THE EFFECT OF VOCAL FRY ON SPEECH INTELLIGIBILITY

By

Kaleigh Susan Cammenga

A THESIS

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

Communicative Sciences and Disorders – Master of Arts

2018

ABSTRACT

THE EFFECT OF VOCAL FRY ON SPEECH INTELLIGIBILITY

By

Kaleigh Susan Cammenga

Speech intelligibility is a measurement of the interaction between three components: the speech signal, the transmission channel, and the listener. Anything that interferes with any of these components can affect intelligibility. Vocal fry, though sometimes associated with vocal pathology, is commonly used in non-disordered speech. Speech produced with vocal fry differs from typical (modal) voicing in terms of pitch, volume, and quality. These differences may negatively impact intelligibility. Currently, no direct evidence exists regarding the impact of vocal fry on speech intelligibility. The purpose of the current study was to answer the research question: Does vocal fry affect the intelligibility of spoken words? We hypothesized that single words produced with vocal fry would be less intelligible than single words spoken in modal voice due to the acoustic characteristics and perceptions of vocal fry. To test this hypothesis, words spoken in both vocal fry and modal voice were collected and compiled to produce a standard intelligibility test procedure. Data from 26 listeners who completed the intelligibility test were analyzed in terms of intelligibility score and listening difficulty rating. The mean intelligibility score as a percentage of words correctly identified was 62.08% for words spoken with fry and 64.56% for words spoken with no fry. This difference was statistically significant. The mean listening difficulty rating was 4.98 for words spoken with fry and 4.56 for words spoken with no fry. This difference was also statistically significant. These results suggest that vocal fry does negatively impact speech intelligibility at the single word level. Decreased speech intelligibility may have numerous possible linguistic, social, and economic implications.

This thesis is dedicated to my family and my fiancé.
Thank you for your constant love, prayers, encouragement, and support.

ACKNOWLEDGEMENTS

A special thank you to Dr. Eric Hunter for serving as my thesis advisor. Thank you for sharing your knowledge and passion for voice research with me and for inspiring my newfound love for research. It has been an incredible experience learning under you. An additional thank you to the members of my committee, Dr. Peter Lapine and Dr. Amanda Hampton-Wray, for sharing your wisdom and perspectives with me.

Thank you to Lady Catherine Cantor Cutiva and Pasquale Bottalico for your expertise and assistance in the development and execution of my thesis. Thank you to Mark Berardi and Callan Gavigan for assisting in the recording of the audio samples. Thank you to Kirsten Rockey and Russel Banks for your assistance in data collection.

And finally, a special thank you to Jacob Morrison. Thank you for loving and encouraging me and for taking interest in my research, serving as my sounding board, and offering your experience and knowledge to assist me along the way. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The term vocal register refers to the "perceptually distinct regions of vocal quality that can be maintained over some ranges of pitch and loudness" (Titze, 2000, pp. 282). Three vocal registers are typically most recognizable: modal, falsetto, and vocal fry. These registers differ based on the vibratory patterns, degree of tension, compression, and adductive force of the vocal folds. Modal phonation is used most frequently in conversational speech. It typically ranges in fundamental frequency from 150 to 500 Hz in adult women and from 80 to 450 Hz in adult men. Vocal fold vibration using modal voice is highly periodic, and it is characterized by moderate tension, compression, and adductive force. Falsetto typically occurs at frequencies above modal phonation and is most commonly used in performance voice rather than normal speech. Falsetto voice production is characterized by moderately high tension and compression and high adductive force of the vocal folds. The vocal folds are typically long, thin, stiff, and often bow-shaped during falsetto voicing (Boone, McFarlane, Von Berg, & Zraick, 2014).

The bursts of acoustic energy that occur following each closure of the vocal folds are not individually perceived in the modal and falsetto registers. Therefore, voice produced in these registers is perceived as continuous sound. However, when the fundamental frequency drops below approximately 70 Hz, referred to as the crossover frequency, the individual bursts of acoustic energy and subsequent periods of silence that are produced by each cycle of opening and closing of the vocal folds can begin to be perceived by the auditory system. This cycle of pulses and gaps is what is perceived as vocal fry (Titze, 2000). Michel (1968) determined that vocal fry occurred at fundamental frequencies significantly lower than both modal register and harsh voice, providing further evidence that vocal fry exists as a distinct phonatory register. Previous authors have reported a mean range of 24-52 Hz (mean of 36.4 Hz) for males speaking in vocal fry and a mean range of 18-46 Hz for females speaking in vocal fry (Hollien & Michel, 1968; Michel, 1968).

Blomgren, Chen, Ng, and Gilbert (1998) compared the aerodynamic, physiologic, and perceptual characteristics of speech produced in the modal and vocal fry registers in 10 female and 10 male normal speakers. The results indicated that speech in vocal fry was produced at significantly lower fundamental frequencies than speech produced in the modal register for both males and females. It was also found that measurements of jitter and shimmer were significantly higher in vocal fry compared to modal. Signal to noise ratio was found to be significantly lower in vocal fry compared to modal. Subglottic air pressure was approximately 1.5-3 times higher in modal voice compared to vocal fry.

From the physiological point of view, vocal fry is associated with short, thick vocal folds with a lax covering and mild adductive force. Specifically, the thyroarytenoid muscles and vocal ligaments are shortened, which reduces the rate of vocal fold vibration (Boone, McFarlane, Von Berg, & Zraick, 2014). In order to maintain the pattern of vocal fold vibration associated with vocal fry, a minimum subglottal pressure of approximately 2 cm H2O is needed. The average airflow rate associated with vocal fry ranges from 12 to 20 cc/s (Boone, McFarlane, Von Berg, & Zraick, 2014). Because vocal fry occurs with less subglottal air pressure, it more frequently occurs at the end of sentences or phrases when air pressure naturally drops off. Because of this, some researchers have considered vocal fry to be a linguistic marker of paragraph and sentence boundaries (Wolk, Abdelli-Beruh, & Slavin, 2012); the contextual aspects of vocal fry will be discussed below.

Whitehead, Metz, and Whitehead (1984) investigated the vibratory patterns of the vocal folds during production of vocal fry. The researchers assessed the vibratory patterns of the vocal folds by using high speech laryngeal films during phonation of the English schwa vowel in pulse register by a female speaker. The speaker produced the vocal fry pattern with both a single and double opening and closing pattern. Both patterns produced the perceptual and acoustic characteristics associated with vocal fry (low pitched, popping quality; fundamental frequency within range typically associated with vocal fry). This suggests that vocal fry can be produced with either a single or multiple opening and closing pattern per cycle; however, the crackling sound that typically distinguishes vocal fry is often thought to be associated with this syncopated pattern and secondary beat (Boone, McFarlane, Von Berg, & Zraick, 2014).

Vocal fry was traditionally classified as an abnormal phonation pattern associated with vocal pathology or was thought to be vocally harmful at the least. Because vocal fry is produced with decreased subglottic air pressure, the resulting sound is typically at a lower sound pressure level (dB). Therefore, in order to compensate for this decrease in volume, speakers using vocal fry will typically increase vocal fold tension in order to increase loudness. Because of this increased tension, consistent production of fry has been shown to be harmful to the vocal fold mucosa (Boone, McFarlane, Von Berg, & Zraick, 2014) and could be damaging to the voice (Colton, Casper, & Leonard, 2006).

Although once classified as abnormal and potentially pathological, such classification is now considered inaccurate (Hollien, Moore, Wendahl, & Michel, 1966) as multiple studies have shown that vocal fry is commonly used in non-disordered speech and may be associated with the biological sex of the individual. For example, Wolk, Abdelli-Beruh, and Slavin (2012) examined the prevalence of vocal fry use in young adult female native Standard American-English speakers where they found that a majority of participants used vocal fry during a sentence reading task. Abdelli-Beruh, Wolk, and Slavin (2014) examined the prevalence of vocal fry in young adult males and compared it to the results of Wolk, Abdelli-Beruh, and Slavin (2012). The results indicated that the prevalence of vocal fry in sentence reading was significantly lower for male speakers compared to female speakers. However, Abdelli-Beruh, Drugman, and Red Owl (2016) examined the speech of 25 male and 29 female subjects. Using an algorithm to detect the "creak segments" associated with vocal fry, they found that the frequency with which vocal fry is used is not significantly related to the individual's biological sex. This contradicts previous research suggesting that vocal fry is used more commonly by females.

The prevalence of vocal fry use is thought to be increasing during recent years. Borrie and Delfino (2017) examined a possible explanation for this increased used. They studied the potential role of conversational entrainment in the use of vocal fry during conversation. Conversational entrainment describes the tendency of communication partners to modify their own behaviors to match up with those of the other person during a conversation. The results indicated that speakers used vocal fry significantly more frequently when conversing with a person who used a substantial amount of vocal fry than when conversing with a person who used less vocal fry. In addition, greater similarity between partners in terms of their use of vocal fry was associated with greater communicative efficiency and enjoyment, which may act as positive reinforcement towards the use of vocal fry in conversation.

Not only is vocal fry frequently used in non-disordered speech, a certain style of vocal fry production has been used as a vocal training technique and has even been utilized in vocal therapy. Because this type of vocal fry requires a relaxed larynx, it can be a useful technique for decreasing tension in the larynx allowing speech to be produced with considerably less tension in the vocal folds. Therefore, it has been used in the treatment of hyperfunctional problems in the cases of vocal nodules, functional dysphonia, and spasmodic dysphonia (Boone, McFarlane, Von Berg, & Zraick, 2014). Vocal fry has also been shown to be useful in the treatment of both puberphonia and ventricular phonation because it requires the client to relax and lower the larynx (Boone, McFarlane, Von Berg, & Zraick, 2014).

Vocal fry also seems to serve a linguistic role. For example, vocal fry is most often exhibited at the end of phrases when subglottal air pressure naturally drops. Abdelli-Beruh, Drugman, and Red Owl (2016) found that vocal fry occurred with the greatest frequency at the end of sentences, regardless of the length of the sentence. This suggests that vocal fry may act as a syntactic marker signaling paragraph and sentence boundaries (Wolk, Abdelli-Beruh, & Slavin, 2012). Lee (2015) submits that vocal fry may be used to mark parenthetical segments, such as elaborating on previous expressions or inserting background information into a narrative structure. Thus, vocal fry is a normal register that can be used selectively by speakers who do not have vocal pathologies. Its uses may include organizing information into different levels of importance, signaling meaning, or indicating a level of distance or detachment.

The use of fry in daily communication may come with a social cost. For example, vocal fry has been associated with negative connotations in relation to success in the labor market (Anderson, Klofstad, Mayew, & Venkatachalam, 2014). However, little is known about why the use of vocal fry, with its unique vibratory pattern and sound quality, is identifiable by a listener in a way that would lead to social implications. Since vocal fry is associated with perceptual differences in pitch, volume, and quality, the use of vocal fry may affect the intelligibility of speech. This may be perceived by the listener in terms of increased or decreased listening effort which may then affect the listener's perception of the speaker. To explore this possibility, the relationship between production of vocal fry and speech intelligibility should be examined.

Speech intelligibility has been defined as the degree to which the intended message of the speaker matches the response of the listener across various transmission systems (Schiavetti, 1992). Thus speech intelligibility is a measurement of the interaction between three components: the speaker, the transmission channel, and the listener. Anything that interferes with any of these components can impact the intelligibility of the communication process. Thus, because the production of vocal fry alters the speech signal produced by the speaker, intelligibility of speech might also be altered.

Schiavetti (1992) identified two types of tasks to measure intelligibility: word identification and scaling procedures. Word identification provides a frequency count of the number of correctly identified words which is typically converted into a percentage. This is the most widely used measure in the assessment of speech intelligibility in the speech sciences (Munro, 2013). Scaling procedures involve the listener making a judgment about the talker's intelligibility by identifying their perception of the intelligibility on a rating scale. This provides a direct measurement of the degree to which aspects of the speech signal may impact listening. For the purpose of this study, we will be assessing intelligibility in terms of both percent words correct and listening effort.

Currently, little direct evidence exists regarding the effects of vocal fry on speech intelligibility. However, previous research has examined the effect of various acoustical characteristics of speech and voice on speech intelligibility that may provide insight into the possible relationship between vocal fry and intelligibility. Ryalls and Lieberman (1982) performed a study in which the subjects were asked to identify vowels differing in fundamental frequency. They found that subjects identified vowels produced with a low or average F0 significantly more accurately than vowels produced with a high F0. However, there was no significant difference between vowels produced with an average F0 compared to a low F0 (Ryalls & Lieberman, 1982). Vocal fry is produced with a low fundamental frequency. This suggests that perhaps intelligibility may not differ between vocal fry and modal voice, at least not as a result of a decreased F0.

Bradlow, Torretta, and Pisoni (1996) evaluated the difference in intelligibility between sentences spoken by 20 speakers in terms of various acoustic characteristics. Their findings relevant to the current study showed that mean F0 was not correlated with intelligibility. However, range of F0 was positively correlated with intelligibility so that a wider range of F0 was associated with higher overall intelligibility. The findings of Watson and Schlauch (2008) were consistent with these results. They compared the intelligibility of resynthesized naturally produced sentences in terms of percent words correctly identified between the unmodified signals and low-flattened and median-flattened signals. Sentences were acoustically modified to have flattened intonation at either the median F0 of the original signal or at the average low F0 of the original signal. They found that the unmodified signal was more intelligible than the low-flattened signal and the low-flattened signal was more intelligible than the median-flattened. However, the differences between the low-flattened and median-flattened signals, though statistically significant, were small. This suggests that the specific value of F0 contributes relatively little to the intelligibility of speech. Lack of variation, however, appears to offer a greater contribution to intelligibility differences (Watson & Schlauch, 2008). This may suggest that low F0 associated with vocal fry may not greatly impact intelligibility. However, according to Colton, Casper, & Leonard (2006), lack of flexibility during the production of vocal fry typically results in a monotone voice. This lack of variation in F0 may negatively impact intelligibility.

Ramig (1992) suggested that decreased intelligibility of disordered phonation may be less related to the specific characteristics of the voice and more related to its perception. Disordered phonation is often negatively perceived and considered less effective and acceptable or distracting. This perception of disordered voice then interferes with the listener's ability to receive and comprehend the speech signal. Vocal fry may also be perceived in this way which could lead to decreased intelligibility.

