

The Effects of Conversational Priming on Vocal Fry Use in College Aged Female Students

By

Ashley Rolf, B.S.

*A Thesis  
Submitted in Partial Fulfillment of  
the Requirements for the Degree of*

Master of Science in Communication Sciences and Disorders

---

Sharyl Samargia, Ph.D., CCC-SLP                      Date

---

Lori Swanson, Ph.D., CCC-SLP                      Date

---

Jorge Bonilla, Ph.D.                                      Date

---

Director, Graduate Studies                              Date

University of Wisconsin – River Falls

2017

**TABLE OF CONTENTS**

<b>ABSTRACT</b>	Page 3
<b>CHAPTER 1: Introduction</b>	Pages 4-18
<b>CHAPTER 2: Methodology</b>	Pages 19-22
<b>CHAPTER 3: Results</b>	Pages 23-27
<b>CHAPTER 4: Discussion</b>	Pages 28-36
<b>REFERENCES</b>	Pages 37-41
<b>APPENDIX A: Consent Form</b>	Page 42
<b>APPENDIX B: Participant Screening Form</b>	Page 43
<b>APPENDIX C: Instructions</b>	Page 44

**Abstract:**

*Background:* The use of vocal fry is increasingly common in young adult American females.

Research suggests possible explanations for this trend, however there is no concrete evidence to explain why some individuals use vocal fry more than others. *Purpose:* To determine the impact of auditory vocal fry priming on the use of vocal fry in a young adult female population.

*Methods:* Twenty-eight females completed a total of eight voice recordings (sustained phonation, reading of the rainbow passage, and conversational speech) before and after being primed with two videos – one with an individual who spoke in a resonant voice and the other who spoke in vocal fry. Recordings were analyzed for changes in fundamental frequency, minimum pitch, maximum pitch, jitter, shimmer, noise-to-harmonics ratio, and percentage of vocal fry used.

*Conclusion:* No statistically significant differences were found for any of the dependent variables across conditions.

## **Chapter 1: Voice Anatomy and Physiology of the Larynx**

Vocal fold vibration is the source of the human voice. The vocal folds are made up of five layers of tissue (listed from superficial to deep): epithelium, superficial layer of the lamina propria, intermediate layer of the lamina propria, deep layer of the lamina propria, and the thyroarytenoid muscle. According to the cover-body theory, the top three layers (epithelium, superficial lamina propria, and intermediate lamina propria) make up the cover of the vocal folds, while the bottom two layers (deep lamina propria and muscle) make up the body (Hirano & Kakita, 1985). Some individuals choose to use a three-layer scheme, in which the mucosa is made up of the epithelium and superficial lamina propria; the ligament is made up of the intermediate and deep lamina propria, and the muscle makes up its own layer. The superficial layers of the vocal folds are made primarily of elastin, and gradually become stiffer as they reach the deeper layers resulting in more flexibility in the superficial layers.

The vocal folds are housed within the larynx, a small organ that is made up of various cartilages, muscles, and joints (Sapienza & Ruddy, 2013). The larynx serves the biological functions such as respiration, thoracic fixation, and swallowing, as well as the non-biological function of phonation (Colton, Casper, & Leonard, 2011; Sapienza & Ruddy, 2013). The laryngeal framework is composed of the thyroid cartilage, cricoid cartilage, arytenoid cartilages (2), corniculate cartilages (2) and epiglottis. The thyroid cartilage is the largest component, sits on top of the ring-shaped cricoid cartilage. Both the thyroid and cricoid cartilages are composed of hyaline cartilage, which gives the thyroid and cricoid their flexible and elastic properties (Colton et al., 2011; Sapienza & Ruddy, 2013). The arytenoids are paired movable structures that sit atop the posterior portion of the cricoid cartilage. These true vocal folds have an anterior attachment point on the inner surface of the thyroid cartilage (just below the notch) and a

posterior attachment at the vocal process of the left and right arytenoid cartilages. The paired corniculate cartilages sit on top of the arytenoids, but do not directly influence the vocal folds. Lastly, the epiglottis is a leaf shaped cartilage that attaches to the inner surface of the thyroid cartilage, near the base of the tongue. It functions as part of a three-step process to protect the airway from foreign objects when swallowing (Colton et al., 2011; Sapienza & Ruddy, 2013).

Movement of, and within, the larynx is attributed to the work of many muscles of the larynx that also serve the purpose of suspending the larynx in the neck at rest. These muscles are separated into two groups: extrinsic and intrinsic muscles. Extrinsic muscles, categorized as suprahyoid and infrahyoid, are attached to the hyoid bone and the exterior portion of the larynx. They assist in elevation and depression of the larynx for voice production, singing, and swallowing. Suprahyoid muscles are composed of the geniohyoid, mylohyoid, stylohyoid, and the anterior and posterior belly of the digastric (Sapienza & Ruddy, 2013). Infrahyoid muscles are composed of the omohyoid, sternohyoid, sternothyroid, and thyrohyoid (Sapienza & Ruddy, 2013).

Intrinsic muscles are housed *within* the larynx and connect the vocal folds to the laryngeal framework. These muscles systematically work together to make the vocal folds function properly. The lateral cricoarytenoid muscle connects the muscular process of the arytenoid to the cricoid cartilage, and functions in adducting the vocal folds (Sapienza & Ruddy, 2013). The thyroarytenoid muscle connects the vocal process of the arytenoids bilaterally to the anterior commissure of the thyroid cartilage (Sapienza & Ruddy, 2013). This muscle provides the body of the vocal folds, and serves in shortening the vocal folds, and assists in vocal fold adduction (Sapienza & Ruddy, 2013). The posterior cricoarytenoid connects the posterior portion of the cricoid cartilage to the muscular process of the arytenoid cartilage (Sapienza & Ruddy,

2013). This muscle is the only whose function is to abduct (i.e. open) the vocal folds. The interarytenoid muscles are composed of two parts: transverse and oblique. The transverse interarytenoids muscle connects the vocal folds together via its attachment at the muscular process of both arytenoid cartilages (Sapienza & Ruddy, 2013). The oblique interarytenoids are attached diagonally to both muscular processes of the arytenoids (Sapienza & Ruddy, 2013). The cricothyroid muscle attaches the lateral portion of thyroid cartilage to the lateral portion of the cricoid cartilage. Upon contraction, this muscle serves in stretching the vocal folds to increase pitch (Sapienza & Ruddy, 2013).

The muscles of the larynx are innervated by the Vagus nerve (CN X). The vagus nerve provides both sensory and motor functions to the laryngeal mechanism. This nerve is divided into two main branches: 1) recurrent laryngeal nerve, responsible for providing motor function to all of the intrinsic muscles except cricothyroid and sensation to subglottic region of the larynx, 2) superior laryngeal nerve external, which innervates cricothyroid, and 3) superior laryngeal nerve internal receives sensory information from the supraglottic space. (Colton et al., 2011; Sapienza & Ruddy, 2013).

### **Myoelastic Aerodynamic Theory**

At rest, the vocal folds are positioned within the larynx in the shape of a “V” with the closed end pointed anteriorly and the open end facing posteriorly. This open position allows air to move freely in and out of the lungs and through the oral cavity for breathing. When the vocal folds are used for phonation, there are several theories in place that explain vocal fold vibration and its consequent sound output.

## THE EFFECTS OF CONVERSATIONAL PRIMING ON VF USE IN COLLEGE AGED FEMALES

Prior to 1958, scientists believed the voice was created by neural impulses to the vocal fold muscle (Rubin, 1960; Sapienza & Ruddy, 2013). However, Van den Berg theorized a multifaceted physiologic explanation for phonation called the Myoelastic Aerodynamic Theory of Phonation, where “Myo” refers to contraction of the lateral cricoarytenoid and thyroarytenoid muscles, “Elastic” refers to the elastic properties of the vocal folds that allow them to stretch and return to their original state, and “Aerodynamic” refers to the air pressure and airflow required, which will later be described in detail as the Bernoulli Effect (Sapienza & Ruddy, 2013).

