practicing each automatic and reading speech task (outlined in section 3.1.c and contained in Appendix A). The tutorial and practice were implemented to reduce a novice and/or learning effect by the participants during the main experiment. The tutorial also explained how to properly describe the map (Fig. 3.3) and provided an example different from the three routes used as stimuli. The example used was

> [I]f asked to 'describe how to get from Hillsboro to the Airport via Gateway' you would say… Starting at Hillsboro, I will take the blue line eastbound towards Beaverton. I will pass Beaverton, Pioneer Square, and the Rose Quarter. I will change at the Gateway station to the red line northbound to the Airport. And finally, I will arrive at the Airport".

The participants then practiced describing a route on the map.

Next the participants were introduced to the vocal effort scale (Fig. 3.1). The scale was presented with anecdotal anchors for the extreme values of 1 and 100 as follows: "a vocal effort level of 1 would be quietly talking to someone next to me at home" and "a vocal effort level of 100 would be trying to shout at someone while standing on an airport runway."

Following the tutorial, the participants started the main experiment. Here the participants were instructed to speak a particular speech stimulus (see the pervious section 3.1.b and Appendix A for more information on the speech stimuli) and at specific vocal effort level. Four vocal effort levels were prompted: (1) 2 or "Minimal vocal effort", (2) 13 or "Slight vocal effort", (3) 25 or "Moderate vocal effort", (4) 50 or "Severe vocal effort." The participants completed each speech stimulus (total of nine unique stimuli; three variations of three types) at each vocal effort level (four) for a total of thirty-six trials. The vocal effort scale (Fig. 3.1), speech stimulus, and map (if applicable) were shown to the participant for each trial (see Figs. 3.4 and 3.5). These trials were randomized for each participant. The experiment concluded after the thirty-six trials.

3.1.e Statistical Analysis

The reference used for the semitone calculation for each participant was the average fundamental frequency of the practice tasks.

As mentioned in 3.1.c, SPSS statistical software was used for statistical analysis. First the statistical assumptions of normality, independence, and equal variance were checked. If these assumptions were met, the sample means for each of the five acoustic parameters (F0, F0sd, L, Lsd, CPPS) were compared across the VELs using one-way analysis of variance (ANOVA) tests with an alpha level of 0.05. Post hoc Tukey HSD tests were used to compare the speech production of the VEL pairs. If the equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used. The test-retest reliability of the scale was measured using Pearson's r (which measures the strength of a linear association) of the measured speech levels (SL) across the repeated vocal effort levels of the speech level. Outliers within the dataset were removed case-by-case that were either greater than the sum of the third quartile (Q3) and 1.5 times the interquartile range (IQR) or less than the difference of the first quartile (Q1) and 1.5 times the IQR as described in Equation 6 below.

$$Outlier_x = \{x > Q3 + 1.5 * IQR, x < Q1 - 1.5 * IQR\} \quad \text{Eq. 6}$$

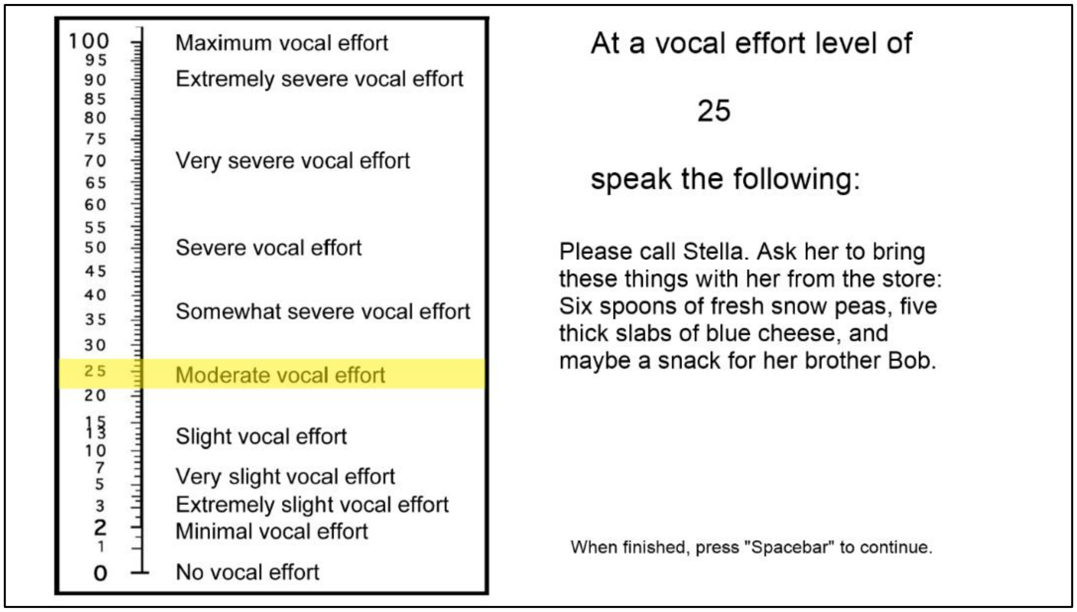

Figure 3.4 Example of the presentation of the vocal effort scale (left), target vocal effort level (upper-right) and speech stimulus (right) during an experimental trial.

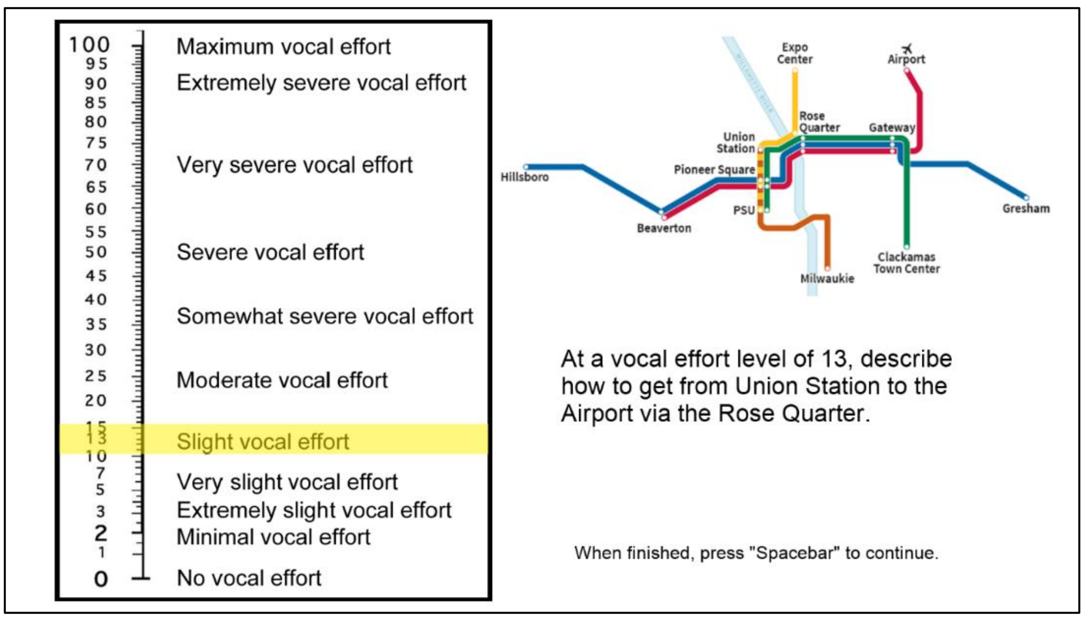

Figure 3.5 Example of the presentation of the vocal effort scale (left) and speech stimulus (right) for a map description task during an experimental trial.

3.3 Experiment 2

The purpose of this experiment was to test hypothesis 2 (H2).

Q2: To what degree are vocal performance and vocal effort related given three equivalent vocal load levels?

H2: The vocal performance and vocal effort will be constant within equivalent load conditions.

The independent variables for this experiment were the vocal load types and levels.

The dependent variables for this experiment were vocal effort ratings (VER) and vocal performance.

Vocal performance was quantified through measurements of the mean and standard deviation of fundamental frequency (F0; F0sd), the mean and standard deviation of speech level (SL; SLsd), and the smoothed cepstral peak prominence (CPPS). Vocal effort was measured through self-ratings on the Borg CR-100 scale. The vocal loads were communicative situations of talker-listener communicative distance, talker loudness goal, and excess background noise. These conditions were selected for this study because they have theoretical vocal intensity equivalences. In other words, a doubling of communicative distance, an increase of background noise of 9 dB (i.e. Lombard effect Bottalico, Passione, Graetzer, & Hunter, 2017), and an increase of a loudness target of 6 dB should all require an increase of 6 dB of vocal intensity to maintain acoustic energy equivalence. These theoretical equivalencies allowed for direct comparison of vocal performance and vocal effort ratings.

In this experiment participants were asked to complete a communicative task with a particular vocal load (communicative distance, loudness goal, or background noise) and then rate their vocal effort level for that task.

3.2.a Participants

With protocol approval of the Michigan State University's Human Research Protection Programs Human Subject's Review Board, this experiment consisted of forty college-age participants (20 males and 20 females). The participants were recruited through an online recruit and scheduling system at Michigan State University in the College of Communication Arts and Sciences. The participants' time was compensated with course credit. The participants were screened for hearing limitations in the same manner as Experiment 1. Following informed consent, the participants proceeded with the experiment.

3.2.b Instrumentation

The participants were recorded with a head-mounted microphone (HMM; omni-directional; Countryman B3) placed approximately 5 cm from the participant's mouth using a digital hand-held recorder (ZOOM, H5 Handy Recorder). A reference microphone (Behringer ECM8000) was placed 50 cm from the speaker's mouth. The reference microphone was absolutely calibrated to 94 dB SPL (relative to 20 μPa). This calibration was used to referentially calibrate the HMM using the same method from 3.2.b. The speech was sampled at 44.1 kHz with 16-bit resolution. The experiment took place in an anechoic chamber (3.4 × 4.6 × 2.4 m, IAC number 107840)—a specialized research facility designed to completely absorb the reflection of sound. The use of this facility was essential to this study as room reverberation plays a significant factor in the perception of background noise and communicative distance, as well as vocal effort (Berardi, Whiting, et al., 2015; Bottalico, 2017; Whiting, Leishman, Eyring, Berardi, & Rollins, 2015).

The stimulus was present using PyschoPy via an external monitor connected to a laptop. All of the user's non-acoustic and non-survey responses were recorded through this program. Schematic for instrumentation is shown in Fig. 3.6.

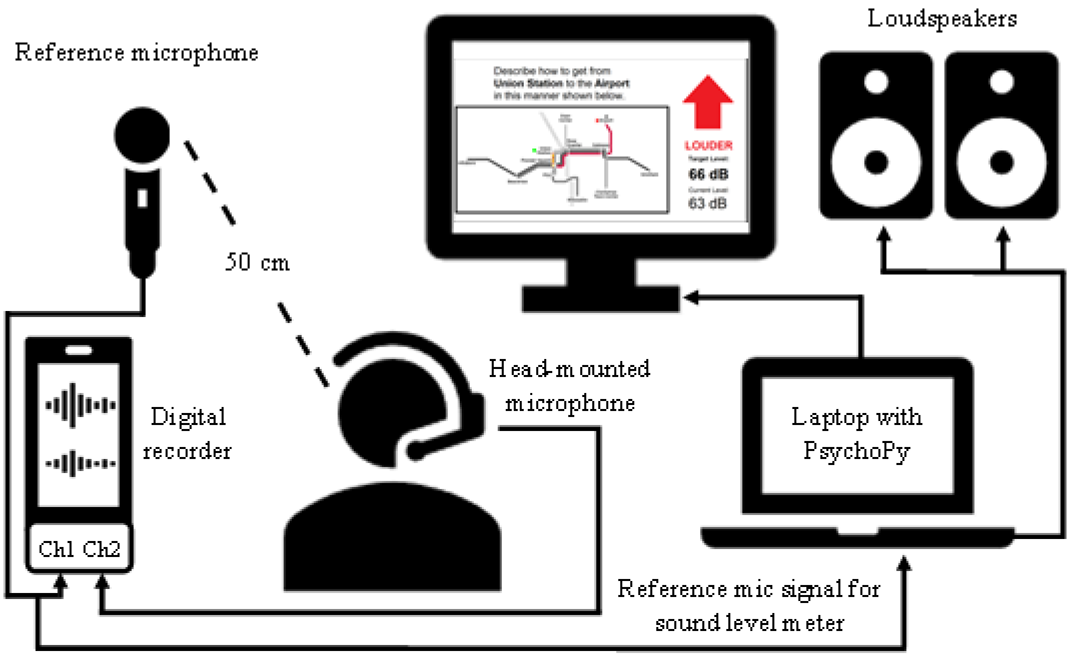

Figure 3.6 Schematic for the instrumentation used in experiment 2

3.2.c Stimuli

This experiment used a map description task to elicit speech in three different types of vocal load conditions with three variations of each condition (for a total of nine conditions). The map description task is similar to the map description task outlined in Experiment 1. The difference in this experiment was that the researcher acts as a communication partner. This means that the participant was instructed that they needed to explain the route on the map in such a way that the researcher could in real time trace the map.

The participant was presented various maps via PsychoPy on a computer screen where all the routes are gray except for the route that must be described (see Fig. 3.7). The research had a stack of gray maps (no colored routes) and traced the described path as the participant explained the route (see researcher's template Fig. 3.8). There were nine different routes such that each route was repeated 3 times but never across the same condition (i.e. each map was not used for the same load condition more than once). (See Appendix B for the maps used and what vocal load conditions for which each was used.)

The three vocal load conditions used in the experiment included communicative distance, loudness goal, and background noise. During this experiment, the participant is seated in the corner of the anechoic chamber and the researcher is seated at 1 meter from the participant except for when communicative distance was manipulated.

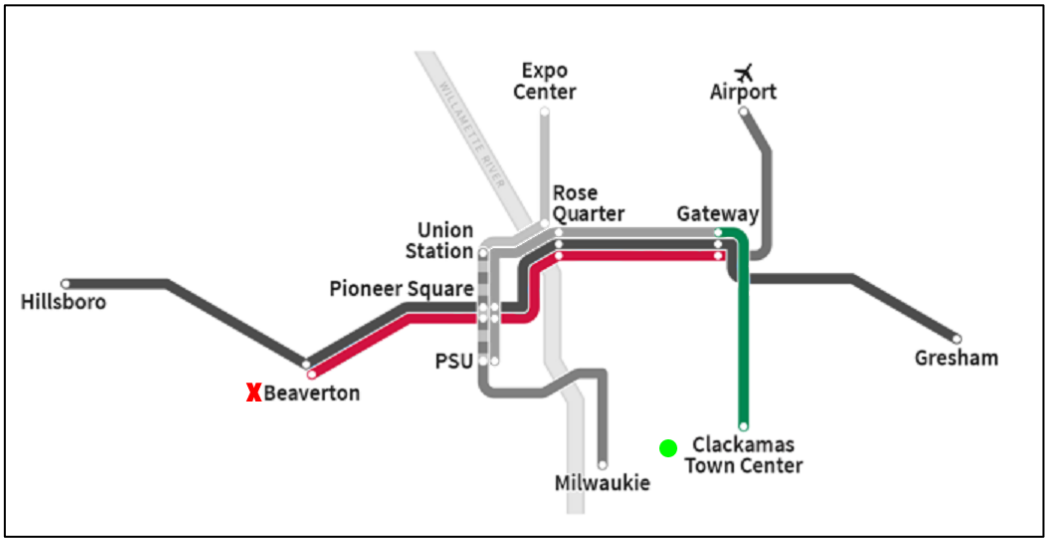

Figure 3.7 Example map for map description task in Experiment 2. This map is instructing the participant to describe the route from Clackamas Town Center to Beaverton via Gateway.

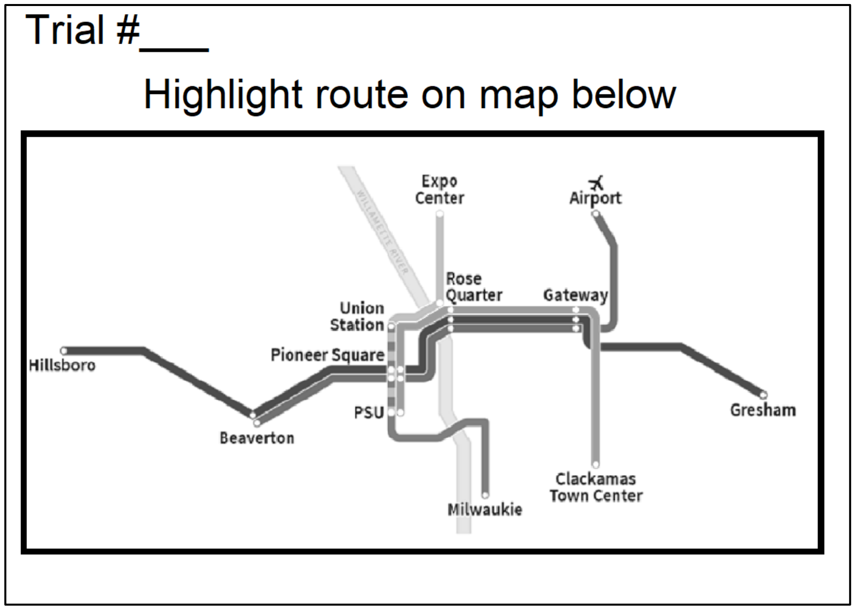

Figure 3.8 Researcher's template for Experiment 2. This template includes a colorless map to be highlighted as the route is described by the participant.

The communicative distance consisted of the researcher moving to three different points in the room along a direct line from the participant (as shown in Fig. 3.9). These three distances were as follows:

- D01; short distance: 1 meter
- D02; moderate distance: 2 meters
- D04; long distance: 4 meters

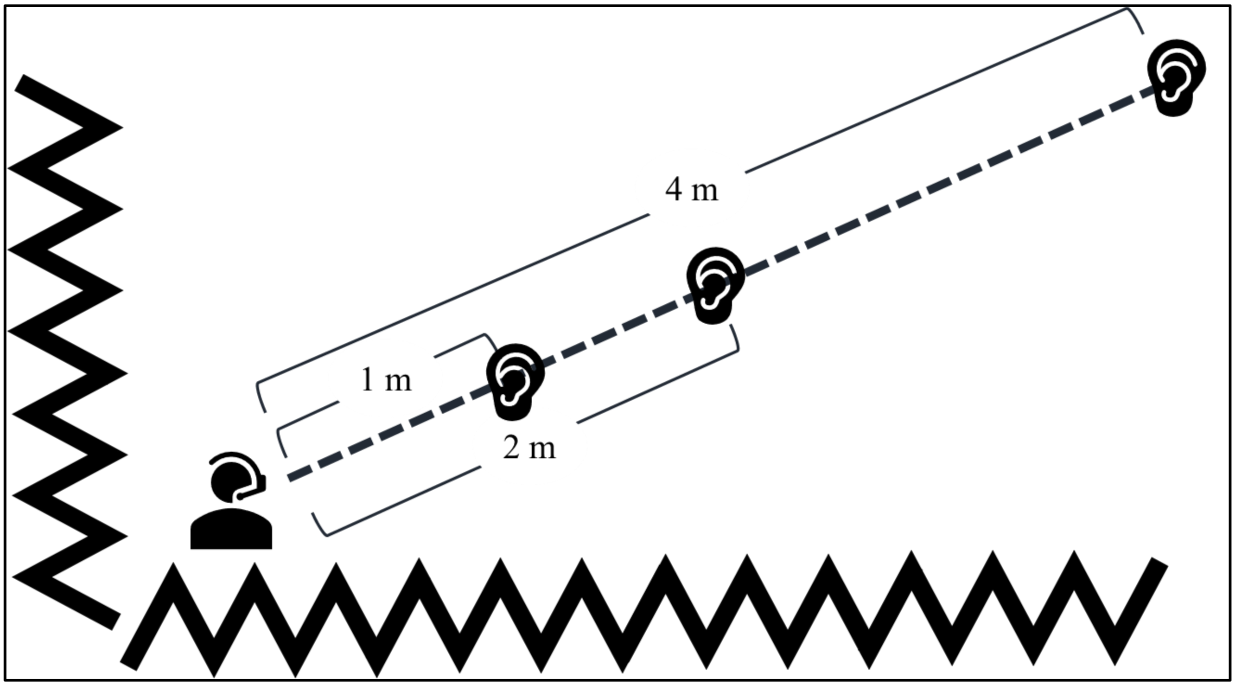

Figure 3.9 Visual description of the communication distance vocal load. The graphic in the bottom left corner with the headset represents the participant, the ears represent the potential locations of the listener during the experiment. The jagged walls are reminders that the experiment was contained in an anechoic chamber.

The loudness goal condition required the participant to speak such that their average voicing level was above a certain intensity threshold. The average voicing level was calculated using a reference microphone placed 1 meter in a direct line from the participant. The intensity thresholds (measured at 1 meters) were as follows:

- L54; low goal: 54 dB
- L60; moderate goal: 60 dB
- L66; high goal: 66 dB

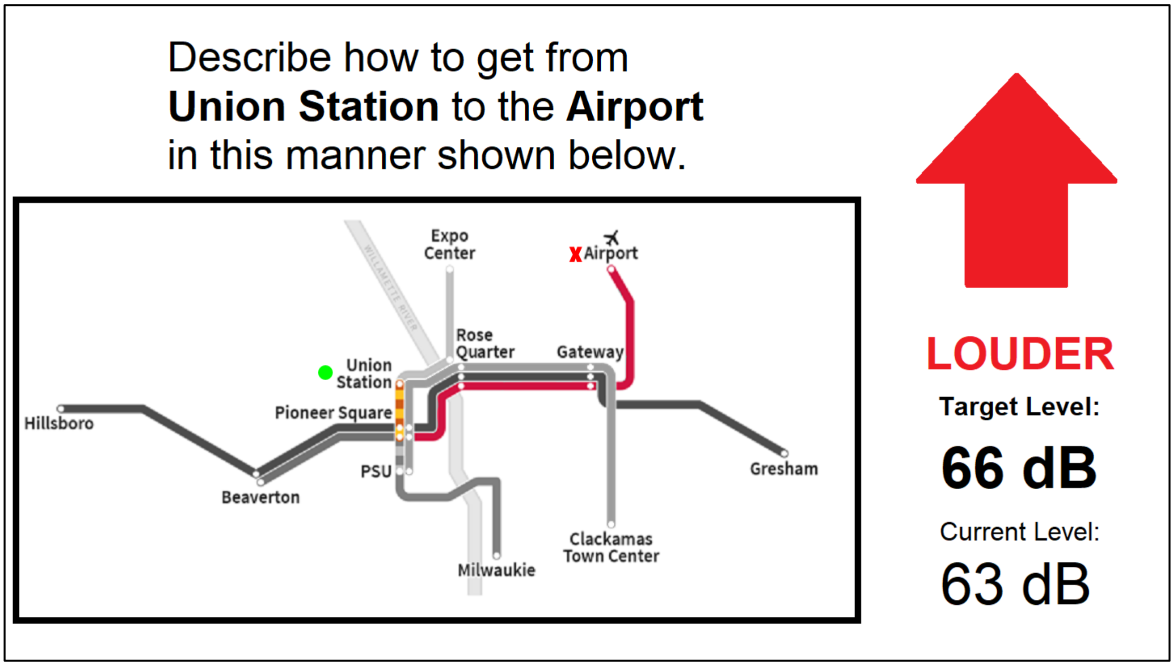

Figure 3.10: Example of map stimulus for the participant in Experiment 2. This particular example would be for the loudness goal condition of 66 dB. The participant is shown their current level and if it is lower than the target, a red arrow (shown in the upper-right) reminds them to talk louder.

The participants were able to see their current average voicing level and target loudness level. A large red arrow pointing up would display on the screen to prompt the participant to increase their speaking intensity (see Fig. 3.10).

For the background noise condition, pink noise played from two loudspeakers placed 30 degrees off axis in a 2-meter arc from the participant (as shown in Fig. 3.11). The background noise levels (measured at participant's location) were as follows:

- N53; low noise: 53 dBA
- N62; moderate noise: 62 dBA
- N71; high noise: 71 dBA

The interval of 9 dB was chosen to allow for a 6-dB vocal intensity increase predicted by the Lombard effect (Bottalico, Passione, et al., 2017). (Note that the Lombard effect has been shown to change slope with different dB ranges, here the dB range was chosen such that the predicted Lombard effect would be consistent with a 6-dB vocal intensity with a 9-dB increase of background noise).

Each of the three vocal load conditions had three variations for a total of nine conditions. Each condition was repeated three times for a total of 27 trials. The trials were presented to each participant randomly.

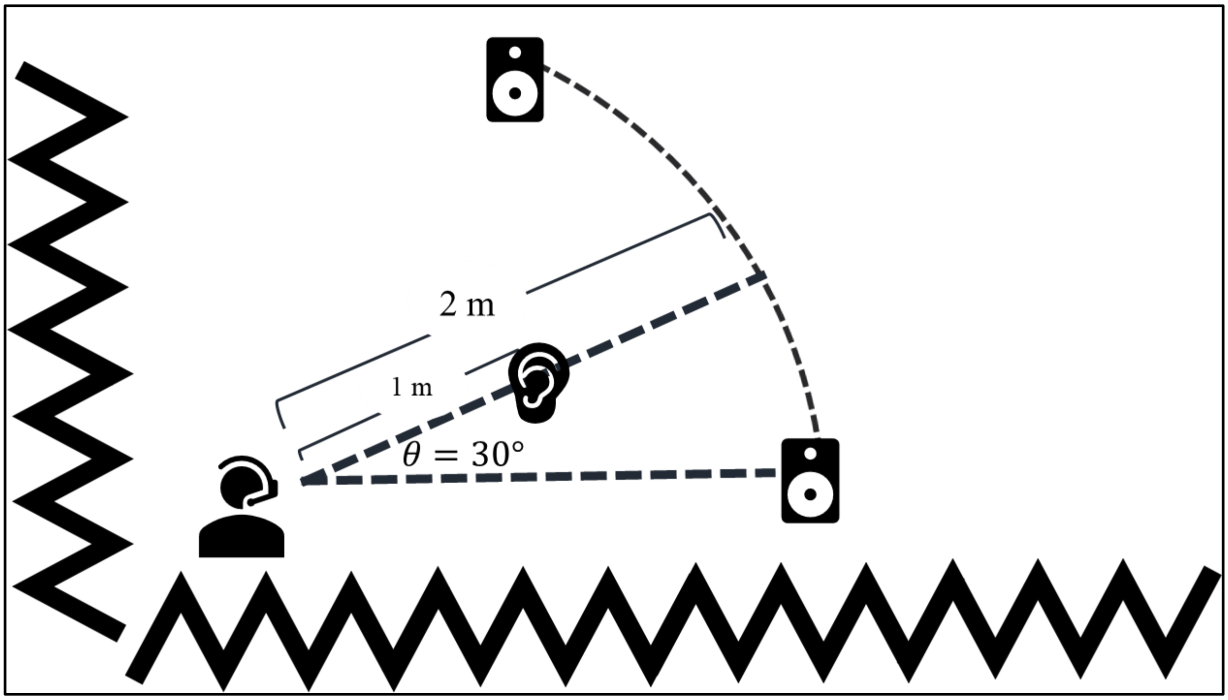

Figure 3.11 Visual description of the background noise vocal load. The graphic in the bottom left corner with the headset represents the participant, the ear represents the location of the listener (1 m). The loudspeakers are shown to be in a 2 m arc from the participant 30 degrees off axis. The jagged walls are reminders that the experiment was contained in an anechoic chamber.

3.2.d Procedure

The participants started with a tutorial to familiarize them with stimuli in the experiment. First the participants were introduced to the vocal effort scale (Borg CR-100). This scale used an anecdotal anchor similar to Experiment 1: “a vocal effort level of 1 would be quietly talking to someone next to me at home” and “a vocal effort level of 100 would be trying to shout at someone while standing on an airport runway.” Then the scale was experientially anchored by having the participants state the days of the week and months of the year at vocal effort levels of (1) 2 or “Minimal vocal effort”, (2) 13 or “Slight vocal effort”, (3) 25 or “Moderate vocal effort”, (4) 50 or “Severe vocal effort.” The purpose of this exercise was to condition the participant to their own sense of vocal effort for four distinct vocal effort levels.

Next the participants were trained on how to perform the map task. This consisted of a demonstration of an example map route. The example used was the same as described in Experiment 1. The participants then practiced describing a map route. Corrective instructions were made by the researcher to ensure clarity and consistency of task performance.

Following the training, the participants were presented with the randomized twenty-seven trials. Each trial consisted of a map route and rating of vocal effort. For each trial, the researcher had a template (Fig. 3.8) to actively record the route description. This was implemented to create a realistic communication scenario. Following the twenty-seven trials, the experiment concluded.

3.2.e Statistical Analysis

The reference used for the semitone calculation was the average fundamental frequency of the baseline condition for a participant which had no loudness goal or background noise and the communication partner was at 1 meter.

As mentioned in 3.1.c, SPSS statistical software was used for statistical analysis. First the statistical assumptions of normality, independence, and equal variance were checked. If these assumptions were met, the sample means for the self-reported vocal effort level (VER) and each of the five acoustic parameters (F0, F0sd, L, Lsd, CPPS) were compared across the vocal load levels using one-way analysis of variance (ANOVA) tests with an alpha level of 0.05. Post hoc Tukey HSD tests were used to provide pair-wise comparison of each load level (e.g. 1 m, 2 m, 4 m in the case of communication distance). If the equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used. Outliers within the dataset were removed case-by-case that were either greater than the sum of the third quartile (Q3) and 1.5 times the interquartile range (IQR) or less than the difference of the first quartile (Q1) and 1.5 times the IQR as described in Equation 6 in 3.1.e.

3.4 Experiment 3

The purpose of this experiment was to test hypothesis 3 (H3).

Q3: To what degree do vocal performance, vocal effort, and/or their interaction change given a combined vocal load of excess background noise over time?

H3: The measured changes in vocal performance, vocal effort, and/or their interaction will change through a vocal load (background noise and prolonged speaking).

The independent variable for this experiment was the point in time across the vocal loading task.

The dependent variables for this experiment were vocal effort ratings (VER) and vocal performance.

Vocal performance was quantified through measurements of the mean and standard deviation of fundamental frequency (F0; F0sd), the mean and standard deviation of speech level (SL; SLsd), and the smoothed cepstral peak prominence (CPPS). Vocal effort was measured through the Borg CR-100 scale. The vocal loads were background noise and prolonged speaking. While the previous experiment measured vocal loading in varied equivalent vocal loads, this experiment investigated the effects of the background noise load and time.

3.3.a Participants

With protocol approval of the Michigan State University's Human Research Protection Programs Human Subject's Review Board, this experiment consisted of forty participants (20 male and 20 females). The participants were recruited through an online recruit and scheduling system at Michigan State University in the College of Communication Arts and Sciences. The participants' time was compensated with course credit. The participants were screened for hearing limitations in the same manner as Experiments 1 and 2. Following informed consent, the participants proceeded with the experiment.

3.3.b Instrumentation

The participants were recorded with a head-mounted microphone (HMM; omni-directional; Countryman B3) placed approximately 5 cm from the participant's mouth. The microphone signal went through a pre-amplifier (Millennia HV-3D) and A/D converter (RME ADI-8 DS) then was recorded using REAPER, a digital audio workstation. A reference sound level meter (SLM; IEC 60651 Type 2) was placed 50 cm from the speaker's mouth. The reference microphone of the SLM was absolutely calibrated to 94 dB SPL (relative to 20 μPa). This calibration was used to referentially calibrate the HMM using the same method from 3.2.b. The speech was sampled at 44.1 kHz with 16-bit resolution.

The stimulus was presented using PyschoPy using an external monitor. All of the user's non-acoustic and non-survey responses were recorded through this program. Schematic for instrumentation is shown in Fig. 3.6.

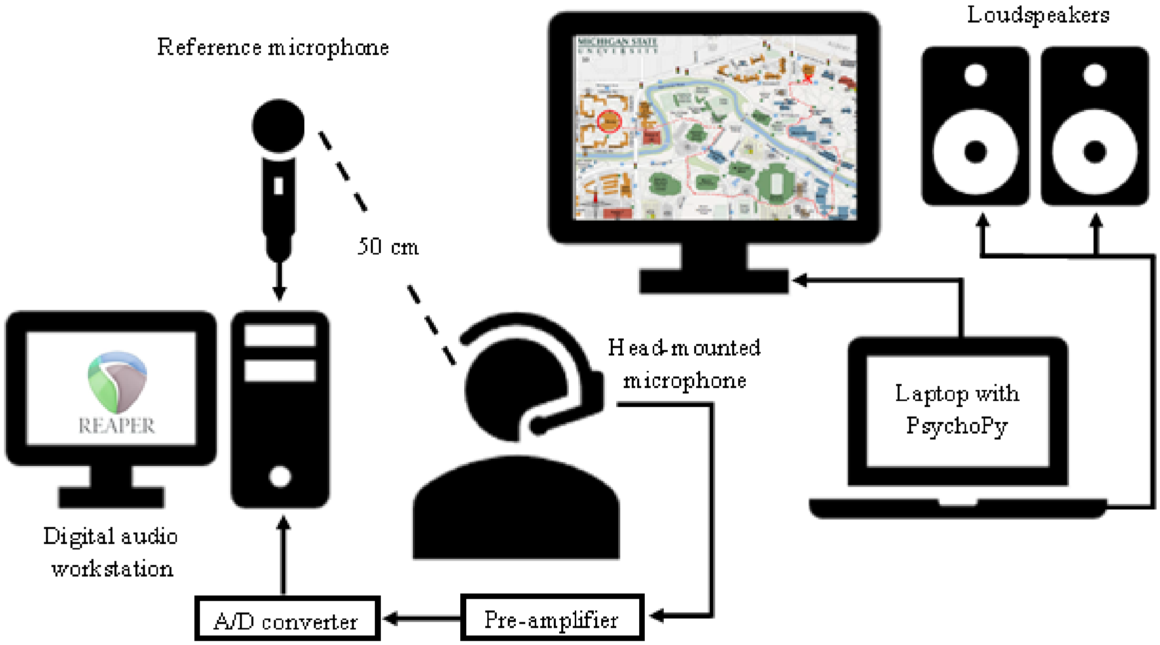

Figure 3.12 Schematic for instrumentation used in experiment 3

3.3.c Stimuli

This experiment had pretest and posttest voice tasks, a vocal loading task, and an intermittent vocal effort rating throughout the vocal loading task. The pretest and posttest voice tasks consisted of the participant reading the Rainbow Passage (Fairbanks, 1960).

During the vocal loading task, the participants described routes on maps. Here the participants were shown a series of maps with a start, path, and end point (see Fig. 3.13). They were asked to describe the map such that someone else could recreate the exact route. The advancement of the maps was self-paced. That is to say, the participant advanced the next page or map when needed.

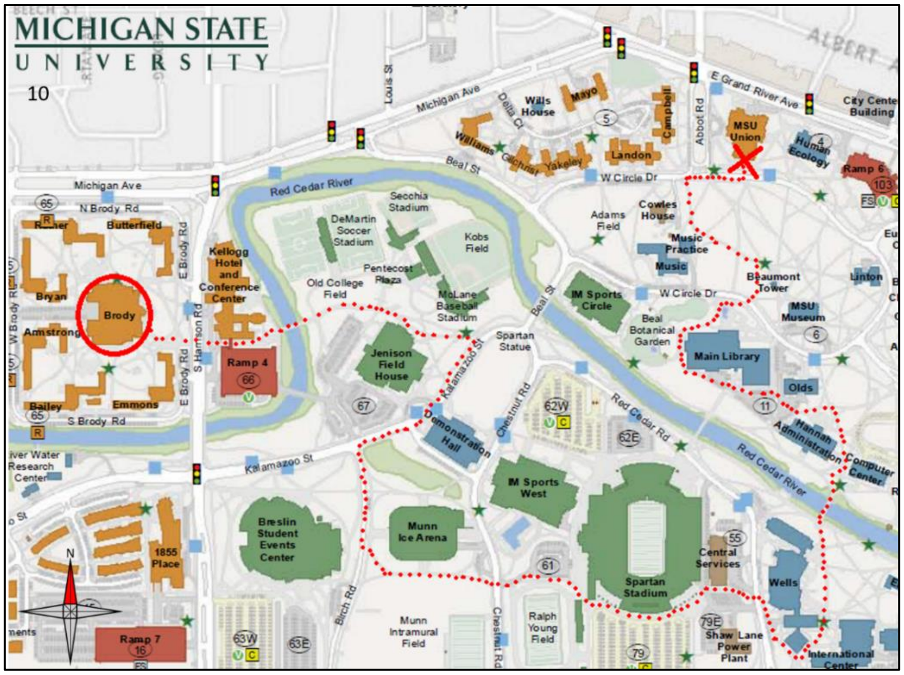

Figure 3.13: Example of a map for Experiment 2. Each map had a compass in one of the corners showing north. Additionally, each map had a starting point denoted by a red circle, an ending point denoted by a red "X", and a dotted line between the two denoting the route to be described.

During the vocal loading task, every five minutes the reading or maps description was interrupted to allow for vocal effort ratings using the Borg CR-100 scale. Then the participants were returned to the last seen map. Additionally, during these breaks in the vocal loading task, participants were cued to drink 30 mL from a small measured cup. This was implemented to prevent dry-mouth sensations from influencing the feeling of vocal effort or fatigue.

3.3.d Procedure

Prior to the main experiment, participants completed the same screenings and tutorials outlined in Experiment 2 (i.e. hearing screening and vocal effort scale training). The participants received additional instruction on how to advance the maps for the vocal loading test. Following the screenings and tutorials, the participants did the pretest voice tasks, including rating their current vocal effort level (Borg CR-100).

Following the pretest voice task, the participants were instructed on how to properly complete the reading or map task. Additionally, they were instructed that they could signal to the researcher to end the vocal loading task if vocal effort or pain in the throat became too elevated.

When the vocal loading task started, background noise in the form of multi-talker speech babble (made of six female and six male North American speakers) was played. Speech babble was chosen because it is quasisteady with a spectrum identical to normal speech creating a more realistic communication situation. Additionally, the presence of this type of noise replicates the design of a clinical VLT by Whitling, Rydell, & Lyberg Åhlander (2015). The background noise started at 45 dBA and gradually increased to 75 dBA over a period of 30 seconds at a rate of 10 dB every 10 seconds (this matches a doubling of perceived loudness every 10 seconds). The background noise persisted through the vocal loading task at 75 dBA until the experiment was either voluntarily terminated or six five-minute intervals (a total of 30 minutes) had been completed. As mentioned above, after each five-minute interval the participants completed an intermittent vocal effort rating and a drink of water.

Following the end of the vocal loading task, the participants did the posttest voice tasks including rating their current vocal effort level (VER) using the Borg CR-100. Following this posttest, the experiment concluded.

3.3.e Statistical Analysis

The reference used for the semitone calculation for each individual was the average fundamental frequency of the Rainbow Passage reading recorded prior to the vocal loading task.

As mentioned in 3.1.c, SPSS statistical software was used for statistical analysis. First the statistical assumptions of normality, independence, and equal variance were checked. If these assumptions were met, the sample means for the self-reported vocal effort level (VER) and each of the five acoustic parameters (F0, F0sd, L, Lsd, CPPS) were compared across each time point of the VLT using one-way analysis of variance (ANOVA) tests with an alpha level of 0.05. Post hoc Tukey HSD tests were used to provide pair-wise comparison of each time point (pre, post, and the six five-minute increments during the loading task). If the equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used. Outliers within the dataset were removed case-by-case that were either greater than the sum of the third quartile (Q3) and 1.5 times the interquartile range (IQR) or less than the difference of the first quartile (Q1) and 1.5 times the IQR as described in Equation 6 in 3.1.e.

In order to investigate the interaction of vocal effort and vocal performance as well as investigate subgroups within the data as predicted by the framework, clustering of vocal effort ratings and acoustic measurements were performed. First, a set of ten features were derived from the vocal effort ratings (shown in Table 3.2). Then the data were clustered into two groups using k-means clustering in SPSS. The clustering models were iterated until a sufficient model was developed with the minimum features. The features were systemically excluded based on feature importance and statistical significance within the model. The two groups were labeled based on the framework's assumptions with how changes in vocal effort would relate to vocal fatigue through vocal loading. One group with a cluster center suggesting vocal effort changes associated with the vocal loading task was labeled the "high vocal load response" group and the other group was labeled the "low vocal load response" group.