Currently, no direct connection between speech produced with vocal fry and intelligibility has been identified. The purpose of the present study is to answer the question: does speech produced with vocal fry impact the intelligibility of spoken words? We hypothesized that words spoken in vocal fry would be less intelligible due the acoustic characteristics and perceptions of vocal fry. We did this by presenting listeners with words spoken in both fry and no fry. Words were spoken by both a male and female speaker and presented with varying levels of signal-to-noise ratio. Intelligibility was measured in terms of the percentage of words correctly identified and the listener's perception of listening difficulty. These results allowed us to quantify the effect of vocal fry on speech

METHODS

PART I: Recording and Preparation of Samples

In preparation for the intelligibility listening task, one male and one female subject were recorded while speaking a list of predetermined words both with and without vocal fry. Both subjects were native English speakers with no reported history of any communication disorders or differences. The recorded words were randomly selected from the Single Word Intelligibility subtest of the Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1984). The Assessment of Intelligibility of Dysarthric Speech (AIDS) is typically used to assess the speech intelligibility of persons with dysarthria. The Single Word Intelligibility subtest contains 50 sets of 12 phonetically similar words. A total of 400 words were randomly selected from the AIDS for this task. The 400 words were then divided into two separate lists, each comprised of 200 different words. One list was recorded by the male subject, and the other list was recorded by the female subject. See Appendix A for the lists of words recorded by each subject. The participants recorded their assigned list twice through, the first time using vocal fry and the second time using modal voice without vocal fry. This produced a total of 800 recorded words. The subjects were instructed to read each word at the end of the carrier phrase, "The next word is." The purpose of this was to simulate the natural production of vocal fry which typically occurs at the end of phrases when subglottic pressure naturally decreases.

The recording procedure was conducted in a sound isolation booth with background noise less than 25 dBA and reverberation time equal to 0.05 seconds. The recordings were acquired with two microphones, a head-mounted Shure B53 microphone and a NTI M2211 Sound Level Meter (NTI Instruments) used as a microphone at a distance of 30 cm. Both microphones were connected to a PC via a Tascam UH 7000 soundboard. Audacity 2.1.2 was utilized as the audio recording and editing software. Once recorded, each word was individually segmented and saved as its own audio file. To ensure that audio quality would not interfere with intelligibility testing, each recorded word was confirmed to have a signal-to-noise ratio (the ration of the speech signal to any background noise present in the recording) higher than 15 dB in all of the octave bands between 125 Hz and 8000 Hz. In addition, the energetic content of each sample was normalized using the voiced portion of each word. This is similar to what has been done in other intelligibility research (Tjaden, Sussman, & Wilding, 2014). See Appendix B for a more detailed description of the normalization process.

Each recorded word was also assessed to ensure that words of poor quality (i.e. poorly recorded samples, samples with interfering background noise, samples with poor vocal fry or intonation) were not included in the intelligibility listening task to minimize the impact of outside factors on intelligibility scores. Three samples were determined to have poor intelligibility due to the segmentation of the sample from the original audio recording (i.e. the sample was cut too close to the beginning of the word). These samples were recut to improve intelligibility and included in the intelligibility listening task.

PART II: Intelligibility Listening Task

The prepared samples were then used to create the intelligibility listening task. Two different versions (Version I and Version II) were created. Each version of the task included a total of 24 test sets. A test set consisted of 16 words spoken by the same speaker (Male or Female), with the same type of voicing (Fry or No Fry), and with the same level of signal-to-noise ratio (No Added Noise-NaN, -12 dB, -6 dB, 0 dB, 6 dB, or 12 dB). There are 24 possible combinations of these three conditions (speaker biological sex, voicing type, and signal-to-noise ratio). Each combination of conditions corresponded to an individual test set. Table 1 contains a list of the conditions for each of the 24 test sets.

Table 1: Combination of conditions for each of the 24 test sets.

Fry		No Fry	
Male	Female	Male	Female
NaN	NaN	NaN	NaN
-12	-12	-12	-12
-6	-6	-6	-6
0	0	0	0
6	6	6	6
12	12	12	12

Two different versions of each test set were created with each containing a different list of 16 words (List A and List B). For each combination of speaker sex and signal-to-noise ratio, List A and List B consisted of the same list of words in both the Fry and No Fry conditions. The two different versions of the intelligibility listening task were created such that if the Fry version of List A was used in one version, then the No Fry version of List A was used in the other version, and vice versa. In addition, if the Fry version of List A was used in one version, then the No Fry version of List B would be used in that same version, and vice versa. For example, for the test set consisting of the conditions of Male speaker, Fry, and signal-to-noise ratio of -6 dB, List A was used in Version I and List B was used in Version II. For the test set consisting of the same conditions (Male speaker, -6 dB), but No Fry instead of Fry, List B was used in Version I and List A was used in Version II. See Appendix C for the list of words presented in each test set. This was done so that every word was presented in both Fry and No Fry in order to counterbalance for the possible effect of the phonological aspects of each word on its intelligibility based on the assumption that any possible impact on intelligibility should affect both Fry and No Fry relatively similarly.

For each version (Version I and Version II), the 24 tests sets were presented in two different pseudorandomization orders. The first order was created utilizing a random sequence generator, and the second order was created by reversing the first order. A pseudorandomization order was only included if no more than two of the 6 dB and/or 12 dB signal-to-noise ratio conditions were presented consecutively. The purpose of this was to minimize the possible effect of listener fatigue on intelligibility scores and listening difficulty ratings by spacing out the tests sets containing the two greatest levels of background noise. Including two pseudorandomization orders in which the second order was a reverse of the first was intended to minimize the effect of listener fatigue as the intelligibility listening task progressed. A decrease in intelligibility of the test sets presented at the end of the intelligibility task due to listener fatigue should be balanced by having those test sets presented at the beginning of the task in the other pseudorandomization order. This created four total intelligibility listening tasks which were labeled Version Ia, Version Ib, Version IIa, and Version IIb. See Appendix D for the order of test sets presented in each version of the intelligibility listening task.

Prior to recruiting participants, IRB approval was obtained from the Michigan State University Human Research Protection Program (Protocol # MCDCR00000104, 17-1218-Intelligibility; approval letter located in Appendix E). Participants were recruited via the Michigan State University College of Communication Arts and Sciences SONA system to participate in the intelligibility listening task. Participants provided informed consent prior to participation (consent form located in Appendix F). Participants were included in the task if they were at least 18 years of age and spoke English as their native language as verified by each participant's SONA prescreen data and verbal confirmation from each participant. The participant's data were only included in the analysis if they had normal hearing, as verified by a hearing screening (using a Beltone Model 120 audiometer) in which hearing thresholds were verified to be within the normal range (≤ 20 dB) in each ear at 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz. Prior to completing the intelligibility listening task, participants were asked to complete a training test set consisting of 16 words in the Male, No Fry condition at a signal-to-noise ratio of -6 dB. The participants then completed the intelligibility listening task.

The intelligibility listening task was presented to each participant via Sennheiser HD-205 headphones. Each test set was presented as an individual audio file on a computer. A separate folder was created for each version of the test containing the audio files for each of the 24 subtests arranged in the order which they appeared in the packet of response forms (audio files and response forms were presented in the randomization orders previously discussed). Participants were instructed to play each audio file in the order presented. Within each file, each of the 16 words were presented at the end of the carrier phrase, "The next word is." The carrier phrase was identical for all presented words. The words were spaced with approximately 8 seconds in between each word to allow participants time to mark their answers. For each test set, participants were asked to indicate the word they believe they heard from a field of 12 phonetically similar words presented on the corresponding test form. Each grouping of 12 words was taken from the Single Word Intelligibility subtest of the AIDS and was presented to the participants in 3 by 4 tables as they appear in the AIDS assessment book. Participants were also asked to rate their perceived listening difficulty on a scale ranging from 1 (extremely easy) to 10 (extremely difficult) for each word. See Appendix G for an example of a test set response form. The total participation time for each participant was between 50 to 60 minutes.

PART III: Analysis Techniques

An intelligibility score and listening difficulty rating were calculated for each test set per participant. The intelligibility score was calculated as a percentage of words correctly identified in each test set. The listening difficulty rating was calculated as the mean of the ratings for each of the 16 words in each test set. The mean and standard deviation of the intelligibility score and listening difficulty rating were calculated and compared for both the Fry and No Fry conditions overall, for the Male and Female speakers, and at each level of signal-to-noise ratio. Comparisons were analyzed using a paired sample T-test. Differences were considered significant if the P value was $< .05$.

RESULTS

Participants

Forty-four participants were recruited via the Michigan State University College of Communication Arts and Sciences SONA system. Of the 44 recruited participants, 29 completed the intelligibility task. Three participants did not meet the inclusion criteria for normal hearing, each having at least one hearing threshold greater than 20 dB as identified on the hearing screening. Therefore, data from 26 participants were included in the analysis. All participants were students of Michigan State University, at least 18 years of age, and native speakers of English. Five participants were male, and 21 were female. Seven participants completed Version Ia of the intelligibility listening task, seven completed Version Ib, six completed Version IIa, and six completed Version IIb. The version of the intelligibility task that each participant completed was assigned based on their order of enrollment such that Participant 1 received Version Ia, Participant 2 received Version IIa, Participant 3 received Version Ib, Participant 4 received Version IIb, Participant 5 received Version Ia, and so forth.

Intelligibility Score

A general comparison of intelligibility scores between the Fry and No Fry (modal) conditions for all of the tests are shown in Table 2. Overall, the mean intelligibility score for the Fry condition was significantly lower than that of the No Fry condition (62.08% to 64.56% respectively, $P < .001$). The mean intelligibility scores of the Fry and No Fry conditions were also compared separately for each speaker sex. For both the Male spoken samples and Female spoken samples, the mean intelligibility score for the Fry condition was lower than that of the No Fry condition. This difference was not significant for the Male spoken samples (63.66% to 64.63% respectively, $P = .255$), but it was significant for the Female spoken samples (60.50% to

64.26% respectively, $P < .001$). Figure 1 shows the comparison between the Fry and No Fry conditions for both the Male and Female speakers.

Table 2: Comparison of mean intelligibility scores between the Fry and No Fry conditions.

	Fry		**No Fry**		
	Mean	SD	Mean	SD	P value
Male speaker	63.66	31.15	64.63	30.91	.255
Female speaker	60.50	31.73	64.26	30.50	< .001*
Overall	62.08	31.48	64.56	30.76	< .001*

**Significant difference*

Table 3 shows the mean intelligibility scores for all of the tests under the Fry and No Fry conditions at each level of signal-to-noise ratio (NaN, -12 dB, -6 dB, 0 dB, 6 dB, and 12 dB). At each level, the mean intelligibility score for the Fry condition was lower than that of the No Fry condition. However, this difference was only significant at the levels of -12 dB (84.13% to 87.02%, $P = .011$) and 12 dB (8.17% to 12.62%, $P = .002$). Figure 2 shows the comparison of mean intelligibility scores between the Fry and No Fry conditions at each level of signal-to-noise ratio.

Table 3: Comparison of mean intelligibility scores between the Fry and No Fry conditions at each level of signal-to-noise ratio.

	Fry		**No Fry**		
	Mean	SD	Mean	SD	P value
NaN	93.87	7.09	95.91	8.70	.052
-12 dB	84.13	7.60	87.02	7.13	.011*
-6 dB	80.53	12.89	82.69	10.37	.366
0 dB	66.95	12.46	67.55	12.44	.783
6 dB	38.82	12.09	41.59	13.97	.231
12 dB	8.17	7.99	12.62	9.53	.002*

NaN = No added noise
**Significant difference*

Differences in mean intelligibility scores were also assessed between the Male and Female speakers. Overall (including both Fry and No Fry conditions), the mean intelligibility score was significantly greater for the Male speaker than for the Female speaker (64.26% to 62.38% respectively, $P = .018$). Under just the Fry condition, the mean intelligibility score was also significantly greater for the Male speaker compared to the Female speaker (63.66% to 60.50% respectively, $P = .007$). However, under the No Fry condition, the mean intelligibility score was only marginally greater for the Male speaker compared to the Female speaker, and this difference was not significant (64.63% to 64.26% respectively, $P = .576$). These differences can also be visualized in Figure 1. See Appendix H for the mean intelligibility score data from all participants.

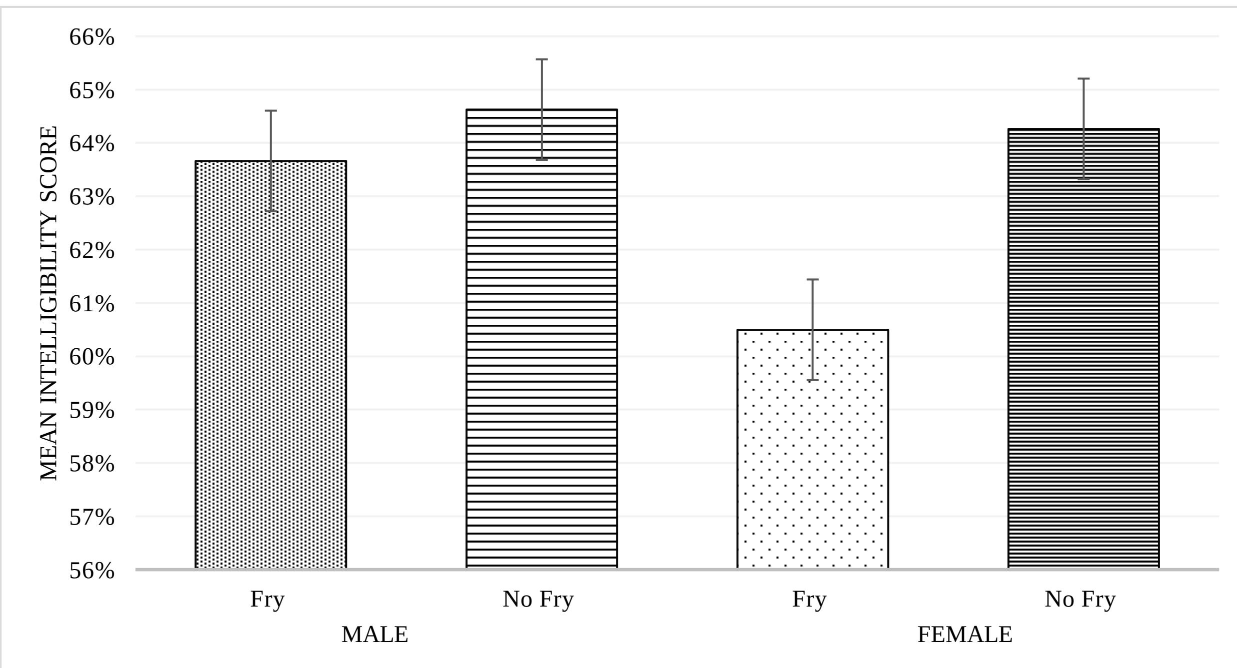

Figure 1: Comparison of mean intelligibility scores between the Fry and No Fry conditions by speaker sex. The first two bars represent the mean intelligibility scores for the Male speaker with the dotted bar representing the Fry condition and the striped bar representing the No Fry condition. The difference between Fry and No Fry for the Male speaker was not statistically significant (63.66% to 64.56%, $P = .255$). The second two bars represent the mean intelligibility scores for the Female speaker with the dotted bar representing Fry and the striped bar representing No Fry. This difference between Fry and No Fry for the Female speaker was statistically significant (60.50% to 64.26%, $P < .001$). Between the male and female speaker, differences were significant for the Fry condition (63.66% to 60.50%, $P = .007$) but not for the No Fry condition (64.63% to 64.26%, $P = .576$). Error bars represent standard error.