This theory holds that vocal fold vibration occurs as an interaction between elasticity and air pressure (Titze, 1980; Sapienza & Ruddy, 2013; Kummer, 2014). The components of this theory state that the vocal folds first adduct and subglottic air pressure is built up until the pressure reaches a threshold that forces the vocal folds to be blown apart laterally from the inferior (bottom) portion of the vocal folds to the superior (top) portion. Once the vocal folds reach maximum excursion, the force associated with the elasticity of the tissue becomes greater than the force associated with the air pressure moving through the glottis, causing the vocal folds to begin to move medially to their adducted position. This medial movement creates constriction in the airway as it narrows, thus causing air velocity to increase. In turn, the pressure exerted on the walls of the vocal folds decreases and causes them to be drawn together. This process, known as the Bernoulli Effect, is the aerodynamic component of this theory (Sapienza & Ruddy, 2013; Kummer, 2014). This cycle repeats, creating the vocal fold vibration that is perceived as sound. This can be measured in cycles per second represented in Hertz (Hz).

The average pitch that we perceive in a person’s voice is referred to the fundamental frequency. The typical fundamental frequency of males is between 85Hz – 155Hz, and between 165Hz – 255Hz for females; however, children have higher fundamental frequencies that can

exceed 300Hz (Stewart, Kling, & Allen, 2016). The difference in normative values is due to the differences in mass of the vocal folds in males (more mass) versus females (less mass). As such, a lower fundamental frequency would be the output of an individual with a more massive laryngeal mechanism, due to the indirect relationship between mass and pitch. Tension also alters fundamental frequency via a direct relationship between tension and pitch, such that a persons pitch increases as tension increases (i.e. the vocal folds are stretched).

While alterations of the vocal folds' mass, length, and tension can create a wide range of frequencies, vocal sounds can be divided into the three registers of modal, falsetto, and pulse (vocal fry). The most commonly used is modal register, which includes a range of fundamental frequency that is used and perceived as a speaker's typical voice (Sapienza & Ruddy, 2013). This register includes a range of fundamental frequencies typically used for speaking and singing including 75-500Hz for males and 130-750Hz for females (Hollien, 1972; Hollien & Michel, 1968). Falsetto register occupies fundamental frequencies that are higher than the modal register; 150-750Hz for males and 220-1700Hz for females (Hollien, 1972; Hollien & Michel, 1968). Vocal fry, commonly known as "pulse" register, occupies frequencies lower than those in the modal registers; 1-70Hz for both males and females (Hollien, 1972; Hollien & Michel, 1968).

### **Vocal Fry**

Vocal fry is a term used to refer to a voice quality characterized by a pulse-like vibratory pattern that is often described as harsh and/or creaky (Hollien, Moore, Wendahl, & Michel, 1966; Yuasa, 2010). Although the term has become more prominent in the literature within the past few decades, its characteristics have been used within the literature much earlier. For individuals who are perceived as using vocal fry, it is more likely to occur at the ends of sentences (Wolk,



Abdelli, & Slavin, 2012), and some have more frequent use than others (Bohm & Shattuck-Hufnagel, 2007). Although vocal fry was once considered a voice disorder on account of its abnormal sound output, research has now concluded that vocal fry can occur in the presence of a vocal pathology but is not pathological in nature (Hollien et al., 1966).

#### *Physiologic Characteristics*

The vibratory pattern of vocal fry consists of a long period of vocal fold adduction followed by short pulses of air within the same cycle of vibration (Childers & Lee, 1991; Sapienza & Ruddy, 2013), giving vocal fry its characteristic “popping” sound output. Research has shown that during the production of vocal fry, the vocal folds remain closed over 75% of the time and, as the frequency decreases, the percentage of closure time increases, suggesting a disproportionate time of glottal closure is a feature of vocal fry (Hollien et al., 1977; Hollien et al., 1966).

#### *Acoustic Characteristics*

Although fundamental frequency varies across gender in modal register, both males and females exhibit decreased fundamental frequencies in vocal fry phonation (Blomgren, Chen, Ng, & Gilbert, 1998; McGlone, 1967; Murry, 1971), with no significant difference in average fundamental frequency values (Blomgren et al., 1998; McGlone, 1967; Hollien & Michel, 1968). Vocal fry can be detected acoustically by increased jitter, shimmer, and noise-to-harmonic ratio values due to aperiodicity and variations in frequency (perturbation) and amplitude (Blomgren et al., 1998).

*Aerodynamic Characteristics*

The term “aerodynamic” refers to the properties of airflow. During vocal fry phonation, airflow rates have been found to significantly decrease for both males and females in comparison to modal register (Blomgren et al., 1998; Murry, 1971). However, airflow rate did not change based on different frequencies produced solely in vocal fry (i.e., frequencies below 70 Hz) (McGlone, 1967). This evidence suggests that frequencies produced in *vocal fry* have no relationship with airflow rate (McGlone, 1967). There is speculation that the Bernoulli Effect is decreased during vocal fry due to the decreased rate of airflow (Chen, Robb, & Gilbert, 2002), however, this has not been empirically proven.

*Air Pressure*

There is controversy reported in the literature in regard to air pressure values for vocal fry. Blomgren et al. (1998) reported that subglottic air pressure was significantly lower in vocal fry phonation when compared to modal register for vowels and contiguous syllables. However, Murry (1971) examined sustained phonation of vowels produced with vocal fry, and found that subglottic air pressure was greater in vocal fry than in modal register.

*Additional gender differences*

Although vocal fry has the same quality among men and women, it differs significantly in its prevalence and in the way it is used. Although the ranges of modal and falsetto registers are significantly higher in frequency for females in comparison to males, both genders demonstrate little to no differences in frequency when speaking in vocal fry (Hollien, 1972). However, it has been a recent noticeable trend that females speak in vocal fry more frequently than men. Yuasa

(2010) studied the differences in the use of vocal fry at the conversational level for males and females of both American and Japanese origin. In this study, 11 American males, 12 American females, and 10 Japanese females were given an emotionally neutral topic and asked to converse with a conversational partner for 10 minutes. A segment of approximately 400 words was randomly extracted from each conversation and used for perceptual analysis. Results showed that 1) American females used vocal fry more than American males, and 2) American females used vocal fry more than Japanese females. However, these findings only speak to conversational/spontaneous speech.

Irons and Alexander (2016) found contrasting results when they compared males and females use of vocal fry for both reading and spontaneous speech. In this study, 16 individuals (10 male, 6 female) who were Standard American English speakers between the ages of 18 to 22 completed various reading and spontaneous speech tasks such as teaching information, giving instructions, and describing their qualifications for a scholarship. The researchers analyzed the amount of vocal fry used both by duration of time spent using vocal fry and also by number of creaky words per second. The researchers found that males differed in the amount of vocal fry produced for both reading and spontaneous speech. The differences in findings suggest that further research may be needed to gain more consistent findings not only across gender but also across varying speech tasks.

### **Communicative Component to Vocal Fry**

There is speculation that the use of vocal fry has a linguistic influence, such that the use of vocal fry may add meaning to a person's statement. The field of psycholinguistics terms vocal fry as "glottalization," and describes it as "perceivably irregular vocal fold vibration," (Bohm &

Shattuck-Hufnagel, 2007). There is literature to suggest that the use of glottalization is somewhat predictable from a communicative standpoint. Researchers have reported the rate of occurrence of glottalization is higher in words at the ends of sentences (Henton & Bladen, 1988; Redi & Shattuck-Hufnagel, 2001; Wolk et al., 2012) and/or at the onset of vowel-initial words (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996). In addition, glottalization is more likely to occur during particular phonemes in anticipation of coarticulation and/or phonetic context (Huffman, 2005) or to add prosodic structure by marking phrase boundaries (Dilley et al., (1996). Although these studies show a communicative component to the use of vocal fry, they do not explain *why* some individuals use vocal fry more frequently than others, or why the prevalence of vocal fry is increasing young adult females.