The vocal performance acoustic metrics were also split into groups. Based on previous literature it is possible that not all of the participants will have a significant change in vocal performance as a result of the vocal loading. Additionally, from the background review (Chapter 2) it is unclear which acoustic measures are most likely to change from the vocal loading. To cluster the participants, a general linear model (GLM) with time (pre and post vocal loading task) as the dependent variable and covariates as the five acoustic parameters was fit for each participant. Participants with both a significant model and at least one significant change in acoustic feature within the model were grouped into a “voice change” group and the others into a “no voice change” group.

Following both the VER and acoustics clustering, four groups were formed from the cross-section of these groups. The vocal performance of the groups was compared pre- and post- vocal loading using independent-samples t-tests for each variable.

Table 3.2 Features used for VER clustering

Feature	Formula
VER Linear Slope	Slope of the linear fit of the six vocal effort ratings during the vocal loading task
VER Linear Fit	The goodness of fit coefficient of the linear fit of the six vocal effort ratings during the vocal loading task
VER Linear Offset	The intercept value of the linear fit of the six vocal effort ratings during the vocal loading task
Noise Load Response	$VER_5 - VER_0$; Difference between vocal effort level after five minutes of vocal loading and the vocal effort level prior to the loading task
Temporal Load Response	$VER_{30} - VER_5$; Difference between vocal effort level after thirty minutes of vocal loading and the vocal effort level after five minutes of vocal loading
Noise Recovery Response	$VER_{post} - VER_{30}$; Difference between vocal effort level after vocal loading task but no background noise and the vocal effort level after thirty minutes of vocal loading
VER Difference	$VER_{post} - VER_0$; Difference between vocal effort level after vocal loading task but no background noise and the vocal effort level prior to the loading task
VER Maximum	max $(VER_5, \dots, VER_{30})$; Maximum vocal effort rating during the vocal loading test
VER Minimum	min $(VER_5, \dots, VER_{30})$; Minimum vocal effort rating during the vocal loading test
VER Range	$VER_{max} - VER_{min}$; Difference between the maximum and minimum vocal effort levels during the vocal loading task

3.3.f Distribution for Collaborative Work

As mentioned in Chapter 2, one limitation in the area of vocal fatigue is the inability to compare across studies. While vocal loading tasks (VLT) are the most commonly used, they vary widely in their methodological approach. Experiment 3 provides an experimental protocol that can be easily reproduced and appropriately varied to support collaborative work in vocal fatigue. Towards this end, the code and relevant examples and explanations of the experimental protocol are available through GitHub.

CHAPTER IV: RESULTS

This chapter provides the results of the three outlined experiments. The results of each experiment are presented separately. Then, for each experiment is included a table of abbreviations and acronyms used for that section, the demographics of the population studied, and the results of the statistical analyses outlined in Chapter III.

4.1 Experiment 1

Table 4.1 Abbreviations for results of experiment 1

Abbreviation	Meaning
F0	Fundamental frequency
F0sd	Standard deviation of fundamental frequency
SL	Speech level
SLsd	Standard deviation of speech level
CPPS	Smoothed cepstral peak prominence
VEL	Vocal effort level
VEL02	Vocal effort level of 2
VEL13	Vocal effort level of 13
VEL25	Vocal effort level of 25
VEL50	Vocal effort level of 50
ST	Semitones
dB	Decibels
R	Pearson's R from a linear regression

4.1.a Demographics

Twenty-two participants consented and completed Experiment 1. One participant was not included due to equipment failure resulting in loss of data. Another participant was excluded due to failure to properly complete the instructions of the protocol. The remaining 20 participants, 10 males and 10 females, were included in these analyses presented below. All of these participants were within normal hearing limits. The participants were all college age and most received course credit as compensation for their participation.

4.1.b Results

The goal of this experiment was to show how vocal performance related with subjective vocal effort from the Borg CR-100 scale. Five acoustic parameters were selected to measure vocal performance, mean fundamental frequency (F0), standard deviation of fundamental frequency (F0sd), speech level (SL), standard deviation of speech level (SLsd), and smoothed cepstral peak prominence (CPPS). There were four vocal effort levels (2, minimal vocal effort; 13, slight vocal effort; 25 moderate vocal effort; 50 severe vocal effort) that the acoustic measures were compared across. Analysis of variance (ANOVA) tests were used to compare the mean speech acoustic measures across the four vocal effort levels. Post hoc Tukey's HSD tests were used to investigate the pair-wise differences across the four vocal effort levels. If equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used. The results are separated by acoustic measure.

Fundamental Frequency

The measured F0 met the assumptions for normality and independence. However, equal variance cannot be assumed. For each vocal effort level (VEL) the F0 (M; SD) is as follows (summarized in Table 4.2 with 95% confidence interval, minimum, and maximum): VEL02; minimal vocal effort (M = -0.09 ST; SD = 1.44 ST), VEL13; slight vocal effort (M = 0.43 ST; SD = 1.29 ST), VEL25; moderate vocal effort (M = 1.32 ST; SD = 1.72), and VEL50; severe vocal effort (M = 2.93 ST; SD = 1.93). Figure 4.1 shows boxplots of the F0 across the vocal effort levels.

Table 4.2 Descriptive statistics for F0 and VEL in experiment 1

Fundamental Frequency (ST)

VEL	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
2	-0.09	1.44	-0.31	0.14	-4.16	5.25
13	0.43	1.29	0.24	0.62	-3.01	5.34
25	1.32	1.72	1.07	1.57	-2.92	6.11
50	2.93	1.93	2.62	3.24	-1.33	6.56

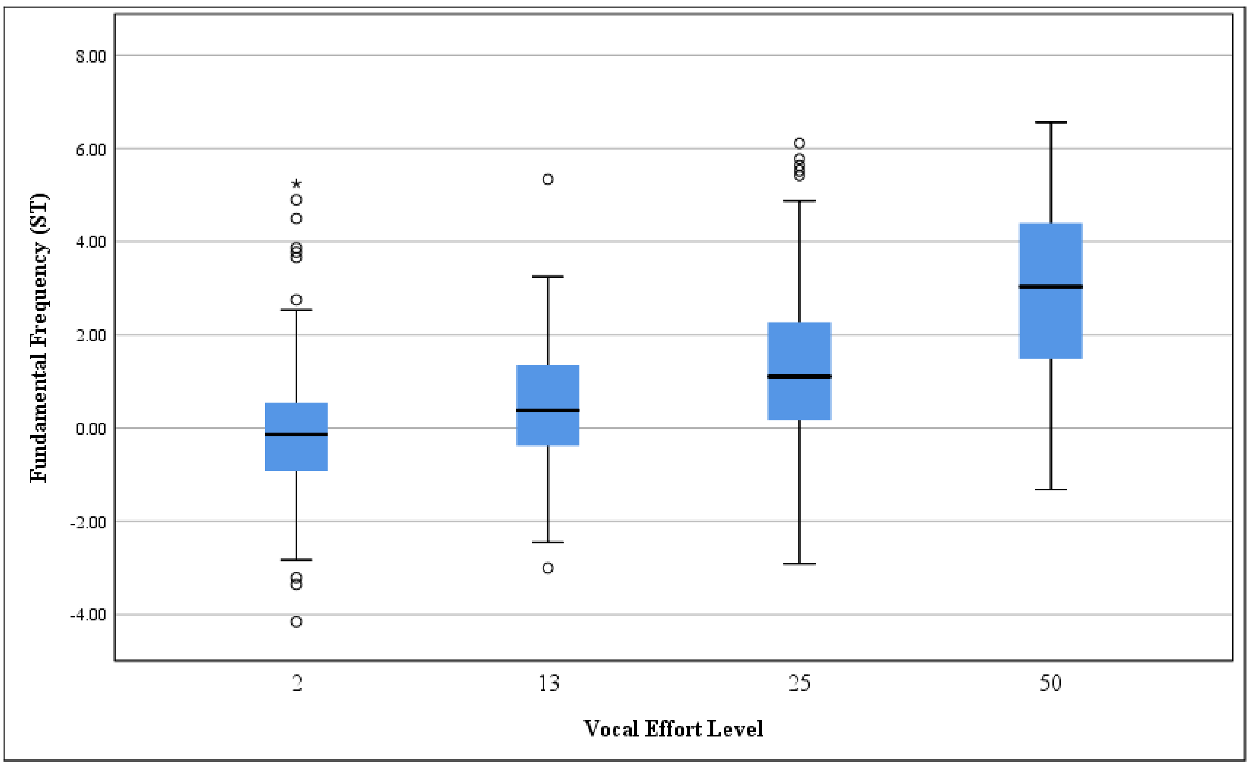

Figure 4.1 Boxplot of F0 and VEL for experiment 1

There was a significant ($p < 0.001$) main effect of F0 on VEL. The post hoc tests show that F0 was different across each VEL (summarized in Table 4.3). There was an increase ($p = 0.004$) in F0 of 0.52 ST from VER02 to VER13, an increase ($p < 0.001$) of 1.41 ST from VER13 to VER25, and an increase ($p < 0.001$) of 3.02 ST from VER25 to VER50.

Table 4.3 Multiple comparison statistics for F0 and VEL for experiment 1

(I) VEL	(J) VEL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
2	13	-0.52	0.15	0.004	-0.91	-0.12
	25	-1.41	0.17	< 0.001	-1.86	-0.96
	50	-3.02	0.19	< 0.001	-3.54	-2.51
13	2	0.52	0.15	0.004	0.12	0.91
	25	-0.89	0.16	< 0.001	-1.32	-0.47
	50	-2.51	0.18	< 0.001	-3.00	-2.02
25	2	1.41	0.17	< 0.001	0.96	1.86
	13	0.89	0.16	< 0.001	0.47	1.32
	50	-1.61	0.20	< 0.001	-2.15	-1.08
50	2	3.02	0.19	< 0.001	2.51	3.54
	13	2.51	0.18	< 0.001	2.02	3.00
	25	1.61	0.20	< 0.001	1.08	2.15

Fundamental Frequency Standard Deviation

The measured F0sd met the assumptions for normality and independence. However, equal variance cannot be assumed. For each VEL the F0sd (M; SD) is as follows (summarized in Table 4.4 with 95% confidence interval, minimum, and maximum): VEL02; minimal vocal effort (M = 1.83 ST; SD = 0.83 ST), VEL13; slight vocal effort (M = 2.13 ST; SD = 1.02 ST), VEL13; moderate vocal effort(M = 2.03 ST; SD = 0.89 ST), VEL50; severe vocal effort (M = 2.16 ST; SD = 0.82 ST). Figure 4.2 shows boxplots of the F0sd across the vocal effort levels.

Table 4.4 Descriptive statistics for F0sd and VEL for experiment 1

Fundamental Frequency Standard Deviation (ST)

VEL	Mean	Std. Deviation	95% Confidence Interval for Mean: Lower Bound	95% Confidence Interval for Mean: Upper Bound	Minimum	Maximum
2	1.83	0.83	1.70	1.96	0.74	4.91
13	2.13	1.02	1.98	2.29	0.69	4.88
25	2.03	0.89	1.89	2.17	0.76	4.98
50	2.16	0.82	2.03	2.29	0.93	4.82

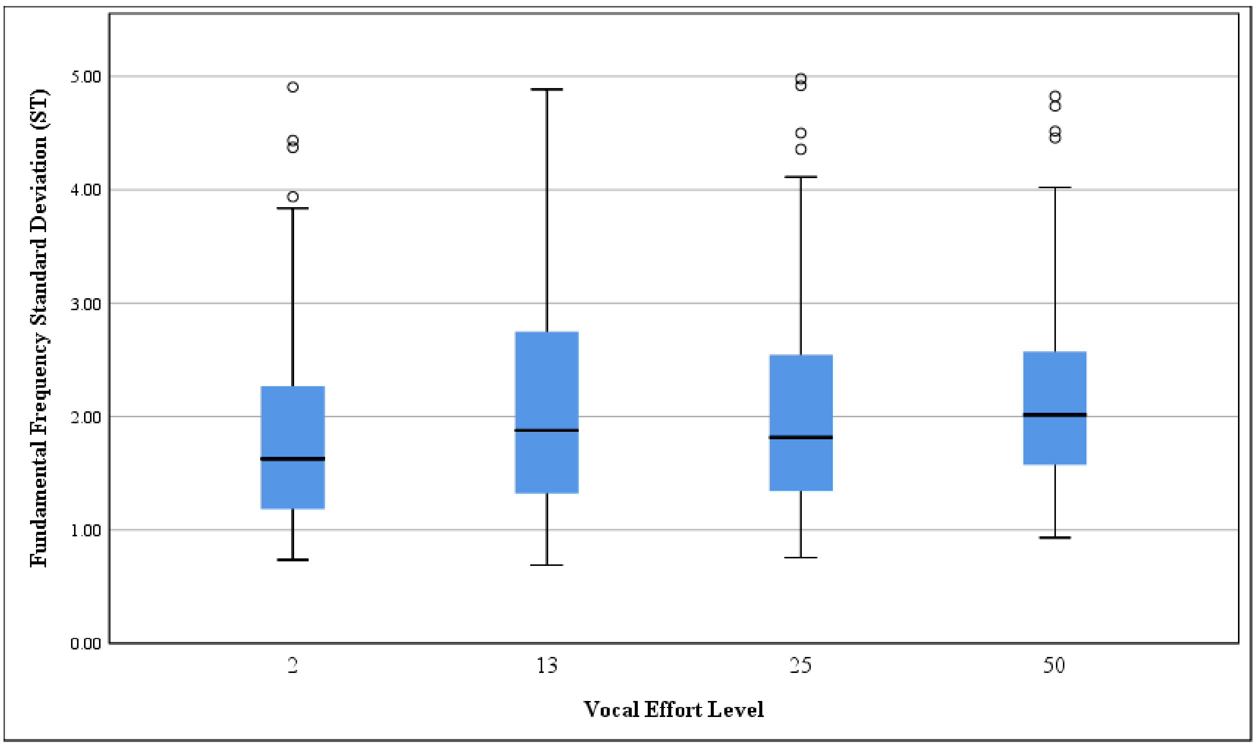

Figure 4.2 Boxplot of F0sd and VEL for experiment 1

There was a significant ($p < 0.001$) main effect of F0sd on VEL. The post hoc tests results are summarized in Table 4.5. There was an increase ($p = 0.019$) in F0sd of 0.31 ST from VER02 to VER13 and an increase ($p = 0.002$) of 0.34 ST from VER02 to VER50.

Table 4.5 Multiple comparison statistics for F0sd and VEL for experiment 1

(I) VEL	(J) VEL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2	13	-0.31	0.10	0.019	-0.58	-0.03
	25	-0.20	0.10	0.208	-0.46	0.05
	50	-0.34	0.09	0.002	-0.58	-0.09
13	2	0.31	0.10	0.019	0.03	0.58
	25	0.11	0.11	0.902	-0.18	0.39
	50	-0.03	0.10	1.000	-0.30	0.24
25	2	0.20	0.10	0.208	-0.05	0.46
	13	-0.11	0.11	0.902	-0.39	0.18
	50	-0.13	0.10	0.654	-0.39	0.12
50	2	0.34	0.09	0.002	0.09	0.58
	13	0.03	0.10	1.000	-0.24	0.30
	25	0.13	0.10	0.654	-0.12	0.39

Speech Level

The measured SL met the assumptions for normality, independence, and equal variance. For each VEL the SL (M; SD) is as follows (summarized in Table 4.6 with 95% confidence interval, minimum, and maximum): VEL02; minimal vocal effort (M = 52.39 dB; SD = 5.46 dB), VEL13; slight vocal effort (M = 56.95 dB; SD = 6.28 dB), VEL25; moderate vocal effort (M = 60.54 dB; SD = 6.94 dB), VEL50; severe vocal effort (M = 66.89 dB; SD = 6.89 dB). Figure 4.3 shows boxplots of the SL across the VELs.

Table 4.6 Descriptive statistics for speech level and VEL for experiment 1

Speech Level (dB)

VEL	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
2	52.39	5.46	51.56	53.23	39.87	63.32
13	56.94	6.28	56.02	57.88	45.35	72.13
25	60.53	6.94	59.52	61.56	47.63	79.50
50	66.89	6.89	65.87	67.91	51.01	82.52

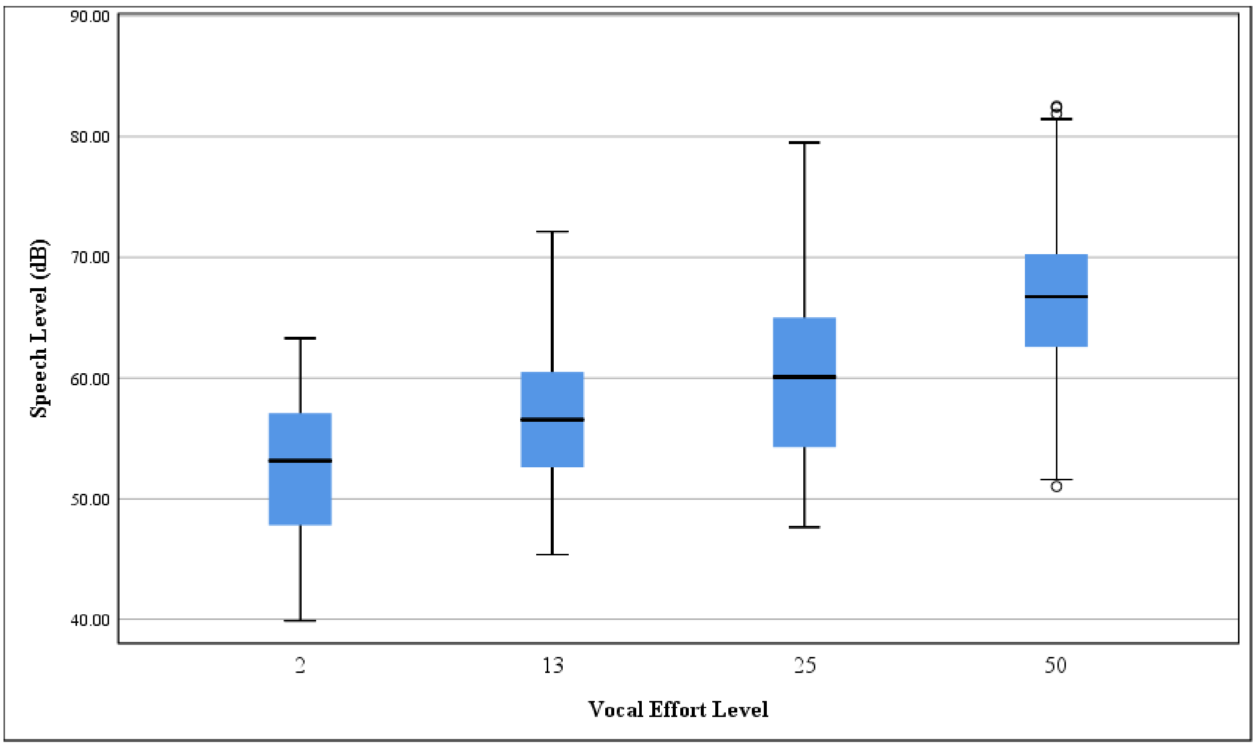

Figure 4.3 Boxplot for SL and VEL for experiment 1

There was a significant (p < 0.001) main effect of SL on VEL. The post hoc tests show that SL was different across each VEL (summarized in Table 4.7). There was an increase (p < 0.001) in SL of 4.56 dB from VEL02 to VEL13, an increase (p < 0.001) of 3.59 dB from VEL13 to VEL25, and an increase (p < 0.001) of 6.35 dB from VEL25 to VEL50.

Table 4.7 Multiple comparison statistics for SL and VEL for experiment 1

		Mean			95% Confidence Interval	
(I) VEL	(J) VEL	Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
2	13	-4.56	0.69	< 0.001	-6.34	-2.77
	25	-8.14	0.69	< 0.001	-9.93	-6.36
	50	-14.50	0.70	< 0.001	-16.29	-12.70
13	2	4.56	0.69	< 0.001	2.77	6.34
	25	-3.59	0.68	< 0.001	-5.34	-1.84
	50	-9.94	0.68	< 0.001	-11.70	-8.18
25	2	8.14	0.69	< 0.001	6.36	9.93
	13	3.59	0.68	< 0.001	1.84	5.34
	50	-6.35	0.68	< 0.001	-8.11	-4.60
50	2	14.50	0.70	< 0.001	12.70	16.29
	13	9.94	0.68	< 0.001	8.18	11.70
	25	6.35	0.68	< 0.001	4.60	8.11

Standard Deviation of Speech Level

The measured SLsd met the assumptions for normality, independence, and equal variance. For each VEL the SLsd (M; SD) is as follows (summarized in Table 4.8 with 95% confidence interval, minimum, and maximum): VEL02; minimal vocal effort (M = 3.43 dB; SD = 0.44 dB), VEL13; slight vocal effort (M = 3.47 dB; SD = 0.4 dB), VEL13; moderate vocal effort(M = 3.51 dB; SD = 0.44 dB), VEL50; severe vocal effort (M = 3.64 dB; SD = 0.39 dB). Figure 4.4 shows boxplots of the SLsd across the VELs.

Table 4.8 Descriptive statistics for SLsd and VEL for experiment 1

Speech Level Standard Deviation (dB)

VEL	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
2	3.43	0.44	3.36	3.49	2.46	4.67
13	3.47	0.40	3.41	3.53	2.60	4.66
25	3.51	0.44	3.44	3.57	2.52	4.63
50	3.64	0.39	3.58	3.69	2.68	4.55

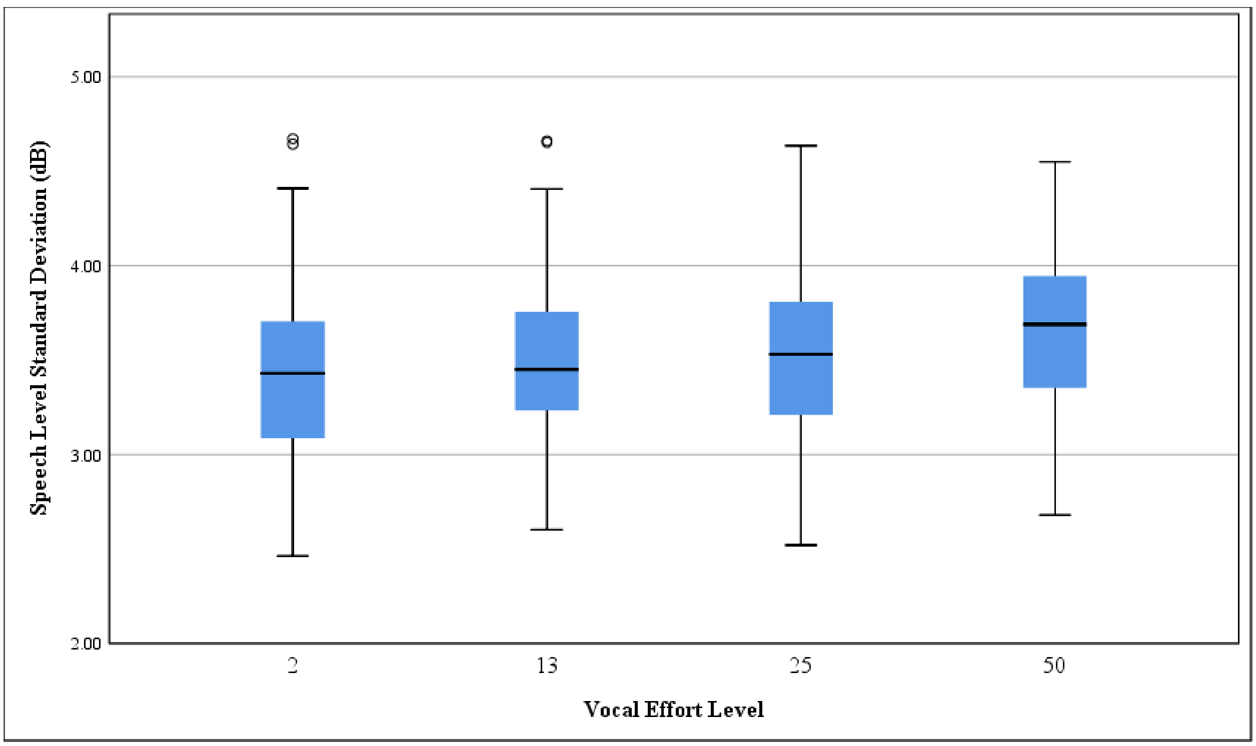

Figure 4.4 Boxplot for SLsd and VEL for experiment 1

There was a significant ($p < 0.001$) main effect of SLsd on VEL. The post hoc tests show that SLsd was only different in the pair-wise comparisons with VEL50 (summarized in Table 4.9). There was an increase ($p < 0.001$) in SLsd of 0.21 dB from VEL02 to VEL50, an increase (p =

0.001) of 0.17 dB from VEL13 to VEL50, and an increase (p =0.021) of 0.13 dB from VEL25 to VEL50.

Table 4.9 Multiple comparison statistics for SLsd and VEL for experiment 1

(I) VEL	(J) VEL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
2	13	-0.04	0.05	0.767	-0.16	0.07
	25	-0.08	0.05	0.256	-0.20	0.03
	50	-0.21	0.05	< 0.001	-0.33	-0.09
13	2	0.04	0.05	0.767	-0.07	0.16
	25	-0.04	0.04	0.812	-0.15	0.08
	50	-0.17	0.04	0.001	-0.28	-0.05
25	2	0.08	0.05	0.256	-0.03	0.20
	13	0.04	0.04	0.812	-0.08	0.15
	50	-0.13	0.04	0.021	-0.24	-0.01
50	2	0.21	0.05	< 0.001	0.09	0.33
	13	0.17	0.04	0.001	0.05	0.28
	25	0.13	0.04	0.021	0.01	0.24

Smoothed Cepstral Peak Prominence

The measured CPPS met the assumptions for normality and independence. However, equal variance cannot be assumed. For each VEL the CPPS (M; SD) is as follows (summarized in Table 4.10 with 95% confidence interval, minimum, and maximum): VEL02; minimal vocal effort (M = 11.99 dB; SD = 2.45 dB), VEL13; slight vocal effort (M = 13.35 dB; SD = 2.44 dB), VEL13; moderate vocal effort(M = 14.56 dB; SD = 2.05 dB), VEL50; severe vocal effort (M = 16.04 dB; SD = 1.53 dB). Figure 4.5 shows boxplots of the CPPS across the VELs.

Table 4.10 Descriptive statistics for CPPS and VEL for experiment 1

Smoothed Cepstral Peak Prominence

VEL	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
2	11.99	2.45	11.59	12.40	7.07	17.09
13	13.35	2.44	12.99	13.71	7.36	18.71
25	14.56	2.05	14.25	14.86	7.65	19.31
50	16.04	1.53	15.82	16.27	12.31	19.18

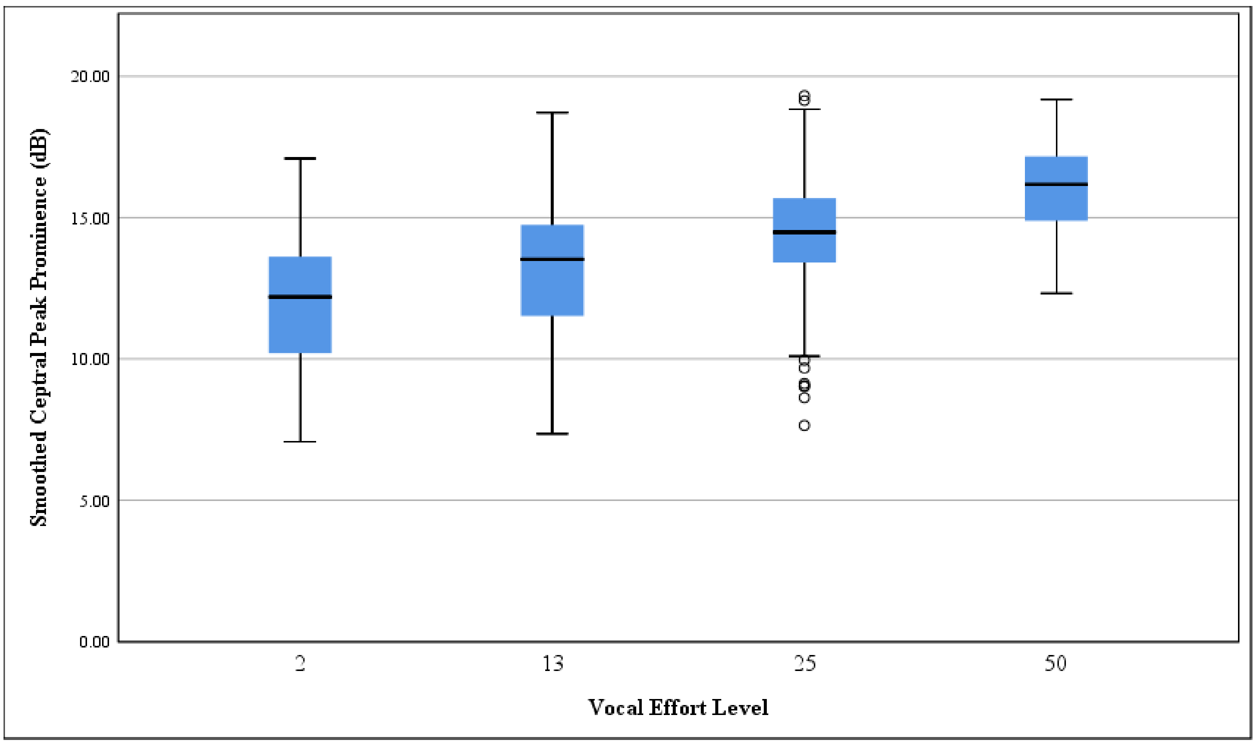

Figure 4.5 Boxplot for CPPS and VEL for experiment 1

There was a significant ($p < 0.001$) main effect of CPPS on VEL. The post hoc tests show that CPPS was different across each VEL (summarized in Table 4.11). There was an increase ($p < 0.001$) in CPPS of 1.36 dB from VEL02 to VEL13, an increase ($p < 0.001$) of 2.56 dB from VEL13 to VEL25, and an increase ($p < 0.001$) of 4.05 dB from VEL25 to VEL50.

Table 4.11 Multiple comparison statistics for CPPS and VEL for experiment 1

(I) VEL	(J) VEL	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
2	13	-1.36	0.28	< 0.001	-2.09	-0.63
	25	-2.56	0.26	< 0.001	-3.24	-1.89
	50	-4.05	0.23	< 0.001	-4.67	-3.43
13	2	1.36	0.28	< 0.001	0.63	2.09
	25	-1.21	0.24	< 0.001	-1.84	-0.57
	50	-2.69	0.22	< 0.001	-3.27	-2.12
25	2	2.56	0.26	< 0.001	1.89	3.24
	13	1.21	0.24	< 0.001	0.57	1.84
	50	-1.49	0.19	< 0.001	-1.99	-0.98
50	2	4.05	0.23	< 0.001	3.43	4.67
	13	2.69	0.22	< 0.001	2.12	3.27
	25	1.49	0.19	< 0.001	0.98	1.99

Test-Retest Reliability of Speech Level

Each participant had significant regression ($p < 0.001$) for SL across the VELs. Descriptive statistics were calculated from the Fischer's z transformed Pearson's r coefficients (R). The mean R was 0.90 (SD = 0.30). Table 4.12 below shows the descriptive statistics for R.

Table 4.12 Descriptive statistics for Pearson's R for linear regression of SL and VEL for experiment 1

	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
R	0.90	0.30	0.87	0.93	0.77	0.96

4.1 Experiment 2

Table 4.13 Abbreviations for results of experiment 2

Abbreviation	Meaning
F0	Fundamental frequency
F0sd	Standard deviation of fundamental frequency
L	Speech level
Lsd	Standard deviation of speech level
CPPS	Smoothed cepstral peak prominence
VER	Vocal effort rating
ST	Semitones
dB	Decibels
dBA	A-weighted decibels
D01	Vocal load of communication distance at 1 meter
D02	Vocal load of communication distance at 2 meters
D04	Vocal load of communication distance at 4 meters
L54	Vocal load of loudness target of 54 dB
L60	Vocal load of loudness target of 60 dB
L66	Vocal load of loudness target of 66 dB
N53	Vocal load of background noise of 53 dBA
N62	Vocal load of background noise of 62 dBA
N71	Vocal load of background noise of 71 dBA

4.2.a Demographics

Forty-eight participants consented for the study. Eleven participants were not included in the analyses. Five were not included for not being native English speakers. One was not included for not passing the hearing screening (HL > 20 dB for 200 Hz and 400 Hz in left ear). The other five participants were not included for either not finishing protocol or not properly following the instructions. A total of 37 participants, 19 males and 18 females, were included in the analyses presented below. All of these participants were within normal hearing limits. The participants were all college age and most received course credit as compensation for their participation.

4.2.b Results

The purpose of this experiment was to show how vocal performance and subjective vocal effort changed with different types of vocal loads. Vocal effort rating (VER) was measured using the Borg CR-100 following each speech task. Five acoustic parameters were selected to measure vocal performance, mean fundamental frequency (F0), standard deviation of fundamental frequency (F0sd), speech level (L), standard deviation of speech level (Lsd), and smoothed cepstral peak prominence (CPPS). Nine vocal load conditions were used including communication distances of 1 meter (D01), 2 meters (D02), and 4 meters (D04), loudness goals of 54 dB (L54), 60 dB (L60), and 66 dB (L66), and background noise levels of 53 dBA (N53), 62 dBA (N62), and 71 dBA (N71). Analysis of variance (ANOVA) tests were used to compare the mean VEL and speech acoustic measures across the variations of each vocal load for each with D01 as a baseline of no load for each. Post hoc Tukey's HSD tests were used to investigate the pair-wise differences across the experimental conditions. If equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used. The results are separated by acoustic measure and type of vocal load.

Vocal Effort Rating and Communication Distance

The measured VER met the assumptions for normality and independence. However, equal variance cannot be assumed. For each vocal load the VER (M; SD) is as follows (summarized in Table 4.14 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = 25.13 ; SD = 16.43), D02; communication distance at 2 meters (M = 28.23 ; SD = 14.67), D04; communication distance at 4 meters (M = 34.59 ; SD = 18.16), Figure 4.6 shows boxplots of the VER across the vocal loads.

Table 4.14 Descriptive statistics for VER and communication distance vocal load for experiment 2

Vocal Effort Rating

Load	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
D01	25.1	16.4	22.0	28.3	2	65
D02	28.2	14.7	25.4	31.0	2	65
D04	34.6	18.2	31.1	38.1	2	75

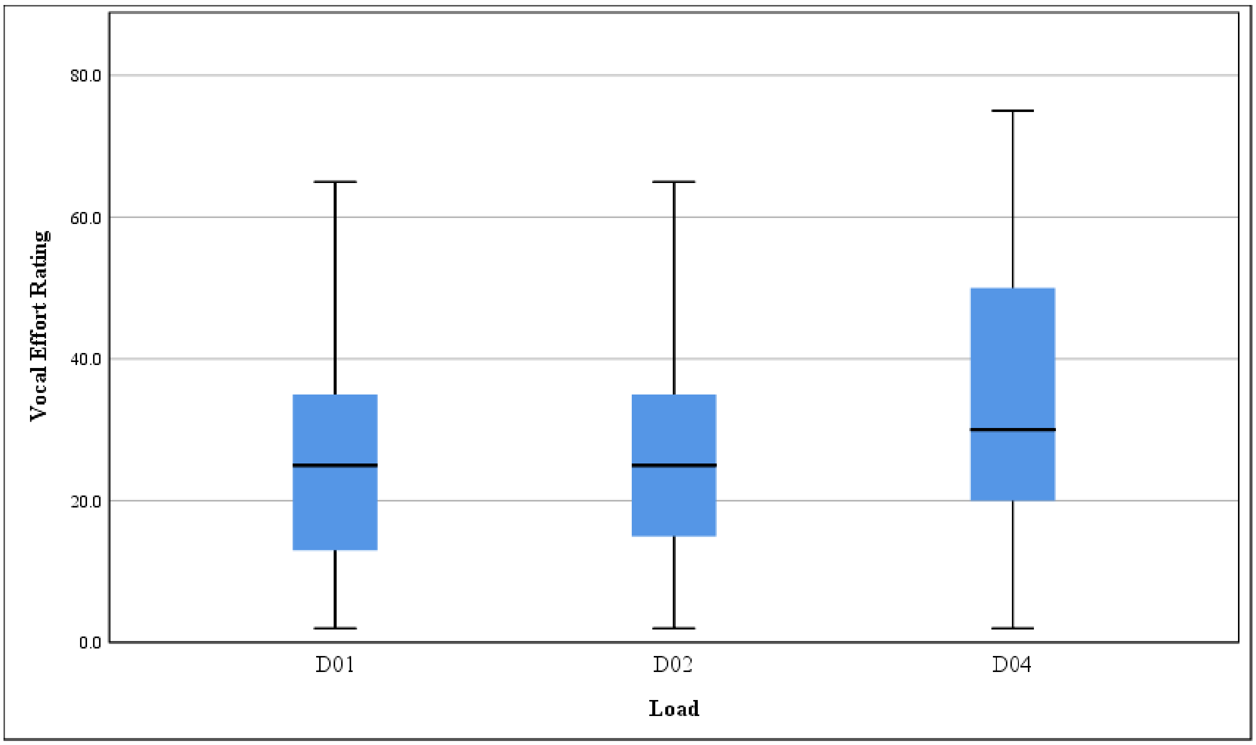

Figure 4.6 Boxplot for VER and communication distance load for experiment 2

There was a significant ($p < 0.001$) main effect of VER on the vocal load of communication distance. The post hoc tests show that VER was only different in the pair-wise comparisons with D04 (summarized in Table 4.15). There was an increase ($p < 0.001$) in VER of 9.5 from D01 to D04 and an increase ($p = 0.017$) of 6.4 from D02 to D04.

Table 4.15 Multiple comparison statistics for VER and communication distance vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D01	D02	-3.1	2.1	0.378	-8.2	2.0
	D04	-9.5	2.4	< 0.001	-15.2	-3.7
D02	D01	3.1	2.1	0.378	-2.0	8.2
	D04	-6.4	2.3	0.017	-11.8	-0.9
D04	D01	9.5	2.4	< 0.001	3.7	15.2
	D02	6.4	2.3	0.017	0.9	11.8

Fundamental Frequency and Communication Distance

The measured F0 met the assumptions for normality, independence, and equal variance. For each vocal load the F0 (M; SD) is as follows (summarized in Table 4.16 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = -0.02 Hz; SD = 0.62 Hz), D02; communication distance at 2 meters (M = 0.37 Hz; SD = 0.74 Hz), D04; communication distance at 4 meters (M = 0.75 Hz; SD = 0.82 Hz), Figure 4.7 shows boxplots of the F0 across the vocal loads.

Table 4.16 Descriptive statistics for F0 and communication distance vocal load for experiment 2

Fundamental Frequency (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	-0.02	0.62	-0.14	0.10	-1.76	2.09
D02	0.37	0.74	0.22	0.51	-1.33	2.25
D04	0.75	0.82	0.58	0.91	-1.75	2.53

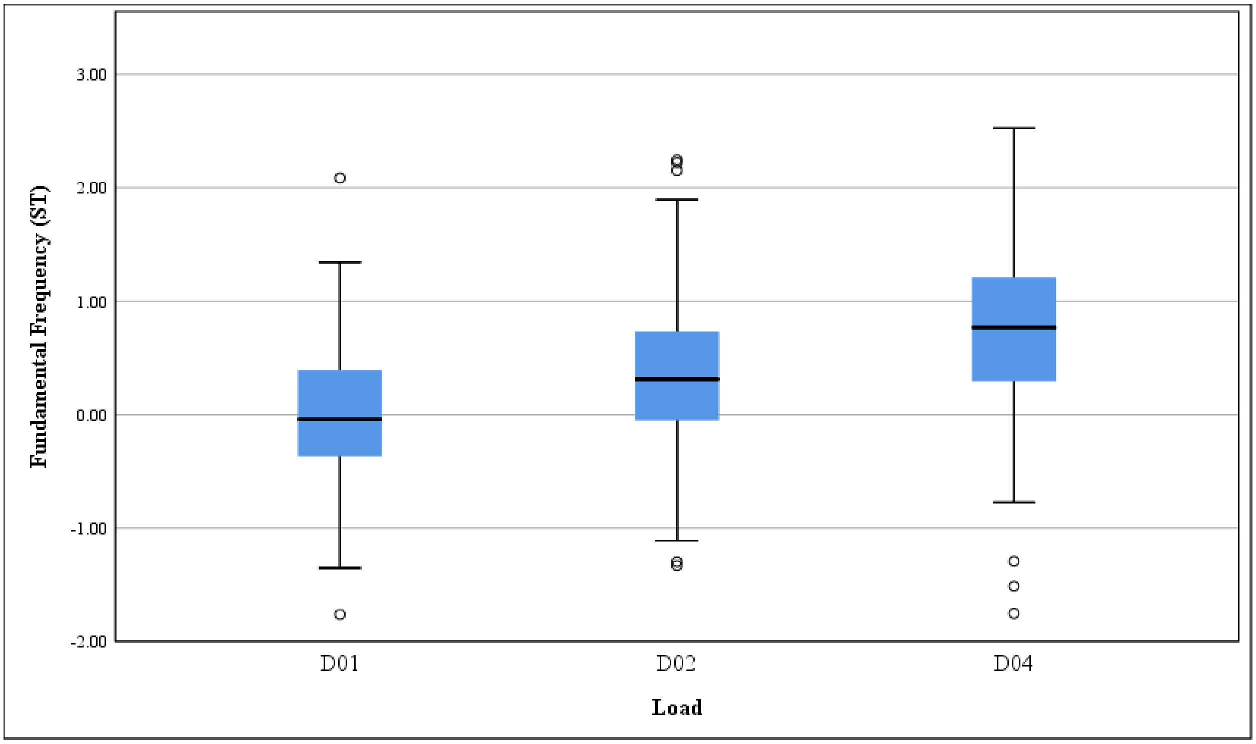

Figure 4.7 Boxplot for F0 and communication distance load for experiment 2

There was a significant ($p < 0.001$) main effect of F0 on the vocal load of communication distance. The post hoc tests show that F0 was only different in the pair-wise comparisons with D04 (summarized in Table 4.17). There was an increase ($p < 0.001$) in F0 of 0.77 ST from D01 to D04 and an increase ($p = 0.001$) of 0.38 ST from D02 to D04.