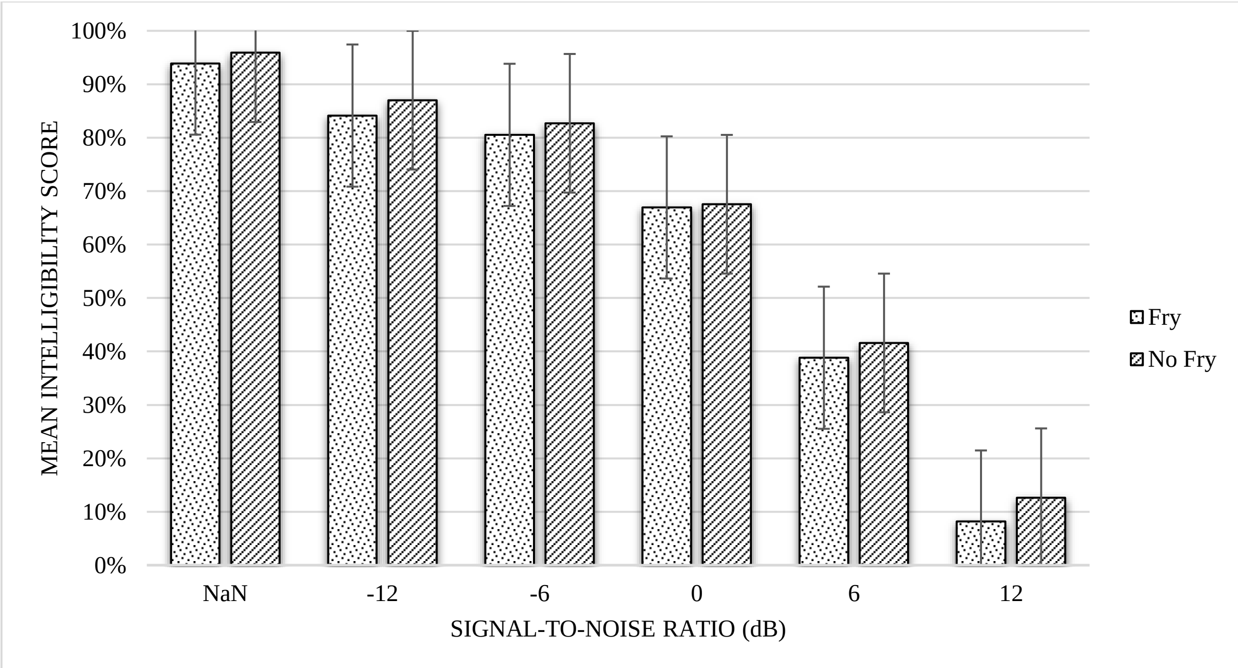

Figure 2: Comparison of mean intelligibility scores between the Fry and No Fry conditions by signal-to-noise ratio. At each level of signal-to-noise ratio the mean intelligibility score was lower for Fry (represented by the dotted bars) compared to No Fry (represented by the striped bars). These differences were only significant at the -12 dB and 12 dB levels (84.13% to 87.02%, $P = .011$; 8.17% to 12.62%, $P = .002$, respectively). Error bars represent standard error.

Listening Difficulty

A comparison of mean listening difficulty ratings between the Fry and No Fry conditions is shown in Table 4. Overall, the mean listening difficulty rating was significantly greater for the Fry condition compared to the No Fry condition (4.98 to 4.56 respectively, $P < .001$). The mean listening difficulty of the Fry and No Fry conditions were also compared separately for each speaker sex. For the Male spoken samples, the mean listening difficulty was significantly greater for the Fry condition compared to the No Fry condition (4.97 to 4.74 respectively, $P = .003$). For the Female spoken samples, the mean listening difficulty was also significantly greater for the Fry condition compared to the No Fry condition (5.00 to 4.42 respectively, $P < .001$). Figure 3 shows the comparison between the Fry and No Fry conditions for both the Male and Female speakers.

Table 4: Comparison of mean listening difficulty ratings between the Fry and No Fry conditions.

	Fry		No Fry		
	Mean	SD	Mean	SD	P value
Male speaker	4.97	3.30	4.74	3.23	.003*
Female speaker	5.00	3.20	4.42	3.24	< .001*
Overall	4.98	3.25	4.56	3.24	< .001*

**Significant difference*

Table 5 shows the mean listening difficulty ratings for all of the tests in the Fry and No Fry conditions for each level of signal-to-noise ratio. At each level, the mean listening difficulty rating for the Fry condition was greater than that of the No Fry condition. This difference was significant at the NaN (1.65 to 1.34, $P = .012$), -12 dB (2.36 to 1.92, $P = .003$), 0 dB (5.09 to 4.43, $P = .011$), and 6 dB levels (7.89 to 7.19, and $P < .001$), but not at the -6 dB (3.05 to 2.74, $P = .110$) or 12 dB levels (9.86 to 9.75, $P = .180$). Figure 4 shows the comparison of mean listening difficulty ratings between the Fry and No Fry conditions at each level of signal-to-noise ratio.

Table 5: Comparison of mean listening difficulty ratings between the Fry and No Fry conditions at each level of signal-to-noise ratio.

	Fry		No Fry		
	Mean	SD	Mean	SD	P value
NaN	1.65	.84	1.34	.53	.012*
-12 dB	2.36	1.15	1.92	.84	.003*
-6 dB	3.05	1.31	2.74	1.33	.110
0 dB	5.09	1.59	4.43	1.71	.011*
6 dB	7.89	1.73	7.19	1.58	< .001*
12 dB	9.86	.34	9.75	.60	.180

NaN = No added noise
**Significant difference*

The mean listening difficulty ratings were also compared between the Male and Female speakers. Overall (including both Fry and No Fry conditions), there was not a significant difference in mean listening difficulty ratings between the Male and Female speakers (4.83 to

4.71, *P* = .096). In just the Fry condition, the difference between the Male and Female speakers was also not significant (4.97 to 5.00, *P*= .756). However, in the No Fry condition, the mean listening difficulty for the Male speaker was significantly greater than the Female speaker (4.74 to 4.42, *P* = .008). These differences can also be visualized in Figure 3. See Appendix I for the mean listening difficulty data from all participants.

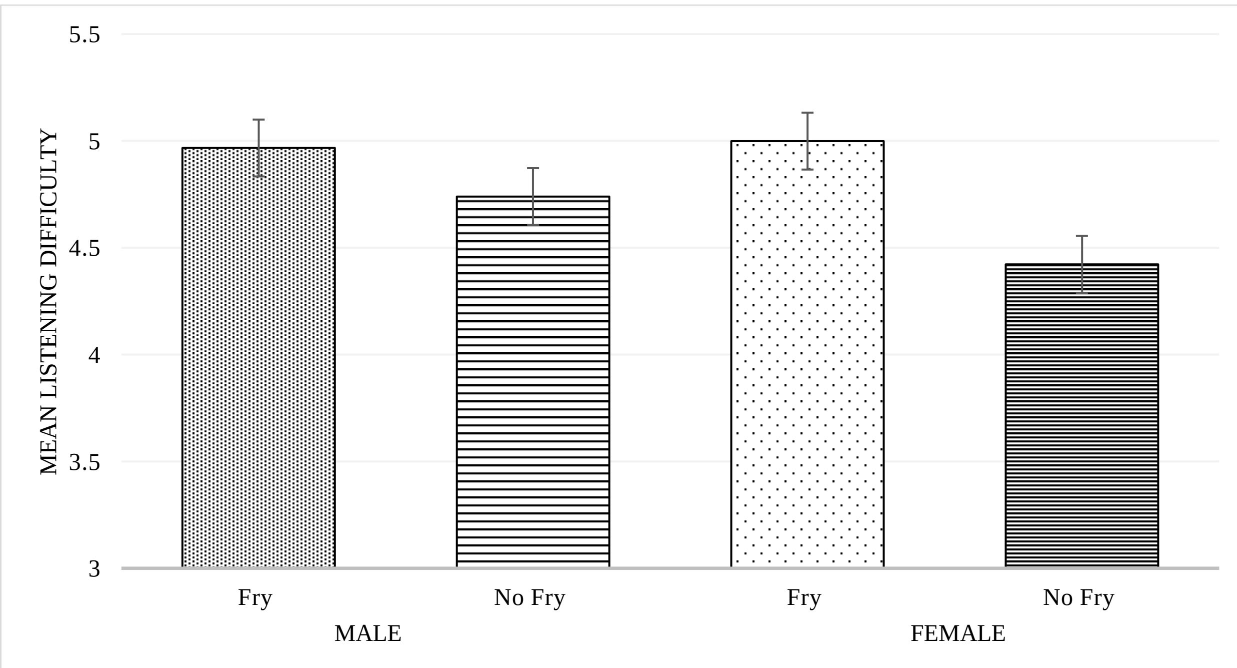

Figure 3: Comparison of mean listening difficulty between the Fry and No Fry conditions by speaker sex. The first two bars represent the mean listening difficulty for the Male speaker, with the dotted bar representing Fry and the striped bar representing No Fry. This difference was statistically significant (4.97 to 4.74, *P* = .003). The second two bars represent the mean listening difficulty for the female speaker, with the dotted bar representing Fry and the striped bar representing No Fry. This difference was also statistically significant (5.00 to 4.42, *P* < .001). Between the male and female speaker, differences were not significant for the Fry condition (4.97 to 5.00, *P*= .756) but were significant for the No Fry condition (4.74 to 4.42, *P* = .008). Error bars represent standard error.

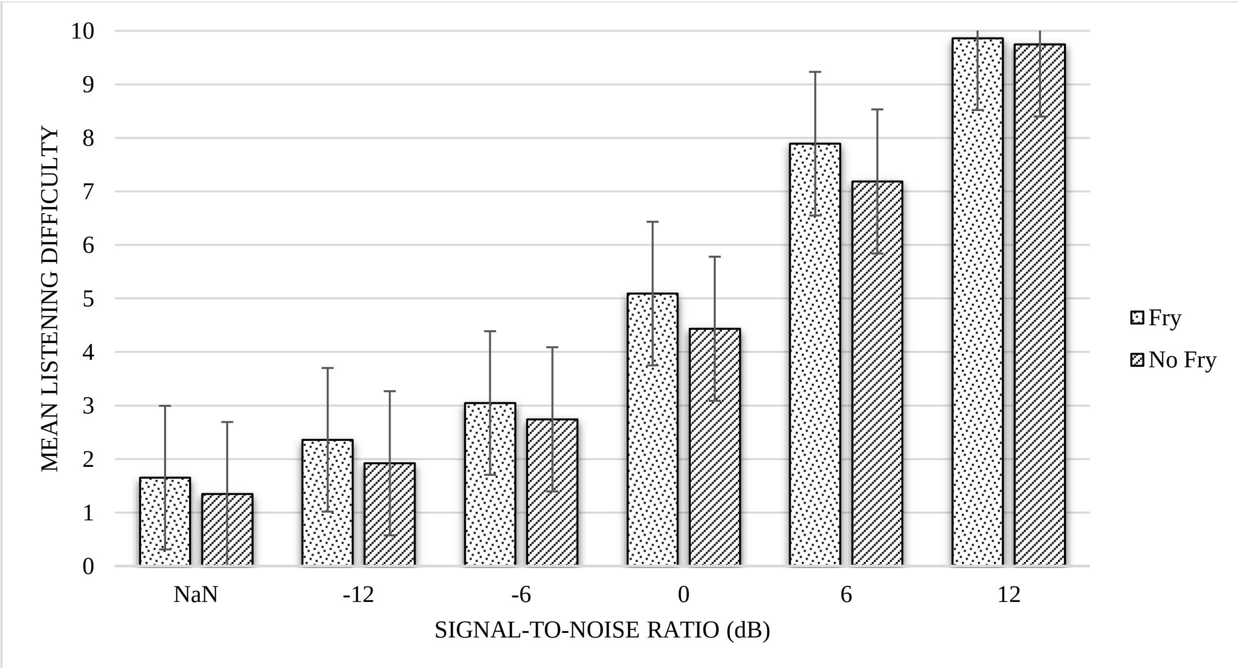

Figure 4: Comparison of mean listening difficulty between the Fry and No Fry conditions by signal-to-noise ratio. At each level of signal-to-noise ratio the mean listening difficulty was greater for Fry (represented by the dotted bars) compared to No Fry (represented by the striped bars). These differences were significant at the NaN (1.65 to 1.34, P = .012), -12 dB (2.36 to 1.92, P = .003), 0 dB (5.09 to 4.43, P = .011), and 6 dB levels (7.89 to 7.19, and P < .001). Error bars represent standard error.

Intelligibility Score and Listening Difficulty: Worst Noise Conditions Removed

Because of the sharp decrease in intelligibility for the test sets presented with the two worst noise conditions (signal-to-noise ratio levels 6 dB and 12 dB), the results of these tests sets were removed and the mean intelligibility scores and listening difficulty ratings were recalculated using the remaining tests. This dramatically reduced standard deviations indicating less variability between each participant's data. Table 6 shows a comparison of the adjusted overall mean intelligibility scores. The adjusted overall mean intelligibility score was significantly lower for the Fry condition compared to the No Fry condition (81.37% to 83.29%, P = .030). For the Male samples, the Fry condition was lower than the No Fry condition; however, this difference was not significant (83.23% to 84.56%, P = .285). For the Female samples, the Fry condition was lower than the No Fry condition, and this difference was significant (79.51% to 82.03%, P = .049). This is the same pattern of significance that was observed in the original analysis. Figure 5 shows the comparison between the adjusted mean intelligibility scores between the Fry and No Fry conditions for both the Male and Female speakers.

Table 6: Comparison of mean intelligibility scores between Fry and No Fry excluding the 6 dB and 12 dB noise conditions.