### **Perception of Vocal Fry**

Although many individuals are unaware when they are speaking in vocal fry, many are able to identify when *others* are speaking in with vocal fry. A sizeable amount (79.8%) of American college students have reported hearing a “creaky voice quality” used frequently by women in the area in which they live (Yuasa, 2010). When asked to listen to sustained /a/ samples in either modal register or vocal fry, listeners correctly differentiated between the different registers with 100% accuracy, suggesting that individuals not trained in voice perception can easily identify differences between these registers (Blomgren et al., 1998). College students have also reported vocal fry to be a frequently selected speaking pattern of many individuals within their age group (Gottliebson, Lee, Weinrich, & Sanders, 2007).

Wolk et al., (2012) examined vocal fry use in native Standard American English (SAE) speakers between the ages of 18-25. Recordings were collected for each participant’s sustained

phonation and a sentence-reading task, which was then analyzed both acoustically and perceptually. Results indicated that the evaluators perceived vocal fry far more often in sentences than in sustained phonation, for which vocal fry was rarely detected. Additionally, they found that approximately two thirds of the population of study used vocal fry. This research suggests that vocal fry appears to be a common trend for college age SAE speakers and that it can be selectively used. Yuasa (2010) speculates that not only is this phenomenon more prevalent among females, it may be used more among young females with higher levels of education.

With its prevalence seemingly increasing in both use and perception, vocal fry has been linked with multiple different connotations. However, there is controversy within the current literature. Yuasa (2010) surveyed 175 college students for their perceptions of two speech samples: one in vocal fry and one in modal register. Listeners rated each sample based on confident/hesitant, formal/casual, educated/uneducated, pretentious/genuine, and aggressive/compliant. Results showed that vocal fry was associated with education, intimacy, genuineness, and non-aggression. Additionally, a majority (78.9%) of listeners stated that they heard women using vocal fry where they live (Yuasa, 2010).

In another study investigating perceptions of vocal fry use, Anderson, Klofstad, Mayew, and Venkatachalam (2014) recruited 800 subjects (400 males, 400 females) ages 18-65 who listened to pairs of recordings of either 7 males or 7 females saying the phrase “thank you for considering me for this opportunity” in (A) their normal voice and (B) in vocal fry. Each participant judged recordings A and B based on which one was more attractive, competent, hireable, educated, and trustworthy (all qualities considered important in the workplace). Results showed that listeners (regardless of their gender) preferred a “normal” sounding voice to vocal fry, but that female listeners rated vocal fry stimuli significantly more negatively than males for

all qualities. Young adult female voices produced in vocal fry were judged to be less trustworthy ( $p < 0.01$ ), less competent ( $p < 0.01$ ), less educated ( $p < 0.01$ ), and less hireable ( $p < 0.01$ ). Although the normal voice was preferred for listeners of all ages, older listeners had more negative judgments of competence. This negative perception of vocal fry by others suggests that individuals who use vocal fry are perceived as less trustworthy, competent, educated, and hireable when judged within the workplace (Anderson et al., 2014). It has also been judged to sound more hesitant and less confident (Yuasa, 2010).

The noted differences in perceptions suggests that there is a contextual influence on the use of vocal fry within the population of young adult, female college students. Although perceptions of vocal fry have been investigated, the role and/or influence of the social environment has not yet been accounted for in the literature.

### **Priming and Social Influences**

Although previous research has focused on the prevalence of vocal fry amongst various age groups and populations, little research has yet been done to explore the cause of this trend. Some existing theories on the etiology of continual vocal fry use in young females suggest that vocal fry stems from sociocultural motivation rather than contextual situations (Yuasa, 2010). The use of “priming,” a term used to describe the process of a stimulus facilitating a response, has the possibility to explain the learned behavior of increased use of vocal fry. Literature in social psychology supports that priming will occur when an individual is exposed to socially relevant stimuli, and that the *prime* can influence or facilitate “impressions, judgments, goals, and actions, often even outside of people’s intention or awareness” (Molden, 2014). Thus, are our young adult females being primed into speaking in more vocal fry?

Wang and Hamilton (2013) propose that individuals have a natural tendency to mimic others' actions. Termed "mimicry," this phenomenon can occur as the result of social priming and is used as a tool to increase social closeness and belonging, resulting in "social affiliation" (Wang & Hamilton, 2013). Studies have shown that mimicry increases in an effort to affiliate with others, especially when there are feelings of exclusion (Lakin & Chartrand, 2003; Lakin, Chartrand, & Arkin, 2008; Wang & Hamilton, 2013), and that conversation partners develop ways to "converge their voices in terms of how they will be best received (Gregory & Webster, 1996).

It is proposed that goals can be subconsciously activated and used to guide our behaviors without our awareness (Bargh, 1990). Termed the "Goal Activation Theory," it is said that a "given prime or stimulus activates a goal, which then automatically leads to pursuit of that goal" (Wang & Hamilton, 2013, p. 1) and that its activation is not dependent on any conscious decision (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001). For example, a person listening to their conversational partner's voice (prime) could activate an internal goal to assimilate and thus alter their behavior when responding.

Social priming and mimicry have been noted among voice literature, in terms of changing voice pitch to influence the perceptions of others. Fraccaro, O'Connor, Re, Jones, DeBruine, and Feinberg (2012) found that lower-pitched voices were perceived as more dominant. Other studies have found that women lowered their pitch and had an increase in "vocal hoarseness" when asked to deliberately change their voice to sound more attractive (Hughes, Farley, & Rhodes, 2010; Hughes, Mogilski, & Harrison, 2014). These studies, along with findings from Yuasa (2010) suggests that younger female college students may be adopting this new linguistic style of vocal fry to create a persona of being more educated, dominant, and attractive in

comparison to their conversational partner. However, there is no empirical evidence to prove this.

Although research pertaining to priming of vocal fry is still in its infancy, there are a few pieces of recent research that suggest that conversational partners may influence the amount of vocal fry a person uses. Borrie and Delfino (2016) studied 20 healthy females (i.e. no history of voice or hearing disorders) from the ages of 18-29 years old in Phoenix, Arizona who were Standard American English speakers. The purpose of their study was to determine 1) whether participants modified their vocal fry use based on their conversational partners use of vocal fry, and 2) whether vocal fry entrainment correlated with measures of communication efficiency and enjoyment.

In this study, participants spoke separately with two conversational partners: one who used substantial amounts of vocal fry, and one who did not. For each conversation, participants were given a pair of pictures and asked to identify 10 differences between the pictures using only their voice. After each task, participants were asked to rate their interaction based on enjoyment. Perceptual measures were used to analyze speech samples, and the percent of vocal fry was determined by dividing the duration of time spent in vocal fry divided by the total duration of time spent speaking. Results revealed that participants used significantly more vocal fry themselves when they were speaking with a conversational partner who used substantial amounts of vocal fry. Moreover, individuals rated their experiences as more enjoyable when their levels of vocal fry use were more closely matched to that of their conversational partner.

These findings correspond with the findings of an unpublished study completed by Rolf, Wojcik, and Hemmerich during which they examined the influence of partner speaking style and media usage on college-age females' use of vocal fry. For this study, 18 healthy females (i.e. no



history of voice disorders) who were between the ages of 18-23 and Standard American English speakers were recruited from a university in Wisconsin. Participants spoke separately with two conversation partners: one who used a resonant voice, and one who used substantial amounts of vocal fry. Participants completed both conditions in random order, however two standard questions were used to prompt each conversation: 1) “What got you interested in your major?” and 2) “What would you like to do for your future career?” Participants were also asked to complete a survey, which addressed their vocal behaviors and media usage.

All recordings were analyzed perceptually using the last 100 syllables of each sample. A percentage of vocal fry was determined by dividing the number of vocal fry syllables by 100. Each participant was asked to rate the amount of media they used for 11 media subcategories (e.g., television shows, movies, social networking, etc.) using an ordinal measurement scale, which was later assigned point values and totaled to obtain a media usage score. Results indicated that vocal fry use was significantly higher when speaking with the peer using vocal fry. Although there was no significant correlation between use of vocal fry and media usage, results suggested that individuals who use more media may be more susceptible to using vocal fry when they are speaking with a conversational partner who also vocal fry frequently. However, it is unknown whether this same relationship would exist when individuals are exposed to different amounts of vocal fry through video as a means of priming.