Table 4.17 Multiple comparison statistics and F0 and communication distance vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (ST)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D01	D02	-0.39	0.10	< 0.001	-0.62	-0.15
	D04	-0.77	0.10	< 0.001	-1.01	-0.53
D02	D01	0.39	0.10	< 0.001	0.15	0.62
	D04	-0.38	0.10	0.001	-0.62	-0.14
D04	D01	0.77	0.10	< 0.001	0.53	1.01
	D02	0.38	0.10	0.001	0.14	0.62

Fundamental Frequency Standard Deviation and Communication Distance

The measured F0sd met the assumptions for normality, independence, and equal variance. For each vocal load the F0sd (M; SD) is as follows (summarized in Table 4.18 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = 2.04 Hz; SD = 0.83 Hz), D02; communication distance at 2 meters (M = 2.23 Hz; SD = 0.92 Hz), D04; communication distance at 4 meters (M = 2.21 Hz; SD = 0.77 Hz), Figure 4.8 shows boxplots of the F0sd across the vocal loads.

Table 4.18 Descriptive statistics for F0sd and communication distance vocal load for experiment 2

Fundamental Frequency Standard Deviation (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	2.04	0.83	1.87	2.20	0.73	4.52
D02	2.23	0.92	2.05	2.41	0.20	4.65
D04	2.21	0.77	2.06	2.37	0.89	4.47

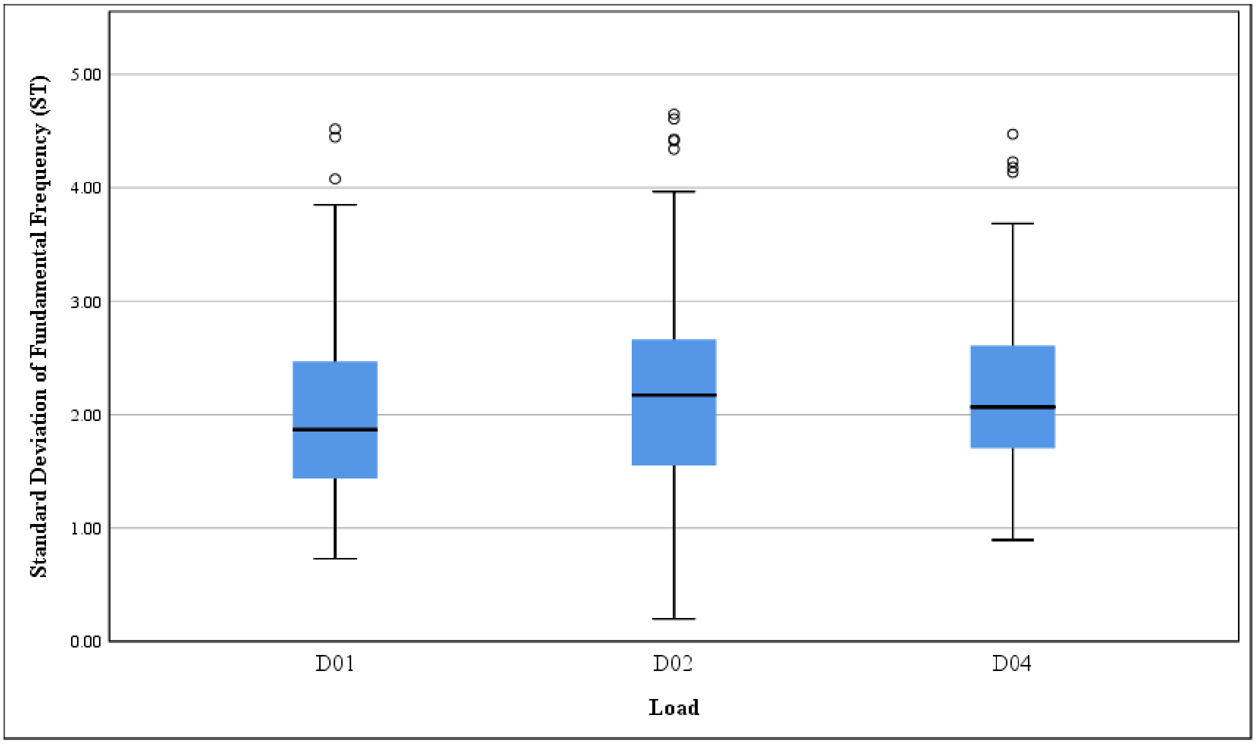

Figure 4.8 Boxplot for F0sd and communication distance load for experiment 2

There was no significant main effect of F0sd on the vocal load of communication distance.

Speech Level and Communication Distance

The measured SL met the assumptions for normality, independence, and equal variance. For each vocal load the SL (M; SD) is as follows (summarized in Table 4.19 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = 65.54 dB; SD = 0.33 dB), D02; communication distance at 2 meters (M = 66.86 dB; SD = 0.33 dB), D04; communication distance at 4 meters (M = 68.36 dB; SD = 0.36 dB), Figure 4.9 shows boxplots of the SL across the vocal loads.

Table 4.19 Descriptive statistics for SL and communication distance vocal load for experiment 2

Speech Level (dB)

			95% Confidence Interval for Mean			
Load	Mean	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
D01	65.54	0.33	64.89	66.19	57.29	74.08
D02	66.86	0.33	66.20	67.52	58.67	73.32
D04	68.36	0.36	67.64	69.07	57.83	76.51

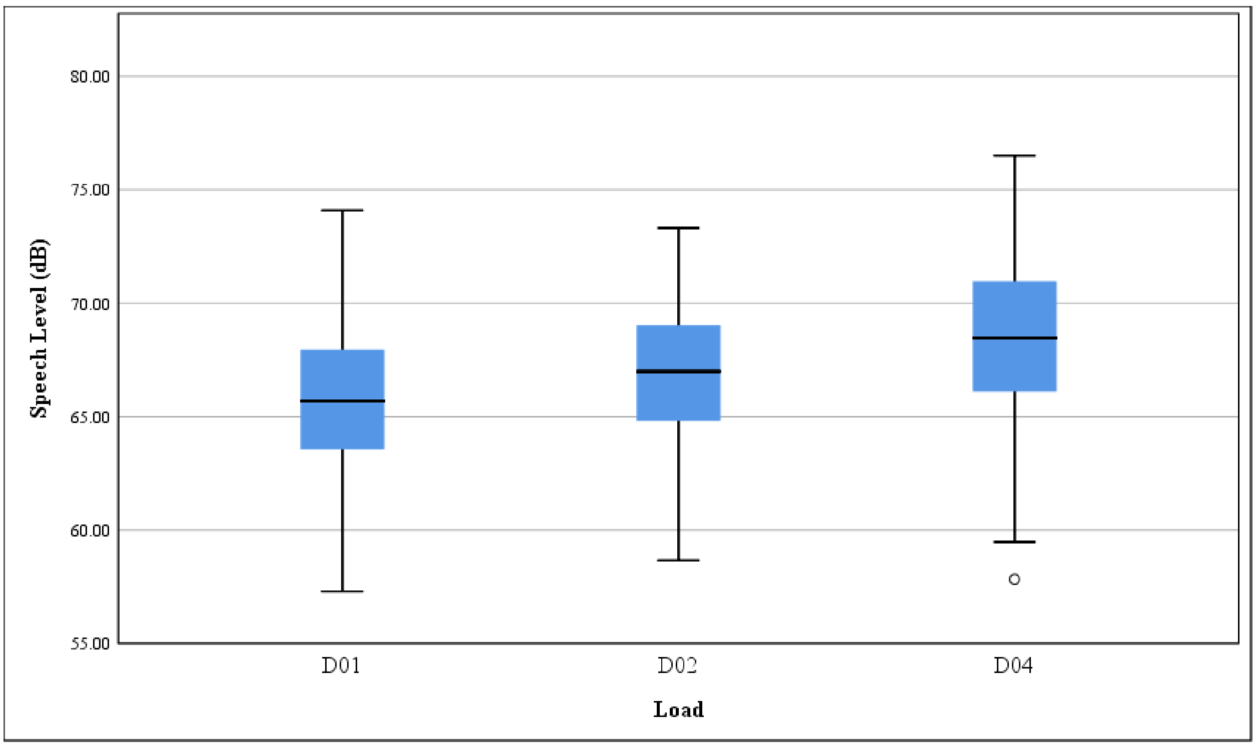

Figure 4.9 Boxplot for SL and communication distance load for experiment 2

There was a significant ($p < 0.001$) main effect of SL on the vocal load of communication distance. The post hoc tests show that SL was different across each vocal load (summarized in Table 4.20). There was an increase ($p = 0.018$) in SL of 1.32 dB from D01 to D02 and an increase ($p = 0.006$) of 1.49 dB from D02 to D04.

Table 4.20 Multiple comparison statistics for SL and communication distance vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D01	D02	-1.32	0.48	0.018	-2.45	-0.19
	D04	-2.81	0.48	< 0.001	-3.95	-1.68
D02	D01	1.32	0.48	0.018	0.19	2.45
	D04	-1.49	0.48	0.006	-2.63	-0.36
D04	D01	2.81	0.48	< 0.001	1.68	3.95
	D02	1.49	0.48	0.006	0.36	2.63

Speech Level Standard Deviation and Communication Distance

The measured SLsd met the assumptions for normality, independence, and equal variance. For each vocal load the SLsd (M; SD) is as follows (summarized in Table 4.21 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = 3.4 dB; SD = 0.28 dB), D02; communication distance at 2 meters (M = 3.43 dB; SD = 0.28 dB), D04; communication distance at 4 meters (M = 3.47 dB; SD = 0.27 dB), Figure 4.10 shows boxplots of the SLsd across the vocal loads.

Table 4.21 Descriptive statistics for SLsd and communication distance vocal load for experiment 2

Speech Level Standard Deviation (dB)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	3.40	0.28	3.35	3.45	2.76	4.14
D02	3.43	0.28	3.37	3.48	2.84	4.14
D04	3.47	0.27	3.41	3.52	2.79	4.05

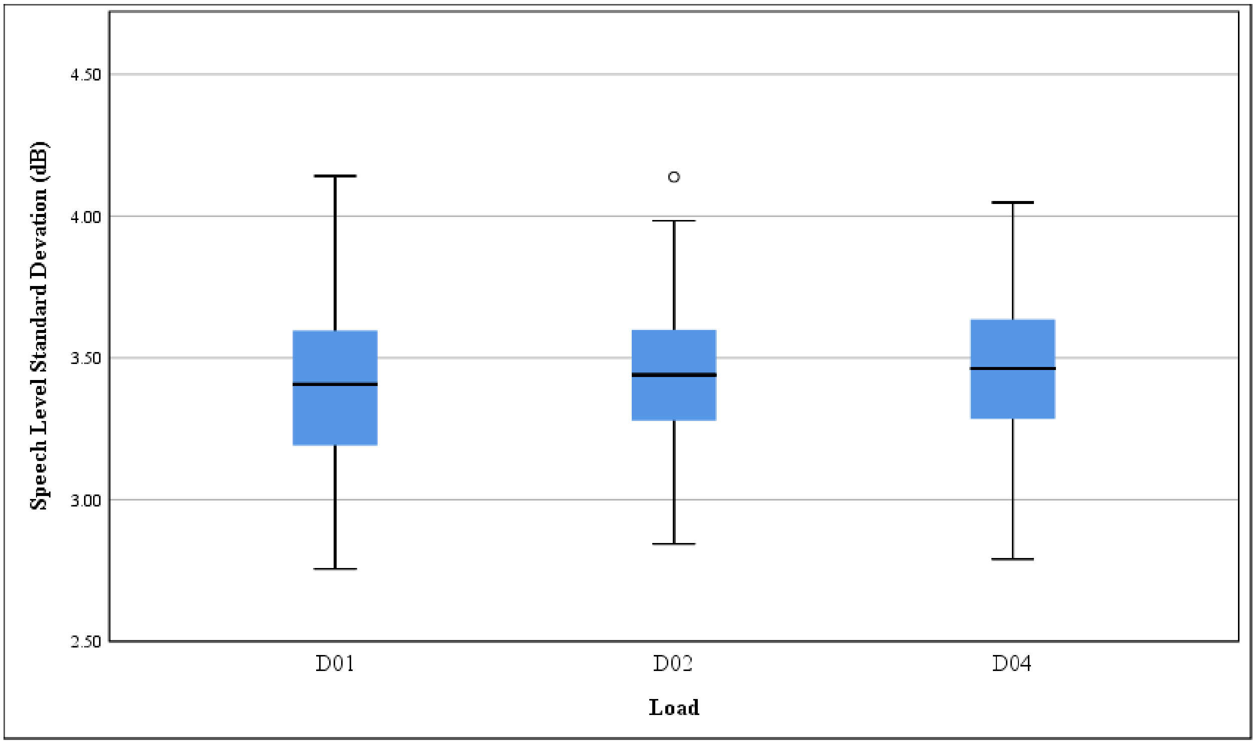

Figure 4.10 Boxplot for SLsd and communication distance load for experiment 2

There was no significant main effect of SLsd on the vocal load of communication distance.

Smoothed Cepstral Peak Prominence and Communication Distance

The measured CPPS met the assumptions for normality, independence, and equal variance. For each vocal load the CPPS (M; SD) is as follows (summarized in Table 4.22 with 95% confidence interval, minimum, and maximum): D01; communication distance at 1 meter/no vocal load (M = 15.3 dB; SD = 1.25 dB), D02; communication distance at 2 meters (M = 15.62 dB; SD = 1.25 dB), D04; communication distance at 4 meters (M = 16.08 dB; SD = 1.23 dB), Figure 4.11 shows boxplots of the CPPS across the vocal loads.

Table 4.22 Descriptive statistics for CPPS and communication distance vocal load for experiment 2

Smoothed Cepstral Peak Prominence (dB)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
D01	15.30	1.25	15.06	15.54	12.08	18.12
D02	15.62	1.25	15.38	15.86	12.06	18.26
D04	16.08	1.23	15.85	16.32	12.49	19.17

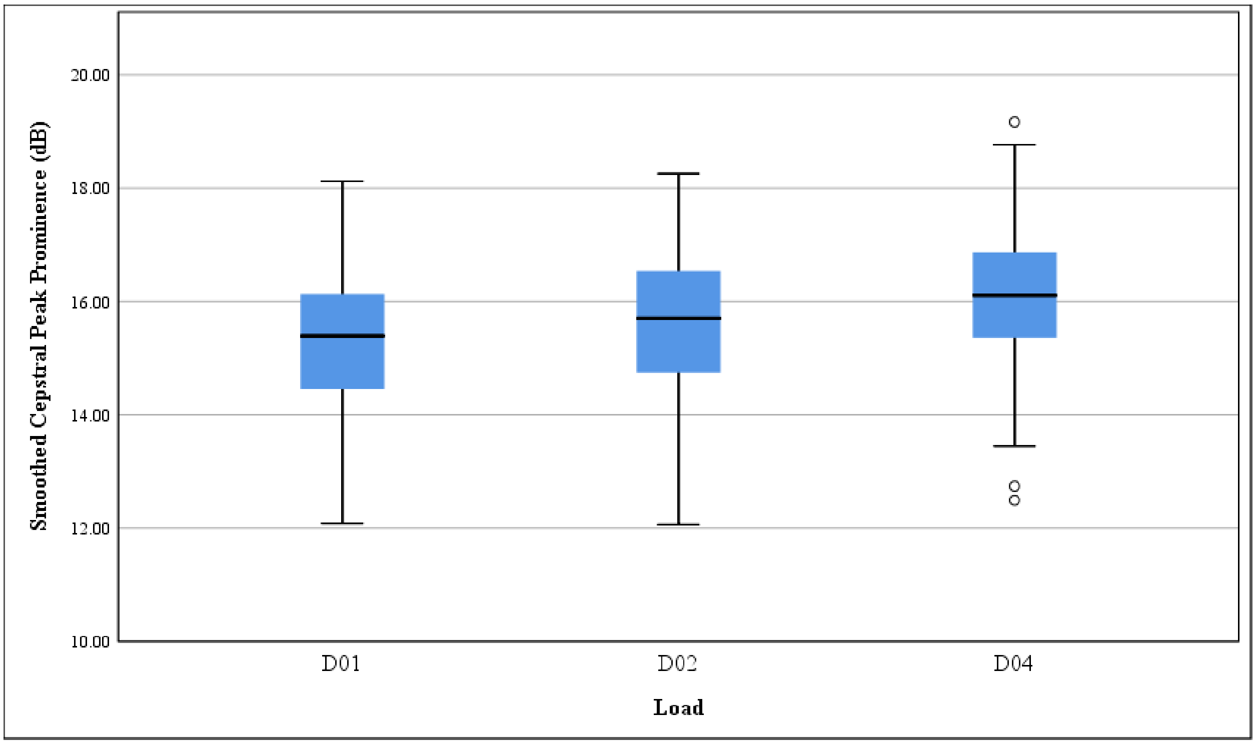

Figure 4.11 Boxplot for CPPS and communication distance load for experiment 2

There was a significant ($p < 0.001$) main effect of CPPS on the vocal load of communication distance. The post hoc tests show that CPPS was only different in the pair-wise comparisons with D04 (summarized in Table 4.23). There was an increase ($p < 0.001$) in CPPS of 0.78 dB from D01 to D04 and an increase ($p = 0.019$) of 0.47 dB from D02 to D04.

Table 4.23 Multiple comparison statistics for CPPS and communication distance vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D01	D02	-0.32	0.17	0.155	-0.72	0.09
	D04	-0.78	0.17	< 0.001	-1.19	-0.38
D02	D01	0.32	0.17	0.155	-0.09	0.72
	D04	-0.47	0.17	0.019	-0.87	-0.06
D04	D01	0.78	0.17	< 0.001	0.38	1.19
	D02	0.47	0.17	0.019	0.06	0.87

Vocal Effort Rating and Loudness Goal

The measured VER met the assumptions for normality, independence, and equal variance. For each vocal load the VER (M; SD) is as follows (summarized in Table 4.24 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 25.13; SD = 16.43), L54; loudness goal of 54 dB (M = 35.43; SD = 16.03), L60; loudness goal of 60 dB (M = 39.06; SD = 15.75), L66; loudness goal of 66 dB (M = 49.81; SD = 16.81). Figure 4.12 shows boxplots of the VER across the loads.

Table 4.24 Descriptive statistics for VER and loudness goal vocal load for experiment 2

Vocal Effort Rating

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	25.1	16.4	22.0	28.3	2	65
L54	35.4	16.0	32.4	38.5	7	80
L60	39.1	15.7	36.1	42.1	10	70
L66	49.8	16.8	46.6	53.0	3	85

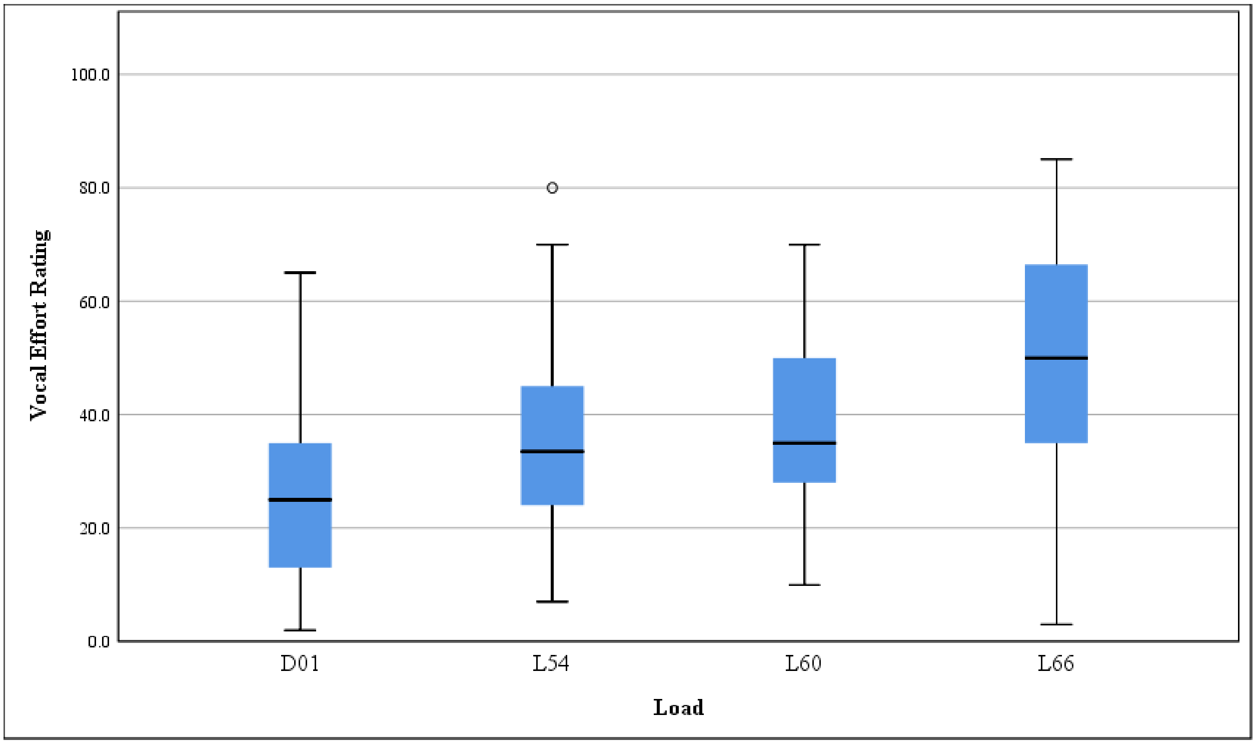

Figure 4.12 Boxplot of VER and loudness goal vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of VER on the vocal load of loudness goal. The post hoc tests show that VER was different in the pair-wise comparisons with D01 and L66 (summarized in Table 4.25). There was an increase ($p < 0.001$) in VER of 10.3 from D01 to L54 and an increase ($p < 0.001$) of 10.7 from L60 to L66.

Table 4.25 Multiple comparisons statistics for VER and loudness goal vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	L54	-10.3	2.2	< 0.001	-16.0	-4.6
	L60	-13.9	2.2	< 0.001	-19.7	-8.2
	L66	-24.7	2.2	< 0.001	-30.4	-18.9
L54	D01	10.3	2.2	< 0.001	4.6	16.0
	L60	-3.6	2.2	0.355	-9.3	2.1
	L66	-14.4	2.2	< 0.001	-20.1	-8.7
L60	D01	13.9	2.2	< 0.001	8.2	19.7
	L54	3.6	2.2	0.355	-2.1	9.3
	L66	-10.7	2.2	< 0.001	-16.5	-5.0
L66	D01	24.7	2.2	< 0.001	18.9	30.4
	L54	14.4	2.2	< 0.001	8.7	20.1
	L60	10.7	2.2	< 0.001	5.0	16.5

Fundamental Frequency and Loudness Goal

The measured F0 met the assumptions for normality and independence. However, equal variance cannot be assumed. For each vocal load the F0 (M; SD) is as follows (summarized in Table 4.26 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = -0.01 ST; SD = 0.69 ST), L54; loudness goal of 54 dB (M = 0.65 ST; SD = 0.98 ST), L60; loudness goal of 60 dB (M = 0.8 ST; SD = 0.92 ST), L66; loudness goal of 66 dB (M = 1.14 ST; SD = 1.18 ST). Figure 4.13 shows boxplots of the F0 across the loads.

Table 4.26 Descriptive statistics for F0 and loudness goal vocal load for experiment 2

Fundamental Frequency (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	-0.01	0.69	-0.14	0.12	-1.93	2.74
L54	0.65	0.98	0.47	0.84	-2.00	2.89
L60	0.80	0.92	0.62	0.98	-1.53	3.23
L66	1.14	1.18	0.91	1.38	-1.76	3.62

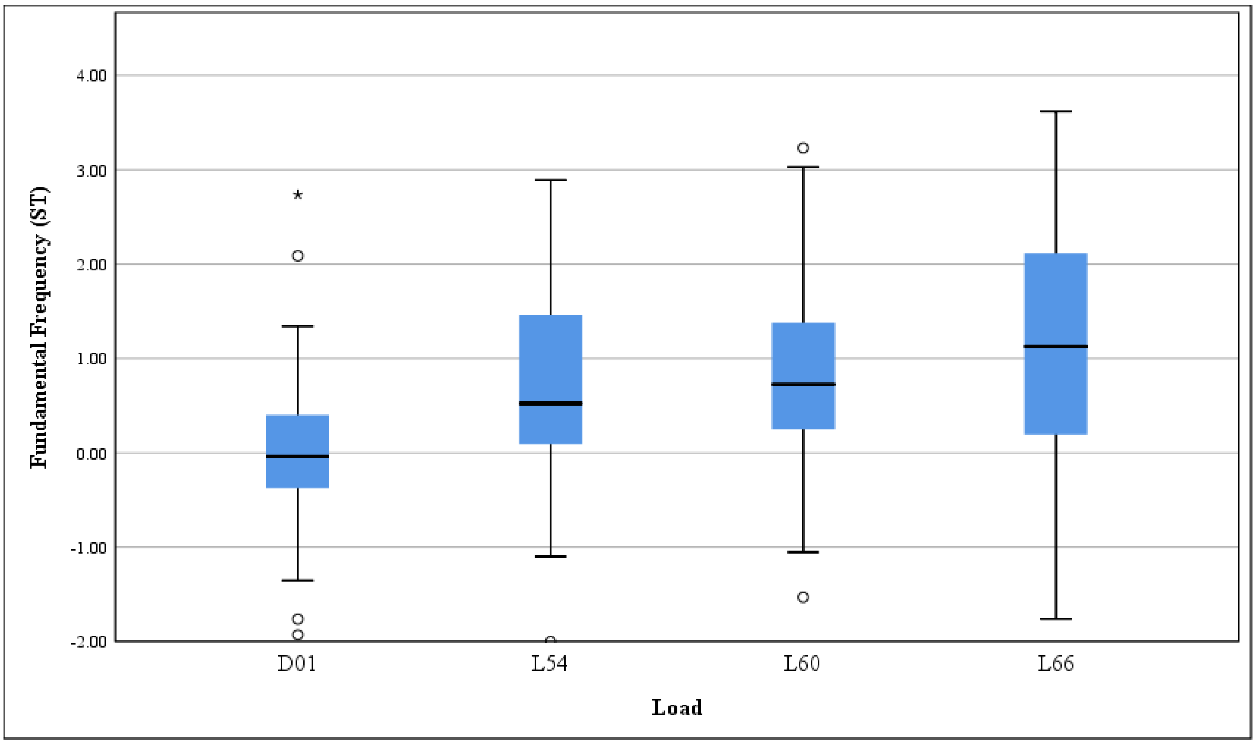

Figure 4.13 Boxplot of F0 and loudness goal vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of F0 on the vocal load of loudness goal. The post hoc tests show that F0 was different in the pair-wise comparisons with D01 and between L54 and L66 (summarized in Table 4.27). There was an increase ($p < 0.001$) in F0 of 0.67 ST from D01 to L54 and an increase ($p = 0.008$) of 0.49 ST from L54 to L66.

Table 4.27 Multiple comparison statistics for F0 and loudness goal vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (ST)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	L54	-0.67	0.11	< 0.001	-0.97	-0.36
	L60	-0.81	0.11	< 0.001	-1.11	-0.52
	L66	-1.16	0.14	< 0.001	-1.52	-0.80
L54	D01	0.67	0.11	< 0.001	0.36	0.97
	L60	-0.14	0.13	0.842	-0.49	0.20
	L66	-0.49	0.15	0.008	-0.89	-0.09
L60	D01	0.81	0.11	< 0.001	0.52	1.11
	L54	0.14	0.13	0.842	-0.20	0.49
	L66	-0.35	0.15	0.117	-0.74	0.05
L66	D01	1.16	0.14	< 0.001	0.80	1.52
	L54	0.49	0.15	0.008	0.09	0.89
	L60	0.35	0.15	0.117	-0.05	0.74

Fundamental Frequency Standard Deviation and Loudness Goal

The measured F0sd met the assumptions for normality, independence, and equal variance. For each vocal load the F0sd (M; SD) is as follows (summarized in Table 4.28 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 2.01 ST; SD = 0.79 ST), L54; loudness goal of 54 dB (M = 2.08 ST; SD = 0.74 ST), L60; loudness goal of 60 dB (M = 2.21 ST; SD = 0.8 ST), L66; loudness goal of 66 dB (M = 2.11 ST; SD = 0.78 ST). Figure 4.14 shows boxplots of the F0sd across the loads.

Table 4.28 Descriptive statistics for F0sd and loudness goal vocal load for experiment 2

Fundamental Frequency Standard Deviation (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
D01	2.01	0.79	1.85	2.17	0.73	4.45
L54	2.08	0.74	1.93	2.22	0.73	4.37
L60	2.21	0.80	2.05	2.37	0.85	4.18
L66	2.11	0.78	1.97	2.26	0.93	4.38

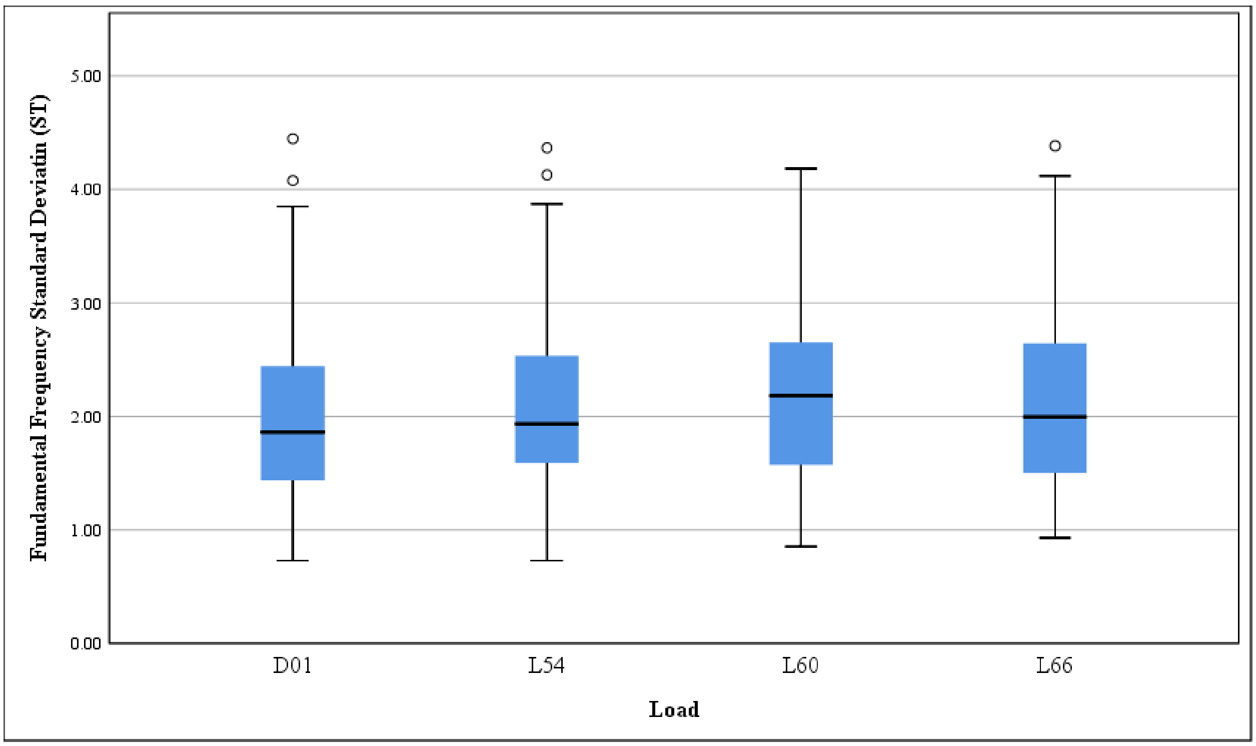

Figure 4.14 Boxplot of F0st and loudness goal vocal load for experiment 2

There was no significant main effect of F0sd on the vocal load of loudness goal.

Speech Level and Loudness Goal

The measured SL met the assumptions for normality, independence, and equal variance. For each vocal load the SL (M; SD) is as follows (summarized in Table 4.29 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 65.54 dB; SD = 3.39 dB), L54;

loudness goal of 54 dB (M = 67.4 dB; SD = 3.72 dB), L60; loudness goal of 60 dB (M = 67.99 dB; SD = 3.6 dB), L66; loudness goal of 66 dB (M = 69.9 dB; SD = 4.27 dB). Figure 4.15 shows boxplots of the SL across the loads.

Table 4.29 Descriptive statistics for SL and loudness goal vocal load for experiment 2

Speech Level (dB)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
D01	65.54	3.39	64.89	66.19	57.29	74.08
L54	67.40	3.72	66.70	68.10	56.93	74.29
L60	67.99	3.60	67.31	68.67	57.41	75.34
L66	69.90	4.27	69.08	70.71	57.48	77.56

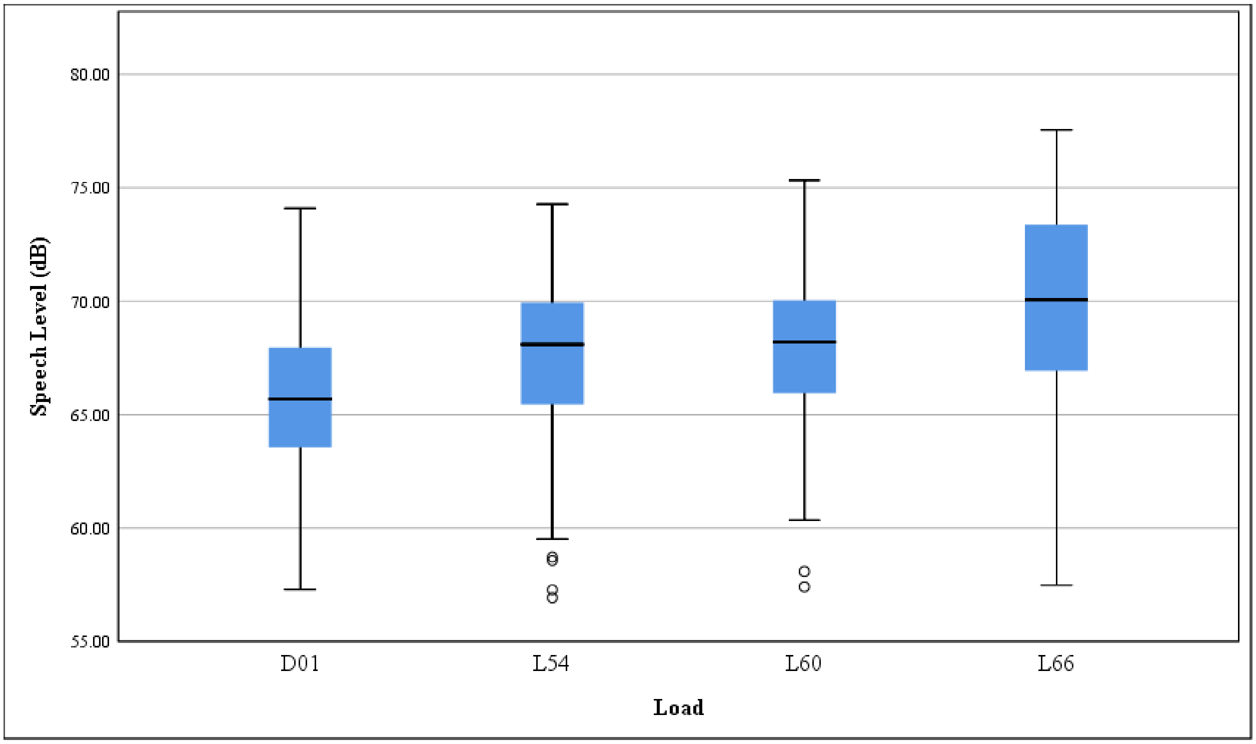

Figure 4.15 Boxplot of SL and loudness goal vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of SL on the vocal load of loudness goal. The post hoc tests show that SL was different in all pair-wise comparisons with D01 and L66 (summarized in Table 4.30). There was an increase ($p = 0.002$) in SL of 1.86 dB from D01 to L54 and an increase ($p = 0.001$) of 1.91 dB from L60 to L66.

Table 4.30 Multiple comparison statistics for SL and loudness goal vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	L54	-1.86	0.51	0.002	-3.17	-0.54
	L60	-2.45	0.51	< 0.001	-3.76	-1.13
	L66	-4.35	0.51	< 0.001	-5.68	-3.03
L54	D01	1.86	0.51	0.002	0.54	3.17
	L60	-0.59	0.51	0.648	-1.89	0.71
	L66	-2.50	0.51	< 0.001	-3.81	-1.19
L60	D01	2.45	0.51	< 0.001	1.13	3.76
	L54	0.59	0.51	0.648	-0.71	1.89
	L66	-1.91	0.51	0.001	-3.22	-0.59
L66	D01	4.35	0.51	< 0.001	3.03	5.68
	L54	2.50	0.51	< 0.001	1.19	3.81
	L60	1.91	0.51	0.001	0.59	3.22

Standard Deviation of Speech Level and Loudness Goal

The measured SLsd met the assumptions for normality, independence, and equal variance. For each vocal load the SLsd (M; SD) is as follows (summarized in Table 4.31 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 3.41 dB; SD = 0.3 dB), L54; loudness goal of 54 dB (M = 3.5 dB; SD = 0.32 dB), L60; loudness goal of 60 dB (M = 3.48 dB; SD = 0.32 dB), L66; loudness goal of 66 dB (M = 3.47 dB; SD = 0.27 dB). Figure 4.16 shows boxplots of the SLsd across the loads.