	Fry		**No Fry**		
	Mean	SD	Mean	SD	P value
Male speaker	83.23	12.05	84.56	12.44	.285
Female speaker	79.51	15.78	82.03	14.93	.049*
Overall	81.37	14.16	83.29	13.80	.030*

**Significant difference*

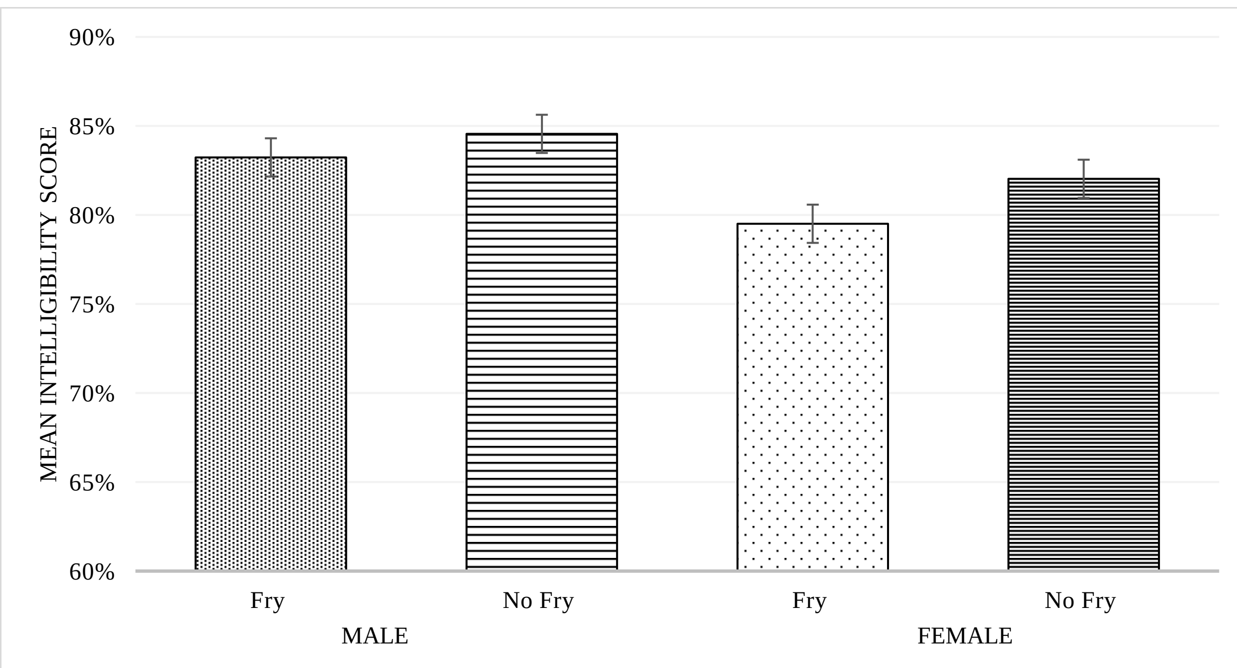

Figure 5: Comparison of adjusted mean intelligibility scores between the Fry and No Fry conditions by speaker sex. The first two bars represent the mean intelligibility scores for the male speaker with the dotted bar representing Fry and the striped bar representing No Fry. This difference was not statistically significant (83.23% to 84.56%, P = .285). The second two bars represent the mean intelligibility scores for the female speaker with the dotted bar representing Fry and the striped bar representing No Fry. This difference was statistically significant (79.51% to 82.03%, P = .049). Error bars represent standard error.

Table 7 shows a comparison of the adjusted overall mean listening difficulty ratings. Overall, the adjusted mean listening difficulty rating was significantly greater for the Fry condition compared to the No Fry condition (3.03 to 2.62, $P < .001$). For the Male samples, the Fry condition was greater than the No Fry condition; however, this difference was not significant (2.90 to 2.74, $P = .176$). For the Female samples, the Fry condition was greater than the No Fry condition, and this difference was significant (3.18 to 2.48, $P < .001$). This pattern of significance differs from the original analysis in that the difference between Fry and No Fry for the Male spoken samples is no longer significant when the worst two noise conditions are removed. Figure 6 shows the comparison between the adjusted mean listening difficulty ratings between the Fry and No Fry conditions for both the Male and Female speakers.

Table 7: Comparison of mean listening difficulty ratings between Fry and No Fry excluding the 6 dB and 12 dB noise conditions.

	Fry		**No Fry**		
	Mean	SD	Mean	SD	P value
Male speaker	2.90	1.66	2.74	1.70	.176
Female speaker	3.18	1.90	2.48	1.62	< .001*
Overall	3.03	1.79	2.61	1.67	< .001*

**Significant difference*

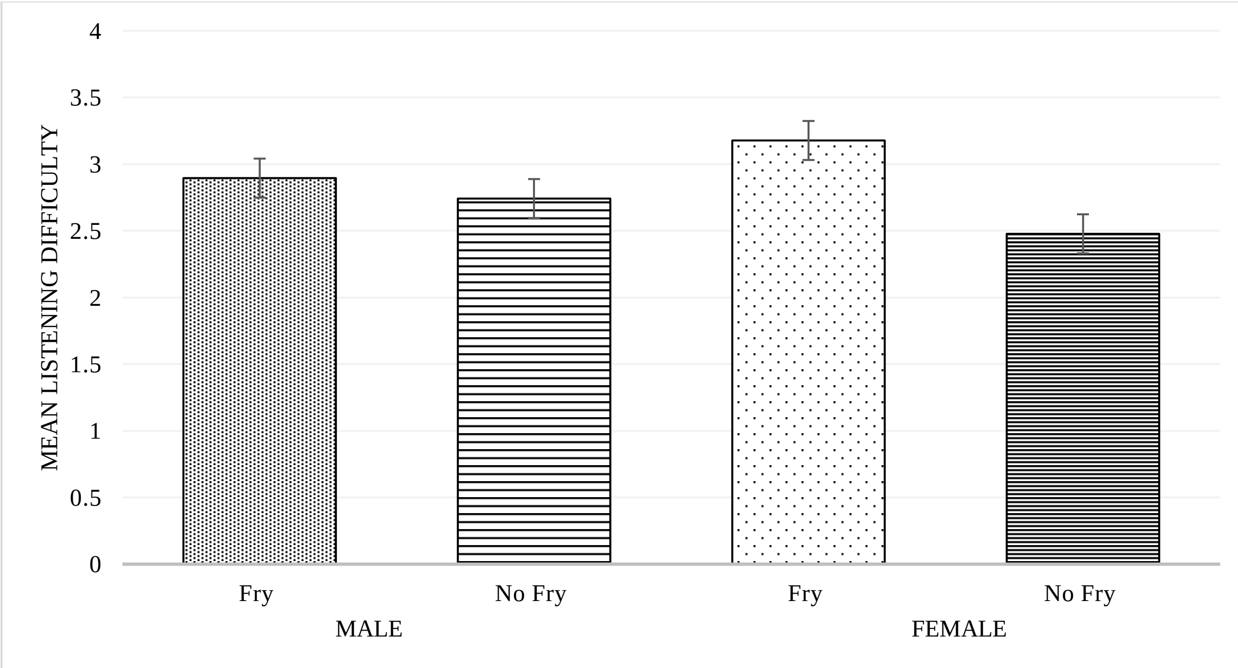

Figure 6: Comparison of adjusted mean listening difficulty between the Fry and No Fry conditions by speaker sex. The first two bars represent the mean listening difficulty for the Male speaker with the dotted bar representing Fry and the striped bar representing No Fry. This difference was not statistically significant (2.90 to 2.74, *P* = .176). The second two bars represent the mean listening difficulty for the female speaker with the dotted bar representing Fry and the striped bar representing No Fry. This difference was statistically significant (3.18 to 2.48, *P* <.001). Error bars represent standard error.

Comparison of Male and Female Listeners

The Fry subtests were also compared between the Male and Female listeners. The comparison of mean intelligibility scores between the Male and Female listeners is shown in Table 8. Overall, there was no significant difference in mean intelligibility score of the Fry subtests between the Male and Female listeners (60.94% to 62.35% respectively; *P* = .756). Mean intelligibility scores were also compared between the Male and Female listeners for the Male speaker and Female speaker conditions separately. For the Male speaker, the mean intelligibility score was greater for the Male listeners compared to the Female listeners; however, this difference was not significant (65.83% to 63.14%, *P* = .673). For the Female speaker, the mean intelligibility score was lower for the Male listeners compared to the Female listeners, but this difference was also not significant (56.04% to 61.56%, *P* = .395). These differences can be visualized in Figure 7.

Table 8: Comparison of mean intelligibility scores of the Fry subtests between the Male and Female listeners.

	Male listeners		**Female listeners**		
	Mean	SD	Mean	SD	P value
Male speaker	65.83	28.58	63.14	31.71	.673
Female speaker	56.04	31.56	61.56	31.68	.395
Overall	60.94	30.50	62.35	31.71	.756

Table 9 shows the comparison of mean listening difficulty ratings of the Fry subtests between the Male and Female listeners. Overall, there was no significant difference in mean listening difficulty rating of the Fry subtests between the Male and Female listeners (4.781 to 5.031 respectively; *P* = .593). Mean listening difficulty ratings were also compared between the Male and Female listeners for the Male speaker and Female speaker conditions separately. For the Male speaker, the mean listening difficulty rating was lower for the Male listeners compared to the Female listeners; however, this difference was not significant (4.881 to 4.988, *P* = .875). For the Female speaker, the mean listening difficulty rating was also lower for the Male listeners compared to the Female listeners, and this difference was also not significant (4.681 to 5.075, *P* = .547). These differences can be visualized in Figure 8.

Table 9: Comparison of mean listening difficulty ratings of the Fry subtests between the Male and Female listeners.

	Male listeners		**Female listeners**		
	Mean	SD	Mean	SD	P value
Male speaker	4.881	3.325	4.988	3.295	.875
Female speaker	4.681	3.192	5.075	3.197	.547
Overall	4.781	3.261	5.031	3.247	.593

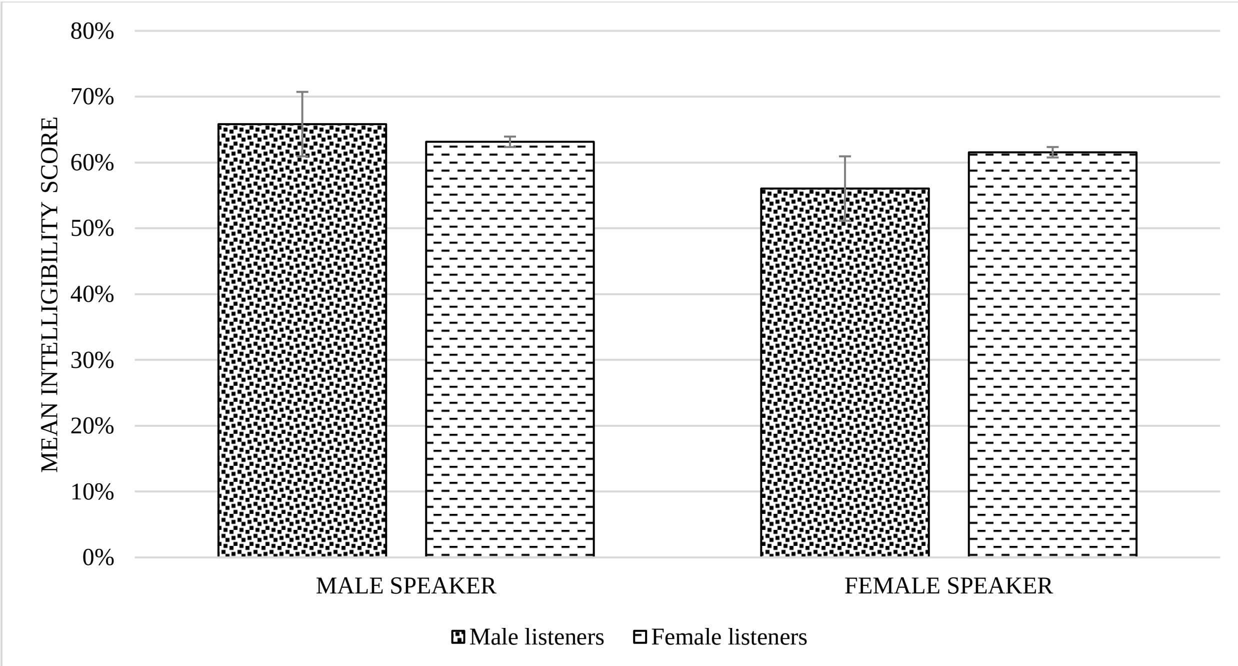

Figure 7: Comparison of mean intelligibility scores between the Male and Female listeners by speaker sex. The first two bars represent the mean intelligibility scores for the Male speaker with the dotted bar representing the Male listeners and the dashed bar representing the Female listeners. This difference was not statistically significant (65.83% to 63.14%, $P = .673$). The second two bars represent the mean intelligibility scores for the Female speaker with the dotted bar representing the Male listeners and the dashed bar representing the Female listeners. This difference was also not statistically significant (56.04% to 61.56%, $P = .395$). Error bars represent standard error.

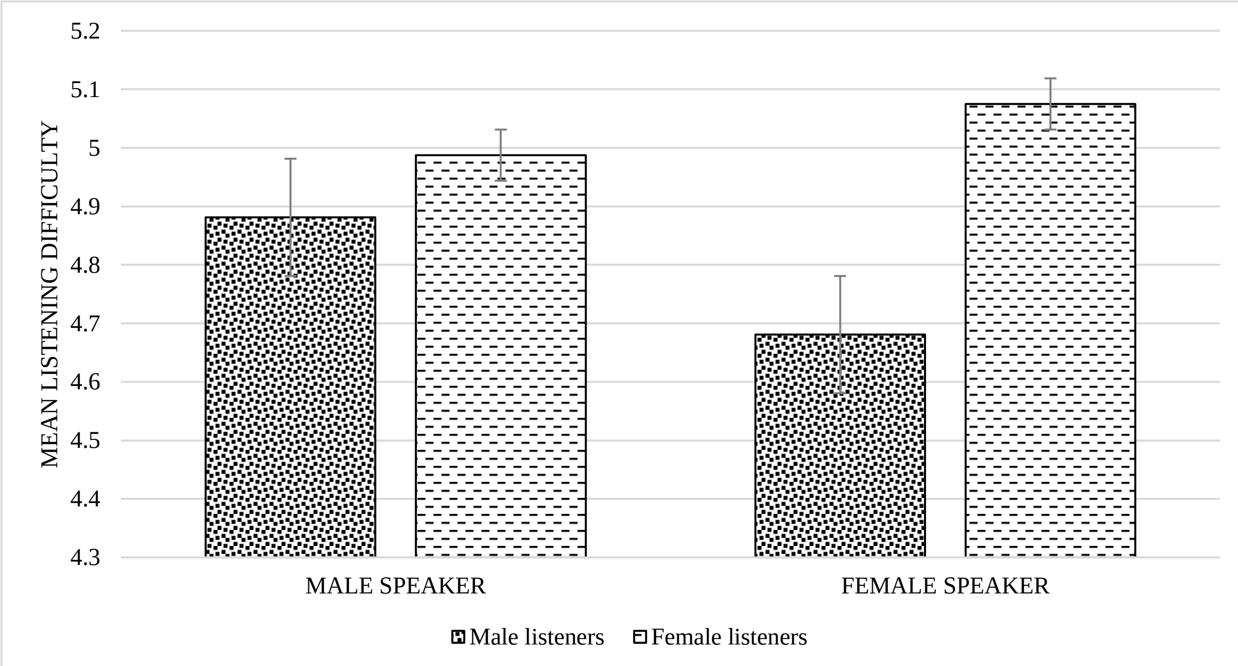

Figure 8: Comparison of mean listening difficulty ratings between the Male and Female listeners by speaker sex. The first two bars represent the mean listening difficulty ratings for the Male speaker with the dotted bar representing the Male listeners and the dashed bar representing the Female listeners. This difference was not statistically significant (4.881 to 4.988, $P = .875$). The second two bars represent the mean listening difficulty ratings for the Female speaker with the dotted bar representing the Male listeners and the dashed bar representing the Female listeners. This difference was also not statistically significant (4.681 to 5.075, $P = .593$). Error bars represent standard error.