Although there has been research done in the area of vocal fry, the influence of priming on the conversational use of vocal fry still remains unclear. A majority of current research is focused on sustained phonation as opposed to conversational speech. Furthermore, there is a need for research to explore the influence of vocal fry when presented auditorily, rather than a face-to-face interaction. Thus, the purpose of this research study is to determine whether the

vocal quality of a speaker influences the use of vocal fry in their conversational partner when the stimulus is presented via video.

**Specific Aim:**

1. To determine the impact of priming on vocal quality (i.e. use of vocal fry) in a young adult female population.

*Hypothesis:*

*(H<sub>o</sub>)=Priming does not influence the use of vocal fry in young adult females.*

*(H<sub>A</sub>)=Priming using vocal fry will significantly increase the use of vocal fry in young adult females. Priming with resonant voice will not result in a significant increase in the use of vocal fry in young females.*

**Independent Variable:** Vocal fry exposure

**Dependent Variables:**

1. Percentage of vocal fry used
2. Fundamental Frequency
3. Jitter
4. Shimmer
5. Noise-to-Harmonics Ratio
6. Minimum and Maximum Pitch

## Chapter 2: Methods

### *Participants*

Participants included 28 female college students between the ages of 18-25 ( $20.68 \pm 1.79$ ) (Table 1) who were Standard American English speakers and had no history of voice disorders (Table 1). A priori power analysis revealed that a sample size of 26 was needed to detect a large effect. All participants were recruited from the Department of Communication Sciences and Disorders at the University of Wisconsin – River Falls (UWRF). Participation in this study was voluntary, and individuals could withdraw from the study at any time. Due to the nature of the study, participants were not told the purpose of the study during recruitment and gathering of data in an effort to avoid bias and changes in participants’ natural behavior. Individuals were informed that the purpose of the study was to obtain normal speech samples for a database. Prior to recruitment, this study was reviewed and approved by the University of Wisconsin-River Falls Institutional Review Board, protocol number W2016-T065.

All participants underwent a screening process to ensure that all inclusionary and exclusionary criteria were met. Criteria included the following: 1) English as primary language, 2) no history of voice disorders, 3) not currently receiving speech therapy, and 4) did not have an upper respiratory infection or cold. Those who stated they were ill were not scheduled for participation until their voice quality had returned to baseline per report of the participant.

**Table 1:** Summary of age of participants

<b>Minimum Age = 18</b>	<b>Maximum Age = 24</b>	<b>Mean Age = 20.68 years</b>
*Ages studied: 18 (n=5), 19 (n=2), 20 (n=6), 21 (n=4), 22 (n=8), 23 (n=1), 24 (n=2)		

*Variables*

The amount of vocal fry used by the conversational partners in each interaction served as the independent variables for this study, as each conversational partner used a standard amount of vocal fry for each interaction. The dependent variables consisted of the following: percentage of vocal fry used (by participants), mean fundamental frequency, jitter (cycle to cycle variations in frequency/pitch), shimmer (cycle to cycle variations in amplitude/intensity), minimum pitch, maximum pitch, and noise-to-harmonics ratio.

*Design and Procedures*

Data collection was completed in the Voice and Speech Science Laboratory at UWRF. All voice samples were recorded using a headset microphone (Plantronics Blackwire C310), with a standard 1-inch microphone to mouth distance. Samples were recorded and analyzed using Audacity on a MacBook computer and the Computerized Speech Lab (Kaypentax). Each session lasted approximately 20 minutes.

Voice samples were recorded for each participant at baseline and after two priming conditions in which the order was randomly assigned: 1) listening to an individual using resonant voice and 2) listening to an individual using vocal fry (Table 2). Baseline data were collected before any dependent variables were introduced. Participants were asked to produce a sustained /a/ three times and read the rainbow passage. After a baseline voice sample was obtained, each participant was primed by watching and listening to a 1.5 – 2 minute video of an individual speaking in resonant voice and in vocal fry in random order. Immediately after watching each video, voice samples were recorded. Voice samples consisted of sustained /a/, reading of the rainbow passage, and conversational speech. Following the first and second conditions,

THE EFFECTS OF CONVERSATIONAL PRIMING ON VF USE IN COLLEGE AGED FEMALES

conversational speech was prompted using two standard questions across all participants: “What got you interested in your major and/or career of choice?” and “What would you like to do for your future career?”

Acoustic analysis was completed using the Computerized Speech Lab (CSL) (Kay Pentax). Fundamental frequency, jitter, shimmer, noise-to-harmonic ratio and pitch range values were obtained for sustained /a/. Reading and conversational speech samples were analyzed for fundamental frequency, pitch range, and percentage of vocal fry used.

Table 2: Breakdown of Conditions

Condition	Task(s)	Variables Obtained
	Sustained /a/	Fo, jitter, shimmer, N/H, pitch range
	Reading of Rainbow Passage	Fo, pitch range, % vocal fry used
	Sustained /a/	Fo, jitter, shimmer, N/H, pitch range
	Reading of Rainbow Passage	
	Conversational Speech Sample	
	Sustained /a/	Fo, jitter, shimmer, N/H, pitch range
	Reading of Rainbow Passage	
	Conversational Speech Sample	

To analyze percentage of vocal fry, a segment containing the last 100-syllables of each conversational sample was chosen and analyzed perceptually to determine the percentage of speech that was produced in vocal fry. Each rainbow passage sample was analyzed by syllables, however these samples were analyzed in its entirety. Group differences for individual samples were analyzed using one-way ANOVA with Bonferroni correction for multiple comparisons. In addition to the researcher, three female undergraduate students were recruited voluntarily to analyze the rainbow passage and conversational samples. All recruited raters underwent a training session with the researcher, during which each rater 1) watched a video describing vocal

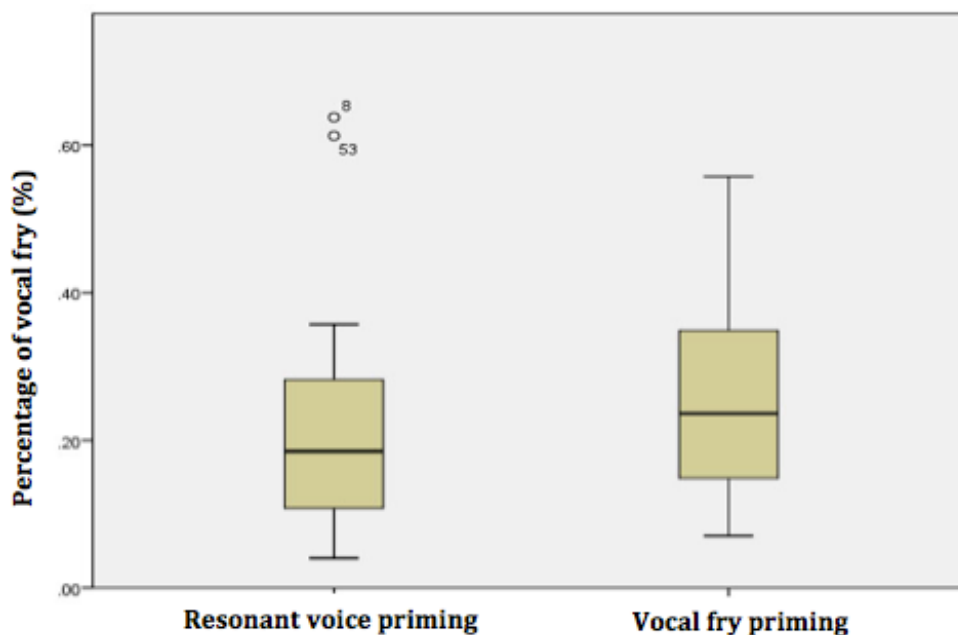
fry and 2) practiced perceptually picking out words with vocal fry for a total of three conversational samples. For the first conversational practice item, the raters listened to the speech sample while simultaneously being told which words and syllables contained vocal fry. For the second practice item, the researcher and rater simultaneously indicated when they heard vocal fry. For the third practice item, the researcher watched and listened as the rater independently indicated when she heard vocal fry. Based on performance on the final practice item, the researcher determined whether raters were ready to begin their independent work. If raters were demonstrating difficulty following the final practice item, additional practice was provided.