Table 4.31 Descriptive statistics for SLsd and loudness goal vocal load for experiment 2

Speech Level Standard Deviation (dB)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
D01	3.41	0.30	3.35	3.46	2.63	4.21
L54	3.50	0.32	3.44	3.56	2.66	4.22
L60	3.48	0.32	3.42	3.54	2.79	4.26
L66	3.47	0.27	3.42	3.52	2.98	4.19

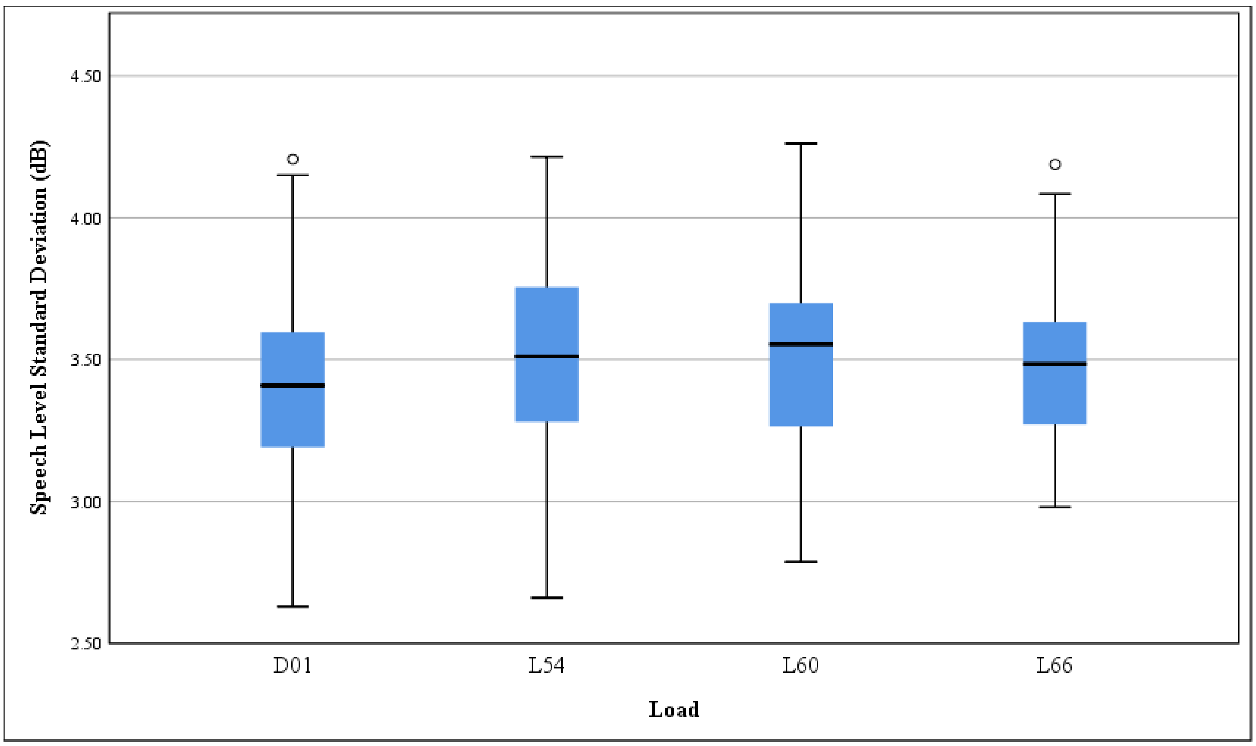

Figure 4.16 Boxplot of SLsd and loudness goal vocal load for experiment 2

There was no significant main effect of SLsd on the vocal load of loudness goal.

Smoothed Cepstral Peak Prominence and Loudness Goal

The measured CPPS met the assumptions for normality, independence, and equal variance. For each vocal load the CPPS (M; SD) is as follows (summarized in Table 4.32 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 15.33 dB; SD = 1.22 dB), L54; loudness goal of 54 dB (M = 15.83 dB; SD = 1.33 dB), L60; loudness goal of 60 dB (M = 15.89 dB; SD = 1.32 dB), L66; loudness goal of 66 dB (M = 16.39 dB; SD = 1.36 dB). Figure 4.17 shows boxplots of the CPPS across the loads.

Table 4.32 Descriptive statistics for CPPS and loudness goal vocal load for experiment 2

Smoothed Cepstral Peak Prominence (dB)

			95% Confidence Interval for Mean			
Load	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
D01	15.33	1.22	15.09	15.57	12.40	18.12
L54	15.83	1.33	15.58	16.08	12.43	19.20
L60	15.89	1.32	15.64	16.14	12.46	19.01
L66	16.39	1.36	16.13	16.64	12.20	19.17

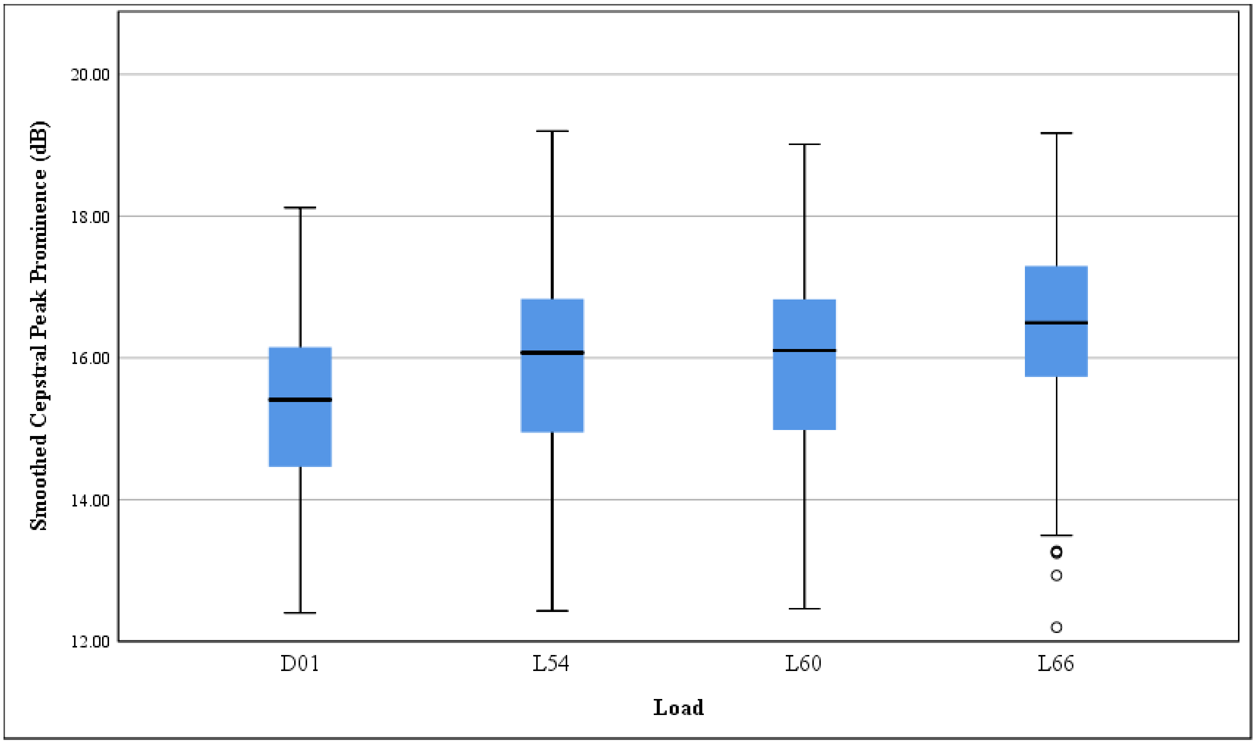

Figure 4.17 Boxplot of CPPS and loudness goal vocal load for experiment

There was a significant ($p < 0.001$) main effect of CPPS on the vocal load of loudness goal. The post hoc tests show that CPPS was different in all pair-wise comparisons with D01 and L66 (summarized in Table 4.33). There was an increase ($p = 0.029$) in CPPS of 0.50 dB from D01 to L54 and an increase ($p = 0.025$) of 0.50 dB from L60 to L66.

Table 4.33 Multiple comparison statistics for CPPS and loudness goal vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	L54	-.50	.18	.0289	-.96	-.04
	L60	-.56	.18	.0099	-1.02	-.10
	L66	-1.06	.18	< 0.001	-1.52	-.60
L54	D01	.50	.18	.0289	.04	.96
	L60	-.06	.18	.9832	-.52	.39
	L66	-.56	.18	.0080	-1.02	-.11
L60	D01	.56	.18	.0099	.10	1.02
	L54	.06	.18	.9832	-.39	.52
	L66	-.50	.18	.0253	-.95	-.04
L66	D01	1.06	.18	< 0.001	.60	1.52
	L54	.56	.18	.0080	.11	1.02
	L60	.50	.18	.0253	.04	.95

Vocal Effort Rating and Background Noise

The measured VER met the assumptions for normality, independence, and equal variance. For each vocal load the VER (M; SD) is as follows (summarized in Table 4.34 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 25.13 ; SD = 16.43), N53; background noise of 53 dBA (M = 31.91 ; SD = 15.68), N62; background noise of 62 dBA (M = 40.19 ; SD = 18.52), N71; background noise of 71 dBA (M = 52.62 ; SD = 17.99). Figure 4.18 shows boxplots of the VER across the loads.

Table 4.34 Descriptive statistics for VER and background noise vocal load for experiment 2

Vocal Effort Rating

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	25.1	16.4	22.0	28.3	2	65
N53	31.9	15.7	28.9	34.9	2	75
N62	40.2	18.5	36.7	43.7	2	90
N71	52.6	18.0	49.2	56.1	13	100

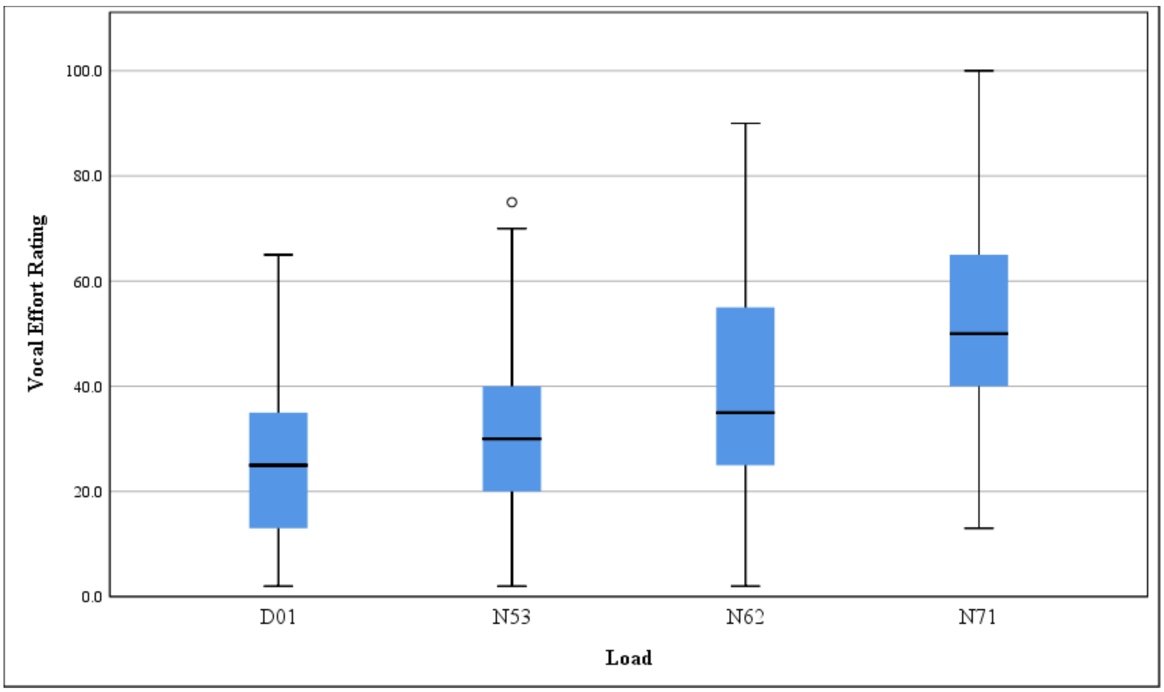

Figure 4.18 Boxplot for VER and background noise vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of VER on the vocal load of background noise. The post hoc tests show that VER was different in all pair-wise comparisons (summarized in Table 4.35). There was an increase ($p = 0.021$) in VER of 6.8 from D01 to N53, an increase ($p = 0.003$) of 8.3 from N53 to N62, and an increase ($p < 0.001$) of 12.4 from N62 to N71.

Table 4.35 Multiple comparison statistics for VER and background noise vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
D01	N53	-6.8	2.3	0.021	-12.8	-0.7
	N62	-15.1	2.3	< 0.001	-21.1	-9.0
	N71	-27.5	2.4	< 0.001	-33.5	-21.4
N53	D01	6.8	2.3	0.021	0.7	12.8
	N62	-8.3	2.3	0.003	-14.3	-2.2
	N71	-20.7	2.3	< 0.001	-26.8	-14.7
N62	D01	15.1	2.3	< 0.001	9.0	21.1
	N53	8.3	2.3	0.003	2.2	14.3
	N71	-12.4	2.3	< 0.001	-18.5	-6.4
N71	D01	27.5	2.4	< 0.001	21.4	33.5
	N53	20.7	2.3	< 0.001	14.7	26.8
	N62	12.4	2.3	< 0.001	6.4	18.5

Fundamental Frequency and Background Noise

The measured F0 met the assumptions for normality and independence. However, equal variance cannot be assumed. For each vocal load the F0 (M; SD) is as follows (summarized in Table 4.36 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = -0.01 ST; SD = 0.69 ST), N53; background noise of 53 dBA (M = 0.65 ST; SD = 0.78 ST), N62; background noise of 62 dBA (M = 1.69 ST; SD = 1.17 ST), N71; background noise of 71 dBA (M = 2.53 ST; SD = 1.36 ST). Figure 4.19 shows boxplots of the F0 across the loads.

Table 4.36 Descriptive statistics for F0 and background noise vocal load for experiment 2

Fundamental Frequency (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	-0.01	0.69	-0.14	0.12	-1.93	2.74
N53	0.65	0.78	0.50	0.80	-1.02	2.81
N62	1.69	1.17	1.46	1.91	-0.54	4.40
N71	2.53	1.36	2.24	2.81	-0.32	5.07

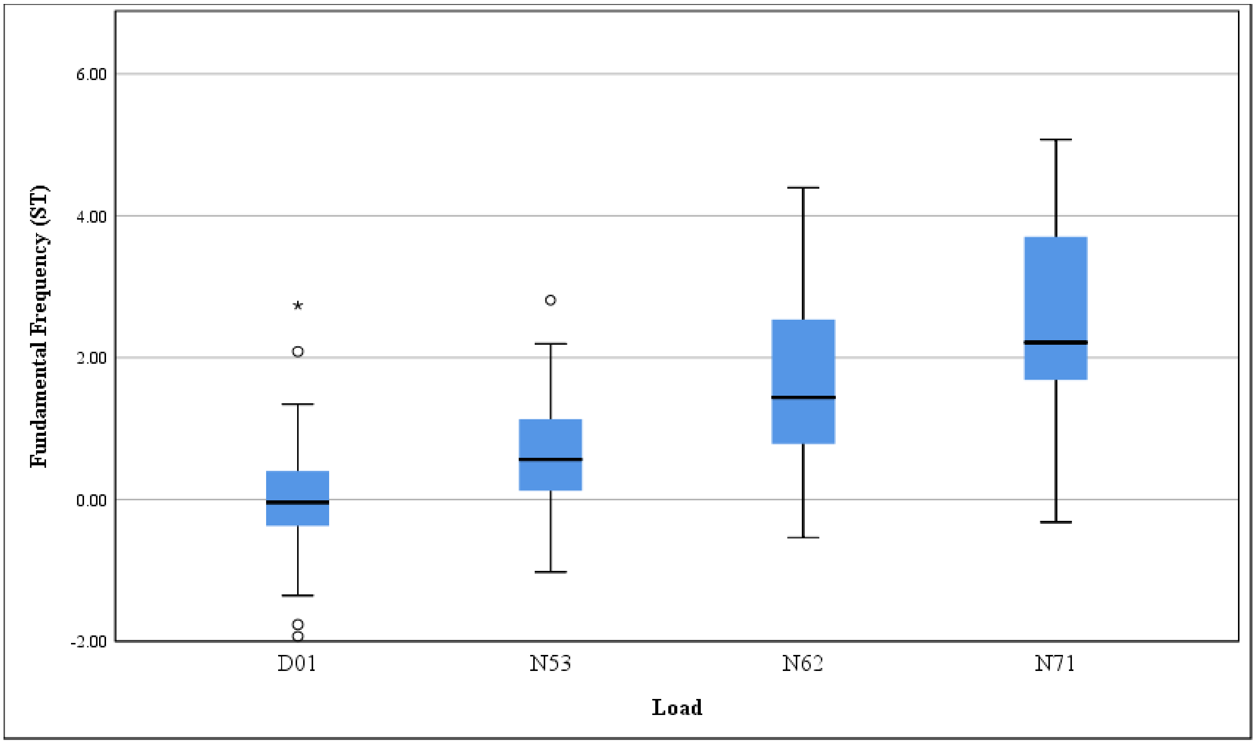

Figure 4.19 Boxplot for F0 and background noise vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of F0 on the vocal load of background noise. The post hoc tests show that F0 was different in all pair-wise comparisons (summarized in Table 4.37). There was an increase ($p < 0.001$) in F0 of 0.66 ST from D01 to N53, an increase ($p < 0.001$) of 1.04 ST from N53 to N62, and an increase ($p < 0.001$) of 0.84 ST from N62 to N71.

Table 4.37 Multiple comparison statistics for F0 and background noise vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (ST)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D01	N53	-.66	.10	< .0001	-.93	-.40
	N62	-1.70	.13	< .0001	-2.05	-1.35
	N71	-2.54	.16	< .0001	-2.96	-2.12
N53	D01	.66	.10	< .0001	.40	.93
	N62	-1.04	.14	< .0001	-1.40	-.68
	N71	-1.87	.16	< .0001	-2.30	-1.45
N62	D01	1.70	.13	< .0001	1.35	2.05
	N53	1.04	.14	< .0001	.68	1.40
	N71	-.84	.18	< .0001	-1.32	-.36
N71	D01	2.54	.16	< .0001	2.12	2.96
	N53	1.87	.16	< .0001	1.45	2.30
	N62	.84	.18	< .0001	.36	1.32

Standard Deviation of Fundamental Frequency and Background Noise

The measured F0sd met the assumptions for normality, independence, and equal variance. For each vocal load the F0sd (M; SD) is as follows (summarized in Table 4.38 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 1.99 ST; SD = 0.76 ST), N53; background noise of 53 dBA (M = 1.99 ST; SD = 0.63 ST), N62; background noise of 62 dBA (M = 2.08 ST; SD = 0.72 ST), N71; background noise of 71 dBA (M = 1.92 ST; SD = 0.7 ST). Figure 4.20 shows boxplots of the F0sd across the loads.

Table 4.38 Descriptive statistics for F0sd and background noise vocal load for experiment 2

Fundamental Frequency Standard Deviation (ST)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
D01	1.99	0.76	1.84	2.14	0.73	4.08
N53	1.99	0.63	1.87	2.12	0.88	3.65
N62	2.08	0.72	1.94	2.22	0.91	4.05
N71	1.92	0.70	1.78	2.05	0.21	3.99

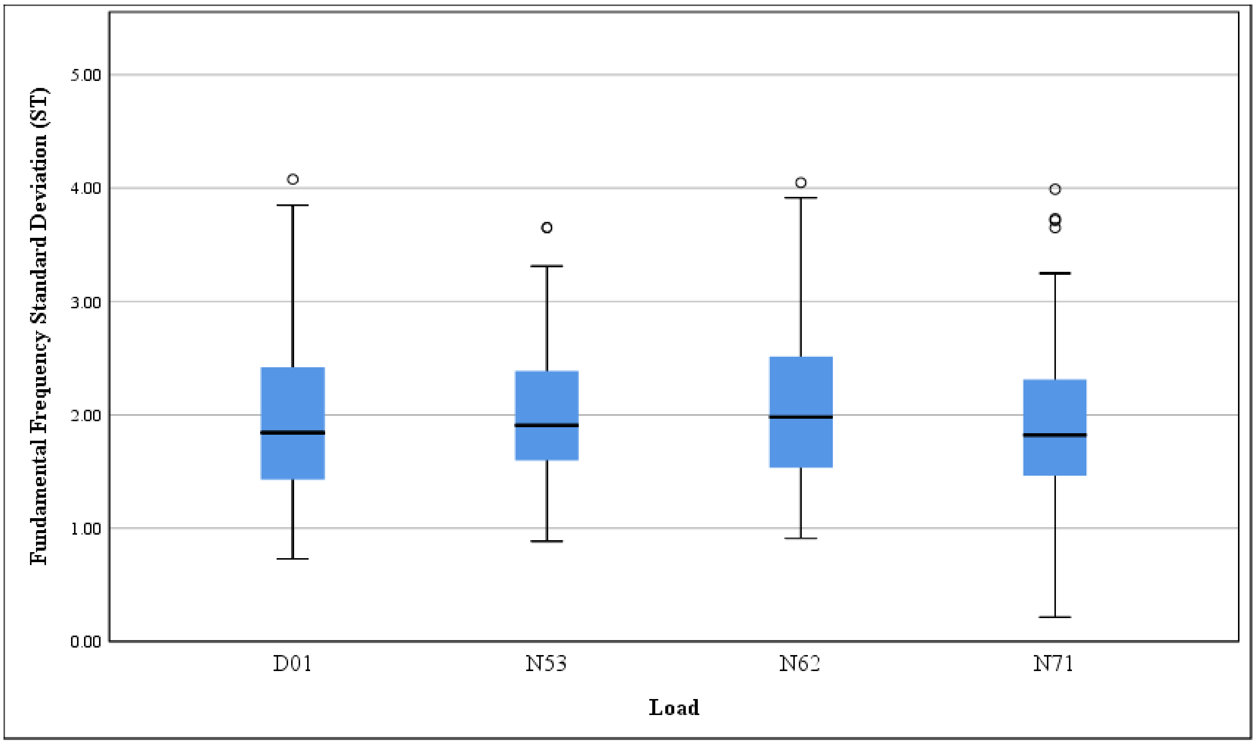

Figure 4.20 Boxplot for F0sd and background noise vocal load for experiment 2

There was no significant main effect of F0sd on the vocal load of background noise.

Speech Level and Background Noise

The measured SL met the assumptions for normality, independence, and equal variance. For each vocal load the SL (M; SD) is as follows (summarized in Table 4.39 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 65.54 dB; SD = 3.39 dB), N53; background noise of 53 dBA (M = 68.21 dB; SD = 3.75 dB), N62; background noise of 62 dBA (M = 71.09 dB; SD = 3.57 dB), N71; background noise of 71 dBA (M = 75.24 dB; SD = 3.41 dB). Figure 4.21 shows boxplots of the SL across the loads.

Table 4.39 Descriptive statistics for SL and background noise vocal load for experiment 2

Speech Level (dB)

			95% Confidence Interval for Mean			
Load	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
D01	65.54	3.39	64.89	66.19	57.29	74.08
N53	68.21	3.75	67.51	68.92	58.07	75.54
N62	71.09	3.57	70.42	71.76	61.47	80.56
N71	75.24	3.41	74.60	75.89	65.49	82.78

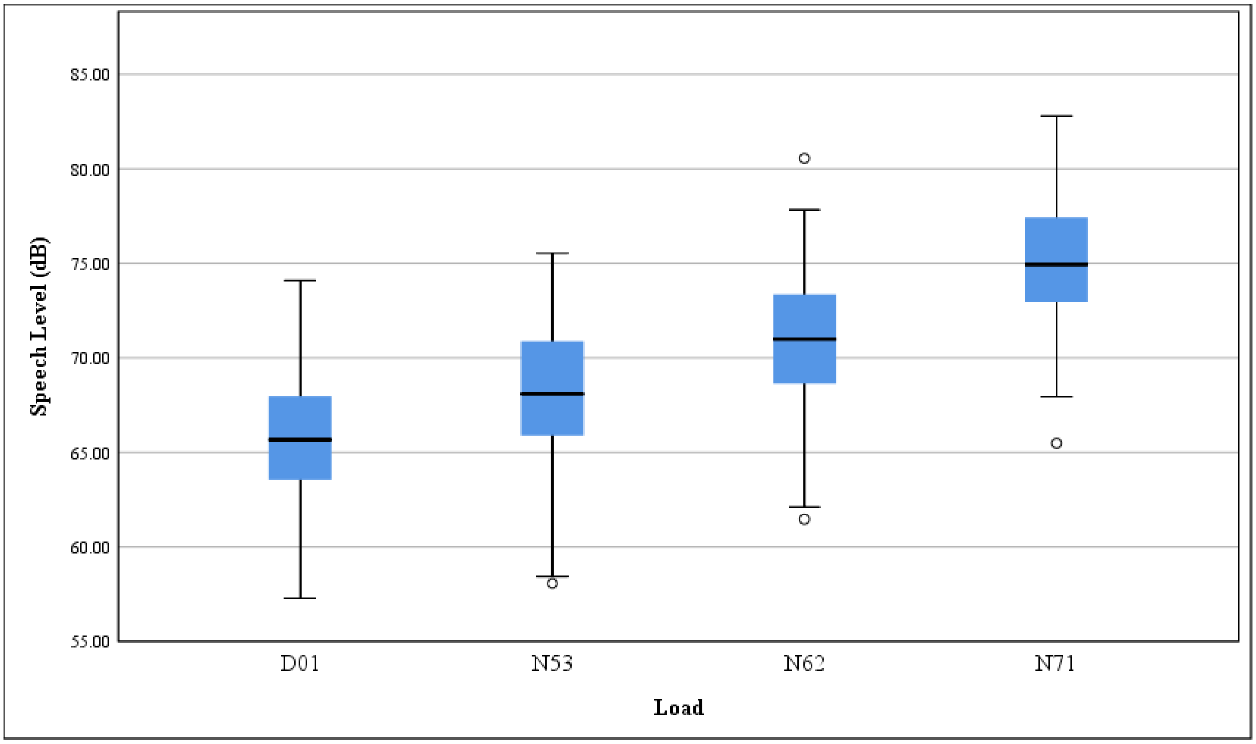

Figure 4.21 Boxplot for SL and background noise vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of SL on the vocal load of background noise. The post hoc tests show that SL was different in all pair-wise comparisons (summarized in Table 4.40). There was an increase ($p < 0.001$) in SL of 2.67 dB from D01 to N53, an increase ($p < 0.001$) of 2.88 dB from N53 to N62, and an increase ($p < 0.001$) of 4.15 dB from N62 to N71.

Table 4.40 Multiple comparison statistics for SL and background noise vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	N53	-2.67	0.48	< 0.001	-3.90	-1.44
	N62	-5.55	0.48	< 0.001	-6.78	-4.31
	N71	-9.70	0.48	< 0.001	-10.94	-8.46
N53	D01	2.67	0.48	< 0.001	1.44	3.90
	N62	-2.88	0.47	< 0.001	-4.10	-1.65
	N71	-7.03	0.48	< 0.001	-8.26	-5.80
N62	D01	5.55	0.48	< 0.001	4.31	6.78
	N53	2.88	0.47	< 0.001	1.65	4.10
	N71	-4.15	0.48	< 0.001	-5.38	-2.93
N71	D01	9.70	0.48	< 0.001	8.46	10.94
	N53	7.03	0.48	< 0.001	5.80	8.26
	N62	4.15	0.48	< 0.001	2.93	5.38

Standard Deviation of Speech Level and Background Noise

The measured SLsd met the assumptions for normality, independence, and equal variance. For each vocal load the SLsd (M; SD) is as follows (summarized in Table 4.41 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 3.41 dB; SD = 0.3 dB), N53; background noise of 53 dBA (M = 3.43 dB; SD = 0.31 dB), N62; background noise of 62 dBA (M = 3.43 dB; SD = 0.31 dB), N71; background noise of 71 dBA (M = 3.32 dB; SD = 0.34 dB). Figure 4.22 shows boxplots of the SLsd across the loads.

Table 4.41 Descriptive statistics for SLsd and background noise vocal load for experiment 2

Speech Level Standard Deviation (dB)

Load	Mean	Std. Deviation	95% Confidence Interval for Mean: Lower Bound	95% Confidence Interval for Mean: Upper Bound	Minimum	Maximum
D01	3.41	0.30	3.35	3.46	2.63	4.21
N53	3.43	0.31	3.37	3.49	2.57	4.22
N62	3.43	0.31	3.37	3.49	2.59	4.04
N71	3.32	0.34	3.25	3.39	2.61	4.08

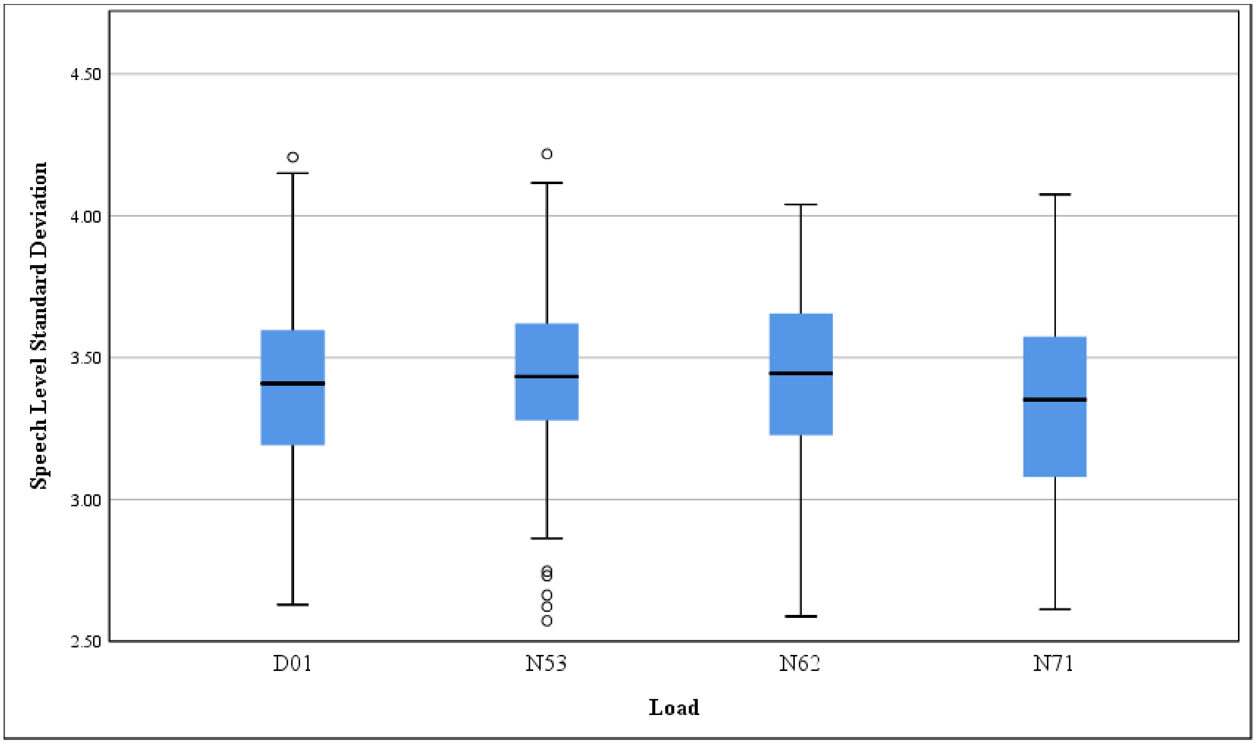

Figure 4.22 Boxplot for SLsd and background noise vocal load for experiment 2

There was no main effect of SLsd on the vocal load of background noise.

Smoothed Cepstral Peak Prominence and Background Noise

The measured CPPS met the assumptions for normality, independence, and equal variance. For each vocal load the CPPS (M; SD) is as follows (summarized in Table 4.42 with 95% confidence interval, minimum, and maximum): D01; no vocal load (M = 15.27 dB; SD = 1.29 dB), N53; background noise of 53 dBA (M = 15.43 dB; SD = 1.44 dB), N62; background noise of 62 dBA (M = 15.07 dB; SD = 1.34 dB), N71; background noise of 71 dBA (M = 14.42 dB; SD = 1.13 dB). Figure 4.23 shows boxplots of the CPPS across the loads.

Table 4.42 Descriptive statistics for CPPS and background noise vocal load for experiment 2

Smoothed Cepstral Peak Prominence (dB)

			95% Confidence Interval for Mean			
Load	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
D01	15.27	1.29	15.02	15.52	11.80	18.12
N53	15.43	1.44	15.16	15.70	11.69	18.24
N62	15.07	1.34	14.82	15.33	11.92	18.09
N71	14.42	1.13	14.20	14.65	11.46	17.35

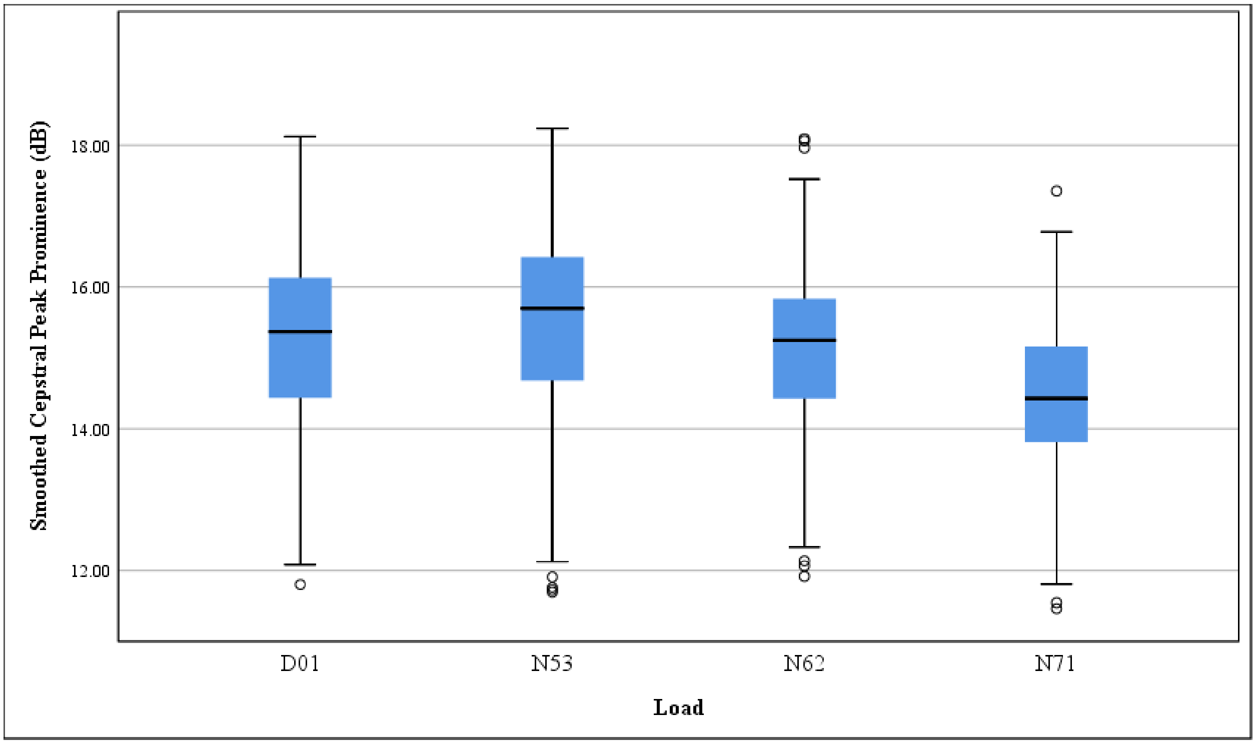

Figure 4.23 Boxplot for CPPS and background noise vocal load for experiment 2

There was a significant ($p < 0.001$) main effect of CPPS on the vocal load of background noise. The post hoc tests show that SL was different in all pair-wise comparisons with N71 (summarized in Table 4.43). There was a decrease ($p < 0.001$) in CPPS of 0.84 dB from D01 to N71, a decrease ($p < 0.001$) of 1.01 dB from N53 to N71, and a decrease ($p = 0.002$) of 0.65 dB from N62 to N71.

Table 4.43 Multiple comparison statistics for CPPS and background noise vocal load for experiment 2

(I) Load	(J) Load	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D01	N53	-0.17	0.18	0.789	-0.62	0.29
	N62	0.20	0.18	0.693	-0.27	0.66
	N71	0.84	0.18	< 0.001	0.37	1.31
N53	D01	0.17	0.18	0.789	-0.29	0.62
	N62	0.36	0.18	0.174	-0.09	0.82
	N71	1.01	0.18	< 0.001	0.55	1.47
N62	D01	-0.20	0.18	0.693	-0.66	0.27
	N53	-0.36	0.18	0.174	-0.82	0.09
	N71	0.65	0.18	0.002	0.18	1.12
N71	D01	-0.84	0.18	< 0.001	-1.31	-0.37
	N53	-1.01	0.18	< 0.001	-1.47	-0.55
	N62	-0.65	0.18	0.002	-1.12	-0.18

4.3 Experiment 3

Table 4.44 Abbreviations for results of experiment 3

Abbreviation	Meaning
F0	Fundamental frequency
F0sd	Standard deviation of fundamental frequency
L	Speech level
Lsd	Standard deviation of speech level
CPPS	Smoothed cepstral peak prominence
VER	Vocal effort rating
ST	Semitones
dB	Decibels
VLT	Vocal loading task
PRE	Time before vocal loading task
POST	Time after vocal loading task
VL05	Time at 5 minutes after start of vocal loading task
VL10	Time at 10 minutes after start of vocal loading task
VL15	Time at 15 minutes after start of vocal loading task
VL20	Time at 20 minutes after start of vocal loading task
VL25	Time at 25 minutes after start of vocal loading task
VL30	Time at 30 minutes after start of vocal loading task

4.3.a Demographics

Forty-two participants consented for the study. Five participants were not included for either not finishing protocol or not properly following the instructions. A total of 37 participants, 18 males and 19 females, were included in the analyses presented below. All of these participants were within normal hearing limits. The participants were all college age and most received course credit as compensation for their participation.

4.3.b Results

The purpose of this experiment was to measure changes in vocal performance and vocal effort through a vocal load of background noise over time to implicate vocal fatigue from the vocal loading. Vocal effort rating (VER) was measured using the Borg CR-100 before (PRE) the vocal loading task (VLT), every five minutes during the vocal loading test (LT05, LT10, LT15, LT20, LT25, LT30), and after (POST) the VLT. The same five acoustic parameters were selected to measure vocal performance, mean fundamental frequency (F0), standard deviation of fundamental frequency (F0sd), speech level (L), standard deviation of speech level (Lsd), and smoothed cepstral peak prominence (CPPS). One-way ANOVA tests were used to compare VER and each vocal performance measure throughout the VLT. Post hoc Tukey's HSD tests were used to investigate the pair-wise differences across each measurement in time. If equal variance could not be assumed, Welch's ANOVA and Tamhane's T2 post hoc tests were used.

Additionally, the data was clustered into high and low vocal effort groups, voice change and no voice change groups, and four groups representing the intersections of these groups. The PRE and POST vocal performance of each group was compared using independent-samples t-tests for each acoustic variable. The results in this section are separated by VER results, the vocal performance results, the VER clustering results, and the final group acoustic results.

Vocal Effort Rating and Vocal Loading

The measured VER met the assumptions for normality and independence. However, equal variance cannot be assumed. For each time point of the vocal loading the VER (M; SD) is as follows (summarized in Table 4.45 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 17.11 ; SD = 13.88), VL05; after 5 minutes of VLT (M = 34.68 ; SD = 17.73), VL10; after 10 minutes of VLT (M = 36.86 ; SD = 19.15), VL15; after 15 minutes of

VLT (M = 40.27 ; SD = 20.51), VL20; after 20 minutes of VLT (M = 42.57 ; SD = 22.48), VL25; after 25 minutes of VLT (M = 45.35 ; SD = 24.69), VL30; after 30 minutes of VLT (M = 47.62 ; SD = 28.22), POST; after the complete VLT (M = 39.84 ; SD = 27.76), Figure 4.24 shows a line graph of the VER across the loading.