DISCUSSION

Overall, the mean intelligibility score of all of the tests was significantly lower for the Fry condition compared to the No Fry condition. This difference remained significant when the two worst noise conditions (signal-to-noise ratio levels 6 dB and 12 dB) were removed and the mean intelligibility scores were recalculated. This suggests that single words spoken in vocal fry are less intelligible than words spoken without vocal fry (modal voicing). In addition, the overall mean listening difficulty rating was significantly greater for the Fry condition compared to the No Fry condition. This difference also remained significant when the 6 dB and 12 dB noise conditions were removed. This suggests, that not only are words spoken in vocal fry less intelligible, the perceived difficulty in understanding these words is also greater. These results confirm our hypothesis.

Differences in intelligibility between vocal fry and modal voicing may have possible social and vocational implications. A study by Anderson, Klofstad, Mayew, and Venkatachalam (2014) compared the perceptions of vocal fry and modal voicing when used in the context of the labor market. Their results suggested that, especially when used by young females, vocal fry was perceived more negatively than modal voicing. In particular, job candidates using vocal fry were perceived as less competent, educated, trustworthy, and hirable when compared to candidates using a normal speaking voice. This leads us to question if perhaps differences in intelligibility or listening difficulty may play a role in the negative social and vocational perceptions of vocal fry.

Another possible implication of the intelligibility differences associated with vocal fry include the potential role vocal fry may play in language processing. Morton and Watson (2001) compared language processing between modal voicing and dysphonic voicing in the classroom setting. Their results showed that recall and inferencing of spoken passages were better when presented in a modal voice compared to a dysphonic voice. Participants also indicated that they perceived the dysphonic voice as unpleasant which resulted in increased listening demands. Although vocal fry is not a dysphonia, it may share some commonalities with dysphonia in how it is perceived (as negative and unpleasant), thus leading to similar implications. In addition, the decreased intelligibility of vocal fry as indicated in the current study may also contribute to differences in language processing. Further research may be warranted to examine the effect of vocal fry on language processing.

When broken down by the speaker's biological sex, a greater difference in mean intelligibility scores between the Fry and No Fry conditions existed for Female speaker compared to the Male speaker. In addition, this difference in mean intelligibility score was only significant for the Female condition and not for the Male condition. This suggests that vocal fry has a greater impact on speech intelligibility when used by a female speaker compared to a male speaker. One possible explanation for this phenomenon is that a greater difference in fundamental frequency exists between vocal fry and modal voicing for female speakers compared to male speakers. Fundamental frequency typically ranges from 150 to 500 Hz in adult women and 80 to 450 Hz in adult men (Boone, McFarlane, Von Berg, & Zraick, 2014). Vocal fry typically occurs at frequencies below 70 Hz (Titze, 2000) with one group of researchers suggesting 24-52 Hz as the typical range for male speakers and 18-46 Hz for female speakers in vocal fry (Hollien & Michel, 1968; Michel, 1968). Thus the difference between vocal fry and modal voicing in terms of fundamental frequency is typically greater for female speakers compared to male speakers. This, perhaps, decreases the naturalness and familiarity of the voice, with a greater impact on the female voice. This greater difference in naturalness of vocal fry for female speakers may contribute to the greater difference in intelligibility as well as the greater likelihood of vocal fry being perceived as negative when used by females as demonstrated in previous studies (Anderson et al., 2014).

When intelligibility scores of vocal fry were compared between the Male and Female listeners for the Male speaker, the mean intelligibility score was greater for the Male listeners compared to the Female listeners. For the Female speaker, the mean intelligibility score was lower for the Male listeners compared to the Female listeners. Although these differences were not statistically significant, this trend may suggest that the intelligibility of vocal fry may be more greatly affected if the speaker and listener are of opposite sex. These results may be limited by sample size as this analysis only included data from 5 Male listeners. Future research should include a greater number of participants to further understand any differences in the interaction between the sex of the speaker and the sex of the listener in terms of speech intelligibility of vocal fry.

Differences in mean intelligibility scores and listening difficulty ratings were not significant at all levels of signal-to-noise ratio. This may suggest that intelligibility may be affected differently by various levels of background noise. This could have potential implications for certain applications of vocal fry, such as individuals using vocal fry in noisy environments, (e.g. teachers speaking in noisy classrooms or employees speaking on the floor of a factory). Competing noise varying in intensity or frequency might affect intelligibility of vocal fry and modal voicing differently due to the different acoustic characteristics of each. To better understand vocal fry and its effect on communication, future research should explore the differences between intelligibility of male and female speakers and between different levels of noise and different types of noise (e.g. spectral differences, modulated, talker babble).

As is the case of all research, there are several potential limiting factors. For example, during the normalization process of the audio samples, the voiced portion of each word was normalized. Since vocal fry voicing is generally at a lower dB level than modal voicing, normalizing the voicing part of the word resulted in an overemphasis of the consonants of the Fry samples. This could have impacted intelligibility by aiding listeners in the identification of words spoken in vocal fry. Despite this, vocal fry was still less intelligible than modal voicing. Perhaps this difference would have been greater had the consonants of the Fry words not been emphasized more than in the No Fry words.

Sample size presents another possible limitation. The analysis was intended to be completed with data from at least 40 participants. Due to a high rate of participant no show, the analysis could only be completed with data from 26 participants. Participants were also all of typical college age, an age group more often associated with the use of vocal fry. Intelligibility of vocal fry may vary for different generations of listeners. Related, since only two talkers provided the recorded samples, dialectal differences unique to these two talkers could also present a possible limitation. For example, one of the words spoken by the male speaker was the word “bag.” However, the word “beg” was one of the options on the response form for this sample. Due to dialectal differences in both production and perception, “beg” could be reasonably chosen as a response. Because each word was included in both the Fry and No Fry conditions, differences in intelligibility due to dialectal differences should have affected both the Fry and No Fry conditions relatively similarly. However, the possibility remains that dialectal differences may have impacted intelligibility scores. Therefore, future research should include more samples from a wider range of talkers and with more listeners.

Finally, the current study examined intelligibility at the single-word level only. Intelligibility in expanded contexts such as at the conversational level may not be as greatly affected by vocal fry due to the added context of the surrounding phrase/sentence. Further research should examine the effect of vocal fry on speech intelligibility in these contexts to better understand the impact of vocal fry on speech intelligibility.

CONCLUSION

The results of this study indicate that speech intelligibility at the single word level is reduced for words spoken in vocal fry compared to words spoken in modal voicing. Reduced intelligibility was greater for the female speaker compared to the male speaker. In addition, perceived listening difficulty was also greater for vocal fry compared to modal voicing, with this difference again greater for the female speaker compared to the male speaker. Reduced intelligibility may contribute to the negative perception of vocal fry. Further research should explore if vocal fry is associated with reduced intelligibility and increased listening difficulty at the phrase or conversational level when context is introduced or if vocal fry affects language processing.

APPENDICES

APPENDIX A:

Words Recorded by the Male and Female Speakers

Table 10: Words recorded by the Male speaker

Arch	Brother	Design	Groan	Mother	Rake	Sheet	Table
Ark	Bumper	Divide	Grove	Naval	Ramp	Ship	Tanner
Author	Burn	Doe	Heart	Near	Rate	Shock	Tarred
Bait	But	Dusk	Hide	Neat	Reach	Shoe	Tart
Ball	By	Dust	Impress	Occur	Red	Short	Thread
Ban	Cane	Eye	Intact	Offer	Refer	Side	Told
Barn	Cape	Faint	Judge	Often	Renter	Sit	Tore
Bat	Car	Fall	Ladle	Option	Rest	Sitter	Tort
Beam	Carp	Fell	Late	Owl	Root	Sitting	Trace
Beat	Cart	File	Law	Pace	Rupture	Slid	Trade
Beep	Case	Flicker	Leaf	Page	Said	Slim	Train
Beer	Cave	Floor	Leak	Pale	Sane	Sold	Troop
Beg	City	Flow	Leap	Par	Sap	Sort	Twine
Below	Claw	Fool	Least	Paste	Scorch	Soul	Warm
Bender	Compress	Fork	Lecture	Pay	Scrap	Sprain	Waste
Best	Contact	Fort	Limb	Pierce	Scratch	Stable	We’re
Bill	Coupe	Full	Lip	Planner	Screech	Stall	Weep
Bin	Crawl	Gag	List	Porch	Scrub	Start	Whip
Bitten	Cross	Game	Loss	Praise	Seer	Store	Wide
Bitter	Dark	Gang	Map	Preach	Sender	Storm	Wife
Blown	Day	Gauge	Mate	Quart	Serve	Strange	Wine
Bone	Dead	Gesture	Meat	Quicker	Shade	Stream	Wipe
Bought	Debate	Globe	Mere	Quit	Shark	Suit	Wire
Bowl	Deer	Glow	Mist	Rage	Shear	Super	Witty
Bread	Define	Grain	Mixture	Rain	Shed	Swarm	Wore

Table 11: Words recorded by the Female speaker

Absurd	Bull	Denture	Grave	Live	Pave	Shell	Teacher
After	Cable	Deny	Great	Loop	Peach	Shore	Term
Ape	Cage	Detract	Green	Low	Pear	Shot	Texture
Art	Cake	Dread	Grind	Make	Phone	Shred	Threat
Awful	Call	Dress	Gross	Mall	Pier	Shut	Timber
Bank	Care	Each	Group	Manner	Pole	Sicker	Tired
Banner	Carve	Ember	Grow	March	Port	Sigh	Torn
Bark	Center	Fable	Harp	Mat	Pour	Sketch	Trait
Bash	Cheer	Fear	Hear	May	Preserve	Slip	Twice
Beast	Closure	Fierce	High	Member	Pretty	So	Twist
Bed	Cloth	Final	Honest	Mold	Race	Space	Vague
Beef	Coal	Fitting	Horn	Mole	Raid	Spade	Wait
Big	Cold	For	Jade	More	Range	Spark	Wake
Bit	Conserve	Form	Just	Musk	Rattle	Sport	Wall
Blow	Contract	Fruit	Kitten	Nap	Ray	Steer	Wart
Boast	Contrast	Funnel	Knit	Nest	Reef	Stiff	Weave
Boat	Cork	Gain	Lease	Obscure	Ride	Street	Weird
Book	Corn	Glass	Leave	Office	Rust	Stretch	While
Boot	Cough	Gleam	Leer	Other	Saddle	Suppress	White
Booth	Court	Glitter	Left	Paid	Sat	Swift	Why
Born	Creature	Grab	Less	Pain	Sauce	Take	Wicked
Boss	Dart	Grace	Lie	Paint	Screen	Tall	Win
Bother	Decide	Grade	Lisp	Panel	Seam	Tap	Wise
Brace	Defeat	Grange	Lit	Park	Sharp	Tape	Wives
Brain	Defend	Grape	Litter	Part	Shelf	Taste	Written

APPENDIX B:

Description of the Normalization Process

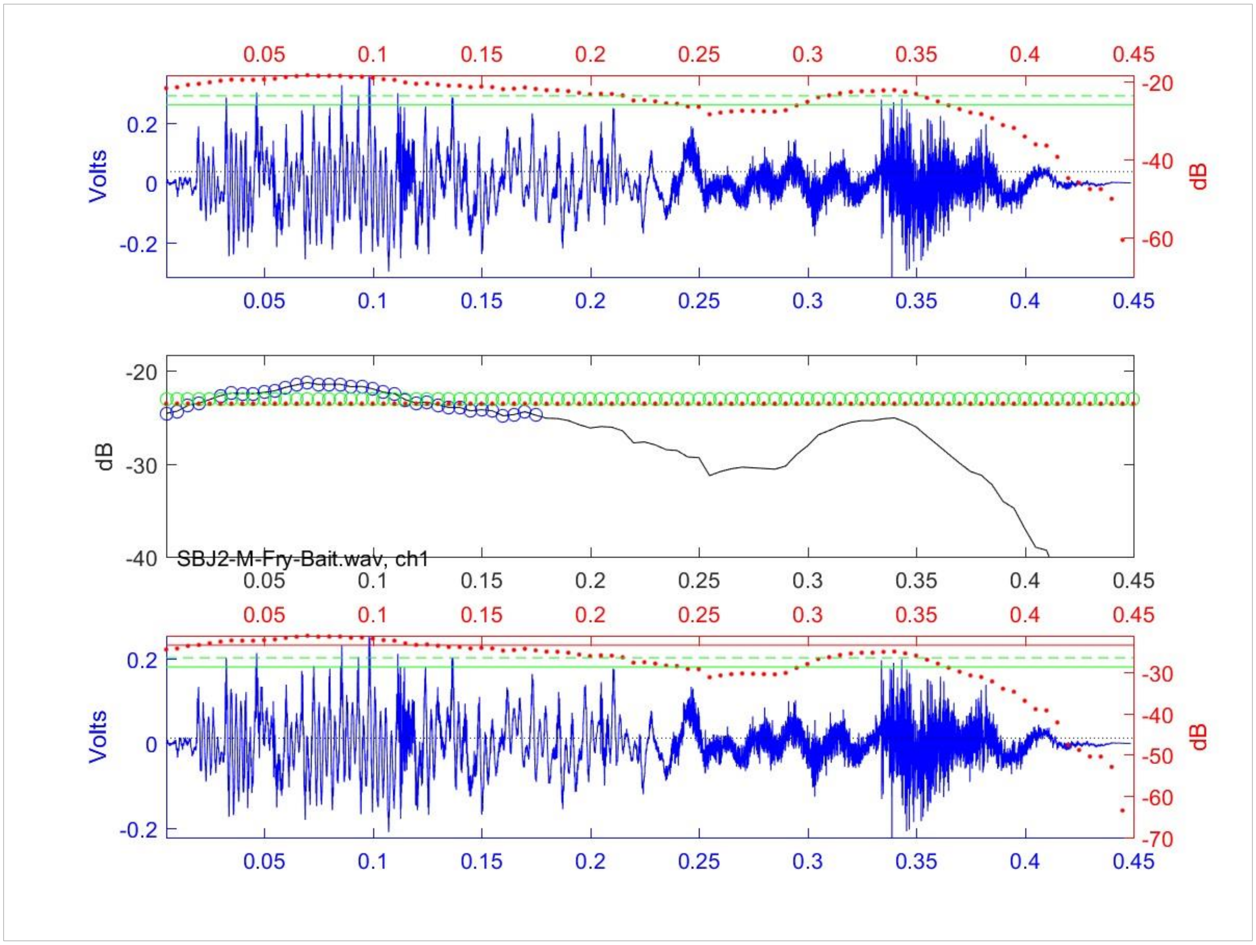

Figure 9: Normalization process example. The blue dots in the center graph represent the estimated vowel area of the sample word "bait."