### Chapter 3: Results

Aim 1: The null hypothesis was supported.

#### *Percentage of vocal fry use*

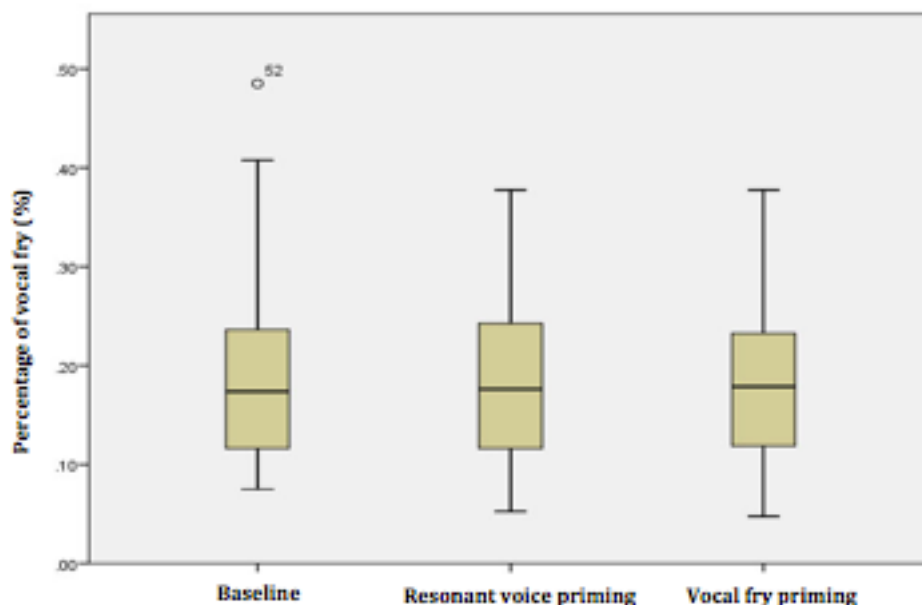
A one-way ANOVA revealed no statistically significant differences in percentage of vocal fry in conversation between resonant priming (0.178% +/- 0.028%) and vocal fry priming (0.236% +/- 0.028%) ( $F=2.14$  (1, 55),  $p=0.15$ ) (Figure 3). However, 23 out of 26 participants increased the amount of vocal fry used in conversation after the vocal fry priming (mean increase of 6%).



**Figure 3.** Mean percentage of vocal fry across all participants during conversation after resonant voice priming and vocal fry priming.

A one-way ANOVA revealed no statistically significant differences in percentage of vocal fry, during reading of the rainbow passage, between conditions ( $F=0.12$ , (2, 83),  $p=0.89$ ). Bonferroni correction for multiple comparisons revealed no significant differences in percentage of vocal fry during reading, between baseline (0.15% +/- 0.01%) and resonant priming (0.147% +/- 0.019%),

( $p=1.0$ ), baseline (0.15% +/- 0.01%) and vocal fry priming (0.142% +/- 0.016%), ( $p = 1.0$ ), or resonant priming (0.147% +/- 0.019%) and vocal fry priming (0.142% +/- 0.016%), ( $p=1.0$ ) (Figure 4).



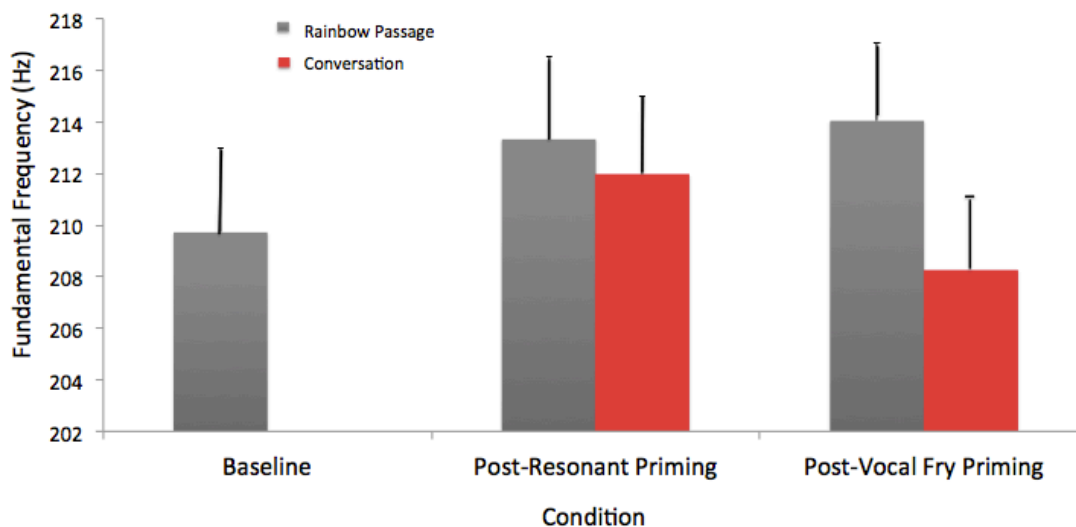
**Figure 4.** Mean percentage of vocal fry used across all subjects at baseline and after both priming conditions during the rainbow passage. One outlier was identified in the baseline condition, however this data was not dropped, as it did not create significance.

An intra-class correlation was conducted using Cronbach's Alpha to determine inter-rater reliability of those rating percentage of vocal fry in both reading passages and conversational passages. Results showed a strong correlation for the percentage of vocal fry detected in the post-resonant priming conversation ( $\alpha=0.886$ ), post-vocal fry priming conversation ( $\alpha=0.800$ ), baseline rainbow passage sample ( $\alpha=0.716$ ), post-resonant priming rainbow passage sample ( $\alpha=0.739$ ), and post-vocal fry priming rainbow passage sample ( $\alpha=0.700$ ).



*Fundamental frequency*

A one-way ANOVA revealed no statistically significant differences in fundamental frequency in *conversation* between resonant priming (211.971Hz +/- 3.544Hz) and vocal fry priming (208.274Hz +/- 3.032Hz) ( $F=0.63$  (1, 55),  $p=0.43$ ). Additionally, no statistically significant differences were found in fundamental frequency *during reading of the rainbow passage* between conditions ( $F=0.53$  (2, 83),  $p=0.59$ ). Bonferroni correction for multiple comparisons revealed no significant differences in fundamental frequency during reading, between baseline (209.72Hz +/- 17.26Hz) and resonant priming (213.329Hz +/- 3.235Hz), ( $p=1.0$ ), baseline (209.72Hz +/- 17.26Hz) and vocal fry priming (214.047Hz +/- 3.015Hz), ( $p=1.0$ ), or resonant priming (213.329Hz +/- 3.235Hz) and vocal fry priming (214.047Hz +/- 3.015Hz), ( $p=1.0$ ) (Figure 5). Other acoustic variables were collected as well, however none reached statistical significance (Tables 3, 4).



**Figure 5.** Mean fundamental frequency across all subjects, during rainbow passage (gray) and conversation (red) at baseline and after both priming conditions. No baseline data were obtained for conversation.