Table 4.45 Descriptive statistics for VER over time for experiment 3

Vocal Effort Ratings

Time	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
PRE	17.1	13.9	12.5	21.7	2	60
VL05	34.7	17.7	28.8	40.6	7	80
VL10	36.9	19.1	30.5	43.2	8	79
VL15	40.3	20.5	33.4	47.1	10	85
VL20	42.6	22.5	35.1	50.1	7	85
VL25	45.4	24.7	37.1	53.6	7	90
VL30	47.6	28.2	38.2	57.0	2	100
POST	39.8	27.8	30.6	49.1	2	100

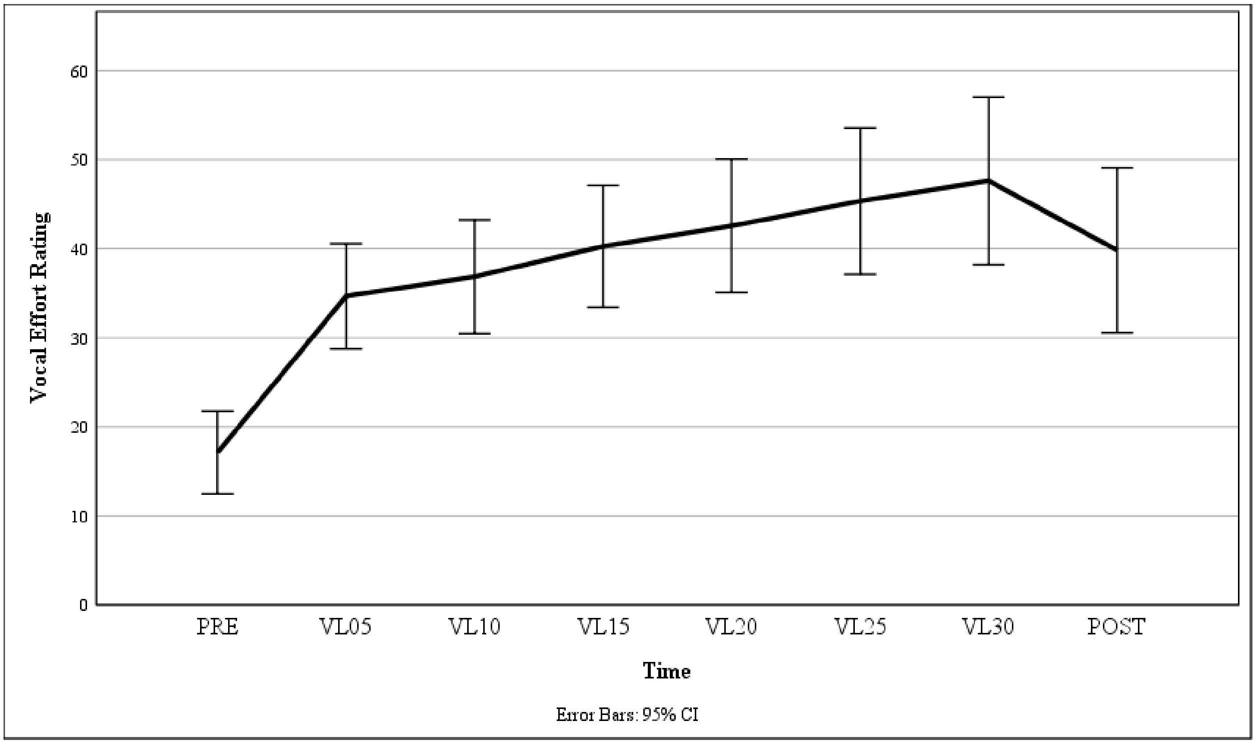

Figure 4.24 Graph of VER over time for experiment 3

There was a significant ($p < 0.001$) main effect of VER across the time of the VLT. Only the VER at time PRE was significantly different the other time points of the VLT (summarized in Table 4.46). There was a significant increase ($p < 0.001$) of 17.6 VER between VL05 and PRE. The were no significant changes of VER after VL05. There was a significant increase ($p = 0.001$) of 22.7 VER between POST and PRE.

Table 4.46 Multiple comparison statistics for VER over time for experiment 3

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
PRE	VL05	-17.6	3.7	< 0.001	-29.6	-5.6
	VL10	-19.8	3.9	< 0.001	-32.4	-7.1
	VL15	-23.2	4.1	< 0.001	-36.4	-9.9
	VL20	-25.5	4.3	< 0.001	-39.6	-11.3
	VL25	-28.2	4.7	< 0.001	-43.5	-13.0
	VL30	-30.5	5.2	< 0.001	-47.5	-13.5
	POST	-22.7	5.1	0.001	-39.5	-6.0
VL30	PRE	30.5	5.2	< 0.001	13.5	47.5
	VL05	12.9	5.5	0.454	-4.9	30.8
	VL10	10.8	5.6	0.821	-7.5	29.0
	VL15	7.4	5.7	0.998	-11.3	26.0
	VL20	5.1	5.9	1.000	-14.2	24.3
	VL25	2.3	6.2	1.000	-17.7	22.2
	POST	7.8	6.5	0.999	-13.3	28.8
POST	PRE	22.7	5.1	0.001	6.0	39.5
	VL05	5.2	5.4	1.000	-12.5	22.8
	VL10	3.0	5.5	1.000	-15.1	21.0
	VL15	-0.4	5.7	1.000	-18.9	18.0
	VL20	-2.7	5.9	1.000	-21.8	16.3
	VL25	-5.5	6.1	1.000	-25.3	14.3
	VL30	-7.8	6.5	0.999	-28.8	13.3

Fundamental Frequency and Vocal Loading

The measured F0 met the assumptions for normality and independence. However, equal variance cannot be assumed. For each time point of the vocal loading the F0 (M; SD) is as follows (summarized in Table 4.47 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = -0.01 ST; SD = 0.78 ST), VL05; after 5 minutes of VLT (M = 2.72 ST; SD = 1.9 ST), VL10; after 10 minutes of VLT (M = 3 ST; SD = 1.96 ST), VL15; after 15 minutes of VLT (M = 2.97 ST; SD = 1.92 ST), VL20; after 20 minutes of VLT (M = 3.06 ST; SD = 1.92 ST), VL25; after 25 minutes of VLT (M = 3.17 ST; SD = 2.04 ST), VL30; after 30 minutes of VLT (M = 3.04 ST; SD = 2.02 ST), POST; after the complete VLT (M = 0.83 ST; SD = 1.34 ST), Figure 4.25 shows a line graph of the F0 across the loading.

Table 4.47 Descriptive statistics for F0 over time for experiment 3

Fundamental Frequency (ST)

Time	Mean	Std. Deviation	95% Confidence Interval for Mean: Lower Bound	95% Confidence Interval for Mean: Upper Bound	Minimum	Maximum
PRE	-0.01	0.78	-0.11	0.09	-1.84	3.31
VL05	2.72	1.90	2.41	3.04	-1.08	7.52
VL10	3.00	1.96	2.68	3.32	-0.42	8.14
VL15	2.97	1.92	2.64	3.29	-1.00	8.28
VL20	3.06	1.92	2.74	3.38	-0.41	7.94
VL25	3.17	2.04	2.83	3.51	-0.23	7.82
VL30	3.04	2.02	2.70	3.38	-0.89	8.09
POST	0.83	1.34	0.65	1.01	-2.90	7.29

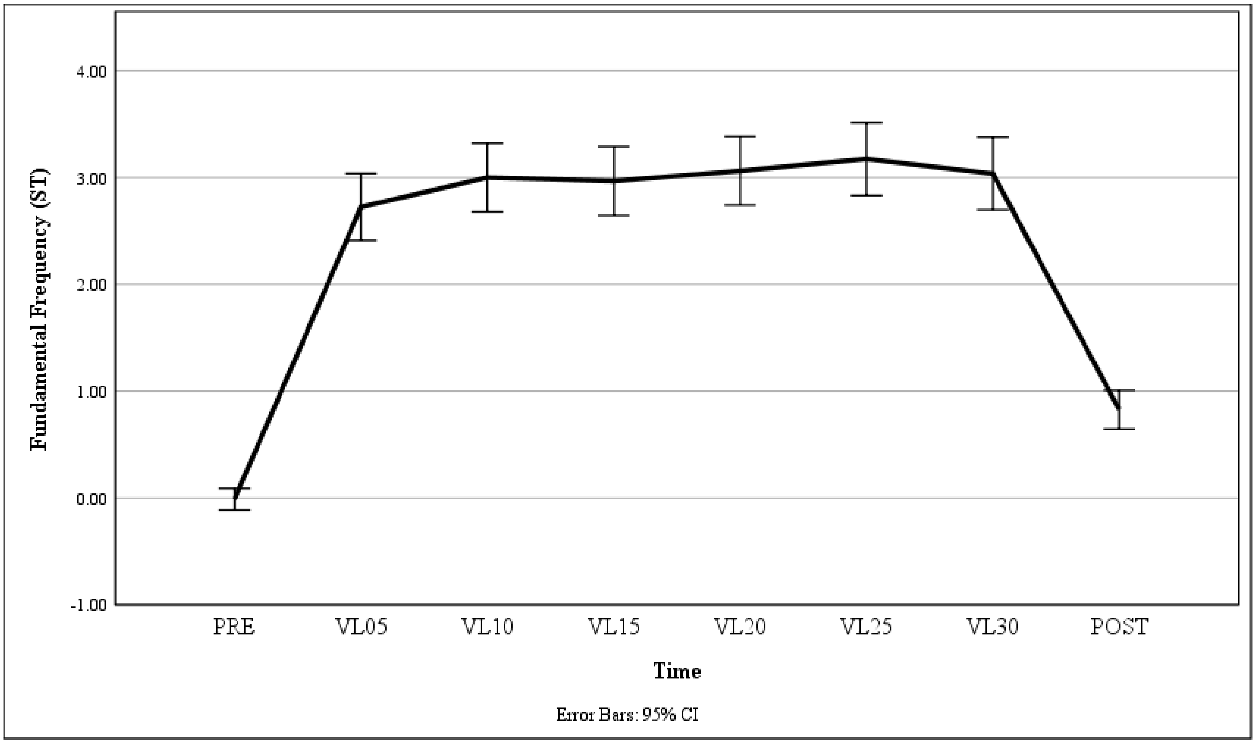

Figure 4.25 Graph of F0 over time for experiment 3

There was a significant ($p < 0.001$) main effect of F0 across the time of the VLT. There was a significant increase ($p < 0.001$) in F0 of 2.74 ST from PRE to VL05 (summarized in Table 4.48). There were no significant changes of F0 across the VLT. There was a significant increase ($p < 0.001$) of 0.84 ST from PRE to POST. There was a significant decrease ($p < 0.001$) of 2.21 ST from VL30 to POST.

Table 4.48 Multiple comparison statistics for F0 over time for experiment 3

(I) Time	(J) Time	Mean Difference (I-J) (ST)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
PRE	VL05	-2.74	0.17	< 0.001	-3.26	-2.21
	VL10	-3.01	0.17	< 0.001	-3.55	-2.47
	VL15	-2.98	0.17	< 0.001	-3.52	-2.44
	VL20	-3.07	0.17	< 0.001	-3.61	-2.54
	VL25	-3.19	0.18	< 0.001	-3.76	-2.62
	VL30	-3.05	0.18	< 0.001	-3.62	-2.48
	POST	-0.84	0.11	< 0.001	-1.17	-0.51
VL30	PRE	3.05	0.18	< 0.001	2.48	3.62
	VL05	0.31	0.23	0.997	-0.42	1.05
	VL10	0.04	0.24	1.000	-0.71	0.78
	VL15	0.07	0.24	1.000	-0.68	0.82
	VL20	-0.03	0.24	1.000	-0.77	0.72
	VL25	-0.14	0.24	1.000	-0.90	0.63
	POST	2.21	0.19	< 0.001	1.59	2.82
POST	PRE	0.84	0.11	< 0.001	0.51	1.17
	VL05	-1.90	0.18	< 0.001	-2.48	-1.32
	VL10	-2.17	0.19	< 0.001	-2.76	-1.58
	VL15	-2.14	0.19	< 0.001	-2.73	-1.55
	VL20	-2.23	0.19	< 0.001	-2.82	-1.65
	VL25	-2.35	0.20	< 0.001	-2.96	-1.73
	VL30	-2.21	0.19	< 0.001	-2.82	-1.59

Standard Deviation of Fundamental Frequency and Vocal Loading

The measured F0sd met the assumptions for normality, independence, and equal variance. For each time point of the vocal loading the F0sd (M; SD) is as follows (summarized in Table 4.49 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 1.91 ST; SD = 0.83 ST), VL05; after 5 minutes of VLT (M = 2.3 ST; SD = 0.76 ST), VL10; after 10 minutes of VLT (M = 2.44 ST; SD = 0.77 ST), VL15; after 15 minutes of VLT (M = 2.49 ST;

SD = 0.73 ST), VL20; after 20 minutes of VLT (M = 2.52 ST; SD = 0.81 ST), VL25; after 25 minutes of VLT (M = 2.53 ST; SD = 0.78 ST), VL30; after 30 minutes of VLT (M = 2.51 ST; SD = 0.76 ST), POST; after the complete VLT (M = 2.16 ST; SD = 0.81 ST), Figure 4.26 shows a line graph of the F0sd across the loading.

Table 4.49 Descriptive statistics for F0sd over time for experiment 3

Fundamental Frequency Standard Deviation (ST)

Time	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
PRE	1.91	0.83	1.79	2.03	0.55	4.61
VL05	2.30	0.76	2.17	2.42	0.99	4.55
VL10	2.44	0.77	2.31	2.57	1.08	4.47
VL15	2.49	0.73	2.36	2.61	1.18	4.73
VL20	2.52	0.81	2.38	2.66	1.01	4.57
VL25	2.53	0.78	2.40	2.66	1.17	4.68
VL30	2.51	0.76	2.39	2.64	1.10	4.76
POST	2.16	0.81	2.04	2.28	0.88	4.68

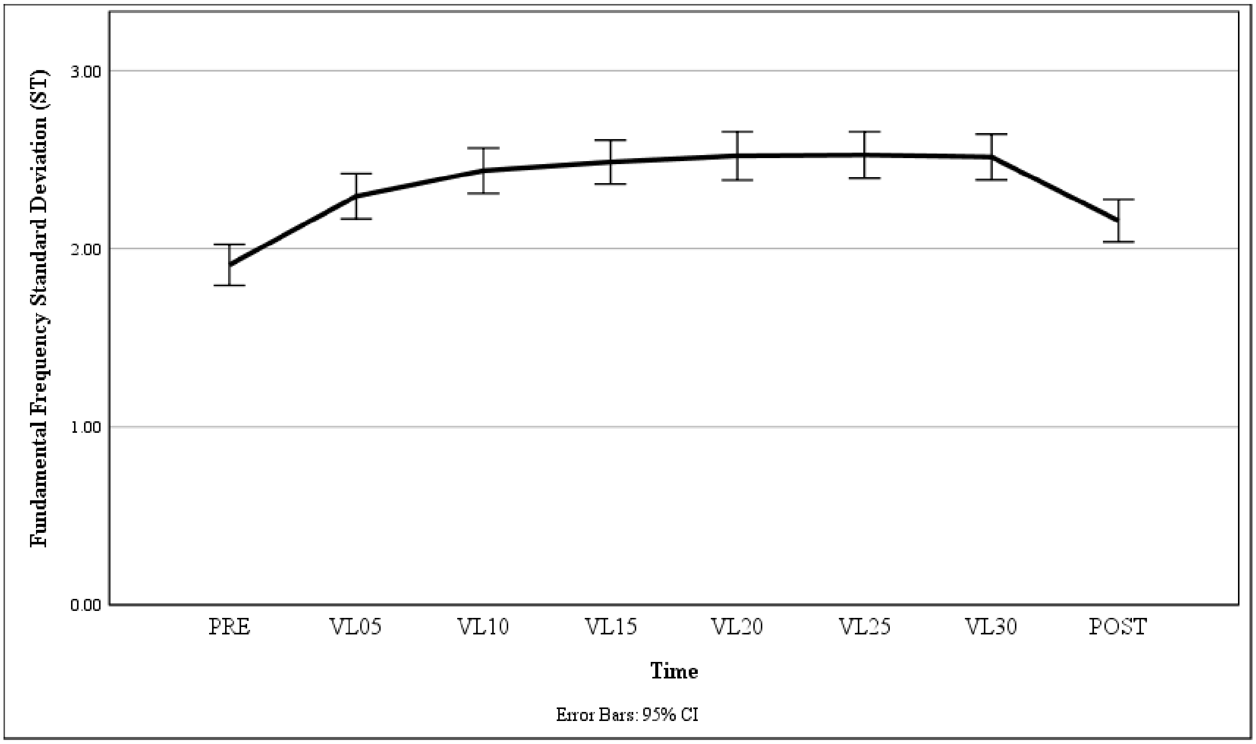

Figure 4.26 Graph of F0sd over time for experiment 3

There was a significant ($p < 0.001$) main effect of F0sd across the time of the VLT. There was an increase ($p < 0.001$) in F0sd of 0.6 ST from PRE to VEL05, no change across the VLT, and a decrease ($p = 0.002$) of 0.36 ST from VEL30 to POST (summarized in Table 4.50). There was also an increase ($p = 0.043$) of 0.25 ST from PRE to POST.

Table 4.50 Multiple comparison statistics for F0sd over time for experiment 3

(I) Time	(J) Time	Mean Difference (I-J) (ST)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
PRE	VL05	-0.39	0.09	< 0.001	-0.65	-0.12
	VL10	-0.53	0.09	< 0.001	-0.79	-0.27
	VL15	-0.58	0.09	< 0.001	-0.84	-0.31
	VL20	-0.61	0.09	< 0.001	-0.87	-0.35
	VL25	-0.62	0.09	< 0.001	-0.88	-0.35
	VL30	-0.60	0.09	< 0.001	-0.87	-0.34
	POST	-0.25	0.08	0.043	-0.49	0.00
VL30	PRE	0.60	0.09	< 0.001	0.34	0.87
	VL05	0.22	0.09	0.285	-0.07	0.50
	VL10	0.08	0.09	0.993	-0.21	0.36
	VL15	0.03	0.09	1.000	-0.26	0.32
	VL20	-0.01	0.09	1.000	-0.29	0.28
	VL25	-0.01	0.09	1.000	-0.30	0.27
	POST	0.36	0.09	0.002	0.09	0.63
POST	PRE	0.25	0.08	0.043	0.00	0.49
	VL05	-0.14	0.09	0.773	-0.41	0.13
	VL10	-0.28	0.09	0.032	-0.55	-0.01
	VL15	-0.33	0.09	0.006	-0.60	-0.06
	VL20	-0.36	0.09	0.001	-0.63	-0.09
	VL25	-0.37	0.09	0.001	-0.64	-0.10
	VL30	-0.36	0.09	0.002	-0.63	-0.09

Speech Level and Vocal Loading

The measured SL met the assumptions for normality, independence, and equal variance. For each time point of the vocal loading the SL (M; SD) is as follows (summarized in Table 4.51 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 56.23 dB; SD = 5.26 dB), VL05; after 5 minutes of VLT (M = 66.99 dB; SD = 4.91 dB), VL10; after 10 minutes of VLT (M = 67.45 dB; SD = 4.79 dB), VL15; after 15 minutes of VLT (M = 67.57 dB;

SD = 4.79 dB), VL20; after 20 minutes of VLT (M = 67.53 dB; SD = 4.67 dB), VL25; after 25 minutes of VLT (M = 67.42 dB; SD = 5.11 dB), VL30; after 30 minutes of VLT (M = 66.86 dB; SD = 4.99 dB), POST; after the complete VLT (M = 57.46 dB; SD = 5.13 dB), Figure 4.27 shows a line graph of the SL across the loading.

Table 4.51 Descriptive statistics for SL over time for experiment 3

Speech Level (dB)

Time	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
PRE	56.23	5.26	55.55	56.91	43.55	68.16
VL05	66.99	4.91	66.18	67.80	55.21	83.28
VL10	67.45	4.79	66.67	68.24	57.13	81.93
VL15	67.57	4.79	66.78	68.37	57.45	83.02
VL20	67.53	4.67	66.77	68.30	55.76	82.62
VL25	67.42	5.11	66.58	68.26	55.54	82.74
VL30	66.86	4.99	66.04	67.69	54.73	81.65
POST	57.46	5.13	56.75	58.17	42.93	70.54

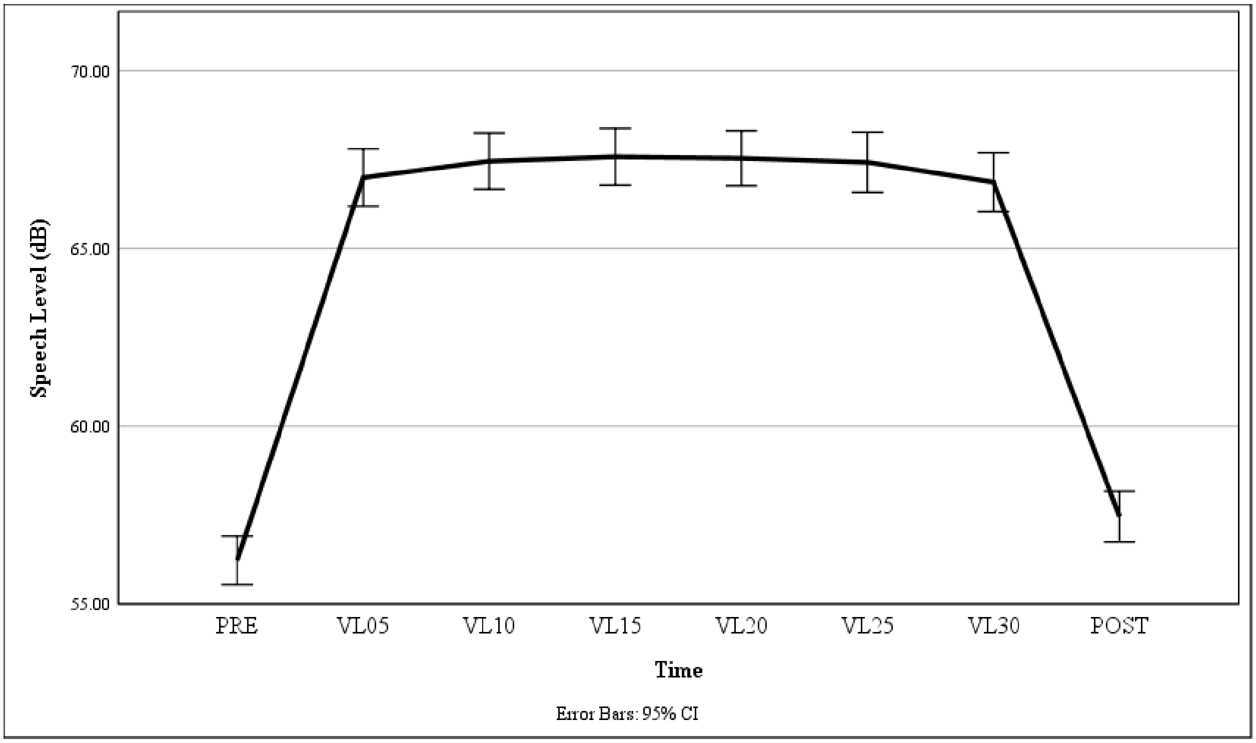

Figure 4.27 Graph of SL over time for experiment 3

There was a significant ($p < 0.001$) main effect of SL across the time of the VLT. There was a significant increase in SL of 10.6 dB ($p < 0.001$) from PRE to VL05 (summarized in Table 4.52). There were no significant changes of SL across the VLT or from PRE and POST. There was a significant decrease in SL of 9.4 dB ($p < 0.001$) from VL30 to POST.

Table 4.52 Multiple comparison statistics for SL over time for experiment 3

(I) Time	(J) Time	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
PRE	VL05	-10.76	0.53	< 0.001	-12.37	-9.16
	VL10	-11.22	0.53	< 0.001	-12.83	-9.62
	VL15	-11.35	0.53	< 0.001	-12.96	-9.73
	VL20	-11.31	0.53	< 0.001	-12.91	-9.70
	VL25	-11.19	0.53	< 0.001	-12.80	-9.59
	VL30	-10.63	0.53	< 0.001	-12.24	-9.02
	POST	-1.23	0.48	0.170	-2.69	0.23
VL30	PRE	10.63	0.53	< 0.001	9.02	12.24
	VL05	-0.13	0.59	1.000	-1.92	1.66
	VL10	-0.59	0.59	0.974	-2.38	1.19
	VL15	-0.71	0.59	0.930	-2.51	1.08
	VL20	-0.67	0.59	0.947	-2.46	1.12
	VL25	-0.56	0.59	0.981	-2.35	1.23
	POST	9.40	0.55	< 0.001	7.75	11.06
POST	PRE	1.23	0.48	0.170	-0.23	2.69
	VL05	-9.53	0.54	< 0.001	-11.19	-7.88
	VL10	-9.99	0.54	< 0.001	-11.64	-8.34
	VL15	-10.12	0.55	< 0.001	-11.77	-8.46
	VL20	-10.07	0.54	< 0.001	-11.73	-8.42
	VL25	-9.96	0.54	< 0.001	-11.61	-8.31
	VL30	-9.40	0.55	< 0.001	-11.06	-7.75

Speech Level Standard Deviation and Vocal Loading

The measured SLsd met the assumptions for normality and independence. However, equal variance cannot be assumed. For each time point of the vocal loading the SLsd (M; SD) is as follows (summarized in Table 4.53 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 3.56 dB; SD = 0.32 dB), VL05; after 5 minutes of VLT (M = 3.78 dB; SD = 0.22 dB), VL10; after 10 minutes of VLT (M = 3.77 dB; SD = 0.19 dB), VL15; after 15 minutes of VLT (M = 3.82 dB; SD = 0.21 dB), VL20; after 20 minutes of VLT (M = 3.81 dB; SD = 0.21 dB), VL25; after 25 minutes of VLT (M = 3.8 dB; SD = 0.21 dB), VL30; after 30 minutes of VLT (M = 3.79 dB; SD = 0.23 dB), POST; after the complete VLT (M = 3.63 dB; SD = 0.33 dB), Figure 4.28 shows a line graph of the SLsd across the loading.

Table 4.53 Descriptive statistics for SLsd over time for experiment 3

Speech Level Standard Deviation (dB)

			95% Confidence Interval for Mean			
Time	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
PRE	3.56	0.32	3.51	3.61	3.00	4.31
VL05	3.78	0.22	3.74	3.81	3.22	4.37
VL10	3.77	0.19	3.74	3.81	3.11	4.20
VL15	3.82	0.21	3.79	3.85	3.26	4.45
VL20	3.81	0.21	3.78	3.84	3.29	4.38
VL25	3.80	0.21	3.77	3.84	3.33	4.32
VL30	3.79	0.23	3.75	3.83	3.25	4.50
POST	3.63	0.33	3.58	3.68	3.01	4.38

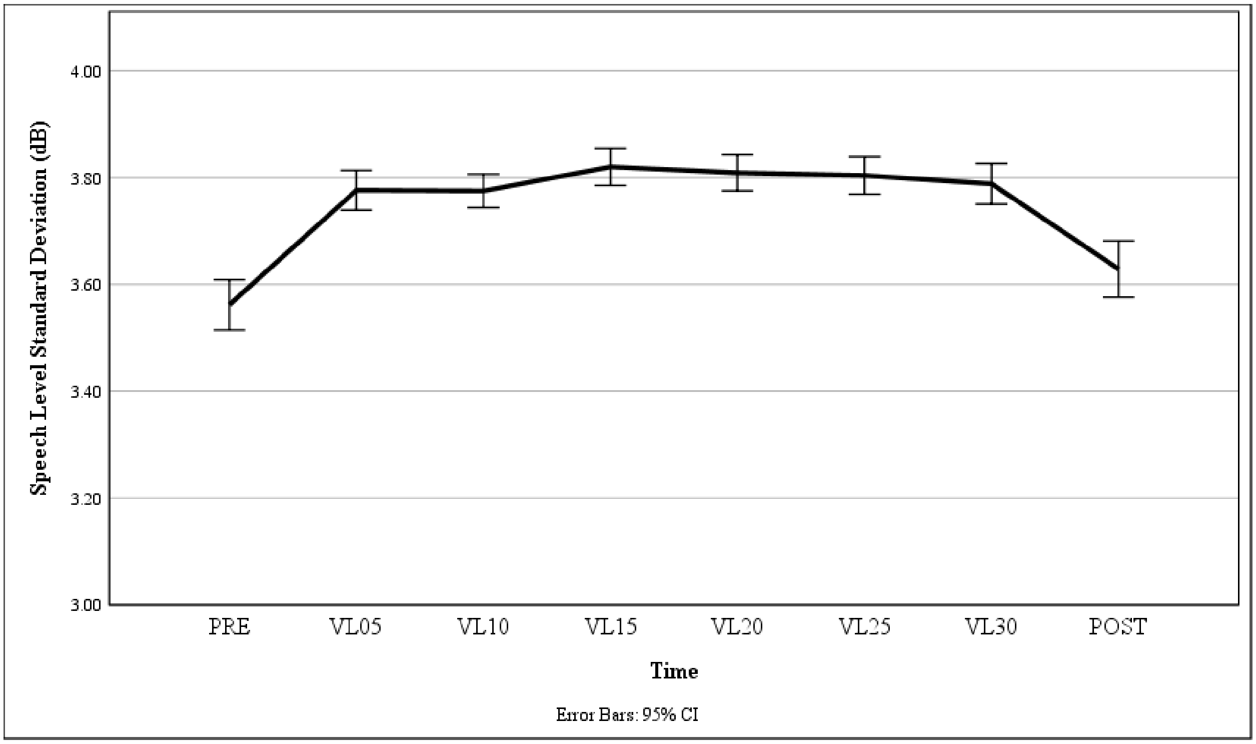

Figure 4.28 Graph of SLsd over time for experiment 3

There was a significant ($p < 0.001$) main effect of SLsd across the time of the VLT. There was a significant increase in SLsd of 0.23 dB ($p < 0.001$) from PRE to VL05 (summarized in Table 4.54). There were no significant changes of SLsd across the VLT or from PRE and POST. There was a significant decrease in SLsd of 0.16 dB ($p < 0.001$) from VL30 to POST.

Table 4.54 Multiple comparison statistics for SLsd over time for experiment 3

(I) Time	(J) Time	Mean Difference (I-J) (dB)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
PRE	VL05	-0.21	0.03	< 0.001	-0.31	-0.12
	VL10	-0.21	0.03	< 0.001	-0.30	-0.12
	VL15	-0.26	0.03	< 0.001	-0.35	-0.16
	VL20	-0.25	0.03	< 0.001	-0.34	-0.15
	VL25	-0.24	0.03	< 0.001	-0.34	-0.15
	VL30	-0.23	0.03	< 0.001	-0.32	-0.13
	POST	-0.07	0.04	0.838	-0.18	0.05
VL30	PRE	0.23	0.03	< 0.001	0.13	0.32
	VL05	0.01	0.03	1.000	-0.07	0.10
	VL10	0.01	0.02	1.000	-0.06	0.09
	VL15	-0.03	0.03	0.999	-0.11	0.05
	VL20	-0.02	0.03	1.000	-0.10	0.06
	VL25	-0.02	0.03	1.000	-0.10	0.07
	POST	0.16	0.03	< 0.001	0.06	0.26
POST	PRE	0.07	0.04	0.838	-0.05	0.18
	VL05	-0.15	0.03	< 0.001	-0.25	-0.04
	VL10	-0.15	0.03	< 0.001	-0.24	-0.05
	VL15	-0.19	0.03	< 0.001	-0.29	-0.09
	VL20	-0.18	0.03	< 0.001	-0.28	-0.08
	VL25	-0.18	0.03	< 0.001	-0.28	-0.07
	VL30	-0.16	0.03	< 0.001	-0.26	-0.06

Smoothed Cepstral Peak Prominence and Vocal Loading

The measured CPPS met the assumptions for normality and independence. However, equal variance cannot be assumed. For each time point of the vocal loading the CPPS (M; SD) is as follows (summarized in Table 4.55 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 13.39 dB; SD = 1.37 dB), VL05; after 5 minutes of VLT (M = 13.57 dB; SD = 1 dB), VL10; after 10 minutes of VLT (M = 13.53 dB; SD = 1.08 dB), VL15; after 15 minutes of VLT (M = 13.51 dB; SD = 1.07 dB), VL20; after 20 minutes of VLT (M = 13.49 dB; SD = 1.03 dB), VL25; after 25 minutes of VLT (M = 13.7 dB; SD = 0.99 dB), VL30; after 30 minutes of VLT (M = 13.61 dB; SD = 0.94 dB), POST; after the complete VLT (M = 13.44 dB; SD = 1.42 dB), Figure 4.29 shows a line graph of the CPPS across the loading.

Table 4.55 Descriptive statistics for CPPS over time for experiment 3

Smoothed Cepstral Peak Prominence (CPPS)

			95% Confidence Interval for Mean			
Time	Mean	Std. Deviation	Lower Bound	Upper Bound	Minimum	Maximum
PRE	13.39	1.37	13.20	13.58	10.28	16.26
VL05	13.57	1.00	13.40	13.74	10.85	15.53
VL10	13.53	1.08	13.35	13.70	10.51	15.43
VL15	13.51	1.07	13.34	13.69	10.57	15.42
VL20	13.49	1.03	13.32	13.66	10.41	15.68
VL25	13.70	0.99	13.53	13.87	10.38	15.80
VL30	13.61	0.94	13.45	13.77	10.97	15.21
POST	13.44	1.42	13.23	13.65	10.38	16.41

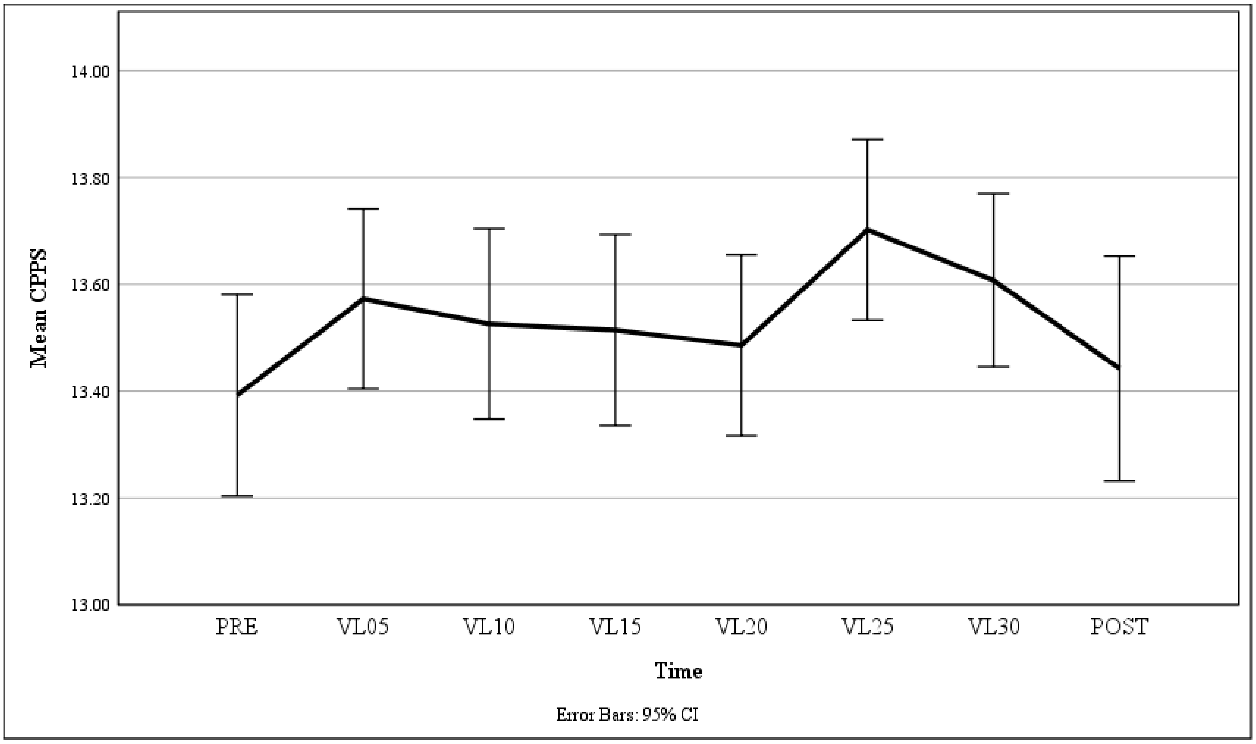

Figure 4.29 Graph of CPPS over time for experiment 3

There was no main effect of CPPS across the time of the VLT.

Vocal Load Response Clustering

In order to detect potential participants groups that differ in vocal effort as a response to the vocal loading, k-means clustering was used on the feature set from 3.4.e. Two features were determined to be sufficient to cluster the data. These features were the noise load response (NLR) and the temporal load response (TLR). Both features were statistically significant (NLR: $p = 0.003$; TLR: $p < 0.001$) features in the k-means clustering (Table 4.56). This significance was obtained through F tests and is used only for descriptive purposes as the clusters are chosen to maximize the differences across the cases in the clusters. Independent samples t-tests confirm that means of the two clusters for NLR ($p < 0.001$) and TLR ($p = 0.001$) are significantly different (Table 4.57). These two features were not correlated meeting the assumptions of orthogonality (Table 4.58). There were 14 participants in cluster 1 (CL1) and 24 participants in cluster 2 (CL2). The center of CL1 was NLR of 26.4 and TLR of 32.9. The center of CL2 was NLT of 12.2 and TLR of 0.8. Since both features relate to vocal load responses and are higher in CL1, CL1 was relabeled to high vocal load response (HVLR) and CL2 was relabeled to low vocal load response (LVLR). Table 4.59 summarizes the count and centers of the two clusters and Figure 4.30 shows the data separated by cluster including cluster centers.

Table 4.56 F tests for feature significance in k-means clustering of VER for experiment 3

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
NLR	1739.954	1	164.146	35	10.600	.003
TLR	8993.050	1	213.338	35	42.154	< 0.001

Table 4.57 Independent samples t-test for comparison of means between LVLR and HVLR clusters in experiment 3

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
NLR	.940	.339	5.902	35	.000	20.461	3.467	13.423	27.498
TLR	1.512	.227	3.469	35	.001	21.730	6.264	9.014	34.447

Table 4.58 Bivariate correlation between features NLR and TLR for experiment 3

		TLR
NLR	Pearson Correlation	.083
	Sig. (2-tailed)	.625

Table 4.59 Cluster information including number of cases and centers for NLR and TLR for experiment 3

		Cluster Center	
Group	N	NLR	TLR
LVLR	23	12.2	0.8
HVLR	14	26.4	32.9

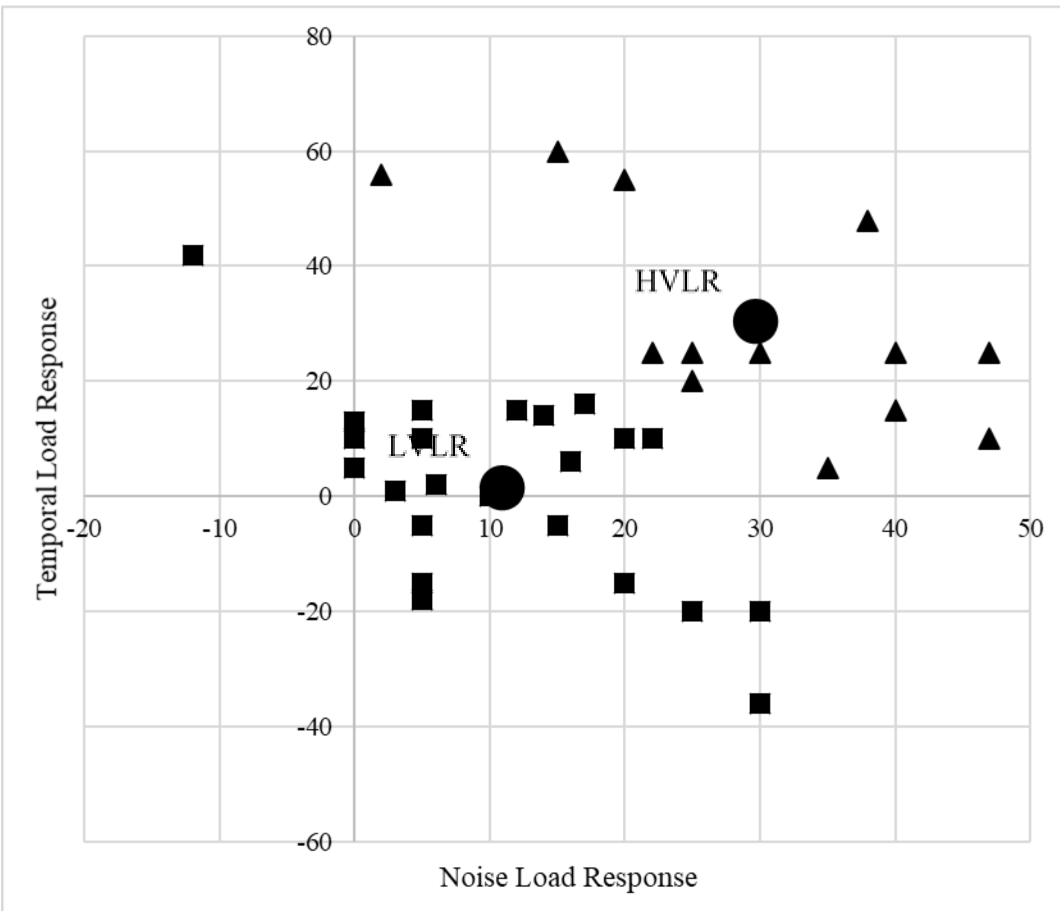

Figure 4.30 Scatter plot of clustered data with cluster centers. Square markers are the LVLR group, triangle markers are the HVLR group, and the circle markers are the cluster centers as labeled.

The analyses for VER and vocal loadings were repeated with the two groups. The measured VER for HVLR met the assumptions for normality and independence. However, equal variance cannot be assumed. For each time point of the vocal loading the VER for HVLR (M; SD) is as follows (summarized in Table 4.60 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 15.43 ; SD = 9.9), VL05; after 5 minutes of VLT (M = 41.79 ; SD = 12.4), VL10; after 10 minutes of VLT (M = 45.07 ; SD = 16.14), VL15; after 10 minutes of VLT (M = 53.93 ; SD = 17.23), VL20; after 10 minutes of VLT (M = 60.86 ; SD = 15.82), VL25; after 10 minutes of VLT (M = 67.43 ; SD = 14.84), VL30; after 10 minutes of VLT (M = 74.71 ; SD = 16.05), POST; after the complete VLT (M = 61.86 ; SD = 29), Figure 4.31 shows a line graph of the VER across the loading.

