2016

The Perception of Creaky Voice: Does Speaker Gender Affect our Judgments?

Kaitlyn E. Lee
University of Kentucky, kaitlyn.elizabeth.lee@gmail.com
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.032

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Kaitlyn E. Lee, Student

Dr. Kevin B. McGowan, Major Professor

Dr. Gregory Stump, Director of Graduate Studies

THE PERCEPTION OF CREAKY VOICE: DOES SPEAKER GENDER AFFECT OUR JUDGMENTS?

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Arts in the
College of Arts and Science
at the University of Kentucky

By

Kaitlyn Elizabeth Lee
Lexington, KY

Directors: Dr. Kevin B. McGowan, Professor of Linguistics
Lexington, KY
2016

ABSTRACT OF THESIS

THE PERCEPTION OF CREAKY VOICE:
DOES SPEAKER GENDER AFFECT OUR JUDGMENTS?

This study focuses on the phonetics of creaky voice saliency and the perceptual sociolinguistic indexes that are evoked during creaky voice use. This study consists of two experiments: the first a listener judgment based Likert scale, the second an AXB study. The first experiment used modal and creaky voice statement-of-fact tokens to determine whether the speaker is or isn't x characteristic (intelligent, feminine, educated, masculine, hesitant, and confident). This study found that both male and female speakers were found to be less intelligent, less educated, less feminine, more masculine, less confident, and more hesitant when using creaky voice phonation as compared to the modal register. Participants also rated male and female speakers as statistically different. During the second experiment the participants listened to continuums that went from modal register to extreme creaky voice (based on F0 levels). Participants performed an AXB task to determine ability at distinguishing levels of creaky voice along the continuum. This study found that participants were less able to correctly detect the level of creaky voice in the female speaker for the lower half of the continuum when compared to the male speaker.

KEYWORDS: Phonetics, Vocal fry, Creak, Glottalization, Gender

Kaitlyn Elizabeth Lee

04/25/2016

THE PERCEPTION OF CREAKY VOICE:
DOES SPEAKER GENDER AFFECT OUR JUDGMENTS?

By

Kaitlyn Elizabeth Lee

_____Kevin B. McGowan_____
Director of Thesis

_____Greg Stump_____
Director of Graduate Studies

_____4/25/2016_____

ACKNOWLEDGMENTS

First, I would like to thank my thesis chair, Dr. Kevin B. McGowan, for his tireless support and guidance. Without his dedication and patience, this project would not be what it is today. His depth of knowledge and his ability to explain complex programs and theories has made my thesis year one of immense academic growth. In addition, his professional support and recommendations have been invaluable during my transition from UK to IU.

I would also like to thank Dr. Rusty Barrett who helped push my cohort and me in the beginning stages of our theses. Your flexibility and personalized assistance was crucial to our success.

Dr. Jennifer Cramer has been an irreplaceable mentor and role model throughout my years in the program, and even before. From helping me as a clueless (crying) undergrad to advising me on Ph.D. applications, Dr. Cramer has shown me what it means to be a successful, female in academia who, seemingly effortlessly, balances her work and family life perfectly. I hope to one-day reach the level of success that you have already achieved.

In addition, thank you to the rest of the Linguistics faculty who have all positively shaped my experience in the program, and helped me to become the student that I am today. I also could not have survived without the support from my cohort. Thanks for being there through the good and bad.

Finally, thank you to my family and boyfriend who have supported me unconditionally. To my parents, Denny and Joetta, who have always stood by me; no matter where I've traveled or what challenge I've accepted, your unwavering love and acceptance have allowed me to take risks that would otherwise have seemed overwhelming.

To my brother, Colin, who has been my constant role model, encouraging me to push my boundaries and reach for opportunities I thought were unobtainable, your dedication to the pursuit of knowledge has kept me focused on my academic goals for many years. Lastly, thank you to my boyfriend, Derek, who joined me at the beginning of this crazy ride. Thank you for your patience, understanding, and jokes during the many stressful weeks.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

Section 1: Introduction 1
- Experiment One: 2
- Experiment Two: 3

Section 2: Literature Review 3
- Previous Methods 11

Section 3: Participants & Volunteers 13

Section 4: Experiment One 15
- Methods 15
  - *Creating tokens* 15
  - *Data Collection Process* 16
- Results 17
- Discussion 22
- Problems 27

Section 5: Experiment Two 28
- Methods 28
  - *Creating tokens* 28
  - *Data Collection Process* 30
- Results 30
- Discussion 32
- Problems 34

Section 6: Conclusions 35

Section 7: Future 37

Appendix A: Experiment One Sentences 39

Appendix B: Participant Survey 40

Appendix C: Experiment One Screenshot 41

REFERENCES 42

VITA 45

LIST OF TABLES

Table 1: Likert Scale averages for experiment one sentences categorized by speaker, voice style and question. 17
Table 2: F0 frequencies for the “pay” tokens at the beginning and end of the vowel, separated by speaker and category. 29

LIST OF FIGURES

Figure 1.1, Listener participants average ages categorized by gender and overall..... 14

Figure 1.2, Participants gender categorized by ethnicity and overall 14

*Figure 2.**1** Likert Scale averages for experiment one sentences categorized by speaker, voice style, and question.*.....18

*Figure 2. **2**: Likert scale averages for experiment one sentences for the male speaker, categorized by voice type and question*……...…….21

*Figure 2. **3**: Likert scale averages for experiment one sentences for the female speaker, categorized by voice type and question.*.....22

Figure 3.1, Proportions of correct responses for experiment two, AXB study, categorized by speaker and series, with linear trend lines for each speaker included. 32

Section 1: Introduction

This study aims to look at perceptual judgments of indexical fields and the salience in the use of creaky voice phonation in the modal voice register. This study consists of two experiments: first, a Likert scale judgment task, second, an AXB task. Through these two experiments this study will analyze perceptual differences between the two phonation types and between speaker genders. The findings of this research will add to the larger corpus of research looking at female speech and the stigmas that exist around young female speech. But first, this paper will give a little more information on the specific terms and concepts.

Creaky voice has historically been considered one of the vocal registers that we have access to. Hollien defined a vocal register as, "…a series or range of consecutive voice frequencies which can be produced with nearly identical voice quality" (1972, p. 1). Recently, creaky voice has been reconsidered as a phonation type instead of a vocal register (Ladefoged & Johnson, 2011). This distinction helps to highlight the naturalness of creaky voice, and distances it from being perceived as a vocal pathology that needs to be fixed. Importantly, creaky voice also has several synonyms as well: vocal fry, glottalization, pulse register, laryngealization, and glottal fry. While I will stick to the term creaky voice throughout the paper, works that have been cited here may use other terms and I will keep their original terminology. Researchers who study the topic have agreed that while terminology may vary, there is no measurable difference between the phenomena described by these various terms.

Creaky voice is characterized by "…compressed and thick vocal folds, resulting in slow vibration, hence low fundamental frequency (f0), and low air flow rates" (Podesva, 2013, p. 429). Perceptually, creaky voice is easily perceived and was even perceived with

100% accuracy in one study, even better than the "normal" modal register (Blomgren, Chen, Ng, & Gilbert, 1998). It is a low voice, with creak, or pulses throughout the utterance. Creaky voice is often, and indeed, usually, used in conjunction with the modal register, slipping back and forth between the two, thus prompting its reclassification as a phonation type instead of a register.

As stated above, this study will contribute to the debate over creaky voice indexicalities (Podesva, 2013) and use by young women. Because we have control over the use of this phonation type, creaky voice, it has been manipulated and reinforced through application over time to have certain indexical meanings. In other words, we use the phonation type creaky voice to create meaning in our utterances. This is obviously not the only way in which we can create meaning, however it is highly manipulable, and can index a variety of meanings. This allows creaky voice to be a highly productive phonation type and also at times a confusing one. Other factors must be considered in addition to the phonation when looking at narrow indexical meanings.

Experiment One:

This study uses judgment-based evaluations of creaky voice. Listeners heard token sentences spoken either in creaky phonation or modal register, and rated the speakers on various indexical categories: hesitant, confident, intelligent, educated, masculine, and feminine. This experiment has 5 hypotheses.

1. Speakers will be rated as more masculine when speaking with creaky voice and more feminine when speaking in the modal register.
2. Speakers will be rated as less hesitant and more confident when using creaky voice.

3. Speakers will be rated as more educated and intelligent when using creaky voice.
4. Overall the male speaker will be rated more preferably than the female speaker.
5. The judgment differences between modal register and creaky voice will be more pronounced in the female speaker than the male speaker.

Experiment Two:

This section examines at what point creaky voice becomes salient in auditory perceptions, and if there is a difference perceptually between male speaker tokens and female speaker tokens. Speakers listened to AXB tasks with both a male and a female speaker, to compare various levels of creak. This experiment has 2 hypotheses.

1. Listeners will be more perceptive (i.e. more accurate) of the modal register to creaky voice phonation transition in the female speaker than the male speaker.
2. Listeners will lose perceptiveness in the lowest creaky categories' series.

The hope of this research project is to shed light on a subject that has been controversial. Creaky voice has been found to be both harmful and helpful to young women, and while too much attention is given to the actions of young women, I am hoping that this attention will be for the best in promoting women's use of a feature that they find productive.