<b>Table 3. ANOVA Results for acoustic measures (PRP = Post-Resonant Priming, PVFP = Post-Vocal fry Priming)</b>			
<b>Variable</b>	<b>Condition</b>	<b>ANOVA</b>	<b>Bonferroni Correction for Multiple Comparisons</b>
			Baseline x PRP: $p=1.0$
			Baseline x PVFP: $p=1.0$
			PRP x PVFP: $p=1.0$
	Conversation	F=0.93 (55, 1), $p=0.34$	PRP x PVFP: $p=1.0$
			Baseline x PRP: $p=1.0$
			Baseline x PVFP: $p=1.0$
			PRP x PVFP: $p=1.0$
	Conversation	F=2.02 (55, 1), $p=0.16$	PRP x PVFP: $p=1.0$
			Baseline x PRP: $p=1.0$
			Baseline x PVFP: $p=1.0$
			PRP x PVFP: $p=1.0$
			Baseline x PRP: $p=1.0$
			Baseline x PVFP: $p=1.0$
			PRP x PVFP: $p=1.0$
			Baseline x PRP: $p=1.0$
			Baseline x PVFP: $p=1.0$
			PRP x PVFP: $p=1.0$

THE EFFECTS OF CONVERSATIONAL PRIMING ON VF USE IN COLLEGE AGED FEMALES

**Table 4.** Comparison of group mean and standard deviation across dependent variables.

<b>Variables</b>	<b>Post – resonant priming mean +/- SD</b>	<b>Post – Vocal Fry Priming mean +/- SD</b>
<b>Fo Sustained</b>	218.267Hz +/- 4.083Hz	218.501Hz +/- 4.206Hz
<b>Fo Rainbow</b>	213.329Hz +/- 3.235Hz	214.047Hz +/- 3.015Hz
<b>Fo Conversation</b>	211.971Hz +/- 3.544Hz	208.274Hz +/- 3.032Hz
<b>Min Pitch Conversation</b>	81.363Hz +/- 2.783Hz	78.185Hz +/- 1.774Hz
<b>Max Pitch Conversation</b>	396.215Hz +/- 14.659Hz	367.294Hz +/- 14.10Hz
<b>Min Pitch Rainbow</b>	86.384Hz +/- 4.652Hz	90.161Hz +/- 5.172Hz
<b>Max Pitch Rainbow</b>	373.568Hz +/- 15.821Hz	361.823Hz +/- 15.084Hz
<b>% Vocal Fry Rainbow</b>	0.147% +/- 0.019%	0.142% +/- 0.016%
<b>% Vocal Fry Conversation</b>	0.178% +/- 0.028%	0.236% +/- 0.028%
<b>Jitter (Sustained ah)</b>	0.009% +/- 0.005%	0.009% +/- 0.005%
<b>Shimmer (Sustained ah)</b>	0.049% +/- 0.002%	0.049% +/- 0.003%
<b>N/H (Sustained ah)</b>	0.161 +/- 0.01	0.156 +/- 0.006

#### **Chapter 4: Discussion**

Participants' use of vocal fry in sustained phonation, reading, and/or in conversation does not appear to be influenced by auditory vocal fry priming alone. These findings reject the proposed hypothesis that a significant change in the amount of vocal fry used would be observed when participants were primed with greater amounts of vocal fry. Results of this study should be considered preliminary, as all participants were 1) female, 2) Standard American English speakers, 3) college students, and 4) recruited from a single geographical location.

Although participants receiving vocal fry priming showed a lower average fundamental frequency and higher percentage of vocal fry use in conversation when compared to resonant priming, these results did not reach statistical significance. Interestingly, although participants did not have significant change in vocal fry use, all participants had vocal fry detectable in their connected speech samples. This information supports the notion that vocal fry is a selectable phonational register for individuals without vocal pathology (Hollien et al., 1966; Hollien et al., 1968). Even though all participants used vocal fry in their speech, there were differences observed in the amount of vocal fry detected across participants. This supports the literature stating that vocal fry is a frequently selected speaking pattern in the young adult female population (Gottliebson, et al., 2007; Yuasa, 2010).

Results suggest the video method used in this study may not be sensitive enough to reveal significant changes in behavior. However, the video method more so reflects the way in which an individual would perceive vocal fry in the media (i.e. through watching and listening to a person speak with no required contingent response) as opposed to in-person conversation. Research has suggested that the behavior of speaking in vocal fry is observed and learned through media use by observing famous figures and current trends (Wolk et al., 2012; Yuasa,

2010). However, when comparing media usage scores with amount of vocal fry used, no significant correlations were present (Rolf, Wojcik, & Hemmerich, Unpublished). In addition, their results suggested that although there is no significant correlation between media usage and production of vocal fry, individuals who use more media are more susceptible to using more vocal fry when they are speaking with a conversational partner who also uses vocal fry.

Although media influence was not a purpose of this study, the video method was used as a vehicle for auditory priming in order to ensure consistent amounts of vocal fry priming across participants while also allowing for individuals to “see” their conversational partner in a similar manner that they would in a face-to-face conversation. Interesting, the results of the current study support the above media findings, as no significant increases in vocal fry use were demonstrated when vocal fry was presented through auditory means (which mirrors the vocal fry priming through media).

The absence of significant change through the video method suggests that in-person conversation between individuals may be more influential for vocal fry priming, as an individual’s susceptibility to be primed may be more influenced by conversational turns that are contingent upon one another, as opposed to a single question and answer format used in the current study. In addition to the frequency of priming exposures, the intensity/duration of priming also may act as a contributing factor in vocal fry usage. Individuals may need a longer duration of priming to result in significant change. Due to the nature of the current study, participants were only primed for 1.5-2 minutes. However, the literature using an in-person conversational model of priming allowed for contingent conversational turns which may have allowed for more frequent, albeit shorter, bursts of priming and thus more opportunities to assimilate to their conversational partner. In this way, participants would receive shorter bursts

of priming over a longer period of time as opposed to one long burst of priming over a shorter period of time. In other words, a single exposure of vocal fry priming of a long duration may not influence a speaker to use vocal fry, as the frequency of priming received may influence the behavior of the respondent.

With regards to vocal patterns within conversation, research has also suggested that women alter their vocal pitch depending on the social implications of their conversational partner (Hughes et al., 2010). The same study by Rolf, Wojcik, and Hemmerich supported this speculation, as results showed that the individual's amount of vocal fry used in conversation increased when speaking with a conversational partner that demonstrated a casual, relaxed disposition in comparison to an individual who demonstrated a very professional disposition. Results of the current study suggest that the video method used may not be the influencing factor for vocal fry priming, as this method lacks the social component of conversation. This idea is also supported by Borrie and Alexander (2016), as they found that the regulation of social differences (i.e. assimilating to the behavior of one's communication partner by using more vocal fry) correlated with a higher level of enjoyment. With use of the video method, participants may not feel the need to regulate differences in voice characteristics due to the lack of social implication.

Since results showed a strong intraclass correlation, it is reasonable to say that vocal fry can be detected auditorily. Thus, perceptual detection is a useful tool to use in addition to other acoustic analysis of vocal fry, but only if individual's have a trained ear to detect vocal fry in connected speech. In the present study, one rater showed slightly inconsistent detection of vocal fry in comparison to the other raters. This shows that, while vocal fry can be detected auditorily, some may need more training than others for consistency. Individuals who were trained to detect

vocal fry were exposed to a greater amount of vocal fry than the participants, which supports the notion that a greater amount of auditory vocal fry priming may be necessary to cause significant change.

Even though significance was not found for perceptual measures, these measurements seem to show a more noticeable difference in the amount of vocal fry used as opposed to acoustic measurements. Although acoustic measurements have shown to be an important component of voice analysis in the past, recent research has begun to question its value. Brockmann-Bauser and Drinnan (2011) looked specifically at the validity of jitter and shimmer measurements and found that they were not indicative of the presence of vocal pathology or perceptual hoarseness. Results of the current study support this finding, as no significant differences in jitter and shimmer were found across any/all conditions.

### *Limitations*

There were several limitations to this study. First, participants were students recruited within the communication sciences and disorders major and may have had previous exposure to education about vocal fry. Students were not asked whether or not they knew what vocal fry was in order to ensure that participants did not suspect vocal fry to be a variable in the study.

Next, no baseline measures were taken for conversational samples. Although the resonant priming could serve as a baseline, as they were primed using a voice without vocal fry, this can only be said for participants who received the resonant condition first. Additionally, these results can only speak to the effects of one exposure to vocal fry priming. Future research should investigate the effects of multiple episodes of vocal fry priming throughout the day or over a

series of weeks to examine if dose/frequency has any influence on an individual's use of vocal fry.