A first attempt at normalizing all of the word files involved equalizing the average amount of acoustic energy of the entire word. Because the words differed in vowel length, number of vowels (e.g. bait, divide), and number of consonants (e.g. quit, contact), equalizing the average acoustic energy of each word resulted in words that were perceptually unequal in energy. Instead, the average acoustic energy of only the voicing portion of each word (vowels and voiced consonants) was normalized as opposed to the entire word. This is similar to was has been done in other research in intelligibility (Tjaden, Sussman, & Wilding, 2014). To do this, a custom MATLAB script was used to obtain the dB contour, fundamental frequency contour, and pitch strength contour for each word. Next, voicing segments of the word were identified using the time information from the fundamental frequency contour. From this, the average dB of the voicing segment was estimated. An example of this is shown by the voicing area of the word "bait" in Figure 9. The voicing area is indicated by the blue circles along the dB contour shown in the second graph. For each word, the average acoustic energy for this voicing area was then normalized to -22 dB. This resulted in a potential advantage for words in the Fry condition. Vocal fry is less intense by nature, therefore increasing the energy of the entire word during the normalization process also increased the intensity of consonants in words in the Fry condition more than the words in the No Fry condition. This could have increased in the intelligibility of words spoken in vocal fry by overemphasizing the consonants which could have assisted listeners in correctly identifying the word. However, the results indicated that the Fry condition was still less intelligible than the No Fry condition despite this effect.

APPENDIX C:

Lists of Words Used in Each Test Set

Table 12: List of words used in each combination of conditions.

	Male Fry		Male No Fry		Female Fry		Female No Fry	
	List A	**List B**	**List A**	**List B**	**List A**	**List B**	**List A**	**List B**
NaN	Arch	Barn	Arch	Barn	Absurd	Banner	Absurd	Banner
	Beg	Bitten	Beg	Bitten	Big	Boot	Big	Boot
	Bread	Cane	Bread	Cane	Brain	Care	Brain	Care
	Cave	Crawl	Cave	Crawl	Coal	Corn	Coal	Corn
	Define	Eye	Define	Eye	Defeat	Dress	Defeat	Dress
	Flow	Game	Flow	Game	Final	Gain	Final	Gain
	Grain	Intact	Grain	Intact	Grade	Grind	Grade	Grind
	Leak	List	Leak	List	High	Lease	High	Lease
	Mixture	Often	Mixture	Often	Lisp	Make	Lisp	Make
	Par	Preach	Par	Preach	Member	Nest	Member	Nest
	Rake	Renter	Rake	Renter	Paint	Pear	Paint	Pear
	Sap	Seer	Sap	Seer	Preserve	Ray	Preserve	Ray
	Shed	Side	Shed	Side	Sauce	Shore	Sauce	Shore
	Sold	Start	Sold	Start	Sketch	Sport	Sketch	Sport
	Super	Thread	Super	Thread	Swift	Teacher	Swift	Teacher
	Train	Weep	Train	Weep	Torn	Wake	Torn	Wake
-12 dB	Ark	Bat	Ark	Bat	After	Bark	After	Bark
	Below	Bitter	Below	Bitter	Bit	Booth	Bit	Booth
	Brother	Cape	Brother	Cape	Bull	Carve	Bull	Carve
	City	Cross	City	Cross	Cold	Cough	Cold	Cough
	Design	Fall	Design	Fall	Defend	Each	Defend	Each
	Fool	Gang	Fool	Gang	Fitting	Glass	Fitting	Glass
	Groan	Judge	Groan	Judge	Grange	Gross	Grange	Gross
	Leap	Loss	Leap	Loss	Honest	Leave	Honest	Leave
	Mother	Option	Mother	Option	Lit	Mall	Lit	Mall
	Paste	Quart	Paste	Quart	Mold	Obscure	Mold	Obscure
	Ramp	Rest	Ramp	Rest	Panel	Phone	Panel	Phone
	Scorch	Sender	Scorch	Sender	Pretty	Reef	Pretty	Reef
	Sheet	Sit	Sheet	Sit	Screen	Shot	Screen	Shot
	Sort	Store	Sort	Store	Slip	Steer	Slip	Steer
	Swarm	Told	Swarm	Told	Take	Term	Take	Term
	Troop	Whip	Troop	Whip	Trait	Wall	Trait	Wall
-6 dB	Author	Beam	Author	Beam	Ape	Bash	Ape	Bash
	Bender	Blown	Bender	Blown	Blow	Born	Blow	Born
	Bumper	Car	Bumper	Car	Cable	Center	Cable	Center
	Claw	Day	Claw	Day	Conserve	Court	Conserve	Court
	Divide	Fell	Divide	Fell	Denture	Ember	Denture	Ember
	Fork	Gauge	Fork	Gauge	For	Gleam	For	Gleam
	Grove	Ladle	Grove	Ladle	Grape	Group	Grape	Group
	Least	Map	Least	Map	Horn	Leer	Horn	Leer
	Naval	Owl	Naval	Owl	Litter	Manner	Litter	Manner
	Pierce	Quicker	Pierce	Quicker	Mole	Office	Mole	Office
	Rate	Root	Rate	Root	Park	Pier	Park	Pier
	Scrap	Serve	Scrap	Serve	Race	Ride	Race	Ride
	Ship	Sitter	Ship	Sitter	Seam	Shred	Seam	Shred
	Soul	Storm	Soul	Storm	So	Stiff	So	Stiff
	Table	Tore	Table	Tore	Tall	Texture	Tall	Texture
	Twine	Wide	Twine	Wide	Twice	Wart	Twice	Wart

Table 12 (cont'd)

	Male Fry		Male No Fry		Female Fry		Female No Fry	
	List A	**List B**	**List A**	**List B**	**List A**	**List B**	**List A**	**List B**
0 dB	Bait	Beat	Bait	Beat	Art	Beast	Art	Beast
	Best	Bone	Best	Bone	Boast	Boss	Boast	Boss
	Burn	Carp	Burn	Carp	Cage	Cheer	Cage	Cheer
	Compress	Dead	Compress	Dead	Contract	Creature	Contract	Creature
	Doe	File	Doe	File	Deny	Fable	Deny	Fable
	Fort	Gesture	Fort	Gesture	Form	Glitter	Form	Glitter
	Heart	Late	Heart	Late	Grave	Grow	Grave	Grow
	Lecture	Meat	Lecture	Meat	Jade	Left	Jade	Left
	Neat	Pace	Neat	Pace	Live	March	Live	March
	Planner	Quit	Planner	Quit	More	Other	More	Other
	Reach	Rupture	Reach	Rupture	Part	Pole	Part	Pole
	Scratch	Shade	Scratch	Shade	Raid	Rust	Raid	Rust
	Shock	Sitting	Shock	Sitting	Sharp	Shut	Sharp	Shut
	Sprain	Strange	Sprain	Strange	Space	Street	Space	Street
	Tanner	Tort	Tanner	Tort	Tap	Threat	Tap	Threat
	Warm	Wife	Warm	Wife	Twist	Weird	Twist	Weird
6 dB	Ball	Beep	Ball	Beep	Awful	Bed	Awful	Bed
	Bill	Bought	Bill	Bought	Boat	Bother	Boat	Bother
	But	Cart	But	Cart	Cake	Closure	Cake	Closure
	Contact	Debate	Contact	Debate	Contrast	Dart	Contrast	Dart
	Dusk	Flicker	Dusk	Flicker	Detract	Fear	Detract	Fear
	Full	Globe	Full	Globe	Fruit	Grab	Fruit	Grab
	Hide	Law	Hide	Law	Great	Harp	Great	Harp
	Limb	Mere	Limb	Mere	Just	Less	Just	Less
	Occur	Page	Occur	Page	Loop	Mat	Loop	Mat
	Porch	Rage	Porch	Rage	Musk	Paid	Musk	Paid
	Red	Said	Red	Said	Pave	Port	Pave	Port
	Screech	Shark	Screech	Shark	Range	Saddle	Range	Saddle
	Shoe	Slid	Shoe	Slid	Shelf	Sicker	Shelf	Sicker
	Stable	Stream	Stable	Stream	Spade	Stretch	Spade	Stretch
	Tarred	Trace	Tarred	Trace	Tape	Timber	Tape	Timber
	Waste	Wine	Waste	Wine	Vague	While	Vague	While
12 dB	Ban	Beer	Ban	Beer	Bank	Beef	Bank	Beef
	Bin	Bowl	Bin	Bowl	Book	Brace	Book	Brace
	By	Case	By	Case	Call	Cloth	Call	Cloth
	Coupe	Deer	Coupe	Deer	Cork	Decide	Cork	Decide
	Dust	Floor	Dust	Floor	Dread	Fierce	Dread	Fierce
	Gag	Glow	Gag	Glow	Funnel	Grace	Funnel	Grace
	Impress	Leaf	Impress	Leaf	Green	Hear	Green	Hear
	Lip	Mist	Lip	Mist	Knit	Lie	Knit	Lie
	Offer	Pale	Offer	Pale	Low	May	Low	May
	Praise	Rain	Praise	Rain	Nap	Pain	Nap	Pain
	Refer	Sane	Refer	Sane	Peach	Pour	Peach	Pour
	Scrub	Shear	Scrub	Shear	Rattle	Sat	Rattle	Sat
	Short	Slim	Short	Slim	Shell	Sigh	Shell	Sigh
	Stall	Suit	Stall	Suit	Spark	Suppress	Spark	Suppress
	Tart	Trade	Tart	Trade	Taste	Tired	Taste	Tired
	We're	Wipe	We're	Wipe	Wait	White	Wait	White

APPENDIX D:

Order of Test Sets for Each Version of the Intelligibility Listening Task

Table 13: Order of test sets as presented in Version Ia and Version Ib of the intelligibility listening task. Each test set is identified by the sex of the speaker (Male or Female), voicing type (Fry or No Fry), level of signal-to-noise ratio (NaN, -12 dB, -6 dB, 0 dB, 6 dB, or 12 dB), and word list (A or B) that was presented in that test set.

	Version Ia				**Version Ib**			
	Speaker Sex	Voicing	SNR	Word List	Speaker Sex	Voicing	SNR	Word List
1.	Female	Fry	0 dB	A	Female	No Fry	6 dB	B
2.	Female	No Fry	-12 dB	B	Female	No Fry	12 dB	B
3.	Female	Fry	-12 dB	A	Female	Fry	NaN	A
4.	Female	Fry	6 dB	A	Male	Fry	-6 dB	A
5.	Female	No Fry	NaN	B	Male	Fry	0 dB	A
6.	Female	Fry	12 dB	A	Female	No Fry	0 dB	B
7.	Male	No Fry	0 dB	B	Male	No Fry	6 dB	B
8.	Female	No Fry	6 dB	B	Male	No Fry	NaN	B
9.	Male	No Fry	-6 dB	B	Male	Fry	12 dB	A
10.	Male	Fry	6 dB	A	Male	Fry	-12 dB	A
11.	Male	No Fry	-12 dB	B	Male	No Fry	12 dB	B
12.	Female	Fry	-6 dB	A	Male	Fry	NaN	A
13.	Male	Fry	NaN	A	Female	Fry	-6 dB	A
14.	Male	No Fry	12 dB	B	Male	No Fry	-12 dB	B
15.	Male	Fry	-12 dB	A	Male	Fry	6 dB	A
16.	Male	Fry	12 dB	A	Male	No Fry	-6 dB	B
17.	Male	No Fry	NaN	B	Female	No Fry	6 dB	B
18.	Male	No Fry	6 dB	B	Male	No Fry	0 dB	B
19.	Female	No Fry	0 dB	B	Female	Fry	12 dB	A
20.	Male	Fry	0 dB	A	Female	No Fry	NaN	B
21.	Male	Fry	-6 dB	A	Female	Fry	6 dB	A
22.	Female	Fry	NaN	A	Female	Fry	-12 dB	A
23.	Female	No Fry	12 dB	B	Female	No Fry	-12 dB	B
24.	Female	No Fry	6 dB	B	Female	Fry	0 dB	A

SNR = Signal-to-noise ratio
NaN = No added noise

Table 14: Order of test sets as presented in Version IIa and Version IIb of the intelligibility listening task. Each test set is identified by the sex of the speaker (Male or Female), voicing type (Fry or No Fry), level of signal-to-noise ratio (NaN, -12 dB, -6 dB, 0 dB, 6 dB, or 12 dB), and word list (A or B) that was presented in that test set.