Section 2: Literature Review

In 1968, Hollien and Wendahl, some of the first linguists to study creaky voice, wrote, "Vocal fry (1) is a normal mode of laryngeal production; (2) it consists of a register of very low fundamental frequencies, and (3) it consists of a train of relatively discrete laryngeal pulses with nearly complete dampening between successive glottal excitations" (p. 506). This is the definition that has been the foundation of all creaky voice studies since

its publication. Over the years is has been cited by many, although, sometimes in various, updated iterations. Indeed, Hollien himself expanded his original definition in 1972 saying that, "the pulse register occupies the lowest range of phonation along the fundamental frequency continuum" (2). Not only is vocal fry "very low" but now it is the "lowest." He expands his original third point further saying that, "for the pulse register, this type of plot yields a rapidly opening and closing motion of the vocal folds and a very long closing period" (19). This creaky vocal fold motion and characterization is very unique when compared to the modal phonation vocal fold movement, "relatively rapid onset followed by a brief open period with longer period closing and a short closed time" (Hollien, 1972, p. 19). It is clear not only from its unique low phonation but also the vocal fold movements that the categorization of creaky voice as a voice style separate from modal voicing is uncontroversial.

In more recent studies, creaky voice has been characterized similarly, but labeled as a phonation, such as the following: "Acoustically, vocal fry lies in the lower end of the fundamental frequency (F_0) spectrum, below the frequency range typical of the modal register" (Wolk, Abdelli-Beruh, Slavin, 2012, p. 122). In recent articles, the foundational qualities of creaky voice are sometimes considered understood and thus are written in parenthesis with little elaboration, "…creaky voice (produced with relatively compressed and thick vocal folds, resulting in slow vibration, hence low fundamental frequency (f0), and low air flow rates)" (Podesva, 2013, p. 429). This shift from register to phonation type helps to illustrate the ease of transition between the two styles. But, for now lets take Hollien's original, foundational three components and analyze them more thoroughly.

First, "Vocal fry is a normal mode of laryngeal production" (Hollien, 1968).

Hollien included this first and foremost because until that time, creaky voice was often considered a laryngeal disorder, which must be corrected by a speech pathologist. In the same publication he also says, "it is illogical to assume that this phenomenon is associated exclusively – or even primarily – with voice disorders" (p. 506). Although he does not going into tremendous detail, it is clear that he believes that because humans without pathologies can intentionally and unintentionally use vocal fry, it has a place in normal speech pathologies. Today this viewpoint is held by speech pathologists. Wolk, Abdelli, & Slavin say, "Speakers with normal laryngeal functioning have the option to switch from the modal register to vocal fry at a given moment. Naturally, an individual with vocal pathology and disordered laryngeal function will lack this ability to control vocal fry" (2012, p. 112). Again the shift to calling creaky voice a phonation has helped reinforce this idea of normal laryngeal functioning.

Second, "Vocal fry consists of a register of very low fundamental frequencies" (Hollien, 1968). Indeed, the creaky voice phonation distinguishes itself in the modal register by manipulating the fundamental frequencies (F_0) to be lower than that of any other phonation or register. "Average frequencies of vocal fry vibration, extrapolated from the literature range from approximately 20-70 Hz, with a mean of approximately 50 Hz" (Blomgren, Chen, Ng, & Gilbert, 1998, p. 2649). When considering creaky voice F_0 to modal voicing, "men typically produce vocal fry with an F0 that is one octave lower than their modal F0, and women produce vocal fry with an F0 that is two octaves below their modal F0" (Chen, Robb,& Gilbert, 2002, p. 822). Chen, Robb, & Gilbert had a slightly narrower F0 range, "35-50 Hz" (2002). In addition to F_0 differences, Moosmüller (2001) found that "creaky vowels produced by women show a lower second formant as compared

to the same vowel produced by the same speaker using modal phonation" (p. 100). However, this same lowering was negligible among men. This is perhaps the most obvious effect of creaky voice phonation, and is used as the most salient feature when listening to speech samples.

Finally, "Vocal fry consists of a train of relatively discrete laryngeal pulses with nearly complete dampening between successive glottal excitations" (Hollien, 1968). Hollien went on to say in 1972 that during creaky voice " the vocal folds are very thick and the ventricular folds appear to come in contact with (or load) the true vocal folds" (p. 19). This thickening of the vocal cords and the ventricular folds are what give creaky voice its distinctive sound and phonation. This thickening is achieved by both men and women equally well, and is what allows men and women to reach the same F_0 range. There are competing views about whether this thickening is easier for men or women. "Henton and Blandon speculate that female speakers may not need to compress their vocal cords as forcefully as males to attain the bunching because their shorter vocal cord length naturally contributes to the bunching and thickness required for creaky voice production" (Yuasa, 2010, p. 321). However, Chen, Robb, & Gilbery suggest that after careful EGG analysis that while F_0 may be similar there is less change for men to reach these levels and that "these conditions appear to favor dicrotic dysphonia in men more so than in women" (2002). There are also conflicting reports about how many pulses may contribute to the creaky voice sound. However, in 1984 Whitehead et. al. "using high-speed laryngeal photography found that the vocal fold vibratory patters in vocal fry can be associated with single, double, or triple opening and closing gestures followed by a lengthy closed period" (as cited in Blomgren, Chen, Ng, & Gilbert, 1998, p. 2650). Perceptually, single, double,

and triple opening and closings produce the same vocal register and are indistinguishable from one another. While it is not possible to detect these pulses auditory, this is driving mechanism for creating the low F0s that are the characteristic quality of creaky voice. Without this unique vocal fold pattern, we would not have creaky voice phonation.

Voice quality and specifically "glottal fry/ glottalization serves an array of linguistic (eg phonological), pragmatic (eg turn-taking), and metalinguistic (EG emotions) purposes in American English" (Abdelli-Beruh, Wolk, & Slavin, 2013, p. 185). We consciously and subconsciously use language to index a variety of meanings, like those discussed above. Podesva (2015) states, "…voice quality functions as a rich phonetic resources through which speakers can display affect and take stances in interaction" (p. 178). Creaky voice can often index meanings that aren't gender or even identity related. Wiener & Chartrand (2014) found that, "…a person who speaks with a creaky voice is judged to be less warm and less competent than a person who speaks with other voice type" (p. 512). Interestingly, and probably well known by all parents of teenage children, Grivičić & Nilep (2004) found that one can use, "…*yeah* with creaky voice to indicate passive recipiency and either a dispreference to continue the current topic, or a disalignment with the primary speaker" (p. 1). This is not surprising, since it is hard to convey emotion while using creaky voice phonation. It is used as a way to distance oneself emotionally from a topic. Again, as Podesva (2013) states, quoting Eckert (2008), "…an indexical field is a prepresentation of the range of social meaning, as well as their inter-relations, that can be activated in the social contexts in which variables are used" (p. 435). As has not been discussed yet, creaky voice can be, and perhaps, most often, is associated with gender and character indexes that are tied, for better or worse, to gender.

Creaky voice was initially, during the earliest studies, considered a male feature. In 1983, Ohala, who studied communication in humans and other animals, described that lower pitches were seen as a sign of dominance and authority, while higher pitches classified a speaker/ producer as a member of a subordinate group. It is unsurprising then that studies have shown that women are more attracted to men with lower pitched voices, and are more likely to pay attention when they speak. Wiener & Chartrand (2014) studied creaky voice ad efficacy and found that, "female listeners who heard a creaky male voice were much more interested in Optus (a fictitious product) than were those who heard a creaky female voice" (p. 514). In addition, Collins (2003) studied women's perceptions on male attractiveness based off of voice samples and found that, "Men with voices in which there were closely spaced, low-frequency harmonics were judged as being more attractive, older, and heavier, more likely to have a hairy chest, and of a more muscular body type" (p. 773). Or in other words, men with lower voices were perceived to be hyper-masculine. Podesva elaborates the low pitch hyper-masculinity, from solely low pitch to creaky voice specifically, saying, "The association between creaky voice and masculinity is iconic, in the sense that the low pitch acoustically characterizing creaky voice is interpreted as resembling masculinity…this iconic association can be reworked at higher orders of indexicality and forges a connection to stances associated with masculinity, such as toughness" (2015, p. 115). This is a key theory in creaky voice phonation study; creaky voice is iconic for masculine indexes. This is what drives young female speakers, as we will see below, to use creaky voice in their own speech. Young female speakers are trying to index this iconic masculinity and its related indexes such as confidence, intelligence, and authority.

Traditionally, as was shown above, creaky voice has been characteristic of male speech, however, as was mentioned above, more recent studies have found that young female speakers are using creaky voice more frequently than males (Abdelli-Beruh, Wolk & Slavin 2013; Yuasa, 2010; Wolk, Abdelli-Beruh, & Slavin, 2012; Podesva, 2013; Pennock, 2005; etc). While no one is certain when young women started using creaky voice, it is certain now that they are. The effect on listeners' perceptions when creaky voice is used, and used by different genders specifically, is the piece that researchers, including this author, are still trying to ascertain.

In Abdelli-Beruh, Wolk, & Slavin's 2013 study on creaky voice in young, American men, they found that, "when asked to read a passage, a marked gender difference in the prevalence of vocal fry emerges: the rate is about four times higher for female speakers than for male speaker" (p. 187). Wolk, Abdelli-Beruh, & Slavin (2012) found that women used creaky voice most often at the ends of utterances. Yuasa (2010), who has the most recent, thorough research, with the soundest methodology, found that American women are more than twice as likely to use creaky voice than American men, and similarly more than twice as likely to use creaky voice compared to Japanese women. American women used creaky voice during 12.4% of their utterances, without any prompting or awareness that creaky voice was the research matter. Yuasa did nothing to elicit creaky voice. Yuasa 2010 also included a question to her listener participants asking whether they " had heard creaky voice frequently used by American women in their regions" and 78.9% of her respondents said yes. Without a doubt, creaky voice phonation has become a young female characteristic, with famous female celebrities such as Kim Kardashian, Katy Perry, and Zooey Deschanel driving the public images of a young female creaky voice user.