The measured VER for LVLR met the assumptions for normality, independence, and equal variance. For each time point of the vocal loading the VER for LVLR (M; SD) is as follows (summarized in Table 4.60 with 95% confidence interval, minimum, and maximum): PRE; before VLT (M = 18.13 ; SD = 15.95), VL05; after 5 minutes of VLT (M = 30.35 ; SD = 19.28), VL10; after 10 minutes of VLT (M = 31.87 ; SD = 19.42), VL15; after 10 minutes of VLT (M = 31.96 ; SD = 17.95), VL20; after 10 minutes of VLT (M = 31.43 ; SD = 18.35), VL25; after 10 minutes of VLT (M = 31.91 ; SD = 19.2), VL30; after 10 minutes of VLT (M = 31.13 ; SD = 19.99), POST; after the complete VLT (M = 26.43 ; SD = 16.37), Figure 4.31 shows a line graph of the VER across the loading separated by the cluster groups.

Table 4.60 Descriptive statistics for VER across time for HVLR and LVLR clusters for experiment 3

Vocal Effort Rating

Cluster	Time	Mean	Std. Deviation	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
HVLR	PRE	15.4	9.9	9.7	21.1	3.0	30.0
	VL05	41.8	12.4	34.6	48.9	13.0	65.0
	VL10	45.1	16.1	35.8	54.4	12.0	70.0
	VL15	53.9	17.2	44.0	63.9	35.0	85.0
	VL20	60.9	15.8	51.7	70.0	25.0	85.0
	VL25	67.4	14.8	58.9	76.0	45.0	90.0
	VL30	74.7	16.0	65.5	84.0	55.0	100.0
	POST	61.9	29.0	45.1	78.6	2.0	100.0
LVLR	PRE	18.1	15.9	11.2	25.0	2.0	60.0
	VL05	30.3	19.3	22.0	38.7	7.0	80.0
	VL10	31.9	19.4	23.5	40.3	8.0	79.0
	VL15	32.0	17.9	24.2	39.7	10.0	79.0
	VL20	31.4	18.4	23.5	39.4	7.0	83.0
	VL25	31.9	19.2	23.6	40.2	7.0	84.0
	VL30	31.1	20.0	22.5	39.8	2.0	85.0
	POST	26.4	16.4	19.4	33.5	2.0	75.0

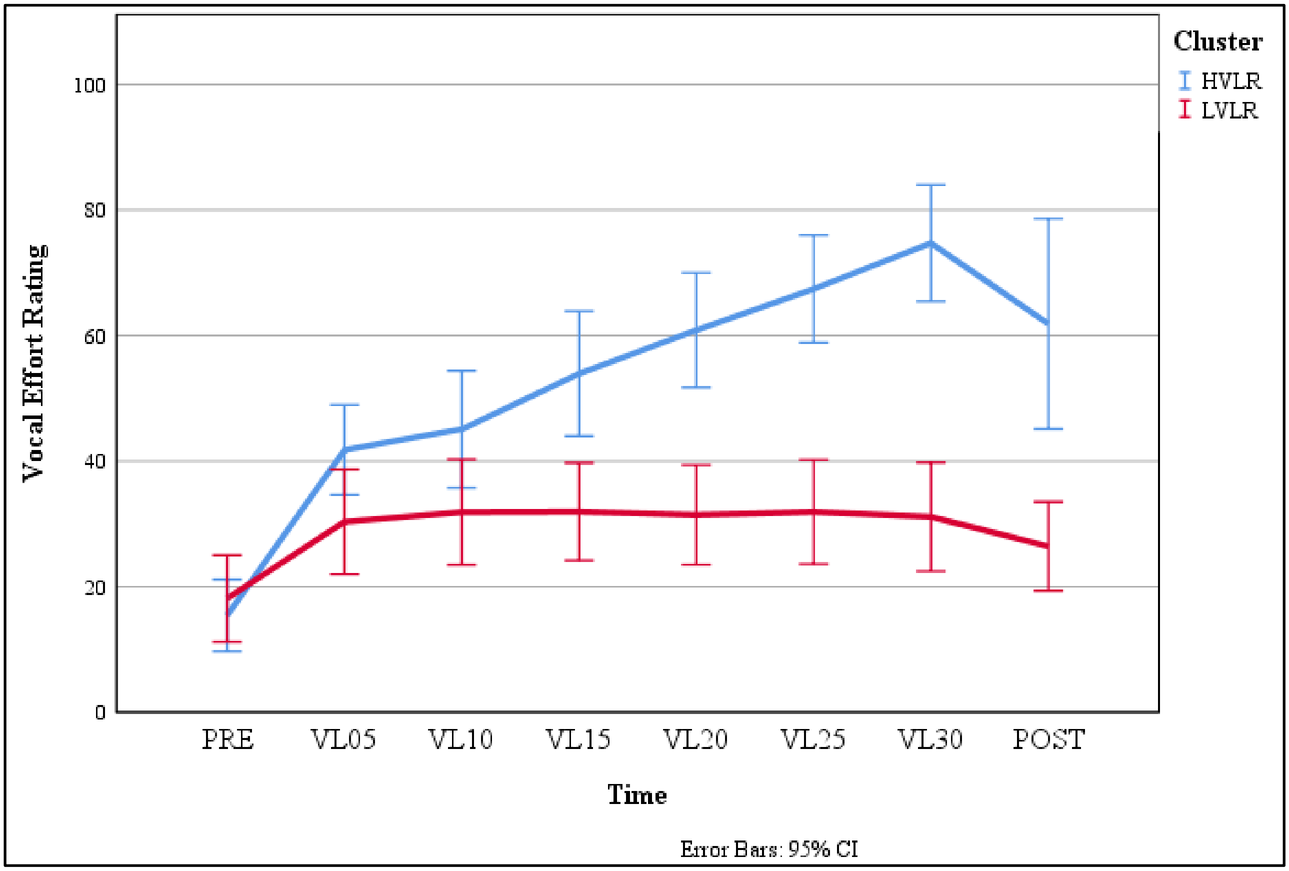

Figure 4.31 Line graph of VER over time separated clusters HVLR and LVLR for experiment 3

There was a significant ($p < 0.001$) main effect of VER across the VLT for HVLR. There was no significant effect of VER across the VLT for LVLR. For HVLR, there was an increase of VER of 26.4 ($p < 0.001$) from PRE to VL05, an increase of 32.9 ($p < 0.001$) from VEL30 to VEL05, and an increase of 46.4 ($p = 0.001$) between PRE and POST (summarized in Table 4.61).

Table 4.61 Multiple comparison statistics for VER across time for HVLR for experiment 3

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
PRE	VL05	-26.4	4.2	< 0.001	-41.2	-11.6
	VL10	-29.6	5.1	< 0.001	-47.6	-11.7
	VL15	-38.5	5.3	< 0.001	-57.5	-19.5
	VL20	-45.4	5.0	< 0.001	-63.1	-27.7
	VL25	-52.0	4.8	< 0.001	-68.8	-35.2
	VL30	-59.3	5.0	< 0.001	-77.2	-41.4
	POST	-46.4	8.2	0.001	-77.0	-15.9
VEL30	PRE	59.3	5.0	< 0.001	41.4	77.2
	VL05	32.9	5.4	< 0.001	14.0	51.9
	VL10	29.6	6.1	0.001	8.5	50.7
	VL15	20.8	6.3	0.075	-1.1	42.6
	VL20	13.9	6.0	0.570	-7.0	34.8
	VL25	7.3	5.8	0.999	-13.0	27.6
	POST	12.9	8.9	0.993	-18.9	44.6
POST	PRE	46.4	8.2	0.001	15.9	77.0
	VL05	20.1	8.4	0.558	-10.8	51.0
	VL10	16.8	8.9	0.879	-15.0	48.5
	VL15	7.9	9.0	1.000	-24.2	40.0
	VL20	1.0	8.8	1.000	-30.7	32.7
	VL25	-5.6	8.7	1.000	-37.0	25.8
	VL30	-12.9	8.9	0.993	-44.6	18.9

There was not a significant difference in VER between HVLR and LVLR at PRE or VL05. There were significant differences between HVLR and LVLR for the other times (summarized in Table 4.62). In each of these cases, the VER of HVLR were higher than LVLR. There was a

VER difference of 13.2 (p = 0.04) at VL10, 22.0 at VL15 (p = 0.001), 29.4 at VL20 (p < 0.001), 35.5 at VL25 (p < 0.001), 43.6 at VL30 (p < 0.001), and 35.4 at POST (p < 0.001).

Table 4.62 Independent sample comparison of VER across time between cluster groups HVLR and LVLR for experiment 3

	Mean Difference (HVLR-LVLR)	Std. Error Difference	95% Confidence Interval of the Difference		Sig. (2-tailed)
			Lower	Upper	
PRE	-2.7	4.7	-12.3	6.9	0.573
VL05	11.4	5.8	-0.3	23.2	0.056
VL10	13.2	6.2	0.6	25.8	0.040
VL15	22.0	6.0	9.8	34.1	0.001
VL20	29.4	5.9	17.4	41.4	< 0.001
VL25	35.5	6.0	23.3	47.7	< 0.001
VL30	43.6	6.3	30.8	56.4	< 0.001
POST	35.4	7.4	20.3	50.5	< 0.001

Acoustic Voice Change Clustering

There were 16 participants with significant (p < 0.05 for the model and for at least one acoustic covariate) general linear models (GLM) comparing the PRE-POST differences of the five vocal performance measures. The 21 participants without significant models were classified as a no voice change group (NVC) and the other 16 participants were classified as a voice change group (YVC). The average goodness-of-fit coefficient, Pearson's r, for YVC was 0.93 (SD = 0.27), since the GLM of NVC were not significant, therefore no goodness-of-fit coefficients are presented (summarized in Table 4.63)

Table 4.63 Descriptive statistics for Pearson's R for general linear model for PRE-POST acoustic changes in voice change group (YVC) for experiment 1

	Mean	Std. Deviation	Minimum	Maximum
R	0.93	0.27	0.84	0.98

Vocal Load Response and Acoustic Voice Change Clustering

The cross-sections of the clusters developed for vocal load response and acoustic voice change formed four groups, low vocal load response and no voice change (LVLR-NVC), low vocal load response and voice changes (LVLR-YVC), high vocal load response and no voice change (HVLR-NVC), and high vocal load response and voice changes (HVLR-YVC). There were 15 participants (10 males and 5 females) in LVLR-NVC, 8 participants (2 males and 6 females) in LVLR-YVC, 6 participants (3 males and 3 females) in HVLR-NVC, and 8 participants (3 males and 5 females) in HVLR-YVC (summarize in Figure 4.32).

Voice Change	Yes (YVC)	8	8
	No (NVC)	15	6
		Low (LVLR)	High (HVLR)
		Vocal Load Response	

Figure 4.32 Number of participants in each of the four vocal load response-voice change groups

The group descriptive statistics for PRE and POST the vocal loading task for vocal effort rating (VER), fundamental frequency (F0), fundamental frequency standard deviation (F0sd), speech level (SL), speech level standard deviation (SLsd), and smoothed cepstral peak prominence (CPPS) are summarized in Table 4.64.

Table 4.64 Descriptive statistics for all five acoustic voice measures across PRE-POST for each vocal load response-voice change groups for experiment 3

Time	Group		VER	F0 (ST)	F0sd (ST)	SL (dB)	SLsd (dB)	CPPS (dB)
PRE	LVLR-NVC	Mean	17.3	-0.01	2.01	57.12	3.54	13.15
		Std. Deviation	16.2	0.71	0.90	5.28	0.40	1.38
		Minimum	2	-1.31	0.90	45.68	2.44	9.66
		Maximum	60	2.03	4.61	68.16	4.76	16.75
	LVLR-YVC	Mean	19.8	-0.09	1.86	53.87	3.36	13.78
		Std. Deviation	16.4	0.76	1.04	3.91	0.39	1.48
		Minimum	3	-1.33	0.55	44.79	2.51	9.44
		Maximum	55	2.12	6.14	64.24	4.19	16.26
	HVLR-NVC	Mean	11.3	-0.01	1.88	53.20	3.35	12.31
		Std. Deviation	10.1	0.65	0.99	4.52	0.44	2.09
		Minimum	3	-1.20	0.76	44.26	2.32	8.61
		Maximum	30	1.73	5.01	59.98	4.27	15.75
	HVLR-YVC	Mean	18.5	0.00	2.03	57.44	3.55	13.54
		Std. Deviation	9.1	0.86	0.88	3.70	0.43	1.25
		Minimum	3	-1.84	0.69	51.90	2.70	10.16
		Maximum	30	2.21	5.61	65.73	4.70	16.09

Table 4.64 (cont'd)

POST	LVLR-NVC	Mean	26.1	0.83	1.98	57.92	3.56	13.55
		Std. Deviation	19.2	0.91	0.74	4.57	0.40	1.14
		Minimum	2	-1.53	0.88	44.83	2.54	10.07
		Maximum	75	2.60	4.46	66.18	4.60	16.10
	LVLR-YVC	Mean	27.1	0.04	2.22	55.56	3.39	13.32
		Std. Deviation	10.1	0.95	1.02	4.31	0.49	2.28
		Minimum	7	-1.84	0.92	45.38	2.50	8.99
		Maximum	35	2.10	5.32	64.28	4.56	17.56
	HVLR-NVC	Mean	56.5	1.08	2.37	55.53	3.55	13.03
		Std. Deviation	28.0	0.86	0.98	5.77	0.41	2.20
		Minimum	10	-0.70	1.00	45.38	2.49	9.53
		Maximum	89	2.48	4.83	66.52	4.18	16.18
	HVLR-YVC	Mean	65.9	0.79	2.46	58.91	3.82	13.00
		Std. Deviation	31	1.13	0.88	2.75	0.49	1.35
		Minimum	2	-2.22	1.26	54.84	2.84	9.73
		Maximum	100	2.59	5.48	67.52	5.04	14.92

Only F0 was significantly different from PRE to POST for LVLR-NVC, HVLR-NVC, and HVLR-YVC. Group LVLR-NVC had a POST F0 increase ($p < 0.001$) of 0.85 ST, group HVLR-NVC had a POST increase ($p < 0.001$) of 1.09 ST, and group HVLR-YVC had a post increase ($p < 0.001$) of 0.78 ST. In addition to F0, group HVLR-YVC had a significant POST F0sd increase ($p = 0.022$) of 0.42 ST, a POST SL increase ($p = 0.029$) of 1.48 dB, a POST SLsd increase of 0.28 dB ($p = 0.004$), and a POST CPPS decrease of 0.54 dB ($p = 0.046$). These comparisons are summarized in Table 4.65.

Table 4.65 Independent samples t-test comparing each of the five acoustic measures across PRE-POST for each vocal load response-voice change group for experiment 3

		Sig. (2-tailed)	Mean Difference (POST-PRE)	Std. Error Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LVLR-NVC	F0	< 0.001	0.85	0.13	-1.10	-0.60
	F0sd	0.866	-0.02	0.13	-0.24	0.29
	SL	0.284	0.80	0.75	-2.28	0.67
	SLsd	0.752	0.02	0.06	-0.15	0.11
	CPPS	0.056	0.39	0.20	-0.80	0.01
LVLR-YVC	F0	0.482	0.13	0.18	-0.49	0.23
	F0sd	0.112	0.36	0.22	-0.80	0.09
	SL	0.057	1.70	0.88	-3.45	0.05
	SLsd	0.771	0.03	0.09	-0.21	0.16
	CPPS	0.250	-0.46	0.40	-0.33	1.25
HVLR-NVC	F0	< 0.001	1.09	0.19	-1.48	-0.70
	F0sd	0.066	0.50	0.26	-1.02	0.03
	SL	0.081	2.33	1.31	-4.95	0.29
	SLsd	0.118	0.20	0.13	-0.45	0.05
	CPPS	0.190	0.73	0.55	-1.83	0.37
HVLR-YVC	F0	< 0.001	0.78	0.19	-1.17	-0.40
	F0sd	0.022	0.42	0.18	-0.79	-0.06
	SL	0.029	1.48	0.67	-2.80	-0.15
	SLsd	0.004	0.28	0.09	-0.46	-0.09
	CPPS	0.045	-0.54	0.27	0.01	1.07

CHAPTER V: DISCUSSION

This chapter provides the interpretation, implications, limitations, and future recommendations from the findings presented in Chapter IV. The chapter is outlined as follows, for each experiment a review of the hypotheses and research questions will be followed by a discussion of the interpretation of the significant findings, implications of how these support or do not support the research hypotheses, the limitations of the experimental methods, and future recommendations for further exploration. Following the discussions for the individual experiments is a general discussion from the perspective of all of the experimental results and the central hypothesis, H0.

H0: **The changes in vocal performance, vocal effort, and/or their interaction through a vocal load will implicate vocal fatigue.**

5.1 Experiment 1

The purpose of this experiment was to test hypothesis 1 (H1).

Q1: Can perceived vocal effort be measured reliably and if so, how does vocal performance in terms of vocal intensity change with vocal effort?

H1: Vocal performance in terms of vocal intensity will be distinct for each vocal effort level and be consistent within and across participants.

In context of the central hypothesis (H0), Experiment 1 aims to validate the use of the Borg CR-100 scale as an instrument to measure vocal effort to be used in future work of vocal effort through vocal loading. Since vocal effort is a psychophysical phenomenon, vocal production as the physical manifestation of vocal effort should directly relate to the psychological sensations of vocal effort. Therefore, the instrument will be considered valid if the voice production for different vocal effort levels are distinct and repeatable.

5.1.a Interpretations and Implications

There were significant increases of fundamental frequency (F0), speech level (SL), and smoothed cepstral peak prominence (CPPS) across the four cued vocal effort levels (VEL02 - minimal vocal effort, VEL13 - slight vocal effort, VEL25 - moderate vocal effort, VEL50 - severe vocal effort; from Borg CR-100 scale on Fig. 3.1). It was expected that SL would increase with vocal effort level. This expectation comes from previous experiments reporting increases of SL with vocal effort (Cushing et al., 2011; Rosenthal et al., 2014; Skinner et al., 1997). Additionally, the acoustic definition for vocal effort is the speech level at 1 meter (ISO, 2002), which supports the expectation of a direct connection between SL and vocal effort. Since vocal intensity measured by the speech level is distinct across the VELs, the first part of hypothesis 1 (H1, "Vocal performance in terms of vocal intensity will be distinct for each vocal effort level … across the participants") is accepted.

The significant changes of F0 and CPPS suggest that vocal intensity is not the only way talkers adjust their voices in response to increased vocal effort. Previous work (Jessen, Köster, & Gfroerer, 2005; McKenna & Stepp, 2018) has shown similar differences in F0 between conversational and raised vocal effort speech. Likewise, CPPS has been shown to increase with raised vocal effort (McKenna & Stepp, 2018; Rosenthal et al., 2014). These results further support that the effort levels elicited using the four levels from the Borg CR-100 scale relates directly with voice production associated with vocal effort.

There were significant changes in F0sd and SLsd with VEL50. This finding suggests that the speech changes associated with higher levels of vocal effort may go beyond changes due to slight and moderate vocal effort. This is a reasonable assumption as these measures have been shown to differ across severe conditions of the voice (V. Wolfe & Martin, 1997).

The test-retest reliability of SL was strong (R = 0.90). This finding supports the second part of hypothesis 1 (H1, "Vocal performance in terms of vocal intensity will be … consistent within participants"). This implies that a talker who self-calibrates to the scale, as was done in the experiment, would be able to reliably repeat VEL at about the same segmentation distance. As a result, a talker should be able to experientially anchor themselves to the scale to reliably rate their vocal effort for a range of voice tasks.

The results of this experiment validate the use of the Borg CR-100 scale to measure perceptual vocal effort. The VELs from the scale were distinct and repeatable in voice production which supports the psychophysical nature of vocal effort.

5.1.b Limitations

First, the study is limited through having only four vocal effort levels elicited. Although they were distinct, they only represent half the scale. Additionally, the VELs in between the ones in the study may be interesting. Second, the population consists of college-age adults. Although this allows for a more homogenous population to study, the generalizability of the results may be limited. While the differences in the groups are significantly different, this study cannot provide normative or expected values of vocal production with these levels of vocal effort due to the small sample size (N = 20). Lastly, the vocal effort scale was anecdotally anchored, in other words the participants were provided with an example of the extremes of the scale. Presumably if the participant had not experienced or could not properly imagine the example presented, then the scale would be different for that individual.

5.1.c Future Recommendations

The first recommendation for future applications of these findings is to scale up the experiment by the hundreds. As stated above (5.1.b) the study does not provide any normative values and does not provide differences within different populations. Future work can explore the various population factors (biological sex, age, voice impairment, hearing impairment, etc.) that could affect vocal production and vocal effort. Additionally, population normative values could benefit clinical assessment of vocal effort, as evaluation of normative deviation is a common clinical instrument in diagnosis and therapy progress. Since multiple acoustic measures of vocal performance related to VEL, future analyses could investigate the relative relationship between these measures and their contribution to the differences across the VELs.

5.2 Experiment 2

The purpose of this experiment was to test hypothesis 2 (H2).

Q2: To what degree are vocal performance and vocal effort related given three equivalent vocal load levels?

H2: The vocal performance and vocal effort will be constant within equivalent load conditions.

Experiment 2 contributes to understanding of how vocal effort and vocal performance are affected by vocal loading from different vocal loads. More specifically, this experiment tests whether three equivalent loads (vocal loads that should maintain acoustic energy equivalence) will elicit three distinct and equivalent vocal effort levels and vocal performance metrics.

5.2.a Interpretations and Implications

Communication Distance Vocal Load

For the condition of communication distance vocal load there were significant increases with distance for both F0 and SL. Only the most extreme condition, D04, saw increases in VER and CPPS. The changes in F0 were not necessarily expected, but the effects are small (about 0.4 ST per doubling of distances). There were expected differences in SL since it is natural to speak louder to someone further away, but the effects are much smaller than anticipated (about 1.4 dB per doubling of distance). The inverse-square law states that a doubling of distance from the sound source results in an attenuation of 6 dB of sound intensity. Similarly, by adjusting the measured SL for the distances (the calibration was at 50 cm, so the distances are 6 dB reduction at each doubling) it would be expected to see that the SL at 1 meter is 59.5 dB, at 2 meters is 54.9 dB, and at 4 meters is 50.4 dB. The 60 dB SPL at 1 meter is consistent with the ISO standard of normal vocal effort (ISO, 2002). Additionally, the VER at 1 meter was 25.1 which is "moderate vocal effort" on the Borg CR-100 scale. This connection further illustrates the utility of the Borg CR-100 scale in measurements of vocal effort as the normal vocal effort from the ISO is equivalent to the moderate vocal effort of the Borg CR-100 scale.

The lower distance-adjusted SL values suggest that individuals adjust beyond the normal (D01 is this case) enough to meet the needs of the new communication situation instead of making each situation equally intelligible. In other words, the perceived loudness would have decreased with distance (since the SL increased less than 6 dB per doubling of distance).

While there was an SL difference for the distance conditions, talkers did not experience the effort of the production the same way as there was only a significant change in VER between the baseline D01 and the extreme D04 conditions. However, the distance load response may have been surpassed by the "closeness" effect, a social phenomenon; Traunmüller and Eriksson (2000, 2011) noted a "closeness" effect in speech production where "differences between 0.3 m and 1.5

m was relatively small" as a result of a "habitual floor effect: Talkers appear to retain their habitual vocal effort for this [close] range of distances [between a communication partner] and increase it appreciably only when clearly required." This effect was also observed for CPPS which suggests that voice quality improvements are made to communicate at longer distances beyond the habitual floor.

Loudness Goal Vocal Load

For the condition of loudness goal vocal load there were significant increases from baseline to the three goal levels for VER, F0, SL, and CPPS. The first two goal levels (L54 and L60) were not significantly different in any parameter. The most extreme case of L66 has significant increases from L54 and L60 for VER, F0, SL, and CPPS. These findings suggest that there is a habitual floor effect with loudness goal of around L60. As participants needed to meet a loudness goal that was excessive to their habitual loudness, there were voice production accommodations beyond increase of SL.

Since the average SL for the baseline condition was 59.5 dB at 1 meter, the L54 and L60 should not require more effort than baseline. However, there is a difference between the baseline condition of D01 and the lower loudness goal conditions (L54 and L60) of VER, F0, SL, and CPPS. This suggests that although more vocal effort was not needed to meet the acoustic intensity requirements of the goal, participants perceived a vocal load associated with the goal. This has implications that individuals may adapt to a perceived vocal demand whether or not actual voice production changes are needed to communicate. One example of this is in telecommunication settings. If two individuals are communicating orally through an online medium (e.g. phone call, video conferencing), one may increase vocal effort and change their vocal production because they perceive a vocal load in the communication situation whether or not the other person is having difficulties hearing them. If the inverse is also true (individuals not reducing effort when a load is removed) then this has implications in amplification use in schoolteachers. Schoolteachers that use amplification but continue to perceive a vocal load would likely not benefit from the gains provided by the amplification system.

Background Noise Vocal Load

For the vocal load condition of background noise there were significant differences between all load conditions for VER, F0, and SL. The parameters of VER, F0, and SL followed similar patterns as the previous loads of increasing with load severity. However, the background noise did not have a habitual floor effect. Changes of F0 and SL to accommodate for background noise is called the Lombard effect and has been well studied. There was a prediction of 6 dB differences across the three vocal loads per the Lombard slope found in Bottalico 2017. This was not the case for this data. The average Lombard slope was 0.39 (dB of SL per 1 dBA of noise level) as opposed to the 0.65 per 1 dBA of noise level previously reported. However, the slope in the present study is between the two slopes identified by Bottalico (0.24, 0.65) and also fall within the range of Lombard slopes reported by Lazarus (0.3—0.6 dB per 1 dB of noise level). There two possible explanations of these difference. (1) The listener was 1 meter away as opposed to 2.5 m which might flatten the Lombard slope. (2) The speech tasks in the present study were map descriptions and not reading tasks. This additional cognitive load may have resulted in a dampened Lombard effect. There was also a Lombard effect of F0 resulting in an average change of 0.10 ST per dBA of noise level.

Another finding is that CPPS had significantly decreased in the extreme noise condition (N71). This is remarkable because in all previous scenarios (including experiment 1) CPPS has increased with vocal effort and loads. This suggests that there are vocal adaptations inherently different in voice response to background noise load than other loads. It could be the case that a goal of vocal quality is associated with increased vocal effort. This would be supported by experiment 1, where CPPS increased with effort when there was no vocal load present. Additionally, the vocal loads of communication distance and loudness goal could have innate vocal quality goal implications. This is opposed to background noise where the vocal quality goal is superseded by the Lombard effect. The lack of significant differences in CPPS for the lower background noise conditions could be a result of competing demands of voice quality and overcoming the noise interference, which is overcome when the noise becomes excessive. It reasonable that this is a learned reflex to optimize effort while communicating in background noise.

All Vocal Load Conditions

The findings support that vocal effort and vocal performance change with vocal load when the load is beyond habitual voice use. This validates the use of the Borg CR-100 scale and the acoustic measures to be used in experiment 3 to show changes in vocal effort and vocal production as a result of changing vocal load level. However, evidence was not found that supports hypothesis 2 (H2, “The vocal performance and vocal effort will be constant within equivalent load conditions”). An explanation for why the communication distance was not equivalent with the other vocal loads is that it was too limited in extent (room dimensions) to be at the same level as the other two. The inequivalence between loudness goal and background noise may relate to vocal ability, specifically previous experience with a vocal load. The loudness goal vocal load is not common and therefore most participants do not have trained reflex to the load. Conversely, speaking in background noise is extremely common resulting in an establish reflex (Lombard effect). This leads to the conclusion that experience with a vocal load may result in more consistent responses to the vocal loading (response to the vocal load) as evidenced by the differences in variance for the vocal effort ratings and acoustic measurements for the background noise load. Therefore, it is recommended to use a vocal load of background noise (and not loudness goal) for the most consistent vocal load responses, supporting the use of this load in experiment 3.

5.2.b Limitations

The most obvious limitation is that communication distance was limited to an extent of 4 meters. This was a space limitation. The other three vocal load conditions only had three variations and, in the case of loudness goal, the first two approached equivalence due to the habitual floor. Similar to experiment 1, the population consists of college-age adults limiting the generalization of the results. Another possible limitation is that the room was an anechoic chamber. This is not a typical acoustic space and it could have made the participants disoriented. In particular, the background noise from loudspeakers sounds different in a space without reverberance due to the lack of surface reflections. As a result, the loudspeakers sound like headphones (only hearing the direct sound) which could have affected the perception of the noise (some participants commented on this effect).

5.2.c Future Recommendations

In general, the study accomplished its design. Further work could be done to test larger extents of the vocal load conditions or investigate various factors that contributed to the results. For example, a future study in a much larger anechoic or hemi-anechoic chamber could test the effect of additional distances to see if the expected 6 dB change in SL would start at a larger distance. Another future iteration could include different speaking tasks (such as reading) or more complicated description tasks to observe how the cognitive load could influence the

Lombard response. Finally, testing different types of background noise (pink, talker babble, etc.) or acoustic environments and their effect on vocal loading would be beneficial.

The results suggest that experience with a vocal load may influence the response of the vocal load. This should be further tested through investigating the effect of familiarity and/or a training effect of the vocal load response.

5.3 Experiment 3

The purpose of this experiment was to test hypothesis 3 (H3).

Q3: To what degree do vocal performance, vocal effort, and/or their interaction change given a combined vocal load of excess background noise over time?

H3: The measured changes in vocal performance, vocal effort, and/or their interaction will change through a vocal load (background noise and prolonged speaking).

Using the tools developed in experiments 1 and 2, experiment 3 investigates the changes in vocal effort and vocal performance from prolonged speaking in background noise. If the observed changes in vocal effort, vocal performance and their interaction implicate vocal fatigue, then the central hypothesis can be accepted validating the proposed framework.

5.3.a Interpretations and Implications

The clearest changes in vocal effort and vocal performance are between the time before (PRE) the vocal loading task (VLT) and after 5 minutes (VL05) of the VLT. Here there are significant increases in each of the outcome variables except CPPS (VER, F0, F0sd, SL, SLsd). This is consistent with the Lombard effect observed in experiment 2. In that experiment there was an increase of 71 dBA of background noise, while here there was an increase of 75 dBA of background noise. However, there were no observed changes throughout the VLT. This was not expected. This suggests that there is a change in the manner of voicing to accommodate the noise and that does not change until the noise is removed. VER trended upward through the VLT but was not significant until clustering was performed. There were only PRE-POST differences of VER and F0 (and a small difference in F0sd). The PRE-POST increase of VER was expected and consistent with a majority of previous VLT studies. For example, the increased F0 has been shown in some studies and not in others. One possible explanation is that the increase could be from a warm-up effect. This effect has been shown in college students as a voice change throughout the day (Ben-David and Icht 2016), it has also been shown in schoolteachers across the workday (Rantala, Vilkman, Bloigu 2002). More changes in vocal production were expected between PRE and POST. Prior to clustering, the changes in VER and F0 are consistent with previously reported responses to vocal loading and vocal warmup but do not implicate vocal fatigue.

The VER clustering provided two distinct groups that had significantly different responses in VER to the vocal load. The two features that resulted from the clustering were the noise demand response and temporal demand response. These features are interesting because they directly relate to individual responses to the two different vocal loads presented (background noise load and prolonged speaking load). The fact that the group with high vocal load responses (HVLR) had large changes in VER, while the other group with low vocal load responses (LVLR) had no significant changes in VER suggests that there is a strong individual component in voice change from loading and therefore vocal fatigue. This is consistent with previous work that showed individual differences in vocal fatigue throughout vocal loading. This also provides insight in the contradictory nature of previous attempts for measuring vocal fatigue associated with vocal loading.

Although the VER clustering provided useful information, the second stage of acoustic clustering provides further clarity. Four groups of participants were formed and compared PRE and POST. Three of the four groups showed the same changes in F0 as seen by the aggregate subject pool. It is interesting to note that the group with low vocal load response and significant voice changes (LVLR-YVC) had no significant acoustic voice changes as a group. Further investigation reveals that the individual variation is very high in this group with the direction of the voice changes not being consistent (which would result in an aggregate of no change, e.g. an increase in F0 averaged with a decrease in F0 would result in a group average of no change in F0). It was the case that the acoustic changes for the group associated with high vocal load response and significant voice changes (HVLR-YVC) were all similar in direction resulting in statistically significant changes in all acoustic measures comparing PRE and POST. This finding supports hypothesis 3 (H3; "the measured changes in vocal performance, vocal effort, and/or their interaction will change through a vocal load, background noise load and temporal load") as the interaction between vocal effort and vocal performance created a group with significant changes in vocal performance and vocal effort.

5.3.b Limitations

Similar to the previous two experiments, the population consists of college-age adults limiting the generalization of the results. Although this study had more participants than most other vocal loading studies (Fujiki & Sivasankar, 2017), segmenting the population into four groups greatly lowers the power of the statistics. Still seeing statistical differences with the small groups is notable.

5.3.c Future Recommendations

As stated in the previous section (5.3.b Limitations), the study would benefit from more subjects. The study design and implementation were developed such that this study could be greatly scaled up for many more participants. One element of this is that the code was developed using the free python-based platform PsychoPy. This allows for the study to be run with identical instructions anywhere with a computer and the necessary hardware (microphones, speakers, etc.). Additionally, the presentation program was developed to have auto segmentation protocols allowing for fast processing of the data thus greatly reducing computation cost. In order to allow for this work to be done at a much larger scale, the scripts used to present the VLT are available on GitHub for the download and use of others (https://github.com/markolopolis/sVLT). Using similar VLT designs will allow for comparable work and an effective increase in sample size.

In addition to scaling up the protocol, psychological and physical measurements should be made of the participants to start to investigate potential correlations of the effect of vocal loading and individual characteristics (e.g. personality, vocal experience, etc.). Identifying these features of the participants could reveal potential risk factors for vocal fatigue leading to an enhanced understanding of vocal fatigue, providing the foundation to reduce its prevalence and impact.

5.4 Validation and Application of the Theoretical Framework

The primary motivation of the dissertation was to propose and test a theoretical framework for vocal fatigue. The framework was developed based on a literature review of vocal fatigue and the related concept of vocal effort. The framework is built around the concept that vocal fatigue (more specifically state fatigue) is the physiological and/or perceptual manifestation of a change in the voice that influences an individual’s intrinsic voice physiology, experience, and perception of vocal load for a particular voice-related communication situation which may be a result of vocal loading or vocal effort. Practically this framework models the concept that the vocal fatigue cannot be necessarily determined, but the changes in vocal performance, vocal effort, and/or their interaction through a vocal load can be determined and will implicate vocal fatigue. This is related to the central hypothesis (H0) of the dissertation.

The primary support of this framework is provided in part by H3. Here significant changes of vocal effort and vocal performance were measured through vocal loading as a result of a classifier from the interaction of vocal effort and vocal performance. The second part of validating the framework is whether these changes implicate vocal fatigue. The HVLR-YVC group showed significant changes in vocal effort and vocal performance that are consistent with possible changes associated with vocal fatigue (e.g. increase in variability of F0 and SL, decrease in CPPS). It reasonable to suggest that the participants in this group experience vocal fatigue.

Additionally, this framework provides theoretical relationships between vocal performance and vocal fatigue that provide a possible explanation for the previous work on vocal fatigue that have reported inconclusive or contradictory results. If multiple groups of individuals exist with varying levels of fatigability, then averaging all of these groups would result in null findings. Here subgroups (whose existence supports the framework) reveal a single group consisting of a little over 20% of the total participants to have significant differences in vocal effort and vocal performance. In other words, future applications will benefit from this classification of participants within a VLT that are most likely to have experienced vocal fatigue. This classification can be used as an independent variable to determine possible covariates associated with fatigability.

The other groups from the classification may also be interesting. Typical VLT studies have the assumption that the participants all belong in either LVLR-NVC or HVLR-YVC. In other words, there is a linear relationship assumed between voice production and vocal effort changes across vocal loading. The framework demonstrates that there are other subgroups that could be present in the data. One of these subgroups is a group that had significant voice changes but did not have a high vocal load response (LVLR-YVC). This group is interesting because it may be the case that these individuals are experiencing vocal fatigue but not feeling effort changes in their voice. As a result, they may be the group of individuals that do not take proper vocal rest when experiencing vocal fatigue (like the HVLR-YVC group may). This is consistent with the theory provided by Whitling et al. (2015) that saw a group of participants with extraordinary endurance in the VLT. They concluded that this over endurance group possibly share traits with patients in voice clinics. In other words, the repeated overuse of the voice without regulatory measures could be a risk factor for voice disorders. This may be the more important group to study.

As stated above, application of this framework would include using the classification as an independent variable to study the possible differences in the groups. Although the study lacks statistical power due to a low sample size for four groups, examples are presented below to show the application of the framework. The first example is the biological sex distribution in the four groups (Figure 5.1). The biggest differences are that there are 10:5 males to females in the LVLR-NVC group and 6-2 females to males in the LVLR-YVC group. This is interesting because, as stated above, LVLR-YVC is a group that is possibly associated with a higher risk of voice problems and the biological sex distribution here matches the fact that many more females report chronic voice disorders than males (Hunter, Tanner, & Smith, 2011; Roy et al., 2005).

			Vocal Load Response	
			Low (LVLR)	High (HVLR)
Voice Change	Yes (YVC)	Male	2	3
		Female	6	5
	No (NVC)	Male	10	3
		Female	5	3

Figure 5.1 Distribution of males and female participants across the vocal load response-voice change groups

A final example of applying the framework is an investigation of ratings of vocal fatigue. Prior to completing experiment 3, the participants completed the vocal fatigue index (VFI). Additionally, the participants also rated their perceived vocal fatigue on a visual analog scale (VAS) before and after the VLT. Looking at this data in the context of the groups, it is the case that the groups with voice change (LVLR-YVC and HVLR-YVC) have significantly higher ($p = 0.01$) scores of the second component of the VFI as compared to the group with no voice change (LVLR-NVC and HVLR-NVC). The voice change groups had a mean VFI-2 score of 3.64 (SD = 2.56), while the no voice change group had a mean score of 1.52 (SD = 2.04). The VFI-2 is "related to the physical discomfort associated with voicing." Additionally, the groups with high vocal load response (HVLR-NVC and HVLR-YVC) had a significantly higher ($p < 0.001$) POST vocal fating rating (POST-VFR) than the groups with a low vocal load response (LVLR-NVC and LVLR-HVC). The HVLR groups had a mean VFR of 0.70 (SD = 0.15) and the LVLR group had a mean POST-VFR of 0.39 (SD = 0.25). Additionally, there were no differences comparing PRE-VFR. These two findings show that the HVLR-YVC is contained in two different approaches to quantify vocal fatigue. It also informs how vocal fatigue is experienced differently between the four groups.

Future applications of the framework could provide distinct groups of fatigable individuals to study the factors associated with vocal fatigue. Additionally, the framework models the concept that an individual's potential to fatigue is based on their vocal load response. Using the methods of this dissertation, the changes in vocal load response associated with a vocal load could be measured to test potential benefits of therapeutic interventions for individuals with chronic vocal fatigue.

CHAPTER VI: CONCLUSION

In order to better understand vocal fatigue, a framework was proposed, and several underlying assumptions tested through experimentation. This framework models the concept that vocal fatigue can be implicated through measured changes in vocal effort, vocal performance, and/or their interaction through vocal loading (response to a vocal load). Towards supporting this framework, three experiments were conducted.

Experiment 1 illustrated how the Borg CR-100 scale could be used as a tool to measure perceived vocal effort ratings (VER) and quantify the relationships between vocal performance and vocal effort. The results of this experiment illustrated connections between fundamental frequency (F0), speech level (SL), and smoothed cepstral peak prominence (CPPS) and vocal effort level. Additionally, this experiment showed that the speech levels produced by the participants were reliably repeatable across elicited vocal effort levels.

Experiment 2 explored the relationships between vocal effort, vocal performance, and vocal loads. Using three different types of vocal loads, communication distance, loudness goal, and background noise, it was found that significant changes in VER, F0, SL, and CPPS existed for vocal loads that were beyond habitual communication experiences. In particular, the background noise vocal load showed the largest changes in VER and the acoustic parameters, suggesting a more refined vocal load response than the other vocal loads.