	Version IIa				**Version IIb**			
	Speaker Sex	Voicing	SNR	Word List	Speaker Sex	Voicing	SNR	Word List
1.	Male	No Fry	-6 dB	A	Male	Fry	-6 dB	B
2.	Male	No Fry	-12 dB	A	Male	Fry	0 dB	B
3.	Female	Fry	NaN	B	Female	No Fry	-12 dB	A
4.	Female	No Fry	NaN	A	Male	No Fry	NaN	A
5.	Female	Fry	-12 dB	B	Female	Fry	0 dB	B
6.	Female	No Fry	0 dB	A	Male	Fry	6 dB	B
7.	Female	No Fry	12 dB	A	Male	Fry	NaN	B
8.	Male	Fry	-12 dB	B	Male	No Fry	12 dB	A
9.	Female	No Fry	6 dB	A	Female	No Fry	-6 dB	A
10.	Female	Fry	-6 dB	B	Male	Fry	12 dB	B
11.	Female	Fry	6 dB	B	Male	No Fry	6 dB	A
12.	Female	Fry	12 dB	B	Male	No Fry	0 dB	A
13.	Male	No Fry	0 dB	A	Female	Fry	12 dB	B
14.	Male	No Fry	6 dB	A	Female	Fry	6 dB	B
15.	Male	Fry	12 dB	B	Female	Fry	-6 dB	B
16.	Female	No Fry	-6 dB	A	Female	No Fry	6 dB	A
17.	Male	No Fry	12 dB	A	Male	Fry	-12 dB	B
18.	Male	Fry	NaN	B	Female	No Fry	12 dB	A
19.	Male	Fry	6 dB	B	Female	No Fry	0 dB	A
20.	Female	Fry	0 dB	B	Female	Fry	-12 dB	B
21.	Male	No Fry	NaN	A	Female	No Fry	NaN	A
22.	Female	No Fry	-12 dB	A	Female	Fry	NaN	B
23.	Male	Fry	0 dB	B	Male	No Fry	-12 dB	A
24.	Male	Fry	-6 dB	B	Male	No Fry	-6 dB	A

SNR = Signal-to-noise ratio
NaN = No added noise

APPENDIX E:

Image of IRB Approval Letter

MICHIGAN STATE
UNIVERSITY

Initial IRB Application Approval

October 13, 2017

To: Eric Hunter
1026 Red Cedar Road
Room 113, Oyer Speech & Hearing Building
East Lansing, MI 48824

Re: **IRB#** 17-1218 Category: EXPEDITED 4, 7
Approval Date: September 29, 2017
Expiration Date: September 28, 2018

Title: Speech Intelligibility in Noise: The Effect of Vocal Fry on Speech Intelligibility

The Institutional Review Board has completed their review of your project. I am pleased to advise you that **your project has been approved**.

The committee has found that your research project is appropriate in design, protects the rights and welfare of human subjects, and meets the requirements of MSU's Federal Wide Assurance and the Federal Guidelines (45 CFR 46 and 21 CFR Part 50). The protection of human subjects in research is a partnership between the IRB and the investigators. We look forward to working with you as we both fulfill our responsibilities.

Renewals: IRB approval is valid until the expiration date listed above. If you are continuing your project, you must submit an ***Application for Renewal*** application at least one month before expiration. If the project is completed, please submit an ***Application for Permanent Closure***.

Revisions: The IRB must review any changes in the project, prior to initiation of the change. Please submit an ***Application for Revision*** to have your changes reviewed. If changes are made at the time of renewal, please include an ***Application for Revision*** with the renewal application.

Problems: If issues should arise during the conduct of the research, such as unanticipated problems, adverse events, or any problem that may increase the risk to the human subjects, notify the IRB office promptly. Forms are available to report these issues.

Please use the IRB number listed above on any forms submitted which relate to this project, or on any correspondence with the IRB office.

If we can be of further assistance, please contact us at 517-355-2180 or via email at IRB@msu.edu. Thank you for your cooperation.

c: Kaleigh Cammenga, Lady Catherine Cantor Cutiva, Russell Banks, Mark Berardi, Rachel Burtka

Michigan State University Founded 1855

Office of Regulatory Affairs
Human Research Protection Programs

Biomedical & Health Institutional Review Board (BIRB)

Community Research Institutional Review Board (CRIRB)

Social Science Behavioral/Education Institutional Review Board (SIRB)

4000 Collins Road
Suite 136
Lansing, MI, 48910
(517) 355-2180
Fax: (517) 432-4503
Email: irb@msu.edu
www.hrpp.msu.edu

MSU is an affirmative-action, equal-opportunity employer.

Figure 10: Image of IRB approval letter.

APPENDIX F:

Image of Consent Form

Research Participant Information and Consent Form

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

Study Title: Speech Intelligibility in Noise
Researchers and Title: Dr. Eric Hunter & Kaleigh Cammenga, M.A. Student at Michigan State University
Department and Institution: Department of Communicative Sciences and Disorders at Michigan State University
Address and Contact Information: 113 Oyer, East Lansing 48823, 517.353.8641
Sponsor: Michigan State University

1. PURPOSE OF RESEARCH
You are being asked to participate in this study to help researchers gain a better understanding of speech intelligibility and listening difficulty in different noise conditions.

2. ELIGIBILITY CRITERIA
In order to participate in this study, you must

- be at least 18 years of age
- speak English as your primary language
- have normal hearing

3. ALTERNATIVE OPTIONS
There are no alternative procedures; however, you have the option to discontinue your participation at any time.

4. WHAT YOU WILL DO
First, you will be asked to participate in a brief hearing screening to verify that you are eligible to participate in this study. Next, you will be asked to listen to multiple series of recorded words through headphones and identify the word you heard out of 12 choices. You will also be asked to rate your difficulty in identifying each word you hear on a scale of 1 to 10. Participation will be less than an hour.

5. POTENTIAL BENEFITS
While the program in which you are being asked to participate may not immediately benefit you, this research may benefit others by increasing our knowledge of speech intelligibility and listening difficulty.

6. POTENTIAL RISKS
There is minimal risk involved in this research program and the procedures should cause you no undue discomfort.

7. PRIVACY AND CONFIDENTIALITY
The data for this study are being collected confidentially. Neither the researchers nor anyone else will be able to link data to you. The data for this project will be kept confidential. Data from this study will be stored in a locked cabinet in a locked room or a password protected computer in the locked laboratory. All information will be kept for at least three years after the close of the study. Only trained researchers under the jurisdiction of this project and Human Research Protection Program will have access to the data collected in the study. Information about

This consent form was approved by a Michigan State University Institutional Review Board. Approved 9/29/17– valid through 9/28/18.
This version supersedes all previous versions. IRB # 17-1218.

Figure 11: Image of consent form.

Figure 11 (cont'd)

you will be kept confidential to the maximum extent allowable by law. Although we will make every effort to keep your data confidential there are certain times, such as a court order, where we may have to disclose your data. Identifying information will not be attached to any of your individual responses or recordings when reporting results from the surveys. You will not be asked to give your name or any other information during the recording that will allow you or your place of employment to be identified. All results will be kept in a secured location accessibly only to those involved in the study. The results of this study may be published or presented at professional meetings, but the identities of all research participants will remain anonymous. By participating, you agree to allow audio recordings of your speech.

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

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

9. COSTS AND COMPENSATION FOR BEING IN THE STUDY

If you are enrolled in a course that has an option of credit for research participation, and the course accepts SONA credit, you will have the option to receive MSU SONA credit. For those students enrolled in such courses, students can also find alternative assignments to earn extra credit if they choose not to participate in this research study but wish to earn extra credit. The SONA system awards 1 credit per 1 hour of research participation. Within the SONA system, neither researchers nor individual instructors will know what studies participants are involved in. If your participation is over an hour, which it will likely be, we will compensate you in half hour increments (rounding up) for up to one hour total. If you don't accept SONA credit, no other remuneration is available.

10. THE RIGHT TO GET HELP IF INJURED

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

11. CONTACT INFORMATION

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

- Kaleigh Cammenga, M.A. Student at Michigan State Univ., East Lansing, MI 48824, cammeng2@msu.edu
- Dr. Eric Hunter, Michigan State Univ, 113 Oyer, East Lansing, MI 48823, 517.353.8641, ejhunter@msu.edu

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

12. DOCUMENTATION OF INFORMED CONSENT

This consent form was approved by a Michigan State University Institutional Review Board. Approved 9/29/17– valid through 9/28/18. This version supersedes all previous versions. IRB # 17-1218.

Figure 11 (cont'd)

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

_______________________________________ ______________________________

Signature Date

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

This consent form was approved by a Michigan State University Institutional Review Board. Approved 9/29/17– valid through 9/28/18. This version supersedes all previous versions. IRB # 17-1218.

APPENDIX G:

Test Set Response Form Example

Word #1

Please circle the word you hear:		
Art	March	Park
Bark	Dark	Ark
Lark	Arch	Dart
Heart	Tart	Spark

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #2

Please circle the word you hear:		
Big	Ban	Bash
Bog	Bought	Beg
Boat	Bank	Boot
Bag	Bat	Beat

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #3

Please circle the word you hear:		
Dread	Bed	Ned
Bread	Said	Red
Tread	Rest	Nest
Dress	Best	Dead

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #4

Please circle the word you hear:		
Cave	Cape	Page
Cane	Pave	Case
Tape	Cage	Taste
Cake	Take	Pace

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #5

Please circle the word you hear:		
Divide	Debate	Decide
Design	Defend	Divine
Defeat	Define	Decal
Devise	Deny	Defy

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #6

Please circle the word you hear:		
Glow	Sew	Pole
Flow	Bowl	Low
Coal	Below	Doe
Blow	Blown	Soul

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #7

Please circle the word you hear:		
Grain	Pain	Gauge
Rain	Sane	Train
Brain	Range	Strange
Grange	Rage	Grade

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #8

Please circle the word you hear:		
Leap	Limp	Lip
Leak	Lean	Leave
Least	Lease	Lit
Leaf	Leash	Live

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #9

Please circle the word you hear:		
Texture	Rapture	Creature
Gesture	Preacher	Rupture
Denture	Mixture	Fixture
Lecture	Closure	Teacher

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #10

Please circle the word you hear:		
Carve	Tear	Pour
Tired	Pear	Tarred
Carp	Car	Cart
Care	Tore	Par

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #11

Please circle the word you hear:		
Rate	Trace	Waste
Brace	Wait	Race
Trait	Rake	Praise
Grace	Pace	Wake

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #12

Please circle the word you hear:		
Vamp	Sat	Ramp
Bat	Tap	Vat
Sap	Mat	Nap
Map	Rat	Damp

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #13

Please circle the word you hear:		
Stretch	Said	Shred
Sketch	Shed	Bread
Fed	Dread	Thread
Threat	Red	Dead

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #14

Please circle the word you hear:		
Full	Pole	Bowl
Mold	Cold	Pull
Mole	Sold	Told
Bull	Fold	Soul

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #15

Please circle the word you hear:		
Occur	Refer	Converge
Reserve	Obscure	Conserve
Super	Observe	Deserve
Absurd	Preserve	Serve

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

Word #16

Please circle the word you hear:		
Raid	Rake	Trait
Rain	Trade	Rage
Train	Range	Rave
Rate	Race	Trace

Please rate your difficulty in understanding the previous word:

Easy									Difficult
1	2	3	4	5	6	7	8	9	10

APPENDIX H:

Participant Intelligibility Score Data for All Test Conditions

Table 15: Intelligibility score as a percentage for each signal-to-noise ratio under the Male Fry condition for each subject.

	Male Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	81.25	93.75	93.75	62.5	31.25	6.25
Sub02	93.75	81.25	81.25	81.25	31.25	18.75
Sub03	87.5	93.75	75	81.25	31.25	6.25
Sub04	87.5	81.25	93.75	62.5	68.75	18.75
Sub05	93.75	93.75	87.5	81.25	31.25	25
Sub06	93.75	87.5	93.75	75	56.25	18.75
Sub07	100	87.5	87.5	62.5	18.75	6.25
Sub09	93.75	93.75	93.75	50	25	12.5
Sub10	100	68.75	81.25	81.25	43.75	6.25
Sub11	81.25	93.75	81.25	43.75	50	6.25
Sub12	93.75	93.75	87.5	81.25	37.5	18.75
Sub13	100	93.75	87.5	50	43.75	12.5
Sub14	100	81.25	68.75	75	56.25	18.75
Sub15	81.25	100	75	62.5	18.75	12.5
Sub16	81.25	81.25	68.75	75	50	6.25
Sub17	87.5	87.5	93.75	56.25	18.75	6.25
Sub19	100	93.75	81.25	81.25	37.5	0
Sub20	93.75	81.25	81.25	62.5	25	12.5
Sub22	87.5	81.25	93.75	87.5	43.75	0
Sub23	93.75	75	93.75	87.5	62.5	18.75
Sub24	87.5	93.75	87.5	56.25	25	0
Sub25	87.5	68.75	75	75	6.25	0
Sub27	81.25	87.5	100	81.25	37.5	31.25
Sub28	68.75	87.5	68.75	68.75	37.5	6.25
Sub29	87.5	93.75	100	75	37.5	25
Sub30	93.75	93.75	87.5	75	56.25	0
Mean	*89.90385*	*87.25962*	*85.33654*	*70.43269*	*37.74038*	*11.29808*

Table 16: Intelligibility score as a percentage for each signal-to-noise ratio under the Male No Fry condition for each subject

	Male No Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	100	87.5	75	68.75	25	12.5
Sub02	81.25	81.25	62.5	43.75	18.75	6.25
Sub03	93.75	100	93.75	75	18.75	12.5
Sub04	87.5	93.75	75	75	43.75	25
Sub05	100	100	87.5	93.75	43.75	12.5
Sub06	100	93.75	68.75	62.5	50	18.75
Sub07	93.75	93.75	87.5	75	18.75	0
Sub09	100	93.75	87.5	75	50	6.25
Sub10	100	93.75	68.75	81.25	37.5	18.75
Sub11	81.25	87.5	87.5	62.5	18.75	0
Sub12	93.75	87.5	100	68.75	43.75	12.5
Sub13	93.75	93.75	87.5	75	37.5	18.75
Sub14	100	81.25	62.5	62.5	50	18.75
Sub15	100	100	93.75	93.75	37.5	18.75
Sub16	87.5	81.25	68.75	75	37.5	37.5
Sub17	93.75	93.75	81.25	68.75	25	6.25
Sub19	100	93.75	87.5	81.25	43.75	0
Sub20	93.75	81.25	81.25	68.75	62.5	25
Sub22	100	93.75	81.25	68.75	37.5	0
Sub23	93.75	93.75	81.25	81.25	56.25	18.75
Sub24	87.5	87.5	87.5	68.75	50	18.75
Sub25	87.5	87.5	87.5	62.5	18.75	0
Sub27	100	87.5	62.5	62.5	31.25	25
Sub28	93.75	81.25	75	68.75	43.75	25
Sub29	100	93.75	87.5	56.25	18.75	18.75
Sub30	100	93.75	100	81.25	25	25
Mean	*94.71154*	*90.625*	*81.49038*	*71.39423*	*36.29808*	*14.66346*

Table 17: Intelligibility score as a percentage for each signal-to-noise ratio under the Female Fry condition for each subject