Some research has a more negative view of this innovation while others are less condemning. Anderson, Kolfstad, Mayew, & Vankatachalam investigate whether creaky voice is damaging for young women in a professional setting. They say "…the use of vocal fry is a vocal "fashion trend" that can be off-putting to older generations, in turn damaging the professional image of young women" (2014, p. 1). Specifically, "the results of this analysis show that the negative perception of vocal fry was significantly stronger when listeners judged female voices on trustworthiness, competence, education, and willingness to hire" (p. 3). If this were the case, creaky voice would not be a productive register for women to use intentionally, or unintentionally. However, this study by Anderson et.al. has had its fair share of methodological complaints, specifically the use of participants 'mimicking' vocal fry and manipulating the data so that creaky voice is as salient as possible, which twists the reality of creaky voice and makes it unnatural.

Other studies have found that using creaky voice has benefitted women. Specifically, in Yuasa's 2010 study, when creak samples were taken from natural recordings and played for listener participants, the college-age participants perceived female creaky voice as, "hesitant, nonaggressive, and informal, but also educated, urban-oriented, and upwardly mobile" (p. 315). While hesitant is not necessarily a positive characteristic, we see positive trends for creaky voice indexing with educated, urban-oriented, and upwardly mobile. While these are not directly tied to Podesva and Ohala's creaky voice theory of iconic masculinity, we can see how Yuasa's findings are moving in that positive, authoritative-leader, masculine direction.

Women have often been negatively indexed not only for their language but for other things as well (ex. music, clothes, movies, emotions, etc). Unfortunately, things that are

associated with women tend to become poorly viewed. Eckert (2014) says, "The ultimate legitimate person in the social order, the white upper middle-class male, is slated to be unemotional, rational, focused on "business," and endowed with global and objective knowledge. Women and adolescents, on the other hand, are viewed as emotional, changeable, irrational, trivial, and unobjective" (p. 529). If society or a person holds these beliefs, it is no wonder that female language is indexed negatively. Consequently, when specifically considering language, Eckert (2014) has this to say, "Considerable attention was paid at one point to the use of like as a discourse marker, and to the use of rising intonation on declaratives (dubbed "uptalk")…both were interpreted as hedges and taken to signal the speakers' lack of concern with precision, or unwillingness to take responsibility for their statements…specifically with respect to girls, they are taken to indicate insecurity and an unwillingness to state a forceful opinion" (p. 540). With negative views of female speech such as this, is it any wonder then why women might try to index masculine, strong, confident, etc instead of the weaker, less competent feminine indexes? This desire to index strong (and thus masculine) characteristics is theorized to be driving young female speakers to adopt creaky voice phonation.

Previous Methods

Before describing my methodology, I will first look at the most recent and relevant work for my topic: Yuasa's 2010 study on creaky voice rates and indexical meanings for young women. This study has been well cited, and praised for its methodologies.

In her research, Yuasa had a two-part study – an auditory and acoustic study, as well as a perception survey. For her auditory and acoustic study she had eleven American male and 12 American female speakers all of California dialects. She also had 10 Japanese

female speakers for analysis across languages. She recorded conversations between the speakers about innocuous topics such as food, in order to eliminate extraneous factors such as heightened emotions, which can greatly affect voice style. When she had her samples, she took an even number of tokens from all of her participants and calculated their percent creak (creaky tokens/ total tokens). By holding sample size consistent for all participants, Yuasa is again controlling factors (loquaciousness) that might skew her data. While primarily Yuasa did auditory detection of creaky voice, she also verified her judgments using Praat analysis.

In the second part of her study, the perception survey, 175 American college students from the University of California, Berkeley and the University of Iowa listened to the previously recorded conversations, discussed above. Two samples of similar lengths were drawn from the American female speakers, one with creak, and one without creak. The participants listened to the samples and then answered questions comparing the modal register to the creaky voice. The participants made judgments about "opposite attributes: confident/ hesitant, formal/casual (intimate), educated (knowledgeable)/ uneducated (ignorant), pretentious/ genuine, and aggressive/ compliant" (p. 328). For these judgments Yuasa used a seven point Likert scale from three to negative three, strongly agree, to strongly disagree. She also asked participants how often they hear vocal fry spoken in their community.

Several of these methods were employed in my own research: comparing modal register to creaky voice phonation, natural speech samples, auditory creak detection, non-emotion evoking topics, participant judgment data, Likert scale, and paired characteristics.

Section 3: Participants & Volunteers

This section will discuss both the listener participants of the study and the speaker volunteers whose voices were used to make the stimuli for the study.

First, the volunteers whose voices were used to make the stimuli were two graduate students in the linguistics program at the University of Kentucky, who had control over their modal to creak voice phonation shift, and who were known to use creaky voice often naturally in speech. The male volunteer was a 27 year old, Caucasian male from Missouri, whose native language is English. The female volunteer was the author of the study, a 24 year old, Caucasian female from Kentucky, whose native language is English. Neither of the volunteers had taught, or had extended interactions with the participants, and therefore their voices were new and unfamiliar to all of the participants.

The participants in this study were students from two introduction to linguistics classes at the University of Kentucky. There were 53 participants in total whose data was used in the study. Participants were all undergraduate students who had had minimal education/ training in linguistics at the time of the study. Participants were invited to participate in the research during class time by the primary investigator/ author of the study, and a reminder to sign up for the study was sent out over the class email system by the classroom instructor/ professor. The average age of participants was 20.6 years old and the range of ages was eighteen to forty-three. The median age was 20 and the mode was 20. Participant data was collected using a demographic information survey (see Appendix B).

Figure 1.1 below depicts their average ages relative to one another. The age is written above each category to the tenth decimal place, where participants are grouped by gender: male female, and combined (all participants together).

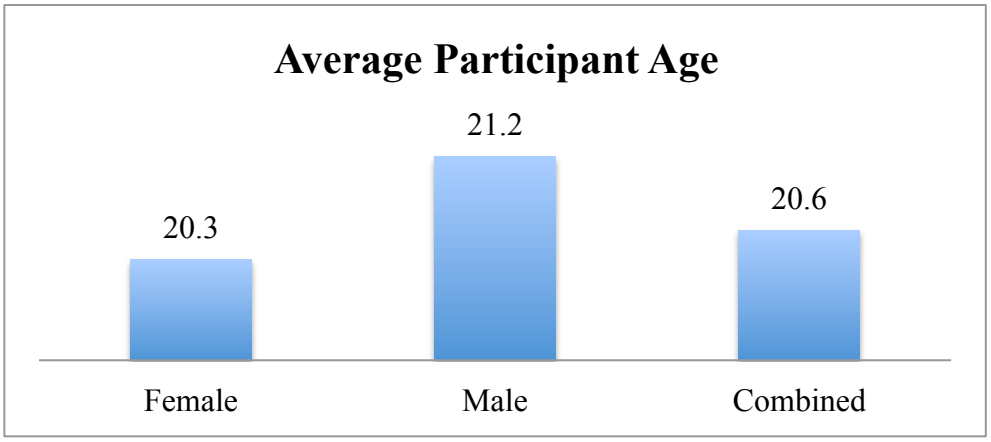

Figure 1.1 Listener participants average ages categorized by gender and overall.

All participants were native English speakers, and none had a second native language. All were from the United States and had lived only in the United States. There were thirty-eight female participants and fifteen male participants. That results in a 72% 28% split in the participant gender distribution. The participants were unaware of what linguistic feature was being tested.

Figure 1.2 below shows the gender categories of participants also separated by ethnicity, where green represents male participants and purple represents female participants.

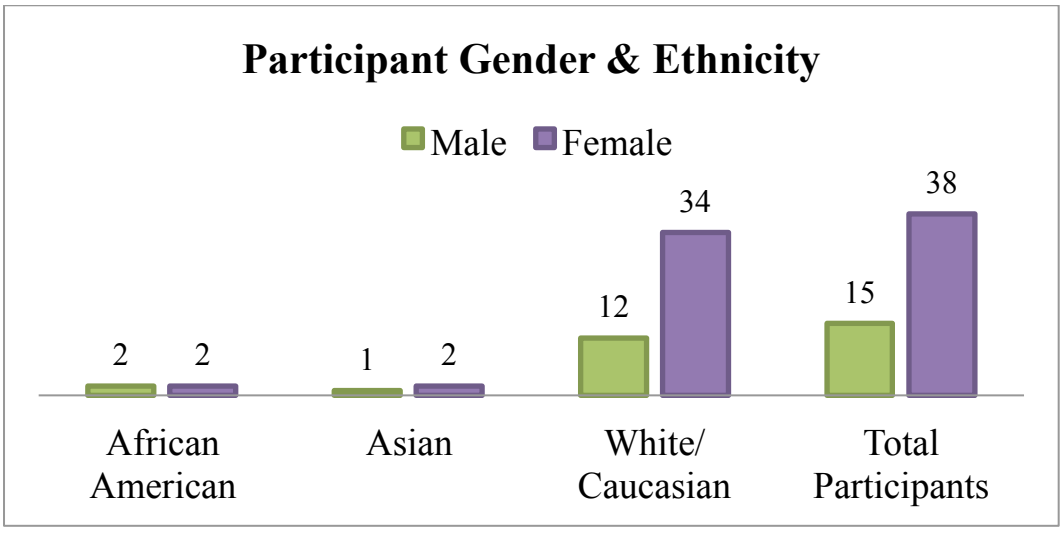

Figure 1.2 Participants gender categorized by ethnicity and overall

Section 4: Experiment One

Methods

Creating tokens

For this experiment, in order to create the stimuli, the volunteer speakers recorded the sixteen sentences found in Appendix A with both the modal register and creaky voice phonation.