Experiment 3 investigated the effects of VER and vocal performance throughout a vocal loading test (VLT) consisting of a temporal load (prolonged speech task of describing complex routes on maps and a background noise load). Initially not many changes could be detected as a result of the VLT, which is consistent with previous research with VLT. The data was then clustered into four groups based on the effort ratings associated with the responses to the vocal loads of noise and time and the acoustic vocal performance changes resulted. This clustering revealed that one group of participants that experienced both high vocal load responses and significant voice changes had [as a group] significant changes in VER and each acoustic parameter of vocal performance (F0 mean and standard deviation, SL mean and standard deviation, and CPPS). This finding suggests that this group of individuals, as opposed to the other groups, experienced the most vocal fatigue, validating the framework.

The proposed framework for the study of vocal fatigue can be applied in future studies to examine potential risks associated with vocal fatigue. This framework provides an analytical approach not previously used to determine fatiguability in individuals. However, future work must be done to expand the capabilities and further test the potential applications of the framework.

APPENDICES

APPENDICES

APPENDIX A: Experiment 1 Stimuli

Automatic Speech Segments:

The alphabet: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z

Count from one to twenty-five: One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty-one, twenty-two, twenty-three, twenty-four, twenty-five

Days of the week and months of the year: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday. January, February, March, April, May, June, July, August, September, October, November, December

Reading Speech Segments:

Marvin Williams is only nine. Marvin lives with his mother on Monroe Avenue in Vernon Valley. Marvin loves all movies. Whenever a new movie is in the area, Marvin is in row one, along the aisle.

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch with its path high above.

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob.

Route Descriptions:

Route A: Describe how to get from Gresham to the Expo Center via Pioneer Square.

Route B: Describe how to get from Union Station to the Airport via the Rose Quarter.

Route C: Describe how to get from the Clackamas Towncenter to Beaverton via Gateway.

APPENDIX B: Experiment 2 Stimuli

Below are the maps and routes used for the communication task in experiment 2 (Figures B.1 through B.12).

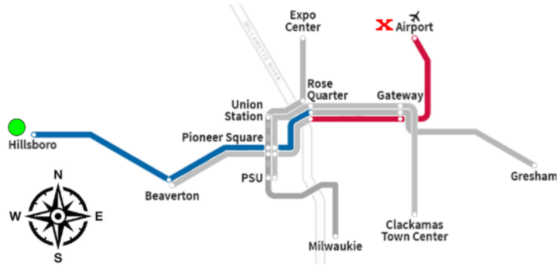

Figure B.1 Map used as example during tutorial

Figure B.2 Map used as practice during the tutorial

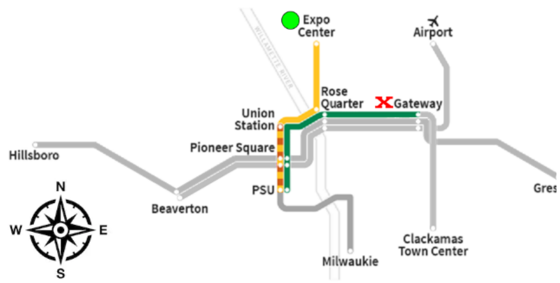

Figure B.3 Map used for D01, L60, and N71

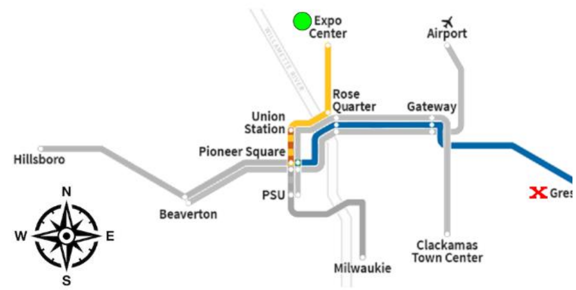

Figure B.4 Map used for D04, L60, and N53

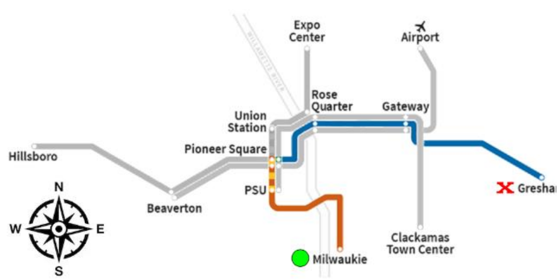

Figure B.5 Map used for D02, L66, and N53

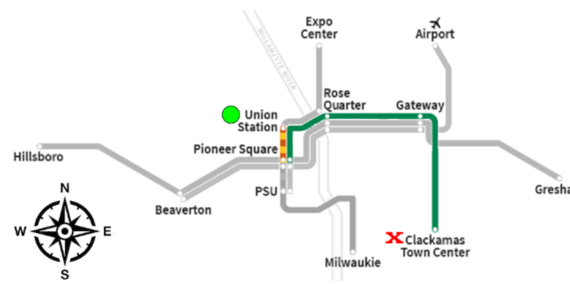

Figure B.6 Map used for D04, L54, and N62

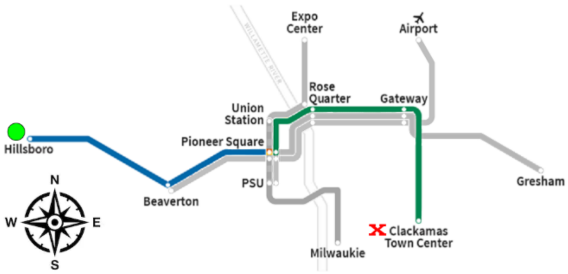

Figure B.7 Map used for D01, L60, and N71

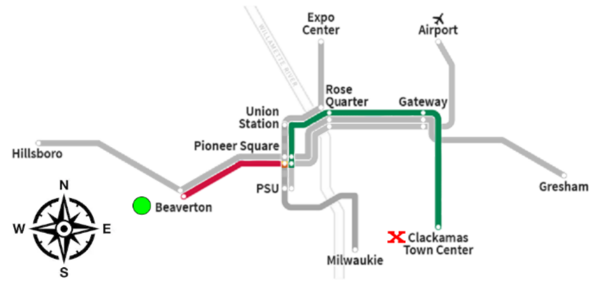

Figure B.8 Map used for D01, L66, and N62

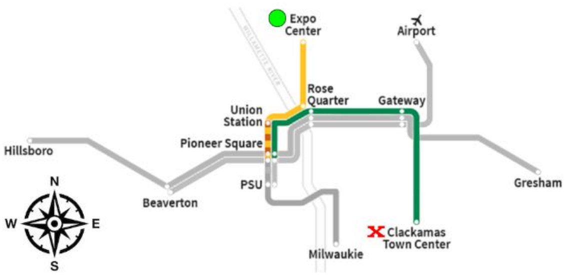

Figure B.9 Map used for D02, L54, N71

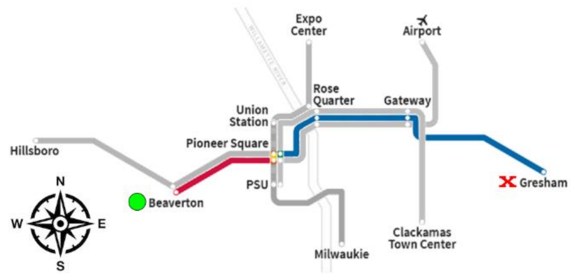

Figure B.10 Map used for D04, L54, N62

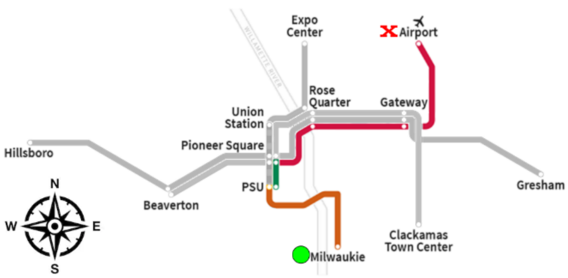

Figure B.11 Map used for D02, L66, and N53

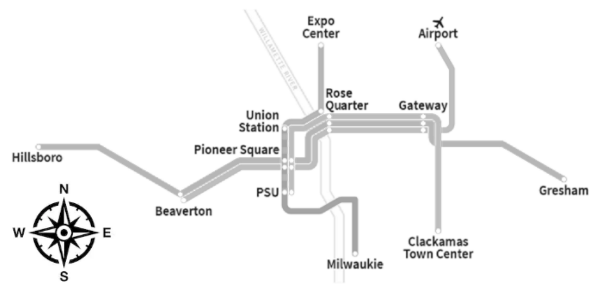

Figure B.12 Map used by communication partner

APPENDIX C: Experiment 3 Stimuli

Rainbow Passage:

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

Maps:

The images used for the map description tasks are below (Figures C.1 through C.25)

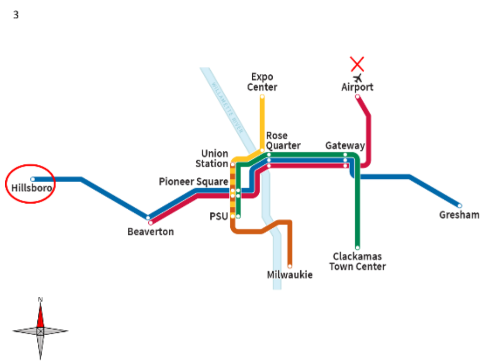

Figure C.1 Image used in the map description task during the vocal loading task for experiment 3

Figure C.2 Image used in the map description task during the vocal loading task for experiment 3

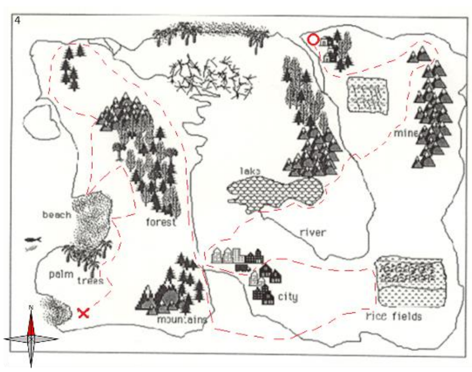

Figure C.3 Image used in the map description task during the vocal loading task for experiment 3

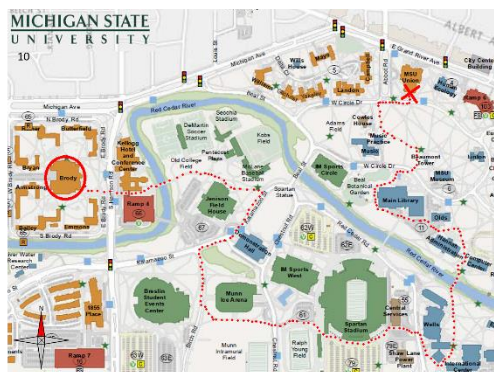

Figure C.4 Image used in the map description task during the vocal loading task for experiment 3

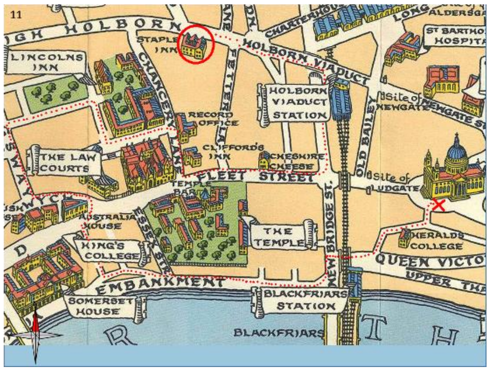

Figure C.5 Image used in the map description task during the vocal loading task for experiment 3

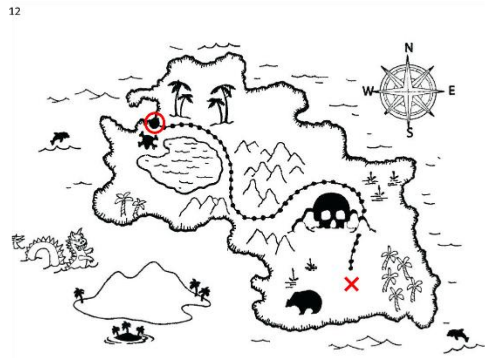

Figure C.6 Image used in the map description task during the vocal loading task for experiment 3

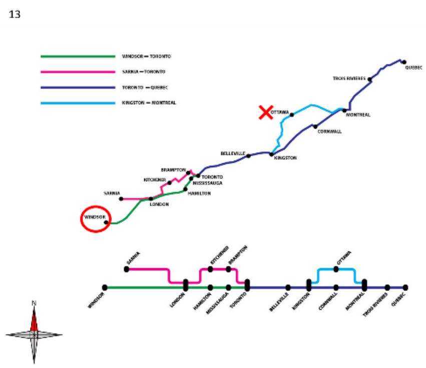

Figure C.7 Image used in the map description task during the vocal loading task for experiment 3

Figure C.8 Image used in the map description task during the vocal loading task for experiment 3

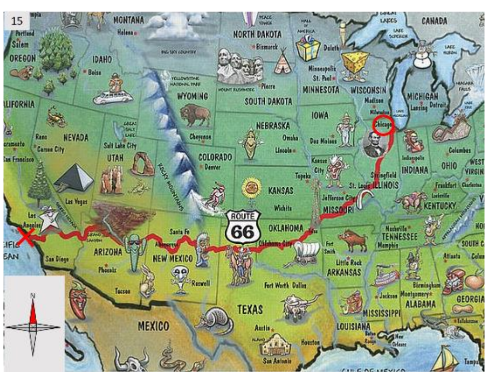

Figure C.9 Image used in the map description task during the vocal loading task for experiment 3

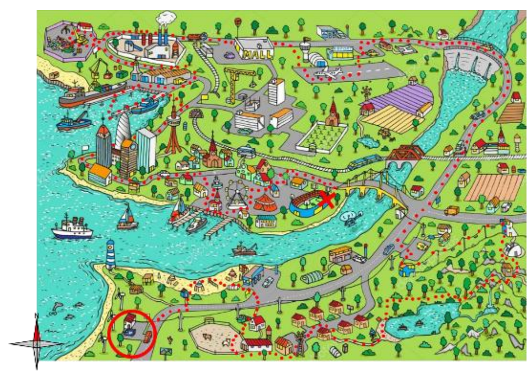
Figure C.10 Image used in the map description task during the vocal loading task for experiment 3

Figure C.11 Image used in the map description task during the vocal loading task for experiment 3

Figure C.12 Image used in the map description task during the vocal loading task for experiment 3

Figure C.13 Image used in the map description task during the vocal loading task for experiment 3

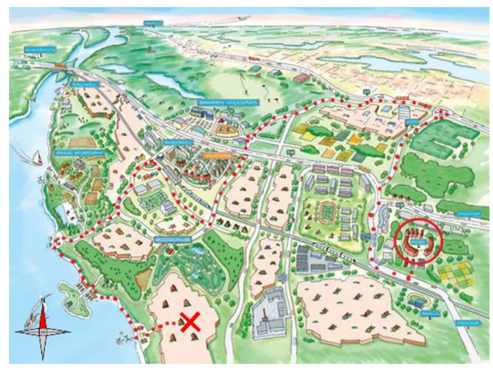
Figure C.14 Image used in the map description task during the vocal loading task for experiment 3

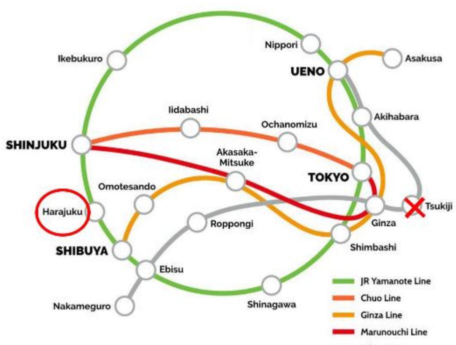

Figure C.15 Image used in the map description task during the vocal loading task for experiment 3

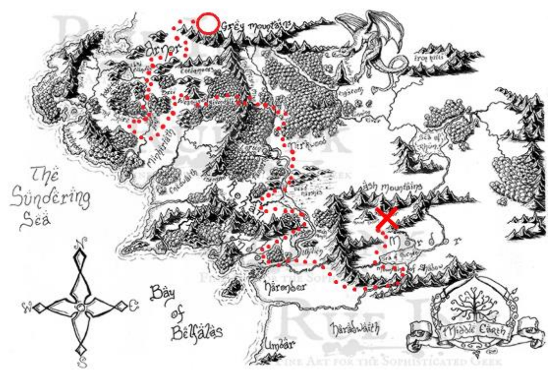

Figure C.16 Image used in the map description task during the vocal loading task for experiment 3

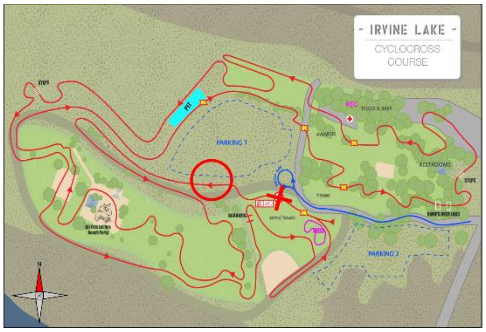

Figure C.17 Image used in the map description task during the vocal loading task for experiment 3

Figure C.18 Image used in the map description task during the vocal loading task for experiment 3

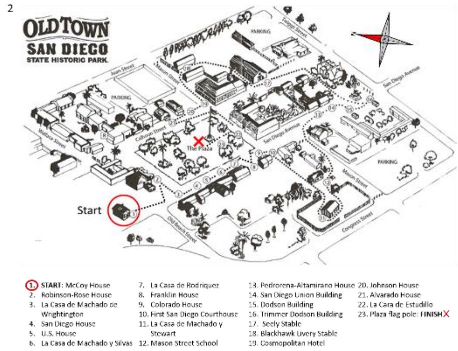

Figure C.19 Image used in the map description task during the vocal loading task for experiment 3

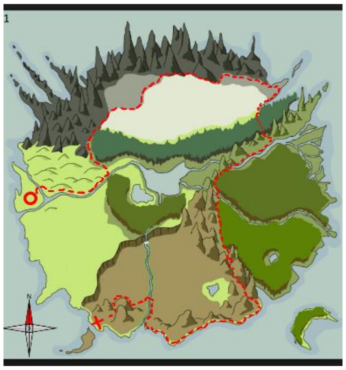

Figure C.20 Image used in the map description task during the vocal loading task for experiment 3

Figure C.21 Image used in the map description task during the vocal loading task for experiment 3

Figure C.22 Image used in the map description task during the vocal loading task for experiment 3

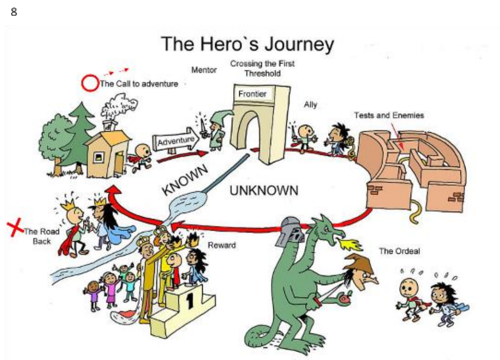

Figure C.23 Image used in the map description task during the vocal loading task for experiment 3

Figure C.24 Image used in the map description task during the vocal loading task for experiment 3

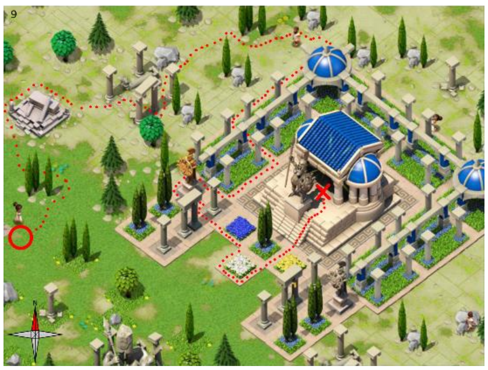

Figure C.25 Image used in the map description task during the vocal loading task for experiment 3

APPENDIX D: Informed Consent Forms

For experiments 1 and 2:

Research Participant Information and Consent Form

You are being asked to participate in a research study. Researchers are required to provide a consent form to inform you about the research study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions you may have.

Study Title: *Acoustic Recording of Voice and Speech Production*
Researcher and Title: Dr. Eric Hunter, Professor
Department and Institution: Department of Communicative Sciences and Disorders at Michigan State University
Address and Contact Information: 113 Oyer, East Lansing 48823, 517.353.8641
Sponsor: Michigan State University

1. PURPOSE OF RESEARCH
You are being asked to participate in a study examining measures of speech function as well as how speech production is affected by a variety of factors. This project is to examine voice production mechanisms with a three part goal of: (1) training students on general voice analysis procedures, (2) collecting pilot speech production data for potentially new ideas in understanding speech production, and (3) testing speech production hypothesis. This study is voluntary and you can stop at any time without penalty.

2. ELIGIBILITY CRITERIA
To participate, it is expected that:

- you are above the age of 18 years old
- you are in good physical and mental health
- you have no previous history of persistent and significant speech or voice complaints

You may be asked to confirm these criteria.

2. ALTERNATIVE OPTIONS
There are no alternative procedures, but you have the option not to participate in this research study.

3. WHAT YOU WILL DO
Participants will be asked to produce a range of voice and speech tasks while being recorded using a variety of sensors to investigate the coordination of speech production mechanisms (like breathing or mouth movement) and the final acoustic output (what can be heard). You may be asked to do a short hearing screen to ensure that your hearing is within normal limits.

If you agree to participate, you will perform several speech related tasks during a session lasting less than an hour. These tasks will be performed in a quiet room that may be unique in that it is built for recording speech. At the beginning of your participation, you will be asked to complete questionnaires about your voice use, vocal health, and personality: these may including the Voice

Handicap Index (VHI), the Vocal-Related Quality of Life (V-RQOL), the Vocal Fatigue Index (VFI), the Reflux Symptom Index (RSI), and the Big Five Inventory (BFI-10). You may also be administered a brief cognitive screening.

Next, we will take audio recordings and make measurements of your voice and speech. You will be asked to perform various speech/voice and breathing tasks. These speech tasks will be performed with recording equipment to ensure that high quality information is obtained from the observation, which will later be used for our analysis of your speech.

During the observation, we will ask you to perform several voice and speech tasks; some of these may be similar to what you would do in normal life while others may be more unique to you (creating funny sounds with your voice). Tasks may include reading printed materials, describing pictures, or problem-solving puzzles. Some of these tasks may be performed while wearing earphones through which you may hear your own voice; your voice may be unprocessed, mixed with noise, or processed to simulate a reverberant room. We may ask you to rate your vocal effort, vocal discomfort, quality of your voice, or general preference for a task. None of these tasks are designed to make you uncomfortable in anyway, though they may be unique and not what you would normally do with your voice in everyday life. These are used measure your speech range profile (which measures your usual speech loudness and pitch) or other speech production parameters.

4. POTENTIAL BENEFITS

While the program in which you are being asked to participate may have no immediate benefit for you, it may benefit others by increasing our knowledge of factors affecting measures of speech and vocal function.

5. POTENTIAL RISKS

There is no know medical risk involved in this research program and the procedures should not cause you any undue discomfort. Perhaps your voice will experience some fatigue at the end of participation.

Equipment which may be used are those found in singing studios, linguistic laboratories, and speech production laboratories. They include such items as microphones, vibration sensors on the neck, straps around the waist/torso to measure breath, and tubes to measure airflow in the mouth. Any tube that would go in your mouth is discarded at the end of your participation. If a sensor is placed on your neck, it will not use adhesive and is low-voltage (the voltage is so low that they cannot be felt). There are no known risks for these devices.

6. PRIVACY AND CONFIDENTIALITY

The data for this study are being collected confidentially. Neither the researchers nor anyone else will be able to link data to you. Data from this study will be stored in a secured location with limited access (locked cabinet in a locked room or a password protected computer in the locked laboratory). All information will be kept for at least three years after the close of the study. Only trained researchers under the jurisdiction of this project and Human Research Protection Program will have access to the data collected in the study. Information about you will be kept confidential to the maximum extent allowable by law. Although we will make every effort to keep your data confidential there are certain times, such as a court order, where we may have to disclose your data. Identifying information will not be attached to any of your individual responses when reporting results from the recordings or surveys. You will not be asked to give your name or any other information during the recording that will allow you or your place of employment to be identified. The results of this study may be published or presented at professional meetings, but the identities of all research participants will remain anonymous. By participating you agree to allow audio recordings which will be used for analysis.

We would like to ask your permission to use your recordings in other ways outside of what is presented above. Please mark below if you allow us to use your recordings: (1) to be presented, usually as an example in a scientific reports or presentations; and/or (2) to allow your recordings to be part of a larger dataset that researchers outside of the research team could access (e.g. public recording repository). In both cases, the recordings would be anonymous.

(1) I agree to allow my anonymous voice recordings to be presented in reports and presentations.
☐ Yes ☐ No Initials____________

(2) I agree to allow my anonymous voice recordings to be part of a larger dataset for others to use.
☐ Yes ☐ No Initials____________

7. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW

Participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time during the process without penalty or loss of benefits to which you are otherwise entitled. You may change your mind at any time and withdraw. <u>You may choose not to answer any question, or to not complete a specific task, or choose to stop participating at any time.</u> Whether you choose to participate or not will have no effect on your grade or evaluation.

8. COSTS AND COMPENSATION FOR BEING IN THE STUDY

As an incentive to participate, subjects/students who participate in this research will be allowed to earn CAS SONA credit for participation. Otherwise, no compensation or remuneration is implied. For those enrolled in courses that allow for CAS SONA credit, you may also find alternative assignments to earn extra credit if you choose not to participate in this research study. The CAS SONA system awards 1 credit per 1 hour of research participation. Neither researchers nor individual instructors will know what studies participants are involved in.

9. THE RIGHT TO GET HELP IF INJURED

If you are injured as a result of your participation in this research project, Michigan State University will assist you in obtaining emergency care, if necessary, for your research related injuries. If you have insurance for medical care, your insurance carrier will be billed in the ordinary manner. As with any medical insurance, any costs that are not covered or in excess of what are paid by your insurance, including deductibles, will be your responsibility. The University's policy is not to provide financial compensation for lost wages, disability, pain or discomfort, unless required by law to do so. This does not mean that you are giving up any legal rights you may have. You may contact Dr. Eric Hunter at 517.353.8641 with any questions or to report an injury.

10. Contact INFORMATION

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact Dr. Eric Hunter, Michigan State Univ, 113 Oyer, East Lansing, MI 48823, 517.353.8641, ejhunter@msu.edu.

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 4000 Collins Rd, Suite 136, Lansing, MI 48910.

11. DOCUMENTATION OF INFORMED CONSENT

Your signature below means that you voluntarily agree to participate in this research study.

__________________________________ ____________________

Signature Date

You will be given a copy of this form to keep.

Consent form for Experiment 3:

Research Participant Information and Consent Form

You are being asked to participate in a research study. Researchers are required to provide a consent form to inform you about the research study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions you may have concerning this project.

Study Title: *Gender, Age and Vocal Effort*
Researcher and Title: Dr. Eric Hunter, Associate Professor
Department and Institution: Department of Communicative Sciences & Disorders at Michigan State University
Address and Contact Information: 113 Oyer, East Lansing 48823, 517.353.8641
Sponsor: Michigan State University

1. PURPOSE OF RESEARCH
You are being asked to participate in this study to help researchers gain a better understanding of how fast the voice gets tired due to a long reading task. In this task you will speak until you are tired of speaking (less than 30 minutes). This study is voluntary.

2. ELIGIBILITY CRITERIA
It is expected that you have no significant vocal complaints and are in good physical and mental health. Persons with a history of recent hospitalization or suffering from any respiratory or oral infections will be excluded from participating. Additionally:

- You must be between 18-70 years of age.
- You must be a native English speaker.
- You will be asked about items which might affect your voice (e.g. hearing, heartburn, smoker)

3. ALTERNATIVE OPTIONS
There are no alternative procedures, but you have the option not to participate in this research study.

4. WHAT YOU WILL DO
We expect that full participation in the study will take between 60-90 minutes. After a brief introduction to the study, you will be asked to participate in a screening process (15-25 minutes) followed by a prolonged speaking task (no more than 30 minutes). Since it is common for your mouth to dry out while speaking for a long time, during the prolonged speaking task, you will be given the opportunity to take regular small drinks of water. However, so that all participants drink the same amount of water, we will use small measured cups and have you drink at a regular intervals. The total amount of water to drink is about the same amount as in a can of soda (less than 20 oz). During the screening process, you may be asked to do some or all of the following:

- Complete questionnaires about your voice and voice use.
- Answer questions about your vocal habits, hydration levels, and history of vocal fatigue.

- Asked about current medications which may affect voice use (e.g. asthma, allergies, heartburn).
- Complete a short hearing screen to ensure that your hearing is within the normal age appropriate ranges.
- Complete a breathing test using a spirometer to measure your lung function; this will be repeated three to five times (our goal is three similar breaths). You will be asked to breathe in deeply and blow into the spirometer followed by an inhalation.
- How you breathe while speaking will be observed, this is done by observation. We may additionally ask you to where a strap around your waist (outside your clothing) which detects when you breath.
- We may ask you to use a scale specifically designed to calculate body water percentage. The scale will also provide other readings such as body weight, body mass, etc. These measurements will only be used as they relate to hydration levels.

After the screening process, you will do a prolonged speech reading task in a soundproof room used for recordings. You will wear a microphone that goes loosely around your neck and a microphone that goes on your head (similar to headphones). We may additionally ask you to wear a strap around your waist (outside your clothing) which detects when you breathe. Before and after the long speaking task, you will be asked to perform some simple vocal tasks, such as a sustained "ah" vowel, pitch glides, and reading a passage. During the prolonged speaking task, you will be asked to read out loud for 30 minutes while noise is played in the room. This will likely result in you speaking at a loud volume. We expect that your voice will get tired before the 30 minutes is up. When you get tired of speaking, you may quit. Most people will go longer than 15 minutes and less than 30 minutes. We are interested to know when people get tired of speaking. Every few minutes you will be reminded to drink a small cup (less than 30 mL) of water.

5. POTENTIAL BENEFITS

While the program in which you are being asked to participate may have no immediate benefit for you, it may benefit others by increasing our knowledge of factors affecting measures of speech and vocal function.

6. POTENTIAL RISKS

There is minimal risk involved in this research program and the procedures should cause you no undue discomfort. The noise level, while annoying, is less than the occupational safety limits. Our goal is to have you speak at a louder volume until your voice is tired. Likely your voice will experience some vocal fatigue, but this should resolve with some nominal vocal rest. In rare cases, extended speaking can result in hoarseness. If you think your voice is getting too tired or if you become uncomfortable with the tasks, you may quit at any time.

Except for the spirometer, other devices to be used are similar to those found in singing studios, linguistic laboratories, and speech production laboratories. They include such items as microphones and surface microphones that go on the neck (to detect speech use in noise). An ear microphone and recorder may be used to record the sound you are surrounded by. While the pulmonary function test is unlikely to cause injury, breathing hard may cause some discomfort.

If there is anything in the screening that does not make you a good subject for our study, you will be remunerated for your time (see below) and no further participation is needed. If you are not healthy enough to participate (for example, if you have a cold or are hoarse from cheering at a sports activity), or if one of the screening procedures indicates that you might not match the level of communication function we are looking for (for example, your hearing is limited compared to peers age matched adults), you may be asked to not participate further. The testing performed in this project is not intended to find abnormalities, the protocol does not diagnose illness and we do not refer to health care providers. Data collected do not comprise a diagnostic or clinical study. Undetected vocal abnormalities are rare but it is possible that the investigators may perceive a vocal abnormality during the initial screening. If this occurs, you will be advised to consult with a licensed physician to determine whether a health examination would be prudent.

7. PRIVACY AND CONFIDENTIALITY

The data recorded for this study will be collected confidentially. Neither the researchers nor anyone else will be able to link data to you. The data for this project will be kept confidential. Data from this study will be stored in a locked cabinet in a locked room or a password protected computer in the locked laboratory. All information will be kept for at least three years after the close of the study. Only trained researchers under the jurisdiction of this project and Human Research Protection Program will have access to the data collected in the study. Information about you will be kept confidential to the maximum extent allowable by law. Although we will make every effort to keep your data confidential there are certain times, such as a court order, where we may have to disclose your data. Identifying information will not be attached to any of your individual responses or recordings when reporting results from the surveys. You will not be asked to give your name or any other information during the recording that will allow you or your place of employment to be identified. All results will be kept in a secure location accessible only to those involved in the study. The results of this study may be published or presented at professional meetings, but the identities of all research participants will remain anonymous. By participating, you agree to allow audio recordings of your speech.

8. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW

Participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. You have the right to say no. You may change your mind at any time and withdraw. You may choose not to answer specific questions or to stop participating at any time. Whether you choose to participate or not will have no effect on your grade or evaluation.

9. COSTS AND COMPENSATION FOR BEING IN THE STUDY

As an incentive to participate, subjects/students who participate in this research will be offered $15 per hour of participation (up to 2 hours or $30) or, if applicable, you can choose to earn extra credit through the MSU SONA software system. If participants are enrolled in a course that allows them to participate in a research study for credit, and the course accepts SONA credit, participants will have the option to receive MSU SONA credit instead of the cash remuneration.

For those enrolled in such courses, students can also find alternative assignments to earn extra credit if they choose not to participate in this research study but wish to earn extra credit. The SONA system awards 1 credit per 1 hour of research participation with a bonus of 0.25 for participating in person (up to 2.25 credits total). Within the SONA system, neither researchers nor individual instructors will know what studies participants are involved in. If your participation is over an hour, which it will likely be, we will compensate you in half hour increments (rounding up) for up to two hours total.

10. THE RIGHT TO GET HELP IF INJURED

In the unlikely event that you are injured as a result of participation in this project, Michigan State University will assist you in obtaining emergency care, if necessary, for your research related injuries. If you have insurance for medical care, your insurance carrier will be billed in the ordinary manner. As with any medical insurance, any costs that are not covered or in excess of what are paid by your insurance, including deductibles, will be your responsibility. MSU's policy is not to provide financial compensation for lost wages, disability, pain or discomfort, unless required by law to do so. This does not mean that you are giving up any legal rights you may have. Please contact Eric Hunter at 517-353-8641 with questions or to report an injury.

11. CONTACT INFORMATION

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher(s):

- Dr. Eric Hunter, Michigan State Univ, 113 Oyer, East Lansing, MI 48823, 517-353-8641, ejhunter@msu.edu
- Mark Berardi, Michigan State Univ, 110 Oyer, East Lansing, MI 48823, 517-353-8641, mberardi@msu.edu

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 4000 Collins Rd, Suite 136, Lansing, MI 48910.

12. DOCUMENTATION of Informed consent

Your signature below means that you voluntarily agree to participate in this research study.

____________________________________ __________________________

Signature

Date

You will be given a copy of this form to keep.

At times, it is useful to use recordings in teaching, presenting research, or future analysis. Therefore, we would like to ask for special permission to use your recordings in those contexts. Your identification would not be associated with the recording. If you do not give permission, it will not affect your ability to participate in the research. If you agree to allow your voice recordings (audio or video) in this way, please indicate:

Yes No Initials____________

BIBLIOGRAPHY

BIBLIOGRAPHY

Altman, K. W., Atkinson, C., & Lazarus, C. (2005). Current and emerging concepts in muscle tension dysphonia: A 30-month review. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2004.03.007

Bastian, R. W., Keidar, A., & Verdolini-Marston, K. (1990). Simple vocal tasks for detecting vocal fold swelling. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(05)80144-4

Bastian, R. W., & Thomas, J. P. (2016). Do Talkativeness and Vocal Loudness Correlate With Laryngeal Pathology? A Study of the Vocal Overdoer/Underdoer Continuum. *Journal of Voice*, *30*(5), 557–562. https://doi.org/10.1016/j.jvoice.2015.06.012

Ben-David, B. M., & Icht, M. (2016). Voice Changes in Real Speaking Situations during a Day, with and Without Vocal Loading: Assessing Call Center Operators. *Journal of Voice*, *30*(2), 247e1-247e11. https://doi.org/10.1016/j.jvoice.2015.04.002

Berardi, M. L., Hunter, E. J., & Leishman, T. W. (2015). Using the stability of vocal onsets to evaluate vocal effort in response to changing acoustical conditions. *J Acoust Soc Am*, *137*(4), 2433.

Berardi, M. L., Whiting, J. K., Eyring, N. G., Rollins, M. K., Jensen, Z. R., Hunter, E. J., & Leishman, T. W. (2015). Acuity to noise and occupational voice risks in simulated acoustical environments. *The Journal of the Acoustical Society of America*, *137*(4), 2393.

Borg, G., & Löllgen, H. (2001). Borg's perceived exertion and pain scales. *Deutsche Zeitschrift Fur Sportmedizin*.

Bottalico, P. (2017). Speech Adjustments for Room Acoustics and Their Effects on Vocal Effort. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2016.10.001

Bottalico, P., Cantor Cutiva, L. C., & Hunter, E. J. (2017). Vocal fatigue in virtual acoustics scenarios. *The Journal of the Acoustical Society of America*, *141*(5), 3541.

Bottalico, P., Graetzer, S., & Hunter, E. J. (2016). Effects of speech style, room acoustics, and vocal fatigue on vocal effort. *The Journal of the Acoustical Society of America*, *139*(5), 2870–2879. https://doi.org/10.1121/1.4950812

Bottalico, P., Graetzer, S., Hunter, E. J., Bottalico, P., Graetzer, S., & Hunter, E. J. (2017). Effects of speech style , room acoustics , and vocal fatigue on vocal effort, *2870*(2016).

Bottalico, P., Passione, I. I., Graetzer, S., & Hunter, E. J. (2017). Evaluation of the starting point of the lombard effect. *Acta Acustica United with Acustica*.

https://doi.org/10.3813/AAA.919043

Boucher, V. J. (2008). Acoustic correlates of fatigue in laryngeal muscles: findings for a criterion-based prevention of acquired voice pathologies. *Journal of Speech Language and Hearing Research*.

Boucher, V. J., Ahmarani, C., & Ayad, T. (2006). Physiologic features of vocal fatigue: Electromyographic spectral-compression in laryngeal muscles. *Laryngoscope*, *116*(6), 959–965. https://doi.org/10.1097/01.MLG.0000216824.07244.00

Boucher, V. J., & Ayad, T. (2010). Physiological Attributes of Vocal Fatigue and Their Acoustic Effects: A Synthesis of Findings for a Criterion-Based Prevention of Acquired Voice Disorders. *Journal of Voice*, *24*(3), 324–336. https://doi.org/10.1016/j.jvoice.2008.10.001

Boyas, S., & Guével, A. (2011). Neuromuscular fatigue in healthy muscle: Underlying factors and adaptation mechanisms. *Annals of Physical and Rehabilitation Medicine*. https://doi.org/10.1016/j.rehab.2011.01.001

Buekers, R. (1998). Are voice endurance tests able to assess vocal fatigue? *Clinical Otolaryngology and Allied Sciences*. https://doi.org/10.1046/j.1365-2273.1998.2360533.x

Calas, M., Verhulst, J., Lecoq, M., Dalleas, B., & Seilhean, M. (1989). Vocal pathology of teachers. *Revue de Laryngologie - Otologie - Rhinologie*.

Cannito, M. P., Doiuchi, M., Murry, T., Woodson, G. E., & York, N. (2012). Perceptual Structure of Adductor Spasmodic Dysphonia and Its Acoustic Correlates. *Journal of Voice*, *26*(6), 818.e5-818.e13. https://doi.org/10.1016/j.jvoice.2012.05.005

Cantor-Cutiva, L. C., Banks, R., Berardi, M., Johnson, B., Clawson, R., Martinez, S., & Hunter, E. (2018). From vocal effort to vocal fatigue. What does the literature say? In *11th International Conference on Voice Physiology and Biomechanics* (pp. 21–22).