	Female Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	93.75	81.25	81.25	62.5	37.5	6.25
Sub02	100	75	25	50	37.5	0
Sub03	93.75	81.25	68.75	68.75	43.75	12.5
Sub04	93.75	81.25	75	56.25	31.25	18.75
Sub05	100	87.5	75	68.75	31.25	0
Sub06	93.75	81.25	87.5	68.75	50	12.5
Sub07	100	81.25	75	37.5	43.75	6.25
Sub09	100	75	50	50	25	12.5
Sub10	100	87.5	93.75	68.75	43.75	6.25
Sub11	100	75	68.75	43.75	37.5	0
Sub12	93.75	81.25	87.5	62.5	37.5	12.5
Sub13	100	93.75	62.5	62.5	37.5	6.25
Sub14	87.5	81.25	75	68.75	37.5	0
Sub15	100	81.25	68.75	93.75	37.5	0
Sub16	100	87.5	87.5	50	37.5	0
Sub17	100	81.25	75	68.75	56.25	0
Sub19	100	75	75	68.75	50	0
Sub20	100	75	87.5	62.5	37.5	0
Sub22	100	81.25	75	68.75	31.25	0
Sub23	100	81.25	75	62.5	50	6.25
Sub24	100	81.25	100	50	43.75	6.25
Sub25	93.75	68.75	87.5	68.75	37.5	0
Sub27	100	93.75	75	62.5	37.5	12.5
Sub28	100	81.25	81.25	87.5	31.25	0
Sub29	93.75	75	75	68.75	31.25	0
Sub30	100	81.25	81.25	68.75	62.5	12.5
Mean	*97.83654*	*81.00962*	*75.72115*	*63.46154*	*39.90385*	*5.048077*

Table 18: Intelligibility score as a percentage for each signal-to-noise ratio under the Female No Fry condition for each subject

	Female No Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	93.75	75	93.75	68.75	50	6.25
Sub02	100	93.75	75	62.5	50	0
Sub03	87.5	87.5	93.75	56.25	43.75	6.25
Sub04	93.75	87.5	87.5	62.5	50	12.5
Sub05	93.75	81.25	100	75	56.25	6.25
Sub06	100	87.5	87.5	56.25	62.5	12.5
Sub07	100	75	62.5	68.75	31.25	12.5
Sub09	93.75	81.25	100	68.75	56.25	6.25
Sub10	100	93.75	93.75	43.75	43.75	0
Sub11	87.5	75	87.5	50	50	6.25
Sub12	100	81.25	81.25	56.25	43.75	18.75
Sub13	93.75	81.25	87.5	81.25	37.5	18.75
Sub14	93.75	68.75	68.75	56.25	43.75	6.25
Sub15	100	81.25	81.25	81.25	56.25	18.75
Sub16	100	93.75	68.75	37.5	37.5	18.75
Sub17	100	81.25	81.25	81.25	50	18.75
Sub19	93.75	81.25	81.25	81.25	62.5	0
Sub20	100	81.25	87.5	56.25	62.5	0
Sub22	100	81.25	87.5	81.25	68.75	0
Sub23	100	93.75	81.25	50	50	25
Sub24	100	87.5	68.75	81.25	25	25
Sub25	93.75	75	93.75	56.25	43.75	0
Sub27	100	87.5	75	43.75	6.25	6.25
Sub28	100	87.5	75	75	43.75	12.5
Sub29	100	87.5	87.5	62.5	37.5	6.25
Sub30	100	81.25	93.75	62.5	56.25	31.25
Mean	*97.11538*	*83.41346*	*83.89423*	*63.70192*	*46.875*	*10.57692*

APPENDIX I:

Participant Listening Difficulty Data for All Test Conditions

Table 19: Mean listening difficulty rating for each signal-to-noise ratio under the Male Fry condition for each subject

	Male Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	3.3125	4.4375	3.8125	5.625	9.6875	10
Sub02	1	3.0625	6.4375	8	10	10
Sub03	1.5	1.4375	1.6875	2.4375	6.9375	10
Sub04	1.75	1.3125	3.1875	4.875	7.875	10
Sub05	1.125	1.4375	3.5625	7	9.875	10
Sub06	1.5	1.875	2.8125	4.8125	8.4375	10
Sub07	1.125	1.4375	1.25	3.0625	7.375	8.5625
Sub09	1.5	1.4375	1.5625	2.75	9	10
Sub10	1	1.3125	1.875	5.5625	9.9375	10
Sub11	1.125	1	1.625	2.9375	4.25	9.875
Sub12	1.875	2.3125	2.875	5.6875	8.125	9.25
Sub13	2	1.533333	2.75	5.0625	7.5625	8.5625
Sub14	1.5625	1	3.5625	6.375	8.1875	10
Sub15	1	1	3.6875	4.25	10	10
Sub16	2.0625	3.8125	4.5625	4.9375	6.25	8.8125
Sub17	1.125	1.4375	1.875	6.6875	9.6875	10
Sub19	3.125	3.1875	4.375	5.5	10	10
Sub20	1.6875	2.4	3.9375	5.6875	8.6875	10
Sub22	2.1875	2.3125	2.9375	4.25	8.875	10
Sub23	1.3125	3.125	2.5625	5.375	8.6875	10
Sub24	1.6875	1.25	2.5	4.3125	8.25	10
Sub25	3	3.3125	3.9375	5.6875	8.875	10
Sub27	1.0625	1.5	1.3125	5.1875	7.5	10
Sub28	3	1.75	4.375	5.875	9.4375	10
Sub29	1.8125	1.625	1.75	3.5625	8.1875	9.375
Sub30	1.375	1.5	1.5625	2.625	8.125	9.5
Mean	*1.723558*	*1.992628*	*2.9375*	*4.927885*	*8.454327*	*9.766827*

Table 20: Mean listening difficulty rating as a percentage for each signal-to-noise ratio under the Male No Fry condition for each subject

	Male No Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	1.25	2.4375	3.5	5.5625	9.0625	9.9375
Sub02	2	3.733333	3.375	7.9375	9.6875	9.9375
Sub03	1.25	1.375	2.25	3.9375	5.133333	9.928571
Sub04	1.125	1.125	1.125	3.4375	5.375	10
Sub05	1	1.25	2.6875	7.0625	8.6875	10
Sub06	1.1875	2.25	5.25	7.25	9	9.8125
Sub07	1.25	1	1.25	2.875	6.0625	9.75
Sub09	1.3125	1.5	1.8125	3.5625	8.6875	10
Sub10	1	1.0625	1.8125	3.375	8.6875	9.625
Sub11	1.0625	1.1875	1.625	1.9375	5.0625	10
Sub12	1.8125	2.9375	4.25	2.9375	6.6875	8.75
Sub13	1.375	1.230769	2.5625	3.8125	7.5625	9.6875
Sub14	1.25	2.375	2.5625	5.125	6.1875	9.6875
Sub15	1.5	1.875	3.9375	7.0625	7.5625	10
Sub16	3.125	3.4375	3.75	5.8125	7.3125	8.1875
Sub17	1	1.375	3.25	4.875	9.75	10
Sub19	1.625	2.625	8.0625	6.625	9.3125	10
Sub20	1.5	1.625	1.9375	3.6875	6.375	10
Sub22	1	2.375	3.125	4.0625	6.375	10
Sub23	1.0625	2.8125	4.1875	4.25	8.75	9.9375
Sub24	1.375	2.5	2.25	4.6875	7.4375	10
Sub25	3	4.1875	4	3.6875	8.5625	10
Sub27	1	2.25	2.625	2.5	5.1875	9.875
Sub28	1	1.6875	2.75	6.8125	7.625	10
Sub29	1.125	1.4375	1.5	3.125	7.75	9.6875
Sub30	1.125	1.0625	2.25	2.4375	5.666667	9.875
Mean	*1.396635*	*2.027465*	*2.987981*	*4.555288*	*7.444231*	*9.79533*

Table 21: Mean listening difficulty rating as a percentage for each signal-to-noise ratio under the Female Fry condition for each subject

	Female Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	2.9375	4.875	4.6875	6.6875	8.5	10
Sub02	2.125	5.75	7.6875	8	9.3125	10
Sub03	5.5625	1.25	1.9375	4.785714	6.0625	9.75
Sub04	1.125	2.1875	3.1875	4.875	6.1875	10
Sub05	1	3.0625	4.25	5.5625	8.9375	10
Sub06	2	4.375	2.5	6.266667	9	10
Sub07	1.1875	1.375	2.5	4.5625	6.75	10
Sub09	1.0625	2.3125	3.4375	4.3125	3.875	9.9375
Sub10	1.0625	1.4375	1.625	7.5625	9.5625	10
Sub11	1	1.625	1.4375	2.3125	3.125	10
Sub12	1.875	3.9375	2.4375	5.375	5.3125	9.8125
Sub13	1	1.75	2.6875	4.9375	6.1875	10
Sub14	1.625	2.0625	2.3125	6.3125	6	10
Sub15	1	1.375	4.3125	8.466667	10	10
Sub16	2.5625	3.5625	3.6875	5.25	7.6875	10
Sub17	1	3	4.375	5.125	8.125	10
Sub19	1	2.8125	5.4375	7.0625	9.375	10
Sub20	1.0625	3	3.3125	3.625	5.375	10
Sub22	1	3.3125	3.0625	7.25	9.6875	10
Sub23	1.1875	5.1875	2.625	6.1875	9.25	10
Sub24	1.125	2	2.9375	5.6875	7.625	9.9375
Sub25	2.1875	2.75	3.875	3.3125	5.5625	10
Sub27	1.0625	1.9375	1.625	2.1875	7.0625	10
Sub28	2	2.4375	2	6.1875	9.3125	10
Sub29	1.25	1.125	2.25	1.875	4.8	9.6875
Sub30	1.0625	2.3125	1.8125	2.8125	7.875	9.75
Mean	*1.579327*	*2.723558*	*3.153846*	*5.253136*	*7.328846*	*9.956731*

Table 22: Mean listening difficulty rating as a percentage for each signal-to-noise ratio under the Female No Fry condition for each subject

	Female No Fry					
	NaN dB	**-12 dB**	**-6 dB**	**0 dB**	**6 dB**	**12 dB**
Sub01	2.1875	2.0625	1.625	4.875	7.8125	10
Sub02	3.125	5	6.5625	6.375	9.5	9.125
Sub03	1.125	1.375	4.8125	2.8	5.0625	10
Sub04	1	2.25	1.875	3.1875	7.125	10
Sub05	1.0625	1.625	2.4375	5.6875	8.25	10
Sub06	1	1.1875	2.25	8.375	8.8125	10
Sub07	1.125	1	1.75	2.25	5.4375	6.375
Sub09	1	1.5	1.375	2.4375	5.625	10
Sub10	1	1.125	2.1875	5.25	7.75	10
Sub11	1.125	1.25	1.9375	2.875	4.75	10
Sub12	1	1.875	2.0625	3.375	7.25	9.4375
Sub13	1.5	1.5625	1.75	4.0625	5	9.4375
Sub14	1.1875	2.0625	2.375	3.9375	7.1875	9.4375
Sub15	1	1.75	3	3.9375	10	10
Sub16	1.6875	2.1875	4.4375	5.8125	6.5625	10
Sub17	1.1875	1.25	3	4.6875	8.3125	10
Sub19	1.3125	2.8125	3.5	5.3125	8.1875	10
Sub20	1	2.0625	2.0625	3.9375	5.9375	10
Sub22	1.0625	2.75	1.875	3.4375	8.125	10
Sub23	1.1875	1.375	1.25	8.5	9.0625	10
Sub24	1.0625	1.375	2.5625	2.0625	5	8.875
Sub25	2.375	2.3125	2.6875	5.75	5.25	10
Sub27	1.125	1.1875	1.9375	5.5	8	10
Sub28	1	1.5	2.125	2.875	5.8125	10
Sub29	1.0625	1.0625	1.25	2.625	3.6875	9.4375
Sub30	1.0625	1.625	2.1875	2.1875	6.625	10
Mean	*1.290865*	*1.8125*	*2.495192*	*4.312019*	*6.927885*	*9.697115*

REFERENCES

REFERENCES

Abdelli-Beruh, N.B., Drugman, T., & Red Owl, R.H. (2016). Occurrence frequencies of acoustic patterns of vocal fry in American English speakers. *Journal of Voice, 30*(6), 11-20.

Abdelli-Beruh, N.B., Wolk, L., & Slavin, D. (2014). 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).

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*, 2649-2658.

Boone, D.R., McFarlane, S.C., Von Berg, S.L., & Zraick, R.I. (2014). *The voice and voice therapy* (9th ed.). Boston, MA: Pearson Education.

Borrie, S.A. & Delfino, C.R. (2017). Conversational entrainment of vocal fry in young adult female American English speakers. *Journal of Voice, 31*(4), 513.e25-513.e32.

Bradlow, A.R., Torretta, G.M., & Pisoni, D.B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*(3), 255-272.

Colton, R.H., Casper, J.K., & Leonard, R. (2006). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (3rd ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Hollien, H. & Michel, J.F. (1968). Vocal fry as a phonational register. *Journal of Speech and Hearing Research, 11*, 600-604.

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.

Lee, S. (2015). Creaky voice as a phonational device marking parenthetical segments in talk. *Journal of Sociolinguistics, 19*(3), 275-302.

Morton, V. & Watson, D.R. (2001). The impact of impaired vocal quality on children's ability to process spoken language. *Log Phon Vocol, 26*, 17-25.

Munro, M.J. (2013). Intelligibility. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-7). Hoboken, NJ: Blackwell Publishing.

Ramig, L.O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson's disease. In R.D. Kent (Ed.) *Intelligibility in speech disorders* (pp. 119-155). Philadelphia, PA: John Benjamins.

Ryalls, J.H. & Liberman, P. (1982). Fundamental frequency and vowel perception. *The Journal of the Acoustical Society of American, 72,* 1631-1634.

Shiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R.D. Kent (Ed.), *Intelligibility in speech disorders* (pp. 11-34). Philadelphia, PA: John Benjamins.

Titze, I.R. (2000). *Principles of voice production* (2nd ed.). Salt Lake City, UT: National Center for Voice and Speech.

Tjaden, K., Sussman, J.E., & Wilding, G.E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*(3), 779-792.

Watson, P.J. & Schlauch, R.S. (2008). The effect of fundamental frequency on the intelligibility of speech with flattened intonation contours. *American Journal of Speech-Language Pathology, 17,* 348-355.

Whitehead, R.L., Metz, D.E., & Whitehead, B.H. (1984). Vibratory patterns of the vocal folds during pulse register phonation. *The Journal of the Acoustical Society of America, 75,* 1293-1297.

Wolk, L., Abdelli-Beruh, N.B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice, 26*(3), 111-116.

Yorkston, K.M. & Beukelman, D.R. (1984). *Assessment of the intelligibility of dysarthric speech*. Austin, TX: Pro-Ed.