These sixteen sentences were chosen because of their factualness, but also because of their mundaneness, and even boringness. Some sentences were found on a website listing animal facts (#TheFactSite, 2015), and some were written by the primary investigator/ author. Because participants were judging the speakers on their intelligence and education among other factors, it was important that the sentences all be similar in content, form, factualness, and interest level, so as not to skew the results. The primary investigator chose appropriate sentences with corroboration from the thesis advisor.

The volunteer speakers recorded the sentences in a Whisper Room sound booth, using an ElectroVoice RE-20 and Steinberg UR-22 usb audio interface. Audacity was used for the initial recording. Audacity was running on Ubuntu Linux with a low latency kernel. The recordings were analyzed and cut into appropriate lengths using Praat (Boersma & Weenink, 2013). Visual, auditory, and measurement analyses were compared in Praat to make sure that the modal register sentences did not contain portions of creaky voice, and that the creaky voice phonation sentences contained a majority of creaky voice over the modal register. Sentences were also held constant between speakers and voice qualities for tone and prosody. Sentences were recorded until appropriate tokens were identified using the criteria above.

Data Collection Process

Participants were asked to listen to sixteen sentences and answer six questions on speaker qualities: confidence, intelligence, femininity, hesitancy, masculinity, and education. These qualities were chosen in part based off of Yuasa's 2010 study, and in part to look at pairing interactions. The six characteristics paired in two opposites and one like pairing: femininity/ masculinity, confidence/ hesitancy, and intelligence/ education, while these characteristics were paired, identical or perfectly opposite results are not expected. Specifically, it was important to ask both about femininity and masculinity separately to see if Podesva and Ohala's theory of creaky voice being iconic for masculinity would be upheld in this study or not.

All six questions were phrased in the following manner: "The speaker sounds _____." A five point Likert scale was used to measure listener judgments. Likert ratings ranged from "Strongly Disagree," "Disagree," "Neutral," "Agree," to "Strongly Agree" (See Appendix C). Participants were asked to listen carefully throughout the entire task, as each sentence was only played once. Participants were not informed that they were listening to tokens of various voice qualities, and similarly, were not informed to focus on any specific quality to make their judgments.

There were four versions of experiment one. Each version contained eight sentences by each speaker. Of these eight sentences, for both the male and female speaker, four were in creaky voice and four were in modal voice. All experiments were created and run in the software called Open Sesame (Mathôt, 2012).

Results

Participants rated both the male speaker and the female speaker on six characteristics. To calculate the data, averages were found for each speaker and each voice quality for the six different characteristics. A linear mixed effects model was fitted with the various characteristics ratings as the dependent variable, and speaker and style (or voice register) as fixed effects. Participant and sentence were also included as random effects with random intercepts and slopes. Statistical analysis was run in R (R Core Team, 2013).

Table 1 below shows the listener participant Likert scale judgment averages. Rows are divided between the male and female speaker and subdivided between voice styles, creak and modal. Columns are characterization/ index question categories. A 0.00 rating represents a unanimous "strongly disagree" ratings across all participants. A 1.00 rating represents a unanimous "disagree" rating, a 2.00 rating represents a unanimous "neutral" rating, a 3.00 rating represents a unanimous "agree" rating, and a 4.00 rating represents a unanimous "strongly agree" rating.

Table 1: Likert Scale averages for experiment one sentences categorized by speaker, voice style and question.

		Educated	Confident	Feminine	Hesitant	Intelligent	Masculine
Male speaker							
	creak	2.80	2.90	0.43	1.14	2.82	3.44
	modal	2.96	3.24	0.52	0.91	3.01	3.36
Female speaker							
	creak	2.67	2.46	3.21	1.58	2.70	0.73
	modal	2.87	2.99	3.50	1.10	3.00	0.47

Figure 2.1 below shows the listener participant Likert ratings. Like *Table One* above, a 0.00 rating on the y-axis of the scale represents unanimous "strongly disagree" ratings across all participants, a 1.00 rating represents a unanimous "disagree" rating, a 2.00 rating represents a unanimous "neutral" rating, a 3.00 rating represents a unanimous "agree" rating, and a 4.00 rating represents a unanimous "strongly agree" rating. Along the x-axis each characteristic/ index question is listed. The bars represent voice quality and speaker gender. The red bars represent the modal register, while the blue bars represent the creaky voice phonation. The boldly shaded bars represent the male speaker, while the paler shaded bars represent the female speaker.

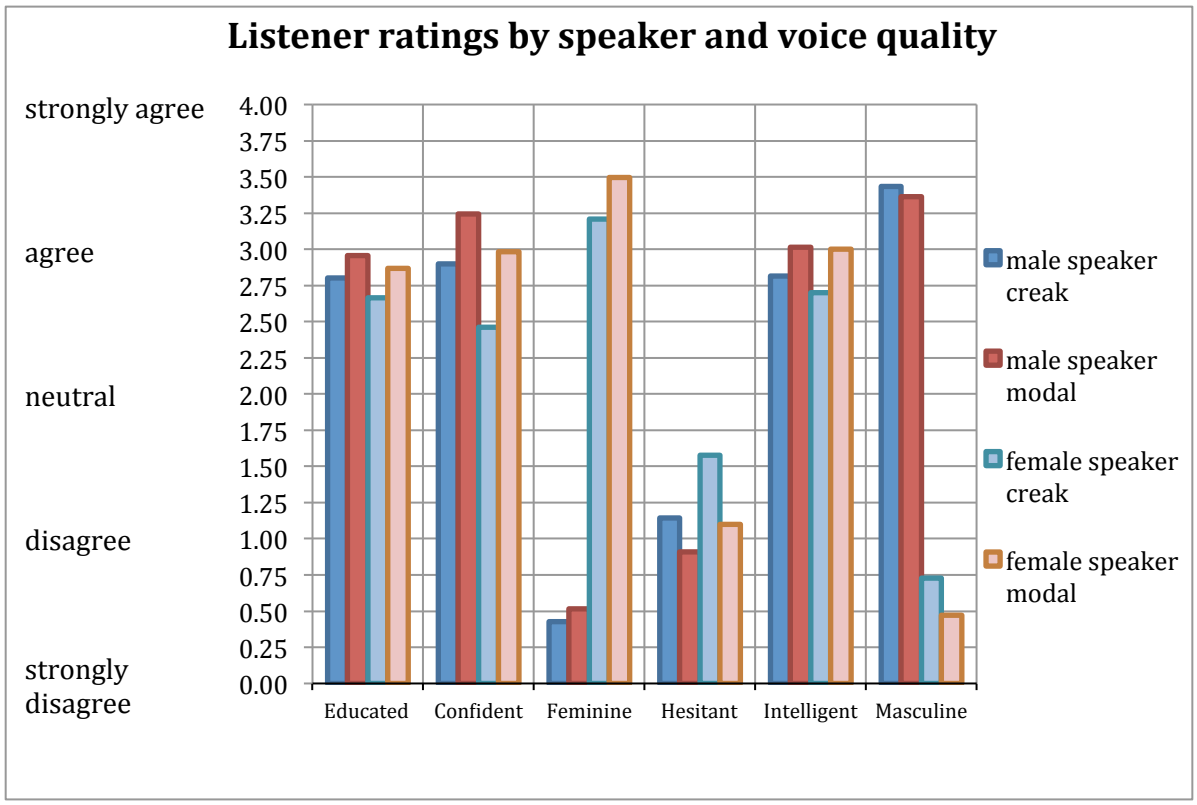

Figure 2.1 Likert Scale averages for experiment one sentences categorized by speaker, voice style, and question.

For the characteristic "educated," the male speaker was rated as less educated when speaking in the creaky voice phonation as compared to the modal register. The same effect was seen with the female speaker. The female speaker is rated as less educated than the male speaker for both voice styles. Both the speaker and style differences are statistically significant using the linear mixed effects model described above. The speaker has a significance rating of (β=0.055; t=2.662; $p<0.01$), while the voice style has a significance rating of (β=-0.087; t=-3.073; $p<0.01$). For this characteristic, there was no significance in the interaction between speaker and style.

For the characteristic "confident," the male and female speaker were rated as less confident when speaking in the creaky voice phonation than in the modal voice register. As seen with the "educated" characteristic, the female speaker is rated as less confident than the male speaker overall. Both of these differences are statistically significant. The speaker has a significance rating of (β=0.174;t=6.927;$p<0.001$), while the voice style has a significance rating of (β=-0.218; t=-4.279; $p < 0.001$). The interaction between speaker and style was approaching significance but was not statistically significant.

For the characteristic "feminine," both speakers, male and female, were rated as less feminine when speaking in the creaky voice phonation as compared to the modal voice register. The female speaker was rated as more feminine than the male speaker. Both of these differences were statistically significant in the mixed linear effects model. The speaker had a significance rating of (β=-1.440; t=-63.444; $p<0.001$), while the voice style had a significance rating of (β=-0.095; t=-4.170; $p<0.001$). In addition, the interaction between speaker and style was also significant at (β=0.050; t=2.207; $p<0.05$). This means that not only was there a difference between the style (creak to modal) and the gender (male

to female) but there was a significant difference between the male creaky rating and the female creaky rating as well.

For the characteristic "hesitant," both speakers male and female were rated more hesitant when using the creaky voice phonation, and the female speaker was rated as more hesitant than the male speaker overall. Both of these differences were statistically significant. The speaker had a significance rating of (β=-0.157; t=-5.220; $p<0.001$), while the voice style had a significance rating of (β=0.176; t=4.587; $p<0.001$). As with the "feminine" characteristic above, the interaction between speaker and style was significant for the "hesitant" characteristic as well, with a significance rating of (β=-0.061; t=-2.034; $p<0.05$).