The instrumentation method used in this study was a possible threat to internal validity as samples were first obtained through a microphone headset and saved onto Audacity, and then recorded onto the Computerized Speech Lab (CSL) for further analysis. This method was used due to technical obstacles and should be considered an inefficient way for data analysis in the future. If this study is to be repeated, researchers might consider recording directly onto the Computerized Speech Lab program.

Due to the nature of the study, reactive arrangements (i.e. “Hawthorne Effect”) was a potential threat to external validity. For all recordings of sustained phonation and the rainbow passage, participants were sitting with the researcher in the same room. However, when participants were watching each video and recording their conversational speech samples, the researcher was not present. This was done in order to ensure that participants were not additionally primed with the researcher’s voice following each video.

Lastly, there were multiple data points that fell outside of the upper and lower boundaries of the data set for the percentage of vocal fry used in the reading of the rainbow passage (one participant) and in conversational speech (two participants). Involvement of these data in analysis did not create significance nor change assumptions, therefore it was determined appropriate to keep these data in the set. Possible explanations for these data could include posture of the participant during recording of each sample, as this was not controlled for within this study. It is known that vocal fry is more likely to occur as a person is running out of air while speaking, therefore if individuals demonstrated poorer posture for one recording, they may have been more likely to use higher amounts of vocal fry without having received any priming.



Although there was a strong agreement between raters for perceptual analysis, there was some variability in the amount of training required for the recruited analysts. If this study is repeated, it may be more efficient to recruit speech language pathologists specializing in voice disorders, or recruit and train speech-language pathology graduate students who have experience with CAPE-V voice analysis. Additionally, it would be helpful to create a more concrete method to train individuals to detect vocal fry consistently in connected speech.

### *Suggestions for Future Research*

Due to the limited research on vocal fry in conversation, there are many different avenues for future research to explore. It would be interesting for future research to investigate relationships between 1) different number of exposures to vocal fry priming, 2) varying durations of vocal fry priming, 3) video priming and face-to-face priming, and 4) personality traits and vocal fry use and/or susceptibility to priming,

Additionally, future research should explore more efficient methods of vocal fry detection and establish best practice for vocal fry measurement. Because many acoustic measurements (i.e. jitter, shimmer, noise-to-harmonics ratio) can only be obtained for sustained phonation, it is difficult to compare the features of vocal fry in sustained phonation and connected speech. Recent research has found instances where individuals have perceptually used vocal fry in connected speech, however their fundamental frequencies increase from baseline rather than decrease (Rolf, Wojcik & Hemmerich, Unpublished). This goes against the pre-determined characteristic of vocal fry being produced in pulse register at frequencies below 70Hz.

Future research may wish to explore the use of Cepstral Analysis to measure and compare Cepstral Peak Prominence (CPP) values of vocal fry production in connected speech versus sustained phonation. Hillenbrand and Houde (1996) state that a cepstrum is “a log power spectrum of a log power spectrum,” and that “... for periodic signals, the first power spectrum shows energy at harmonically related frequencies and the second spectrum displays a strong component corresponding to the regularity of the harmonic peaks” (Hillenbrand & Houde, 1996, p. 314). This method of voice analysis results in Cepstral Peak Prominence (CPP) values, which give information regarding the harmonics in a given speech signal, such that a healthy voice will produce a strong cepstral peak due to its harmonic structure, and a disordered voice will produce a weak cepstral peak due to its poorly defined harmonic structure (Hasanvand, Salehi, & Ebrahimipour, 2016). Used as an alternative to acoustic analysis, CPP will detect roughness and breathiness in both sustained phonation and connected speech (Hasanvand et al., 2016).

Research on Cepstral Analysis has determined that cepstral scores increase with age (Monnappa & Balasubramaniam, 2015) and increase as loudness/intensity increases in vowel production (Awan, Giovinco, & Owens, 2012). Conversely, CPP values have been found to decrease in dysphonic voices (Hasanvand et al., 2016) such as those with vocal nodules (Kumar, Bhat, & Prasad, 2010). When comparing the CPP values in connected speech and sustained phonation, CPP values were significantly different for sustained phonation in comparison to connected speech for both normal and dysphonic voices (Brinca, Batista, Tavares, Goncalves & Moreno, 2014).

However, there is currently limited evidence to support the correlation between the vocal fry and CPP values. Plexico and Sandage (2016) studied the relationship between vocal fry and CPP in reading tasks. In this study, 26 participants (13 male, 13 female) between the ages of 18

to 40 were recorded reading the rainbow passage and analyzed acoustically for fundamental frequency, semitones, CPP, and percent of vocal fry.

The researchers found that while the percent of vocal fry used had a weak relationship with CPP, there was a moderate positive relationship between change in semitones and vocal fry use, as well as change in semitones and sex. Furthermore, a moderate negative relationship was observed between change in semitones and CPP. Although these findings suggest a weak correlation between percentage of vocal fry used and CPP values, the current study's findings can only speak to the relationships between these values during a reading task. Thus, further research is needed to explore the relationship of vocal fry and CPP during sustained phonation and conversational speech tasks.

## **Conclusion**

In sum, a single video exposure to vocal fry priming does not appear to influence a young female's use of vocal fry in reading or in conversation after one exposure through listening alone. Future research should address the influence of vocal fry priming in face to face settings including formal/professional and informal/social to further determine the influence of priming as a potential influencing factor on the use of vocal fry in young females. Future research should also examine the influence of duration, frequency of priming, and severity of vocal fry during priming.

References

- Anderson, R. C., Klofstad, C. A., Mayew, W. J., Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *Plos One*, 9, (5).
- Awan, S. N., Giovinco, A., and Owens, J. (2012). Effects of vocal intensity and vowel type on cepstral analysis of voice. *Journal of Voice*, 26 (5), 670.e16-670.e20.
- Bargh, J. A. (1990). Auto-motives: Preconscious determinants of thought and behavior. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 93-130). New York: Guilford Press.
- Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81, (6), 1014-1027.
- Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech, Language, and Hearing Research*, 1, 227-244.
- Blomgren, M., Chen, Y., Ng, M. L., Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The Journal of the Acoustical Society of America*, 103, (5), 2649-2658.
- Bohm, T., Shattuck-Hufnagel, S. (2007). Utterance-final glottalization as a cue for familiar speaker recognition. In: *Annual Conference of the International Speech Communication Association (Interspeech)*, 2657-2660.
- Borrie, S. A., Delfino, C. R. (2016). Conversational entrainment of vocal fry in young adult American English speakers. *Journal of Voice*, 26 (3), 1-8.
- Brinca, L. F., Batista, A. P. F., Tavares, A. I., Goncalves I. C., and Moreno, M. L. (2014). Use of cepstral analysis for differentiating normal from dysphonic voices: a comparative study

of connected speech versus sustained vowel in european portuguese female speakers.

*Journal of Voice*, 28, (3), 282-286.

Brockmann-Bausser, M., Drinnan, M. J. (2011). Routine acoustic voice analysis: time to think again? *Wolters Kluwer Health*, 19, (3), 165-170.

Chen, Y., Robb, M. P., Gilbert, H. R. (2002). Electrolottographic evaluation of gender and vowel effects during modal and vocal fry phonation. *Journal of Speech, Language, and Hearing Research*, 45, 821-829.

Childers, D. G., Lee, C. K. (1991). Vocal quality factors: analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, 90, 2394-2410.

Colton, R.H., Casper, J.K., & Leonard, R. (2011). *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Baltimore: Williams & Wilkins.

Dilley, L., Shattuck-Hufnagel, S., Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, 24, 423-444.

Fraccaro, P. J., O'Connor, J. J. M., Re, D. E., Jones, B. C., DeBruine, L. M., Feinberg, D. R. (2012). Faking it: deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, 85, 127-136.

Gottliebson, R. O., Lee, L., Weinrich, B., Sanders, J. (2007). Voice problems of future speech-language pathologists. *Journal of Voice*, 21, (6), 699-704.