Cantor-Cutiva, L. C., Bottalico, P., & Hunter, E. (2018). Work-related communicative profile of radio broadcasters: a case study. *Logopedics Phoniatrics Vocology*, *0*(0), 1–14. https://doi.org/10.1080/14015439.2018.1504983

Cantor Cutiva, L. C., & Burdorf, A. (2015). Medical Costs and Productivity Costs Related to Voice Symptoms in Colombian Teachers. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2015.01.005

Carroll, T., Nix, J., Hunter, E., Emerich, K., Titze, I., & Abaza, M. (2006). Objective measurement of vocal fatigue in classical singers: A vocal dosimetry pilot study. *Otolaryngology - Head and Neck Surgery*, *135*(4), 595–602. https://doi.org/10.1016/j.otohns.2006.06.1268

Chang, A., & Karnell, M. P. (2004). Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: A study of vocal fatigue. *Journal of Voice*, *18*(4), 454–466. https://doi.org/10.1016/j.jvoice.2004.01.004

Childers, D. G., Hicks, D. M., Moore, G. P., Eskenazi, L., & Lalwani, A. L. (1990). Electroglottography and Vocal Fold Physiology. *Journal of Speech Language and Hearing Research*. https://doi.org/10.1044/jshr.3302.245

Cho, S. W., Yin, C. S., Park, Y. B., & Park, Y. J. (2011). Differences in self-rated, perceived, and acoustic voice qualities between high- and low-fatigue groups. *Journal of Voice*, *25*(5), 544–552. https://doi.org/10.1016/j.jvoice.2010.07.006

Cipriano, M., Astolfi, A., & Pelegrín-García, D. (2017). Combined effect of noise and room acoustics on vocal effort in simulated classrooms. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.4973849

Claassen, H., & Werner, J. A. (1992). Fiber differentiation of the human laryngeal muscles using the inhibition reactivation myofibrillar ATPase technique. *Anatomy and Embryology*. https://doi.org/10.1007/BF00185983

Colton, R. H., Casper, J. K., & Leonard, R. (2011). *Understanding voice problem: A physiological perspective for diagnosis and treatment: Fourth edition*. *Understanding Voice Problem: A Physiological Perspective for Diagnosis and Treatment: Fourth Edition*. https://doi.org/10.1006/bbrc.2000.3788

Cooper, D. S., & Rice, D. H. (1990). Fatigue resistance of canine vocal fold muscle. *Annals of Otology, Rhinology and Laryngology*.

Cooper, Donald S., & Titze, I. R. (1985). Generation and Dissipation of Heat in Vocal Fold Tissue. *Journal of Speech Language and Hearing Research*. https://doi.org/10.1044/jshr.2802.207

Cushing, I. R., Li, F. F., Cox, T. J., Worrall, K., & Jackson, T. (2011). Vocal effort levels in anechoic conditions. *Applied Acoustics*. https://doi.org/10.1016/j.apacoust.2011.02.011

D’antona, G., Megighian, A., Bortolotto, S., Pellegrino, M. A., Ragona, R. M., Staffieri, A., … Reggiani, C. (2002). Contractile properties and myosin heavy chain isoform composition in single fibre of human laryngeal muscles. *Journal of Muscle Research & Cell Motility*, *23*(3), 187–195.

D’haeseleer, E., Behlau, M., Bruneel, L., Meerschman, I., Luyten, A., Lambrecht, S., … Van Lierde, K. (2016). Factors Involved in Vocal Fatigue: A Pilot Study.

Damsté, P. H. (1970). The phonetogram. *Practica Oto-Rhino-Laryngologica*, *32*(3), 185–187.

Davis, J. M. (1995). Central and peripheral factors in fatigue. *Journal of Sports Sciences*. https://doi.org/10.1080/02640419508732277

De Bodt, M. S., Wuyts, F. L., Van De Heyning, P. H., Lambrechts, L., & Abeele, D. Vanden.

(1998). Predicting vocal outcome by means of a vocal endurance test: A 5-year follow-up study in female teachers. *Laryngoscope*. https://doi.org/10.1097/00005537-199809000-00020

Deliyski, D. D. (1993). Acoustic Model And Evaluation of Pathological Voice Production, (September).

Doellinger, M., & Berry, D. A. (2006). Visualization and Quantification of the Medial Surface Dynamics of an Excised Human Vocal Fold During Phonation. *Journal of Voice*, *20*(3), 401–413. https://doi.org/https://doi.org/10.1016/j.jvoice.2005.08.003

Doellinger, M., Lohscheller, J., McWhorter, A., & Kunduk, M. (2009). Variability of Normal Vocal Fold Dynamics for Different Vocal Loading in One Healthy Subject Investigated by Phonovibrograms. *Journal of Voice*, *23*(2), 175–181. https://doi.org/10.1016/j.jvoice.2007.09.008

dos Santos, A. P., Silverio, K. C. A., Dassie-Leite, A. P., Costa, C. da C., & Siqueira, L. T. D. (2018). Relation Between Musculoskeletal Pain and Voice Self-Assessment in Tele-Operators. *Journal of Voice*, 1–11. https://doi.org/10.1016/j.jvoice.2018.07.006

Drechsel, J. S., & Thomson, S. L. (2008). Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model. *The Journal of the Acoustical Society of America*, *123*(6), 4434–4445. https://doi.org/10.1121/1.2897040

Eadie, T. L., & Stepp, C. E. (2013). Acoustic Correlate of Vocal Effort in Spasmodic Dysphonia, *122*(3), 169–176. https://doi.org/10.1177/000348941312200305

Edstrom, L., Lindquist, C., & Martensson, A. (1974). Correlation between functional and histochemical properties of the intrinsic laryngeal muscles in the cat. *Ventilatory and Phonatory Control Systems*, 392–404.

Enflo, L., Sundberg, J., & McAllister, A. (2013). Collision and Phonation Threshold Pressures Before and After Loud, Prolonged Vocalization in Trained and Untrained Voices. *Journal of Voice*, *27*(5), 527–530. https://doi.org/https://doi.org/10.1016/j.jvoice.2013.03.008

Erickson-Levendoski, E., & Sivasankar, M. (2011). Investigating the Effects of Caffeine on Phonation. *Journal of Voice*, *25*(5), e215–e219. https://doi.org/https://doi.org/10.1016/j.jvoice.2011.02.009

Eustace, C. S., Stemple, J. C., & Lee, L. (1996). Objective measures of voice production in patients complaining of laryngeal fatigue. *Journal of Voice*, *10*(2), 146–154. https://doi.org/10.1016/S0892-1997(96)80041-5

Fabron, E. M. G., Regaçone, S. F., Marino, V. C. de C., Mastria, M. L., Motonaga, S. M., & Sebastião, L. T. (2015). Self-perception, complaints and vocal quality among undergraduate students enrolled in a Pedagogy course. *CoDAS*, *27*(3), 285–291. https://doi.org/10.1590/2317-1782/20152014178

Faham, M., Jalilevand, N., Torabinezhad, F., Silverman, E. P., Ahmadi, A., Anaraki, Z. G., & Jafari, N. (2017). Relationship between Voice Complaints and Subjective and Objective Measures of Vocal Function in Iranian Female Teachers. *Journal of Voice*, *31*(4), 507.e1-507.e6. https://doi.org/10.1016/j.jvoice.2016.10.011

Fairbanks, F. (1960). The Rainbow Passage. In *Voice and Articulation Drillbook*.

Fanchini, M., Ferraresi, I., Modena, R., Schena, F., Coutts, A. J., & Impellizzeri, F. M. (2015). Use of the CR100 Scale for Session Rating of Perceived Exertion in Soccer and Its Interchangeability with the CR10. *International Journal of Sports Physiology and Performance*. https://doi.org/10.1123/ijspp.2015-0273

Fisher, K. V, & Swank, P. R. (1997). Estimating Phonation Threshold Pressure. *Journal of Speech, Language, and Hearing Research*, *40*(5), 1122–1129. Retrieved from http://dx.doi.org/10.1044/jslhr.4005.1122

Ford Baldner, E., Doll, E., & van Mersbergen, M. R. (2015). A Review of Measures of Vocal Effort With a Preliminary Study on the Establishment of a Vocal Effort Measure. *Journal of Voice : Official Journal of the Voice Foundation*, *29*(5), 530–541. https://doi.org/10.1016/j.jvoice.2014.08.017

Fujiki, R. B., Chapleau, A., Sundarrajan, A., McKenna, V., & Sivasankar, M. P. (2017). The Interaction of Surface Hydration and Vocal Loading on Voice Measures. *Journal of Voice*, *31*(2), 211–217. https://doi.org/10.1016/j.jvoice.2016.07.005

Fujiki, R. B., & Sivasankar, M. P. (2017). A Review of Vocal Loading Tasks in the Voice Literature. *Journal of Voice*, *31*(3), 388.e33-388.e39. https://doi.org/10.1016/j.jvoice.2016.09.019

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*. https://doi.org/10.1016/S0892-1997(05)80179-1

Gorham-Rowan, M., Berndt, A., Carter, M., & Morris, R. (2016). The effect of a vocal loading task on vocal function before and after 24 hours of thickened liquid use. *J. Speech Pathol. Ther*, *1*, 1–5.

Gotaas, C., & Starr, C. D. (1993). Vocal Fatigue Among Teachers. *Folia Phoniatrica et Logopaedica*. https://doi.org/10.1159/000266237

Guzmán, M., Malebrán, M. C., Zavala, P., Saldívar, P., & Muñoz, D. (2013). Acoustic Changes of the Voice as Signs of Vocal Fatigue in Radio Broadcasters: Preliminary Findings. *Acta Otorrinolaringologica (English Edition)*. https://doi.org/10.1016/j.otoeng.2013.06.011

Halpern, A. E., Spielman, J. L., Hunter, E. J., & Titze, I. R. (2009). The Inability to Produce Soft Voice (IPSV): A Tool to Detect Vocal Changes in School Teachers. *Logopedics Phoniatrics Vocology*. https://doi.org/10.1080/14015430903062712

Hamdan, A L, Sibai, A., & Rameh, C. (2006). Effect of Fasting on Voice in Women [epub ahead of print] [Record Supplied By Publisher]. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2006.01.009

Hamdan, Abdul Latif, Ziade, G., Kasti, M., Akl, L., Bawab, I., & Kanj, N. (2017). Phonatory Symptoms and Acoustic Findings in Patients with Asthma: A Cross-Sectional Controlled Study. *Indian Journal of Otolaryngology and Head and Neck Surgery*. https://doi.org/10.1007/s12070-016-1035-8

Hast, M. H. (1969). The primate larynx: A comparative physiological study of intrinsic muscles. *Acta Oto-Laryngologica*. https://doi.org/10.3109/00016486909124371

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (2005). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.411872

Hillman, R. E., Holmberg, E. V. A. B., & Perkell, J. S. (1989). OBJECTIVE ASSESSMENT OF V O C A L H Y P E R F U N C T I O N : EXPERIMENTAL FRAMEWORK AND INITIAL RESULTS, *32*(June), 373–392.

Hirano, M. (1981). GRBAS" scale for evaluating the hoarse voice & frequency range of phonation. *Clinical Examination of Voice*, *5*, 83–84.

Hoh, J. F. Y. (2005). Laryngeal muscle fibre types. *Acta Physiologica Scandinavica*. https://doi.org/10.1111/j.1365-201X.2004.01402.x

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1212–1223.

Holmberg, E., Ihre, E., & Södersten, M. (2007). Phonetograms as a tool in the voice clinic: Changes across voice therapy for patients with vocal fatigue. *Logopedics Phoniatrics Vocology*, *32*(3), 113–127. https://doi.org/10.1080/14015430701305685

Horáček, J., Laukkanen, A.-M., Šidlof, P., Murphy, P., & Švec, J. G. (2009). Comparison of Acceleration and Impact Stress as Possible Loading Factors in Phonation: A Computer Modeling Study. *Folia Phoniatrica et Logopaedica*, *61*(3), 137–145. https://doi.org/10.1159/000219949

Hunter, E. J., & Banks, R. E. (2017). Gender Differences in the Reporting of Vocal Fatigue in Teachers as Quantified by the Vocal Fatigue Index. *Annals of Otology, Rhinology and Laryngology*, *126*(12), 813–818. https://doi.org/10.1177/0003489417738788

Hunter, E. J., Bottalico, P., Graetzer, S., Leishman, T. W., Berardi, M. L., Eyring, N. G., … Whiting, J. K. (2015). Teachers and teaching: speech production accommodations due to changes in the acoustic environment. *Energy Procedia*, *78*, 3102.

Hunter, E. J., Maxfield, L., & Graetzer, S. (2019). The Effect of Pulmonary Function on the Incidence of Vocal Fatigue Among Teachers. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2018.12.011

Hunter, E. J., Tanner, K., & Smith, M. E. (2011). Gender differences affecting vocal health of women in vocally demanding careers. *Logopedics Phoniatrics Vocology*. https://doi.org/10.3109/14015439.2011.587447

Hunter, E. J., & Titze, I. R. (2009). Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise. *Annals of Otology, Rhinology and Laryngology*, *118*(6), 449–460. https://doi.org/10.1177/000348940911800608

Ilomäki, I., Kankare, E., Tyrmi, J., Kleemola, L., & Geneid, A. (2017). Vocal Fatigue Symptoms and Laryngeal Status in Relation to Vocal Activity Limitation and Participation Restriction. *Journal of Voice*, *31*(2), 248.e7-248.e10. https://doi.org/10.1016/j.jvoice.2016.07.025

ISO. (2002). ISO 9921:2002(E), Ergonomics-Assessment of Speech Communication. Geneva, Switzerland: ISO.

J., H. E., Catherine, C.-C. L., Eva, van L., Miriam, van M., Devie, N. C., Pasquale, B., … Susanna, W. (2020). Toward a Consensus Description of Vocal Effort, Vocal Load, Vocal Loading, and Vocal Fatigue. *Journal of Speech, Language, and Hearing Research*, *63*(2), 509–532. https://doi.org/10.1044/2019_JSLHR-19-00057

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI): development and validation. *American Journal of Speech-Language Pathology*, *6*(3), 66–70.

Jessen, M., Köster, O., & Gfroerer, S. (2005). Influence of vocal effort on average and variability of fundamental frequency. *International Journal of Speech, Language and the Law*. https://doi.org/10.1558/sll.2005.12.2.174

Jonsdottir, V., Laukkenen, A.-M., & Siiki, I. (2003). Changes in Teachers ’ Speech during a Working Day with and without Electric Sound Amplification. *Folia Phoniatrica e Logopaedica*, *601*, 282–287. https://doi.org/10.1159/000066149

Junqua, J. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.405631

Kagan, L. S., & Heaton, J. T. (2017). The Effectiveness of Low-Level Light Therapy in Attenuating Vocal Fatigue. *Journal of Voice, 31*(3), 384.e15-384.e23. https://doi.org/10.1016/j.jvoice.2016.09.004

Katch KI, K. V. M. W. (2009). *Exercise Physiology: Nutrition, Energy, and Human Performance*. *7th Edition*. https://doi.org/10.1161/01.STR.25.4.793

Kelchner, L., Toner, M. M., & Lee, L. (2006). Effects of prolonged loud reading on normal adolescent male voices. *Language, Speech & Hearing Services in Schools*. https://doi.org/http://dx.doi.org/10.1044/0161-1461(2006/012)

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kitch, J. A., & Oates, J. (1994). The perceptual features of vocal fatigue as self-reported by a group of actors and singers. *Journal of Voice*, *8*(3), 207–214. https://doi.org/10.1016/S0892-1997(05)80291-7

Kitch, J. A., Oates, J., & Greenwood, K. (1996). Performance effects on the voices of 10 choral tenors : acoustic and perceptual findings. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(96)80002-6

Kostyk, B. E., & Rochet, A. P. (1998). Laryngeal airway resistance in teachers with vocal fatigue: A preliminary study. *Journal of Voice*, *12*(3), 287–299. https://doi.org/10.1016/S0892-1997(98)80019-2

Koufman, J. A., & Blalock, D. (1988). Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall Syndrome. *Laryngoscope*. https://doi.org/10.1288/00005537-198805000-00003

Kristiansen, J., Lund, S. P., Persson, R., Shibuya, H., Nielsen, P. M., & Scholz, M. (2014). A study of classroom acoustics and school teachers' noise exposure, voice load and speaking time during teaching, and the effects on vocal and mental fatigue development. *International Archives of Occupational and Environmental Health*, *87*(8), 851–860. https://doi.org/10.1007/s00420-014-0927-8

Lagier, A., Vaugoyeau, M., Ghio, A., Legou, T., Giovanni, A., & Assaiante, C. (2010). Coordination between Posture and Phonation in Vocal Effort Behavior, 195–202. https://doi.org/10.1159/000314264

Laukkanen, A. M., Ilomäki, I., Leppänen, K., & Vilkman, E. (2008). Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers. *Journal of Voice*, *22*(3), 283–289. https://doi.org/10.1016/j.jvoice.2006.10.001

Laukkanen, A. M., & Kankare, E. (2006). Vocal loading-related changes in male teachers' voices investigated before and after a working day. *Folia Phoniatrica et Logopaedica*, *58*(4), 229–239. https://doi.org/10.1159/000093180

Laukkanen, A. M., Mäki, E., & Leppänen, K. (2009). Electroglottogram-based estimation of vocal economy: "Quasi-output-cost ratio." *Folia Phoniatrica et Logopaedica*, *61*(6), 316–322. https://doi.org/10.1159/000252847

Laukkenen, A.-M., Jarvinen, K., Artkoski, M., Waaramaa-Maki-Kulmala, T., Kankare, E., Sippola, S., … Salo, A. (2004). Changes in Voice and Subjective Sensations during a 45-min Vocal Loading Test in Female Subjects with Vocal Training. *Folia Phoniatrica e Logopaedica*.

Leanderson, R., & Sundberg, J. (1988). Breathing for singing. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(88)80051-1

Lehto, L., Laaksonen, L., Vilkman, E., & Alku, P. (2006). Occupational voice complaints and objective acoustic measurements - Do they correlate? *Logopedics Phoniatrics Vocology*, *31*(4), 147–152. https://doi.org/10.1080/14015430600654654

Lehto, L., Laaksonen, L., Vilkman, E., & Alku, P. (2008). Changes in Objective Acoustic Measurements and Subjective Voice Complaints in Call Center Customer-Service Advisors During One Working Day. *Journal of Voice*, *22*(2), 164–177. https://doi.org/10.1016/j.jvoice.2006.08.010

Leong, K., Hawkshaw, M. J., Dentchev, D., Gupta, R., Lurie, D., & Sataloff, R. T. (2013). Reliability of objective voice measures of normal speaking voices. *Journal of Voice*, *27*(2), 170–176. https://doi.org/10.1016/j.jvoice.2012.07.005

Lien, Y.-A. S., Michener, C. M., Eadie, T. L., & Stepp, C. E. (2015). Individual Monitoring of Vocal Effort With Relative Fundamental Frequency: Relationships With Aerodynamics and Listener Perception. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2015_jslhr-s-14-0194

Liénard, J.-S., & Di Benedetto, M.-G. (1999). Effect of vocal effort on spectral properties of vowels. *The Journal of the Acoustical Society of America*, *106*(1), 411–422.

Linville, S. E. (1995). Changes in glottal configuration in women after loud talking. *Journal of Voice*, *9*(1, Mar), 57–65. https://doi.org/https://doi.org/10.1016/S0892-1997(05)80223-1

Lombard, E. (1911). Le signe de l'élévation de la voix (translated from French). *Annales Des Maladies de l'oreille et Du Larynx*. https://doi.org/10.1145/1168987.1169028

Lourenço, B. M., Costa, K. M., & Da Silva Filho, M. (2014). Voice disorder in cystic fibrosis patients. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0096769

Lu, Y., & Cooke, M. (2009). Speech production modifications produced in the presence of low-pass and high-pass filtered noise. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.3179668

Maryn, Y., De Bodt, M., & Roy, N. (2010). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*. https://doi.org/10.1016/j.jcomdis.2009.12.004

Mascarello, F., & Veggetti, A. (1979). A comparative histochemical study of intrinsic laryngeal muscles of ungulates and carnivores. *Basic and Applied Histochemistry*.

Matsushita, H. (1975). The Vibratory Mode of the Vocal Folds in the Excised Larynx. *Folia Phoniatrica et Logopaedica*, *27*(1), 7–18. https://doi.org/10.1159/000263963

Maxfield, L. M., Hunter, E., & Greatzer, S. (2016). The effect of compromised pulmonary function on speech production among female school teachers. *The Journal of the Acoustical Society of America*, *139*(4), 2105. https://doi.org/10.1121/1.4950255

Mccabe, D. J., & Titze, I. R. (2002). Chant Therapy for Treating Vocal Fatigue Among Public School Teachers (2), *11*(November).

McKenna, V. S., & Stepp, C. E. (2018). The relationship between acoustical and perceptual measures of vocal effort. *The Journal of the Acoustical Society of America*, *144*(3), 1643–1658. https://doi.org/10.1121/1.5055234

McKenzie, D. C. (2012). Respiratory physiology: Adaptations to high-level exercise. *British Journal of Sports Medicine*. https://doi.org/10.1136/bjsports-2011-090824

Meynadier, Y., El Hajj, A., Pitermann, M., Legou, T., & Giovanni, A. (2018). Estimating Vocal Effort from the Aerodynamics of Labial Fricatives: A Feasibility Study. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2017.08.010

Mooshammer, C., & Mooshammer, C. (2010). Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German a), *1047*. https://doi.org/10.1121/1.3277160

Munier, C., & Kinsella, R. (2008). The prevalence and impact of voice problems in primary school teachers. *Occupational Medicine*, *58*(1), 74–76. https://doi.org/10.1093/occmed/kqm104

Naghibolhosseini, M., Deliyski, D. D., Zacharias, S. R. C., de Alarcon, A., & Orlikoff, R. F. (2018). Temporal Segmentation for Laryngeal High-Speed Videoendoscopy in Connected Speech. *Journal of Voice*, *32*(2), 256--e1.

Nanjundeswaran, C., Jacobson, B. H., Gartner-Schmidt, J., & Verdolini Abbott, K. (2015). Vocal Fatigue Index (VFI): Development and Validation. *Journal of Voice*, *29*(4), 433–440. https://doi.org/10.1016/j.jvoice.2014.09.012

Nanjundeswaran, C., van Mersbergen, M., & Morgan, K. (2017). Restructuring the Vocal Fatigue Index Using Mokken Scaling: Insights Into the Complex Nature of Vocal Fatigue. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2017.09.008

Neils, L. R., & Yairi, E. (1987). Effects of speaking in noise on vocal fatigue and vocal recovery. *Folia Phoniatrica et Logopaedica*. https://doi.org/10.1159/000265846

Niebudek-Bogusz, E., Kotyło, P., & Śliwińska-Kowalska, M. (2007). Evaluation of voice acoustic parameters related to the vocal-loading test in professionally active teachers with dysphonia. *International Journal of Occupational Medicine and Environmental Health*, *20*(1), 25–30. https://doi.org/10.2478/v10001-007-0001-9

Paes, S. M. (2017). Dosage dependent effect of high - resistance straw exercise in dysphonic and non - dysphonic women, *29*(1), 1–9. https://doi.org/10.1590/2317-1782/20172016048

Pearl Solomon, N., & Stemmle DiMattia, M. (2000). Effects of a vocally fatiguing task and systemic hydration on phonation threshold pressure. *Journal of Voice*, *14*(3), 341–362. https://doi.org/10.1016/S0892-1997(00)80080-6

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Pelegrín-García, D., Smits, B., Brunskog, J., & Jeong, C.-H. (2011). Vocal effort with changing talker-to-listener distance in different acoustic environments. *The Journal of the Acoustical Society of America*, *129*(4), 1981–1990.

Pellicani, A. D., Fontes, A. R., Santos, F. F., Pellicani, A. D., & Aguiar-Ricz, L. N. (2018). Fundamental Frequency and Formants Before and After Prolonged Voice Use in Teachers. *Journal of Voice*, *32*(2), 177–184. https://doi.org/10.1016/j.jvoice.2017.04.011

Pellicani, A. D., Ricz, H. M. A., & Ricz, L. N. A. (2015). Phonatory function after prolonged voice use in brazilian woman. *CoDAS*, *27*(4), 392–399. https://doi.org/10.1590/2317-1782/20152014201

Pette, D., & Staron, R. S. (1990). Cellular and Molecular Diversities of Mammalian Skeletal Muscle Fibers. *Rev. Physiol. Biochem. Pharmacol.* https://doi.org/10.1007/s00198-012-2236-y

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). In *Ear and Hearing*. https://doi.org/10.1097/AUD.0000000000000312

Poburka, B. J., Patel, R. R., & Bless, D. M. (2017). Voice-Vibratory Assessment With Laryngeal Imaging (VALI) form: reliability of rating stroboscopy and high-speed videoendoscopy. *Journal of Voice*, *31*(4), 513--e1.

Potvin, J. R., & Fuglevand, A. J. (2017). A motor unit-based model of muscle fatigue. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1005581

Rantala, L., Paavola, L., Körkkö, P., & Vilkman, E. (1998). Working-day effects on the spectral characteristics of teaching voice. *Folia Phoniatrica et Logopaedica*, *50*(4), 205–211. https://doi.org/10.1159/000021462

Remacle, A., Garnier, M., Gerber, S., David, C., & Petillon, C. (2018). Vocal Change Patterns During a Teaching Day: Inter- and Intra-subject Variability. *Journal of Voice*, *32*(1), 57–63. https://doi.org/10.1016/j.jvoice.2017.03.008

Rosen, C. A., & Simpson, C. B. (2008). *Operative Techniques in Laryngology*. Springer.

Rosenblum, L. D., Fowler, C. A., Gurskis, K., Liberman, A., Marks, L., Mattingly, I., … Summerfield, Q. (1991). Audiovisual Investigation of the Loudness-Effort Effect for Speech and Nonspeech Events, *17*(4), 976–985.

Rosenthal, A. L., Lowell, S. Y., & Colton, R. H. (2014). Aerodynamic and acoustic features of vocal effort. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2013.09.007

Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice Disorders in the General Population : Prevalence , Risk Factors , and Occupational Impact, (November), 1988–1995. https://doi.org/10.1097/01.mlg.0000179174.32345.41

Roy, N., Smith, M. E., Allen, B., & Merrill, R. M. (2007). Adductor Spasmodic Dysphonia Versus Muscle Tension Dysphonia : Examining the Diagnostic Value of Recurrent Laryngeal Nerve Lidocaine Block, *116*(3), 161–168.

Rubin, A. D., Jackson-Menaldi, C., Kopf, L. M., Marks, K., Skeffington, J., Skowronski, M. D., … Hunter, E. J. (2019). Comparison of Pitch Strength With Perceptual and Other Acoustic Metric Outcome Measures Following Medialization Laryngoplasty. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2018.03.019

Rubin, A. D., Praneetvatakul, V., Heman-Ackah, Y., Moyer, C. A., Mandel, S., & Sataloff, R. T. (2005). Repetitive phonatory tasks for identifying vocal fold paresis. *Journal of Voice*, *19*(4), 679–686. https://doi.org/10.1016/j.jvoice.2004.11.001

Sahgal, V., & Hast, M. H. (1974). Histochemistry of primate laryngeal muscles. *Acta Oto-Laryngologica*. https://doi.org/10.3109/00016487409126356

Sampaio, C., & Jos, E. (2012). Vocal Effort and Voice Handicap Among Teachers, *26*(6), 15–18. https://doi.org/10.1016/j.jvoice.2012.06.003

Sandage, M. J., Connor, N. P., & Pascoe, D. D. (2013). Voice function differences following resting breathing versus submaximal exercise. *Journal of Voice*, *27*(5), 572–578.

Sapienza, C. M., Crandell, C. C., & Curtis, B. (1999). Effects of sound-field frequency modulation amplification on reducing teachers' sound pressure level in the classroom. *Journal of Voice*, *13*(3), 375–381. https://doi.org/10.1016/S0892-1997(99)80042-3

Scherer, R. C., Titze, I. R., Raphael, B. N., Wood, R. P., Ramig, L. A., & Blager, R. F. (1987). Vocal fatigue in a trained and an untrained voice user. *Laryngeal Function in Phonation and Respiration*, 533–544.

Searl, J., & Knollhoff, S. (2018). Sense of Effort and Fatigue Associated With Talking After Total Laryngectomy. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/2018_AJSLP-17-0218

Shewmaker, M. B., Hapner, E. R., Gilman, M., Klein, A. M., & Johns, M. M. (2010). Analysis of Voice Change During Cellular Phone Use : A Blinded Controlled Study. *Journal of Voice*, *24*(3), 308–313. https://doi.org/10.1016/j.jvoice.2008.09.002

Shiotani, A., Westra, W. H., & Flint, P. W. (1999). Myosin heavy chain composition in human laryngeal muscles. *Laryngoscope*. https://doi.org/10.1097/00005537-199909000-00030

Shoffel-havakuk, H., Marks, K. L., Morton, M., Iii, M. M. J., & Hapner, E. R. (2019). Validation of the OMNI Vocal Effort Scale in the Treatment of Adductor Spasmodic Dysphonia, (February), 448–453. https://doi.org/10.1002/lary.27430

Sivasankar, M. (2002). Effects of vocal fatigue on voice parameters of Indian teachers. *Indian Journal of Otolaryngology and Head and Neck Surgery*, *54*(3), 245–247. https://doi.org/10.1007/BF02993116

Sivasankar, Mahalakshmi, Erickson, E., Schneider, S., & Hawes, A. (2008). Phonatory Effects of Airway Dehydration: Preliminary Evidence for Impaired Compensation to Oral Breathing in Individuals With a History of Vocal Fatigue. *Journal of Speech Language and Hearing Research*. https://doi.org/10.1044/1092-4388(2008/07-0181)

Skinner, M. W., Holden, L. K., Holden, T. A., Demorest, M. E., Fourakis, M. S., Skinner, M. W., … Demorest, M. E. (1997). by adults with cochlear implants Speech recognition at simulated soft , conversational , and raised-to-loud vocal efforts by adults with cochlear implants, *3766*. https://doi.org/10.1121/1.418383

Sluijter, A. M. C., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, *100*(4), 2471–2485.

Smith, E., Mendoza, M., Barkmeier, J., Lemke, J., Hoffman, H., Smith, E., & Al, E. T. (1998). Spasmodic Dysphonia and Vocal Fold Paralysis : Outcomes of Voice Problems on Work-Related Functioning, *12*(2), 223–232.

Smith, M. E., Roy, N., Wilson, C., & Hypothesis, O. (2006). Lidocaine Block of the Recurrent Laryngeal Nerve in Adductor Spasmodic Dysphonia : A Multidimensional Assessment, (April), 591–595. https://doi.org/10.1097/01.MLG.0000205588.04450.AC

Smith, N. R., Rivera, L. A., Dietrich, M., Shyu, C. R., Page, M. P., & DeSouza, G. N. (2016). Detection of Simulated Vocal Dysfunctions Using Complex sEMG Patterns. *IEEE Journal of Biomedical and Health Informatics*, *20*(3), 787–801. https://doi.org/10.1109/JBHI.2015.2490087

Södersten, M., Granqvist, S., Hammarberg, B., & Szabo, A. (2002). Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings.

Journal of Voice, *16*(3), 356–371. https://doi.org/10.1016/S0892-1997(02)00107-8

Solomon, N. P. (2006). What is orofacial fatigue and how does it affect function for swallowing and speech? *Seminars in Speech and Language*. https://doi.org/10.1055/s-2006-955117

Solomon, N. P. (2008). Vocal fatigue and its relation to vocal hyperfunction. *International Journal of Speech-Language Pathology, 10*(4), 254–266. https://doi.org/10.1080/14417040701730990

Solomon, N. P., Glaze, L. E., Arnold, R. R., & van Mersbergen, M. (2003). Effects of a vocally fatiguing task and systemic hydration on men's voices. *Journal of Voice*, *17*(1), 31–46.

Spencer, M., Siegmund, T., & Mongeau, L. (2008). Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation. *The Journal of the Acoustical Society of America, 123*(2), 1089–1103. https://doi.org/10.1121/1.2821412

Stemple, J. C., Stanley, J., & Lee, L. (1995). Objective measures of voice production in normal subjects following prolonged voice use. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(05)80245-0

Stepp, C. E., Sawin, D. E., & Eadie, T. L. (2012). The Relationship Between Perception of Vocal Effort and Relative Fundamental Frequency During Voicing Offset and Onset. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2012/11-0294)

Švec, J. G., Titze, I. R., & Popolo, P. S. (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.1850074

Tanner, K., Merrill, R. M., Houtz, D. R., Sauder, C., Elstad, M., & Wright-costa, J. (2010). Nebulized Isotonic Saline Versus Water Following a Laryngeal Trained Sopranos, *53*(December).

Tao, C., & Jiang, J. J. (2007). Mechanical stress during phonation in a self-oscillating finite-element vocal fold model. *Journal of Biomechanics, 40*(10), 2191–2198. https://doi.org/10.1016/j.jbiomech.2006.10.030

Tepe, E. S., Deutsch, E. S., Sampson, Q., Lawless, S., Reilly, J. S., & Sataloff, R. T. (2002). A pilot survey of vocal health in young singers. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(02)00093-0

Ternström, S., Bohman, M., & Södersten, M. (2006). Loud speech over noise: Some spectral attributes, with gender differences. *The Journal of the Acoustical Society of America, 119*(3), 1648–1665.

Thomas-kersting, C., & Casteel, R. L. (1989). HARSH VOICE : VOCAL EFFORT

PERCEPTUAL RATINGS AND SPECTRAL NOISE LEVELS OF HEARING-IMPAIRED CHILDREN, *22*, 125–135.

Timmermans, B., De Bodt, M., Wuyts, F., & Van de Heyning, P. (2003). Vocal hygiene in radio students and in radio professionals. *Logopedics Phoniatrics Vocology*, *28*(3), 127–132. https://doi.org/10.1080/14015430310018333

Titze, I., Lemke, J., & Montequin, D. (1997). Populations in the U.S. workforce who rely on voice as a primary tool of trade: A preliminary report. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(97)80002-1

Titze, I. R. (1992). Acoustic interpretation of the voice range profile (phonetogram). *Journal of Speech, Language, and Hearing Research*, *35*(1), 21–34.

Titze, I. R. (1994). Mechanical stress in phonation. *Journal of Voice*. https://doi.org/10.1016/S0892-1997(05)80302-9

Titze, I. R. (1999). Toward occupational safety criteria for vocalization. *Logopedics Phoniatrics Vocology*. https://doi.org/10.1080/140154399435110

Titze, I. R., Hunter, E. J., & Švec, J. G. (2007). Voicing and silence periods in daily and weekly vocalizations of teachers. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.2390676

Traunmüller, H., & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.429414

Traunmüller, H., & Eriksson, A. (2011). Acoustic effects of variation in vocal effort by men , women , and children Acoustic effects of variation in vocal effort by men , women , and children, *3438*(2000). https://doi.org/10.1121/1.429414

Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., … Kavukcuoglu, K. (2016). WaveNet: A generative model for raw audio. *SSW*, *125*.

van Leer, E., & van Mersbergen, M. (2017). Using the Borg CR10 Physical Exertion Scale to Measure Patient-perceived Vocal Effort Pre and Post Treatment. *Journal of Voice : Official Journal of the Voice Foundation*, *31*(3), 389.e19-389.e25. https://doi.org/10.1016/j.jvoice.2016.09.023

Verdolini, K., & Ramig, L. O. (2001). Review: Occupational risks for voice problems. *Logopedics Phoniatrics Vocology*. https://doi.org/10.1080/14015430119969

Verdolini, Katherine, Chan, R., Titze, I. R., Hess, M., & Bierhals, W. (1998). Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges. *Journal of Voice*, *12*(4), 415–423. https://doi.org/https://doi.org/10.1016/S0892-1997(98)80050-7

Verdolini, Katherine, Titze, I. R., & Fennell, A. (1994). Dependence of phonatory effort on hydration level. *Journal of Speech, Language, and Hearing Research*, *37*(5), 1001–1007.

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 et Logopaedica*. https://doi.org/10.1159/000266266

Vilkman, E. (2004). Occupational safety and health aspects of voice and speech professions... 28th World Congress of the International Association of Logopedics and Phoniatrics, Brisbane, Australia, 29th August to 2nd September 2004. *Folia Phoniatrica et Logopaedica*.

Vilkman, Erkki, Lauri, E.-R., Alku, P., Sala, E., & Sihvo, M. (1999). Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *Journal of Voice*, *13*(2), 303–312. https://doi.org/https://doi.org/10.1016/S0892-1997(99)80036-8

Vintturi, J., Alku, P., Lauri, E. R., Sala, E., Sihvo, M., & Vilkman, E. (2001). The effects of post-loading rest on acoustic parameters with special reference to gender and ergonomic factors. *Folia Phoniatrica et Logopaedica*. https://doi.org/10.1159/000052687

Vogel, A. P., Fletcher, J., Snyder, P. J., Fredrickson, A., & Maruff, P. (2011). Reliability, stability, and sensitivity to change and impairment in acoustic measures of timing and frequency. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2009.09.003

Warrick, P., Dromey, C., Irish, J. C., Durkin, L., Pakiam, A., & Lang, A. (2000). Botulinum Toxin for Essential Tremor of the Voice With Multiple Anatomical Sites of Tremor : A Crossover Design Study of Unilateral Versus Bilateral Injection, (August), 1366–1374.

Weinberger, S. H., & Kunath, S. A. (2011). The Speech Accent Archive: Towards a typology of English accents. In *Language and Computers*. https://doi.org/10.1163/9789401206884_014

Welham, N. V., & Maclagan, M. A. (2003). Vocal fatigue: Current knowledge and future directions. *Journal of Voice*, *17*(1), 21–30. https://doi.org/10.1016/S0892-1997(03)00033-X

Wenke, R. J., Goozee, J. V, Murdoch, B. E., & LaPointe, L. L. (2006). Dynamic assessment of articulation during lingual fatigue in myasthenia gravis. *Journal of Medical Speech-Language Pathology*, *14*(1), 13–32.

Whiting, J. K., Leishman, T. W., Eyring, N. G., Berardi, M. L., & Rollins, M. K. (2015). Evaluation of a real-time convolution system for perception of self-generated speech in simulated rooms. *The Journal of the Acoustical Society of America*, *138*(3), 1900.

Whitling, S., Lyberg-Åhlander, V., & Rydell, R. (2017). Recovery From Heavy Vocal Loading in Women With Different Degrees of Functional Voice Problems. *Journal of Voice*, *31*(5), 645.e1-645.e14. https://doi.org/10.1016/j.jvoice.2016.12.012

Whitling, S., Rydell, R., & Lyberg Åhlander, V. (2015). Design of a clinical vocal loading test with long-time measurement of voice. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2014.07.012

Willis, J., Michael, D. D., Boyer, H., & Misono, S. (2015). Prevalence and severity of dysphonia in patients with cystic fibrosis: A pilot study. *Otolaryngology - Head and Neck Surgery (United States)*. https://doi.org/10.1177/0194599815581832

Wingate, J. M., Brown, W. S., Shrivastav, R., Davenport, P., & Sapienza, C. M. (2007). Treatment Outcomes for Professional Voice Users. *Journal of Voice*, *21*(4), 433–449. https://doi.org/10.1016/j.jvoice.2006.01.001

Wolfe, V. I., Long, J., Youngblood, H. C., Williford, H., & Olson, M. S. (2002). Vocal parameters of aerobic instructors with and without voice problems. *Journal of Voice*, *16*(1), 52–60. https://doi.org/10.1016/S0892-1997(02)00072-3

Wolfe, V., & Martin, D. (1997). Acoustic correlates of dysphonia: Type and severity. *Journal of Communication Disorders*. https://doi.org/10.1016/S0021-9924(96)00112-8

Wu, Y. Z., Crumley, R. L., Armstrong, W. B., & Caiozzo, V. J. (2000). New perspectives about human laryngeal muscle: Single-fiber analyses and interspecies comparisons. *Archives of Otolaryngology - Head and Neck Surgery*. https://doi.org/10.1001/archotol.126.7.857

Xue, C., Kang, J., Hedberg, C., Zhang, Y., & Jiang, J. J. (2019). Dynamically Monitoring Vocal Fatigue and Recovery Using Aerodynamic, Acoustic, and Subjective Self-Rating Measurements. *Journal of Voice*, *33*(5), 809.e11-809.e18. https://doi.org/https://doi.org/10.1016/j.jvoice.2018.03.014

Yiu, E. M. L., & Chan, R. M. M. (2003). Effect of hydration and vocal rest on the vocal fatigue in amateur karaoke singers. *Journal of Voice*. Journal of Voice. https://doi.org/10.1016/S0892-1997(03)00038-9

Yiu, E. M. L., Wang, G., Lo, A. C. Y., Chan, K. M. K., Ma, E. P. M., Kong, J., & Barrett, E. A. (2013). Quantitative high-speed laryngoscopic analysis of vocal fold vibration in fatigued voice of young karaoke singers. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2013.06.010

Zacharias, S. R. C., Deliyski, D. D., & Gerlach, T. T. (2018). Utility of laryngeal high-speed videoendoscopy in clinical voice assessment. *Journal of Voice*, *32*(2), 216–220.

Zealer, D. L. (1983). The contractile properties and functions of single motor units of thyroartenoid muscle in the cat. *Vocal Fold Physiology: Contemporary Research and Clinical Issues. College Hill Press, San Diego*.

Zelinka, P., Sigmund, M., & Schimmel, J. (2012). Impact of vocal effort variability on automatic speech recognition. *Speech Communication*, *54*(6), 732–742. https://doi.org/10.1016/j.specom.2012.01.002