For the characteristic "intelligent," both speakers, male and female, were rated less intelligent when using the creaky voice phonation, and the female speaker was rated as less intelligent overall. The only statistically significant variable found with the linear mixed effects modal was voice style with a rating of (β= -0.125; t=-2.940; $p<0.01$). Speaker was not a significant variable for the characteristic "intelligent."

For the final characteristic "masculine," both speakers, male and female, were rated as more masculine when using the creaky voice phonation in comparison to the modal register, and the male speaker was rated as more masculine compared to the female speaker. Both of these differences were significant. The speaker had a significance rating of (β=1.401; t=57.369; $p<0.001$), while voice style was (β=0.082; t=3.366; $p<0.001$). The interaction between speaker and voice style approached significance at ($p<0.06$), but was not statistically significant.

Below is *Figure 2.2,* which depicts the listener participant Likert scale data for the male speaker from *Figure 2.1*. Again, the blue bar represents creaky voice, while the red bar represents the modal register. The y-axis scale is the same as in *Figure 2.1,* as well as the categories for the x-axis.

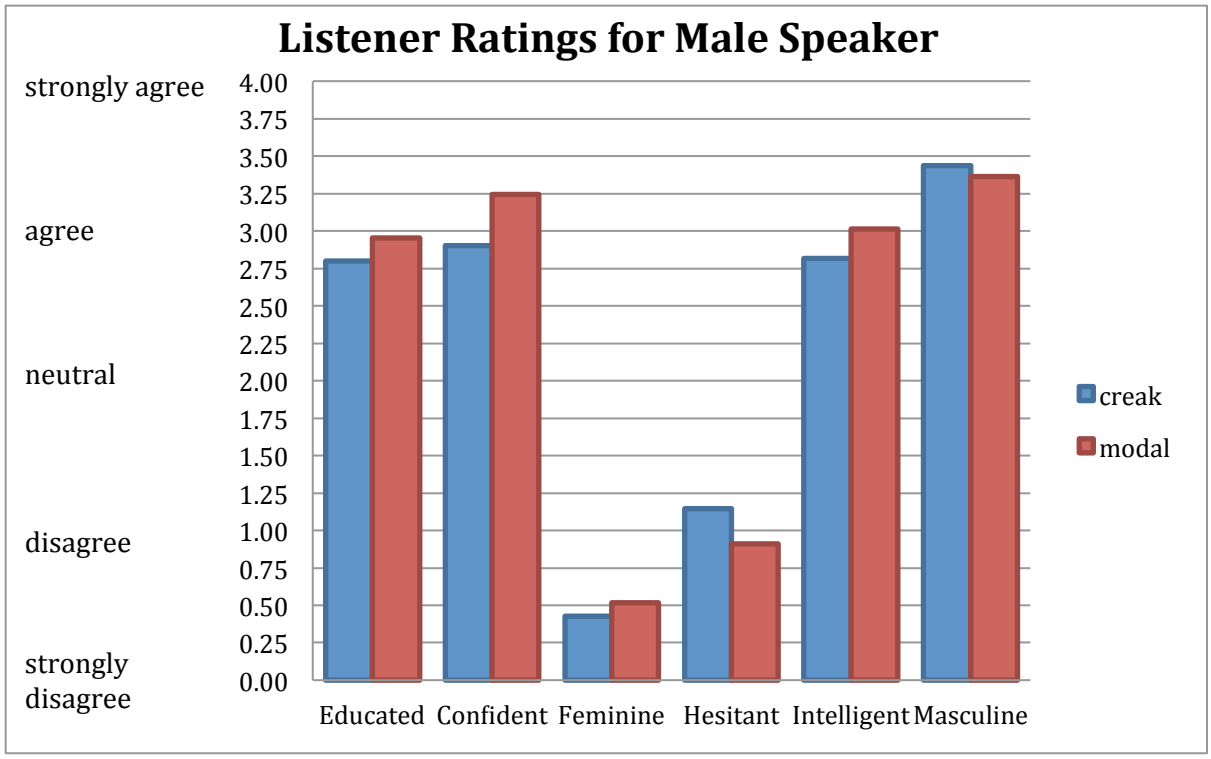

Figure 2. 2: Likert scale averages for experiment one sentences for the male speaker, categorized by voice type and question

Below is *Figure 2.3,* which depicts the listener participant Likert ratings for the female speaker from *Figure 2.1*. The blue bar represents creaky voice, while the red bar represents the modal register. The y-axis scale is the same as in *Figure 2.1,* as well as the categories for the x-axis.

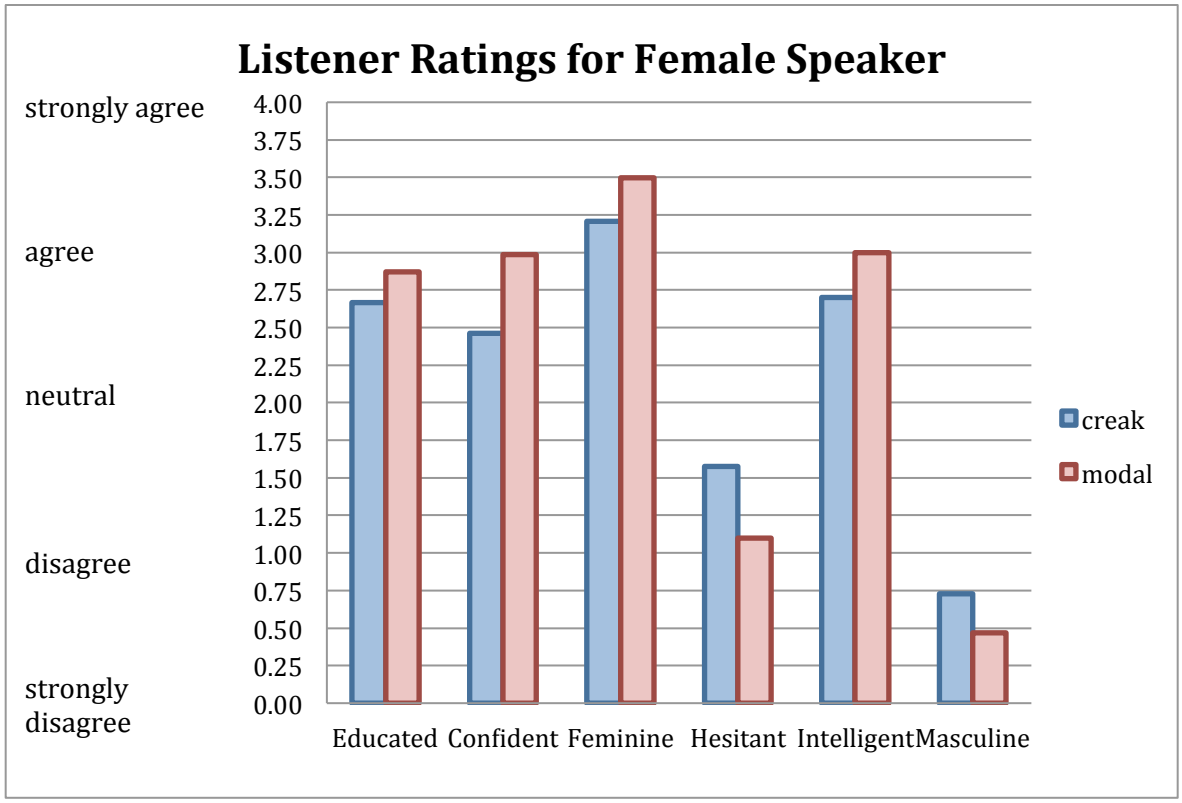

Figure 2. 3: Likert scale averages for experiment one sentences for the female speaker, categorized by voice type and question.

Discussion

As seen in the results section above there are three aspects to analyze in this section: the speaker, the voice style, and the interaction between the two. Every characteristic was significant in at least one of the three aspects and two characteristics (hesitant and feminine) were significant in all three aspects. The hypotheses will also be revisited in this section to see if the predictions were supported.

Hypothesis one stated, "Speakers will be rated as more masculine when speaking in the creaky voice and more feminine when speaking in the modal register" - this study supported this hypothesis. As we saw in the results section, both speakers, male and female were rated as more masculine when using creak register, statistically significantly, and more feminine when using the modal register statistically significantly. Because of the F0

effects in the creak register, these findings are not surprising. The creak register lowers F0 in both males and females, and people perceive lower F0s as being a male characteristic. Therefore, our validation of experiment one, hypothesis one is unsurprising. I will discuss whether this finding supports Ohala and Podesva's iconic masculinity theory later.

When analyzing masculine and feminine within the three aspects, speaker, voice style, and the interaction between the two, we expect to see a significant difference between the speakers since they are of different genders, and indeed we do. Hypothesis one, overlaps with the analysis of voice style. Crucially, when looking at the interaction between speaker and voice style for the characteristics masculine and feminine, we also see a significant interaction for the feminine, but not the masculine. Since these two categories are related it is somewhat surprising that we see significance in one but not the other, although if we look at the masculine significance rating again, it is approaching significance. This difference in the feminine interaction between gender and speaker tells us that the difference from female speaker modal to creaky voice phonation is significantly different from the change in the male speaker modal to creaky voice phonation difference. This means that participants were harsher on the female speaker's femininity when the creaky voice was used, and more lenient on the male speaker's femininity when the modal register was used. This finding supports the idea that women's speech is rated more negatively than male's speech.