Gregory, S. W., Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation in social status perceptions. *Journal of Personality and Social Psychology*, 70, (6), 1231-1240.

- Hasanvand, A., Salehi, A., Ehrahimipour, M. (2016). A cepstral analysis of normal and pathologic voice qualities in Iranian adults: a comparative study. *Journal of Voice*, 30, 1-7.
- Henton, C. & Bladon, A. (1988). Creak as a sociophonetic marker. *Language, speech and mind: studies in honour of Victoria A. Fromkin* (L. M. Hyman and C. N. Li, editors), pp. 3-29. London: Routledge.
- Hillenbrand, J., Houde, R. A. (1996). Acoustic correlates of breathy vocal quality dysphonic voices and continuous speech. *Journal of Speech Language and Hearing Research*, 39, 311-321.
- Hirano, M., & Kakita, Y. (1985). Cover-body theory of vocal fold vibration. In R. G. Daniloff (Ed.), *Speech science: Recent advances* (pp. 1-46). San Diego, CA: College-Hill Press.
- Hollien, H., Moore, P., Wendahl, R. W., Michel, J. F. (1966). On the nature of vocal fry. *Journal of Speech and Hearing Research*, 9, 245-247.
- Hollien, H., Michel, J. F. (1968). Vocal fry as a phonational register. *Journal of Speech and Hearing Research*, 11, 600-604.
- Hollien, H. (1972). On vocal registers. *Communication Sciences Laboratory Quarterly Report*, 10, 1-33.
- Hollien, H., Girard, G. T., Coleman, R. F. (1977). Vocal fold vibratory patterns of pulse register phonation. *Folia Phoniatica*, 29, 200-205.
- Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics*, 33, 335-362.
- Hughes, S. M., Farley, S. D., Rhodes, B. C. (2010). Vocal and physiological changes in response to the physical attractiveness of conversational partners. *Journal of Nonverbal Behavior*,

34, 155-167.

- Hughes, S. M., Mogilski, J. K., Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior*, 38, 107-127.
- Irons, S. T., Alexander, J. E. (2016). Vocal fry in realistic speech: acoustic characteristics and perceptions of vocal fry in spontaneously produced and read speech. *Journal of the Acoustical Society of America*, 140, (4), 3397.
- Kumar, B. R., Bhat, J. S., Prasad, N. (2010). Cepstral analysis of voice in persons with vocal nodules. *Journal of voice*, 24, (6), 651-653.
- Kummer, A. W. (2014). *Cleft palate and craniofacial anomalies: Effects on speech and resonance*. New York: Cengage Learning.
- Lakin, J L., Chartrand, T L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 8, (4), 334-339.
- Lakin, J L., Chartrand, T. L., Arkin, R. M. (2008). I am too just like you: nonconscious mimicry as an automatic behavioral response to social exclusion. *Psychological Science*, 19, (8), 816-822.
- McGlone, R. E. (1967). Air flow during vocal fry phonation. *Journal of Speech and Hearing Research*, 10, 299-304.
- Molden, D. C. (2014). Understanding priming effects in social psychology: what is “social priming” and how does it occur? *Social Cognition*, 32, 1-11.
- Monnappa, D., Balasubramanium, R. K. (2015). Cepstral analysis of voice in healthy aged individuals. *Journal of Laryngology and Voice*, 5, 34-48.
- Murry, T. (1971). Subglottal pressure and airflow measures during vocal fry phonation. *Journal of Speech and Hearing Research*, 14, 544-551.



- Plexico, L. W., Sandage, M. J. (2016). Influence of glottal fry on acoustic voice assessment: a preliminary study. *Journal of voice*, 26, 1-5.
- Redi, L., Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, 29, 407-429.
- Rubin, H. J. (1960). The neurochronaxi theory of voice production – a refutation. *American Medical Association Archives of Otolaryngology*, 71, 913-920.
- Sapienza, C., Ruddy, B. H. (2013). *Voice disorders*. San Diego, CA: Plural Publishing.
- Stewart, C. F., Kling, I. F., Allen, E. L. (2016). *Voice rehabilitation: Testing hypotheses and reframing therapy*. Burlington, MA: Jones & Bartlett Learning.
- Titze, I. (1980). Comments on the Aerodynamic Theory of Phonation. *Journal of Speech and Hearing Research*, 23, 495-510.
- Titze, I. R. (2012). Have workplace and beautiful teeth changed female voice characteristics? *Journal of Singing*, 68, 299-300.
- Van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech Language and Hearing Research*, 1, (3), 227-243.
- Wang, Y., Hamilton, A. F. (2013). Understanding the role of the ‘self’ in the social priming of mimicry. *Plos One*, 8, (4), 1-11.
- Wolk, L., Abdelli, N. B., Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26, (3), 111-116.
- Yuasa, I. P. (2010). Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile american women? *American Speech*, 85, (3), 315-337.

**APPENDIX A: Consent Form**

Participant Name: \_\_\_\_\_

Principal Investigator: **Ashley Rolf, Graduate Student Researcher**  
**Sharyl Samargia, PhD, CCC-SLP Faculty Advisor**

Principal Investigator Phone Number(s): **Ashley Rolf: 715-559-4424**  
**Sharyl Samargia: 715-425-3834**

Title of Project: **Speech Samples of College Female Students**

**Purpose:**

Data collected from this study will be used to create a database of speech samples of college female students who are native English speakers. These samples will be used within the department of Communication Sciences and Disorders as normal comparisons for disordered speech.

**Procedure:**

A microphone headset will be used to capture all speech samples throughout the study. The participant will be asked to produce three recordings of the rainbow passage and three recordings of sustained phonation of the vowel /a/. Between each of these recordings, the participant will be asked to listen to a short video and record their answers. Data collection is expected to take 20 minutes per participant including consent and data recording.

**Time Required:**

The entire process should take about 20 minutes.

**Risks:**

Potential risks include vocal fatigue due to amount of time speaking, feelings of self-consciousness due to being audio-recorded, and/or feelings of embarrassment if data is unintentionally released.

Your rights as a subject:

- a. All confidential information will be locked in a cabinet with only the research team having the ability to access. No identifying information will be used in dissemination and subject codes will be used in place of names.
- b. You may withdraw from the study at any time by stating you would like to be done. There will be no penalty in doing so.
- c. If you have any questions during the entire course of the study, please do not hesitate to contact the Principal Investigators listed above. Once the study is complete, you can also contact either personnel for a copy of the results.

If you have concerns about how you were treated during this study, contact:

- d. Molly Van Wagner, Interim Director of Grants and Research, 101 North Hall, UW-RF, 715/425-3195.

This project has been approved by the UW-River Falls Institutional Research Board for the Protection of Human Subjects, protocol # W2016-T065.

**I have read the above information and willingly consent to participate in this experiment.**

**Signed:** \_\_\_\_\_

**Date:** \_\_\_\_\_



## APPENDIX C: Instructions

### Video 1

Please watch the video on your screen. When you are finished, please record your answer to the following question:

**How did you become interested in your field of choice?**



*Please see the researcher when you are finished. Thank you!*

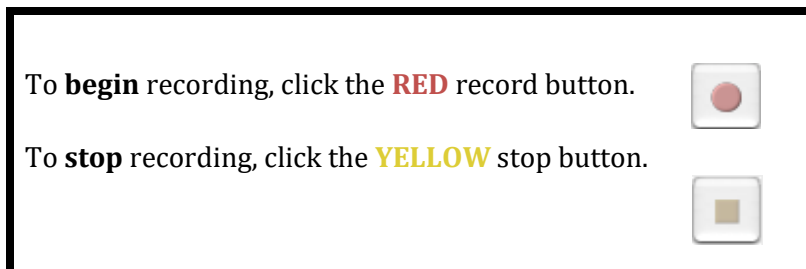
---

### Video 2

Please watch the video on your screen. When you are finished, please record your answer to the following question:

**What are your plans for your future career?**

(Examples: setting you'd like to work in, adults and/or kids, favorite disorders, etc.)



*Please see the researcher when you are finished. Thank you!*