Hypothesis two stated, "Speakers will be rated as less hesitant and more confident when using the creaky voice phonation" – this study failed to support this hypothesis. In fact, this study found the opposite. Participants rated speakers as more hesitant and less confident when using the creaky voice phonation.

When analyzing the characteristics hesitant and confident within the three aspects, speaker, voice style, and the interaction between the two, we expect the results to be opposites of each other since they are ideologically opposites. Indeed this is what we find. When considering the aspect of speaker, as shown in the results section hesitant and confident were statistically significant. The female speaker is consistently rated as more hesitant and less confident than the male speaker. Following the analysis of hypothesis one, participants were harsher on the female speaker, irrelevant to voice style, and solely due to gender. When we analyze the second aspect voice style, as we saw in the results section, creaky voice is rated significantly more hesitant and less confident than the modal register. This shows a negative attitude of the participants and a dis-preference towards hearing the creaky voice phonation. When we look at the interaction between speaker and voice style for the characteristics hesitant and confident, there is a statistical significance for hesitance, but not confidence. This result (for creak use by the female speaker) is congruent with the findings in Yuasa's 2010 study. This finding is interesting because it shows that the participants are willing to judge the female speaker more harshly than the male speaker for using vocal fry, but they don't give the male speaker a larger bump in the confident judgments, an index traditionally related to iconic masculinity, for using the modal register let alone the creaky voice phonation, which actually lowered confidence judgments. This shows that for the listeners, the creaky voice phonation is the deviation from the norm and is therefore more stigmatized.

Hypothesis three stated, "Speakers will be rated as more educated and intelligent when using the creaky voice phonation" – this study failed to support this hypothesis. As reported in the results section above, participants rated both speakers as less educated and

less intelligent when using the creaky voice phonation compared to the modal register, exactly the opposite of what was hypothesized.

When analyzing these characteristics through the three aspects, speaker, voice style, and the interaction between the two, we expect the results to be similar since the characteristics are often closely related. When considering the aspect of speaker we see that only education is significant, and not intelligence. While it is somewhat unexpected that these diverge, it is at least reassuring that they diverge in this way, implying that the female speaker is not inherently less intelligent than the male speaker. When looking at the voice style for education and intelligence we see that both are significant. Both speakers are perceived as less intelligent and less educated when using creaky voice. We do not see a significant interaction between speaker and voice style, showing that participants are not harsher for these characteristics when it comes to rating the female speaker, instead dis-preferring creaky voice for both speakers.

Hypothesis four stated, "Overall the male speaker will be rated more preferably than the female speaker" - this study supported this hypothesis. We have seen this throughout the analysis of hypothesis one through three. Even when the female speaker was rated more feminine than the male speaker, which on the surface level seems positive, the literature and societal judgments have shown that this is not a preferable quality, indexing hesitancy, inability to make a decision, and insecurity.

Hypothesis five stated, "The difference between modal register and creaky voice phonation will be more pronounced in the female speaker than the male speaker" – this study partially supported this hypothesis. As was stated in hypothesis one and two, the interactions between speaker and voice style were statistically significant. Again this

shows that the participants were harsher on the female speaker's use of creaky voice phonation than the male speaker's usage, following the tradition of negative views on female speech.

These results coincide with portions of the literature review, but disagree in other regards. As discussed by Eckert (2014), we see that female speech, regardless of voice modality or phonation, is rated more harshly than male speech. This is not a surprising result. Participants' harsh ratings of female speech are theorized to be the motivation for young women to use creaky phonation, and thus index the iconic masculine features that Ohala (1983) and Podesva (2015) discuss. However, we must discuss Ohala and Podesva's claim that creaky voice phonation is iconic. While this study found that creaky voice significantly affected participants judgments of the speakers', both male and female, masculinity, this was not to the levels that we would expect if creaky voice was iconically tied to masculinity and its corresponding indexes like confidence. In fact, this study didn't find any indication of a relationship between creaky voice phonation and the corresponding indexes of masculine speech that Ohala and Podesva discuss, such as confidence and intelligence. So while Podesva and Ohala claim that creaky voice phonation is iconic for the masculine index, this study did not support this theory. In addition, this study didn't find any of the corresponding masculine indexes, confidence and intelligence, to be iconic. In fact, intelligence was not significantly different for gender. This suggests a need to reevaluate Podesva and Ohala's claim, and to reconsider young females' motivations for using the creaky voice phonation.

Problems

While the experiments were designed to be as consistent and scientifically measurable as possible, there were, potentially, some considerations that were overlooked or perhaps did not function as expected in the experiment. This section will outline the effects these variables might have had on the outcome of the experiment findings.

There are three considerations that should be kept in mind when viewing the results for experiment one.

First, there were only two speakers recording tokens for the two experiments, one female and one male speaker. Therefore, any trends or conclusions must be taken with the understanding that they are based on one speaker of each gender and might not accurately represent the whole picture. It would be better to have multiple speakers per gender from a variety of age groups and ethnicities so that we could generalize our findings on larger portions of the population. Perhaps this is why we see deviation from Yuasa's findings and this study's findings.

Second, the sentences choices should be considered. While they were chosen for their uninteresting content and held consent for topic, it is uncertain how these sentences affected the participant listeners. In the future, pilot tests of sentences to find inter-rater reliability should be carried out before they are chosen for the final study. By doing this preliminary step, we can be sure to rule out any unintended effects as well as outlier sentences that standout from the group.

Finally, the listener participant sample was rather homogenous. Since the participants were pulled from introductory linguistics classes, from a majority white

institution, the participants were majority white and typically eighteen to twenty-one years old. Therefore, the findings must be understood as to represent the beliefs and opinions of this narrow section of the population. Studies (Anderson et.al, 2014; Yuasa, 2010) have shown that younger listeners are more open and accepting of creaky voice in general, and therefore it is possible that, at the very least, an older participation sample's result could vary widely (perhaps more negatively) from this sample group.

Section 5: Experiment Two

Methods

Creating tokens

For this experiment the two volunteer speakers recorded the word "pay" in as many levels of creaky phonation to modal register as possible, while attempting to hold tone constant. After both speakers completed many repetitions, Praat was used to isolate likely tokens and measure F0.

Each token was then matched with another token from the same speaker with similar F0s throughout the diphthong [eɪ], thus creating token pairs. For both speakers, each pair was placed on a continuum from lowest F0 to highest F0 in the creak to modal register range in order to create two continuum, one for the female speaker and one for the male speaker. The continuums were based more prominently from the final F0 formant at the end of the vowel, although the beginning of the vowel was still significant. For each speaker this resulted in a seven-step continuum. These steps were used to create six pairs (step 1_2, step 2_3, etc), which were in turn used for the AXB experiment. Category 7 on the continuums was modal register, Category 6 through 1 on the continuums were creaky voice phonation with 6 being the most modal (highest F0 levels) and 1 being the furthest

from the modal register (lowest F0 levels).

Below in *Table 2* the F0 frequencies are listed for each token separated by category and speaker. The F0 frequency for the beginning and end of the diphthong [eɪ] are represented as such: F0 at the beginning of token vowel/ F0 at the end of token vowel.

Table 2: F0 frequencies for the "pay" tokens at the beginning/ end of the vowel, separated by speaker and category.

		Male Speaker	Female Speaker
Category1	Token1	37/45	162/38
	Token2	32/44	157/38
Category2	Token1	61/28	76/64
	Token2	61/23	77/66
Category3	Token1	107/41	146/68
	Token2	109/51	165/61
Category4	Token1	114/50	89/73
	Token2	113/50	95/71
Category5	Token1	134/45	159/100
	Token2	133/48	169/105
Category6	Token1	149/98	175/55
	Token2	146/97	175/58
Category7	Token1	158/118	184/198
	Token2	166/121	199/196

For both experiments it was important to use natural speech samples for the tokens. Since this study investigates how creaky voice is perceived in everyday life, the samples needed to be realistic examples that could be heard in everyday speech, unlike in Anderson et al (2014). While this may have added some uncontrollable variation in the samples, it is in the hypotheses best interests to prioritize natural speech samples over perfectly controlled samples.

Data Collection Process

After the completion of Experiment One, participants were asked to do an AXB test. In an AXB task, listeners hear 3 tokens in a series, and are tasked with deciding if the second sound (X token) is more like the first (A token) or third (B token) sound in the series. Participants had 4 seconds after hearing the 3 tokens to push a button on the Black Box Toolkit USB response box to indicate their response, A or B. This is a classic psycholinguistic experiment similar to the ABX study in the foundational Liberman, Harris, and Griffith (1957) discrimination of speech sounds study.

All participants took the same version of Experiment Two and heard twenty-eight series of AXB tokens. Each participant heard fourteen AXB series from the male speaker and fourteen AXB series from the female speaker. Each participant heard three series for the first to second step and the sixth to seventh step in the creaky to modal voice continuums, while only hearing two series for the other step pairings (2-3,3-4,4-5, & 5-6).

Again, participants were not informed of the experimental feature. This experiment was also created and run in the software called Open Sesame.

Results

In experiment two, participants responded with which token, first (A) or third (B), in a series of three tokens, the second (X) token sounded like. The data was collected and proportioned for comparison. Statistical analysis was done in R.

When comparing proportion of correct responses, reported in *Figure 3.1* below, it appears that participants were best at perceiving the correct answer between the series using category six and seven on the continuum. This is the switch from the modal register to the

creaky voice phonation.

Overall, it also appears the participants were better at perceiving the correct response in the male speaker. This was shown to be statistically significant. There is a statistical difference between the participants' perceptions of the male and female speaker with the following ratings ($\beta = 0.042$; $t = 2.442$; $p<0.05$). This is especially true the lower down the creaky voice spectrum we go (1_2, 2_3, & 3_4). There was a statistically significant interaction between speaker and category for the 3_4 category with significance ratings as follows ($\beta = 0.091$; $t = 2.267$; $p<0.05$). In addition, the interaction between speaker and category approached significance for the 2_3 category, with significance ratings as follows ($\beta=0.083$; $t = 2.083$; $p = 0.054$).

While not significant, we can see that participants were better at perceiving the correct response for the female speaker in categories 4_5, 5_6, and 6_7.

Below is *Figure 3*, which depicts the AXB proportions of correct responses for each series pairing. The green bar represents the male speaker and the purple bar represents the female speaker. The y-axis represents percent correct, with 1 representing 100% correct, .5 representing 50% correct, and 0 representing 0% correct. The x-axis represents the continuums series that were tested, where 7 is the most modal and 1 is the most creaky. The lines represent the linear trend lines for each speaker.

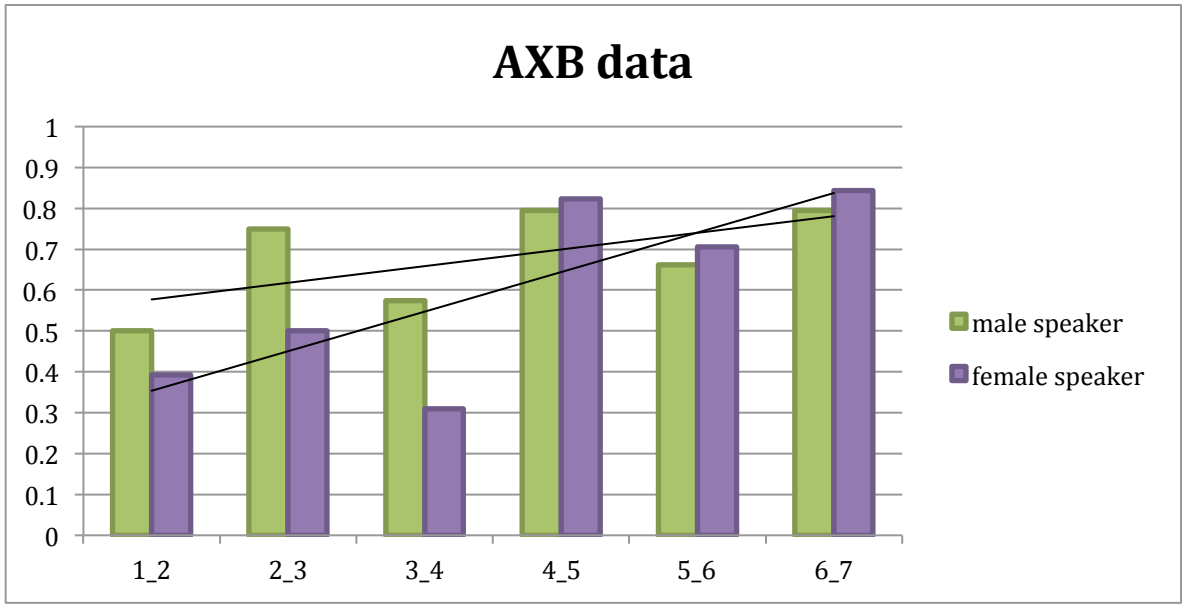

Figure 3.1 :Proportions of correct responses for experiment two, AXB study, categorized by speaker and series, with linear trend lines for each speaker included.

Discussion

For this experiment, as was seen above in the results section, there were two considerations for analysis: significance of speaker, and significance of interaction between speaker and category level. We found that participants judged speakers as statistically significantly different, and that one interaction was statistically significantly different, while another approached significance. While this experiment did not perform exactly as was expected, there are interesting findings to discuss. Overall, it was expected that the data would have a sharp categorical cut off for creaky voice perceptivity, but we do not see this clearly, whether by design fault or the situation being more complicated than originally hypothesized is still unclear at this time.

Hypothesis one for experiment two stated: **"**Listeners will be more perceptive (i.e. more accurate) of the modal to creak register transition in the female speaker than the male speaker" – this was seen in this study, but not statistically supported. While this is observed in the data, the difference was not statistically significant. We did see that overall

listeners were more perceptive of the male speaker than the female speaker, thus having a higher proportion of correct responses. This was not a predicted hypothesis but does give us interesting analysis. Since participants were more accurate with the female speaker's voice closer to the modal register, and less so as the creaky voice increased (F0 dropped), it would be interesting to rerun the study with more categories at the creaky voice phonation/ modal register divide, and expand category pairs into just the modal register to see if participants became statistically significantly more correct with the female speaker than the male speaker in the modal register.

The second hypothesis for experiment two stated: "Listeners will lose perceptiveness at the lowest creak categories" – this study supported this hypothesis. This was observed for both speakers, thus validating this hypothesis. In addition, there was a significant interaction at these lower creaky voice levels (lower F0) between category and speaker, meaning that while listener participants were less likely to answer correctly as creak increased they were significantly worse at this for the female speaker than the male speaker.

This finding was unexpected, but is perhaps the most exciting result from experiment two. The data shows a sharp drop in correct perception of the female speaker between categories 4_5 and 3_4. Perhaps, this is the categorization that was expected, although it is in a different location than anticipated. This finding suggests that participant listeners were able to distinguish the upper F0 limits in the creak register but not the lower registers, at least for the female speaker. There are two possible explanations for this finding; either the anatomical mechanisms of our auditory system do not allow for clear perception at these lowest levels, or the participants do not have a sufficient examples of

female creak at these low levels in their exemplar modals to have clear categories. Since we did not see the clear drop-off in the male speaker's results, I believe that the second explanation fits this data better. In the exemplar modal explanation, based off of Keith Johnson's 2007 theory, we did not see the categorical drop for the male speaker's data because the F0 differences were not as significant in the male speaker as the female speaker, and therefore was more familiar and had more examples in the participants' exemplar phonation modal. This suggests that participants' judgments and interactions with creaky voice phonation can be manipulated. The ability to manipulate and train the salience perception would allow for and explain the highly indexical nature of creaky voice that we see in the literature and in this study; because participants' perception can be trained for saliency, it can be trained for indexical categories.

Problems

During experiment two there are three issues that should be considered when reviewing the findings: variable control, category level matching, and modal categories.

First, to consider variable control, it is important to remember how pairs were made for each category. Each pair was matched based on F0 levels at the beginning and the end of the vowels. Since the vowel in "pay" is a diphthong it was important to consider both points when matching the tokens. Creaky voice is a complex phonation however, and it is possible that there were other variables in the creaky phonation that affected the participants' responses that were not controlled. This is the danger of using natural tokens, but the benefits, natural speech tokens, were more important than knowing for certain that everything was perfectly controlled.

Second, a consideration should be the matching of the categories. Since F0 was the

only method chosen for creating the pairs, it is possible that they were unevenly matched in another consideration. As shown in the literature review, F0 is a measurement that is strongly correlated with creaky voice, but there are other methods that could have been considered such as F2-A1, H2 –H1, etc. Perhaps one of these methods might prove to be a better method for finding pairs, and this is why we didn't see the clear categorization that was expected, particularly in the male speaker's data, and particularly with category 4_5.

Third, all but one of the pairs was with creaky voice, which created unbalanced continuums. Perhaps if more categories of modal phonation were include we would see opposite effects, where the participants were statistically better at choosing the right answer for the female speaker instead of the male speaker. While this does not affect the results from this data, it might present an incomplete picture.

Section 6: Conclusions

Overall, there are several conclusions that can be made about this study. First, participants don't like creaky voice phonation– especially when used by a woman. In experiment one, the creaky voice phonation was consistently rated more negatively for both speakers, except when considering the male speaker's creaky voice masculinity rating. But in every other consideration it was negative. Creaky voice phonation made the speakers: less educated, more hesitant, less intelligent, more masculine (possibly negative for female speaker), less feminine (possibly negative for female speaker), and less confident. These conclusions were unexpected. Unlike Yuasa's positive results where females were rated "educated, urban-oriented, and upwardly mobile," this study shows the opposite trend even precisely using the same categorization "educated." This difference in findings could be because Yuasa used conversational data, but this study used natural data, so it would

surprise me if this was the main reason for the differing results.

This study's findings are also incongruent with research that shows lower F0s are more confident, dominant, authoritative, etc. And not only were these iconic characteristics, confident, dominant, and authoritative, not found for the female speaker when using creaky voice phonation, but, critically, they were not found even for the male speaker. This result leads us to reevaluate Ohala and Podesva's claim that creaky voice is iconic for indexing masculinity. If we must reject creaky voices' iconic masculinity, then we must question the theorized reason why young female speakers are increasing their use of creaky voice. While it had been hypothesized that using creaky voice as a female indexed these positive traditional male indexes, we find that this does not work even for the male speaker, and therefore cannot be the motivation for the increase.

The second conclusion from this research is that the participants really just don't like female speech at all, or, at least, they like it less than male speech regardless of the voice quality. This was an unexpected, but not unsurprising result. As was reported in the literature review, female speech and characteristics are frequently judged more harshly than male characteristics and speech. While this could be the motivation for females using creaky voice phonation – to imitate male speech- as was reported above this study shoes that this method is not a productive strategy.

Third, it seems that while females are using creaky voice phonation more often, perceptually we are not used to hearing extremely low F0s from female speakers. This unfamiliarity could be driving negative perceptions. Perhaps, with more examples in our exemplar modals of creaky voice low F0s for female speakers, perceptions of young female creaky voice will improve.

If your goal is to sound more masculine at any cost, this study does support the idea, although not to an iconic level, that creaky voice will help you index masculinity regardless of your gender. But, creaky voice does not seem to give the original co-indexes that have been traditionally related to low F0, male, masculine speech, namely: confidence, intelligence, and education.

Section 7: Future

Creaky voice is not likely to fade out of speech, or become less productive in presenting character indexes, so it would be interesting to continue with the project, making some of the suggested alteration in the problems section. It would also be interesting to see a longitudinal study over say fifteen or twenty years to look at how these character index judgments evolve with a likely continued creaky voice increase.

In future studies, I would like to include different characteristics with the original indexes for the judgments in experiment one, including the original upwardly mobile, and urban from Yuasa's 2010 study. Perhaps the characteristic chosen for this study, although related to previous indexes, were ill-chosen. Changing the characteristic indexes might yield different results more compatible with recent findings. Also, as was discussed in the problems section, having only one male and one female speaker representing their genders is problematic. In the future having a variety of male and female speakers would yield more reliable natural results. Perhaps this is another reason why this study's results differ from Yuasa's. In addition, perhaps in the future, creaky voice will come to be indexed with a new characterization and the methods should be adjusted for this.

Finally, this current study could be expanded to show firsthand, and not just through the literature review, that the use of creaky voice is natural, ubiquitous, and favors female

speakers. If the participants were held to similar demographics, it would be interesting to look at creaky voice phonation used in classrooms during formal presentations, and informal discussions. If the study were expanded to include more age groups then a different situation would have to be envisioned. Looking at these situational differences might help postulate a new theory as to why young female speakers are adopting the creaky voice phonation.

Appendix A: Experiment One Sentences

1. An aardvark is the first animal listed in the dictionary.
2. Arachnophobia is the fear of spiders.
3. A baby wolf is called a pup.
4. The butterfly is the adult state of the caterpillar.
5. When whales jump out of the water it is called breeching.
6. Cats are the only animal that can purr.
7. A cheetah is the fastest animal on land.
8. An elephant is the worlds largest living land mammal.
9. A baby deer is called a fawn.
10. Fireflies are also known as lightning bugs.
11. The life of a housefly is only fourteen days.
12. Male lions have a mane of fur around their head.
13. Squirrels bury acorns in the fall to feed themselves throughout the winter.
14. Tigers have striped skin and striped fur.
15. A whale is the largest mammal.
16. The fear of animals is called zoophobia.

Appendix B: Participant Survey

Demographic Information

Please answer the following questions to the best of your ability. You may leave questions blank if you do not feel comfortable providing an answer. If you select "other" as an answer, please provide more specific information in the blank provided.

Name: _____

Age: _____

Gender (check one): □ Male □ Female □ Do not wish to Identify

Race/ethnicity (check one):

□ White □ African-America □ Hispanic □ Native American □

Other_____

Birthplace (city and county):

_____

Place of residence (city and county):

_____

Other places lived (include number of years lived in each location):

_____

Native language(s): □ English □ Other: _____

Languages Studied:

_____

Appendix C: Experiment One Screenshot

Below is a screenshot from OpenSesame during Experiment One to show indexical questions/ categories and the Likert rating scales.

REFERENCES

Abdelli-Beruh, N., Wolk, L., & Slavin, D. (2013). Prevalence of Vocal Fry in Young Adult Male American English Speakers . *Journal of Voice , 28* (2), 185-190.

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).

Arana, G. (2013). Creaky voice: Yet another example of young women's linguistic ingenuity. The Atlantic. Retrieved Nov. 4, 2015, from http://www.theatlantic.com/sexes/archive/2013/01/creaky-voice-yet-another-example-of-young-womens-linguistic-ingenuity/267046/

Babel, M.; McGuire, G. (2015). Preceptual Fluency and Judgments of Vocal Aestetics and Sterotypicality. *Cognitive Science*, 39, 766-787

Blomgren, M.; Chen, Y. Ng, M.L.; Gilbert, H. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *Journal of the Acoustical Society of America,* 103(5), 2649-2658

Boersma, P. & Weenink, D. (2013): Praat: doing phonetics by computer [Computer program]. Version 5.4.19, retrieved from http://www.praat.org/

Chen, Y., Robb, M. P., & Gilbert, H. R. (2002). Electroglottographic 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. (1995) Glottal source modeling for voice conversion. *Speech Communication,* 16(2), 127-138

Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour,* 65, 997-1004

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour,* 60, 773-780

Drugman, T., Kane, J., & Gobl, C. (2014). Data-driven detection and analysis of the patterns of creaky voice . *Computer Speech and Language , 28*, 1233-1253.

Eckert, P. (2014). Language and Gender in Adolescence. Editors Ehrlich, S.; Meyerhoff, M.; Holmes, J. In *The Handbook of Language, Gender, and Sexuality*. (529-545). Somerset, NJ: John Wiley & Sons, Ltd

Greer, S. D. (2015). *The Perception of Coolness: Voice Quality and Its Social Uses and Interpretations* . University of Calgary , Department of Linguistics.

Grivičić, T., & Nilep, C. (2004). When Phonation Matters: The Use and Function of *yeah* and Creaky Voice. *Colorado Research in Linguistics,* 17(1), 1-11

Henton, C. G., & Blandon, R. A. W. (1985). Breathiness in Normal Female Speech: Inefficiency Versus Desirability, *Language & Communication,* 5(3), 221-227

Hollien, H., & Wendahl, R. W. (1968). Perceptual Study of Vocal Fry, *Acoustical Society of America (ASA),* 43(3), 506-509

Hollien, H. (1972). On Vocal Registers. *Communication Sciences Laboratory Quarterly,* 1-33

Ishi, C.T.; Sakakibara, K.; Ishiguro, H; Hagita, N. (2008) A method for Automatic Detection of Vocal Fry. *IEEE Transactions on Audio, Speech, and Language Processing,* 16(1) 47-56

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. W. Mullennix (Eds), Talker variability in speech processing (pp. 145–165). San Diego, CA: Academic Press.

Kane, J., Drugman, T., & Gobl, C. (2013). Improved automatic detection of creak . *Computer Speech and Language , 27*, 1028-1047.
Ladefoged, P., & Johnson, K. (2011). A Course in Phonetics. Edition 6. Boston, MA. Wadsworth Cengage Learning
Levon, E. (2015). Integrating Intersectionality in Language, Gender, and Sexuality Research. *Language and Linguistics Compass*, 9(7), 295-308
Liberman, A. M., Harris, K. S., and Griffith, B. C. 1957. The discrimination of speechsounds within and across phoneme boundaries. Journal of ExperimentalPsychology, 54:358–368.
Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. doi:10.3758/s13428-011-0168-7
Mendoza-Denton, N. (2011). The Semiotic Hitchhiker's Guide to Creaky Voice: Circulation and Gendered Hardcore in a Chicana/o Gang Persona . *Journal of Linguistic Anthropology , 21* (2), 261-280.
Moosmüller, S. (2001). The influence of creaky voice on formant frequency changes. *Forensic Linguistics,* 8(1), 100-112
Ohala J. (1983). The Origins of Sound Pattern in Vocal Tract Constraints. In MacNeilage P. (Ed), *The Production of Speech* (189-216). Springer New York
Pennock, B. (2005). The changing voice of women. In *Actas XXVIII Congreso Internacional AEDEAN,* 407-415
Pittam, J.; Gallois, C. (1987) Predicting Impressions of Speakers from Voice Quality: Accoustic and Perceptual Measures. *Journal of Language and Social Psychology,* 5(4) 233-247
Podesva, R.; Callier, P. (2015). Voice Quality and Identity. Applied Linguistics, 35, 173-194
Podesva, R.; Kajino, S. (2014) Sociophonetics, Gender, and Sexuality. Editor Ehrlich, S.; Meyerhoff, M.; Holmes, J. In *The Handbook of Language, Gender, and Sexuality* (103-122). Somerset, NJ: John Wiley & Sons, Ltd
Podesva, R. (2013). Gender and the social meaning of non-modal phonation types, *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society,* 427-448
Podesva, R. (2007) Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics,* 11(4) 478-504
Podesva, R. (2006). *Phonetic Detail in Linguistic Variation: Its Linguistic Significance and Role In the Construction of Social Meaning*. Stanford University, Department of Linguistics
R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Wiener, H. J., & Chartrand, T. L. (2014). The Effect of Voice Quality on Ad Efficacy. *Psycology & Marketing , 31* (7), 509-517.
Winkler, P. (1983) Subjective assessments of Differnet Vocal Timbres: Hoarse, Creaky, and Strained Voices. *Quantitative Linguistics,* 19(19), 289-311.
Wolk, L.; Abdelli-Beruh, N.B.; Slavin, D (2012). Habitual Use of Vocal Fryin Young Adult Female Speakers. *Journal of Voice,* 26(3), e111-e116
Yan, Q., Vaseghi, S., Rentzos, D., & Ho, C.-H. (2007). Analysis and Synthesis of Formant Spaces of British, Australian, and American Accents . *IEEE Transactions On Audio, Speech, And Language Processing , 15* (2), 676-689.

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

#TheFactSite (2015). Animals, Bizarre. 300 Random Animal Facts. Retrieved on 12/10/15, from http://www.thefactsite.com/2010/09/300-random-animal-facts.html Web.

VITA

Kaitlyn Elizabeth Lee

Education:

B.A. in Classical Studies *cum laude*
Centre College, May 2013

Professional Positions Held:

Teaching Assistant, August 2014- May 2016
University of Kentucky, Writing, Rhetoric, and Digital Studies

Scholastic and Professional Honors:

Fulbright English Teaching Assistant, August 2013 – June 2014
Fulbright Program

John C. Young Scholarship, August 2012 – May 2013
Centre College