ACOUSTIC CHARACTERISTICS OF TENSE AND LAX VOWELS ACROSS SENTENCE POSITION IN CLEAR SPEECH

by

Lindsay Roesler

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

Master of Science
in Communication Sciences and Disorders

at
The University of Wisconsin-Milwaukee
August 2013

ABSTRACT

ACOUSTIC CHARACTERISTICS OF TENSE AND LAX VOWELS ACROSS SENTENCE POSITION IN CLEAR SPEECH

by

Lindsay Roesler

The University of Wisconsin-Milwaukee, 2013
Under the Supervision of Professor Jae Yung Song

The purpose of this study was to examine the acoustic characteristics of tense and lax vowels across sentence positions in clear speech. Recordings were made of 12 participants reading monosyllabic target words at varying positions within semantically meaningful sentences. Acoustic analysis was completed to determine the effects of Style (clear vs. conversational), Tenseness (tense vs. lax), and Position (sentence-medial vs. sentence-final) on vowel duration, vowel space area, vowel space dispersion, and vowel peripheralization. The results showed speakers had longer durations and expanded vowel spaces in clear speech for both tense and lax vowels. Importantly, the amount of increase was similar for tense and lax vowels suggesting the defining properties of lax vowels (i.e., short duration and centralization) were manipulated in clear speech. A significant main effect of position for lax vowel space expansion showed greater vowel spaces for lax vowels in sentence-medial position in clear speech. Clear speech vowel adaptations appear to be dynamic with both vowel-specific and general transformations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge the people who have made this project possible. My sincerest appreciation to Dr. Jae Yung Song. Your commitment and guidance in helping me to complete this thesis has been unparalleled. You have taught me the value of striving to become a detail-oriented person. Moreover, thank you for the countless laughs we have shared over the many months.

A special thanks to my committee members, Dr. Marylou Gelfer, and Dr. Carol Seery for taking the time to act as mentors, and provide me with such excellent feedback throughout this process.

Distinct recognition also goes to all of the lab assistants in the Phonetics Lab who helped me during this study. You all made the time spent in the lab a little bit sunnier.

Finally, thank you to my family. You always provided encouragement at the exact moments I needed it.

1. Introduction

Clear speech is a speaker dependant adjustment that is produced in communication settings in which a speaker judges an enhanced signal is required for a listener to comprehend the message (Krause & Braida, 2002; Picheny, Durlach, & Braida, 1986, Smiljanic & Bradlow, 2009; Wright, 2004). These adjustments can be produced in an attempt to overcome a noisy environment, or a language barrier, or to communicate with someone who is hard of hearing (Moon & Lindblom, 1994; Smiljanic & Bradlow, 2009). The modifications speakers use to enhance communication are made spontaneously, and are often made unconsciously (Smiljanic & Bradlow, 2009). The use of clear speech as an intelligibility enhancing strategy has also been documented cross-linguistically in the Canadian French, Spanish, and Croatian languages (Smiljanic & Bradlow, 2005; Smiljanic & Bradlow, 2009). The aim of this study is to provide a systematic investigation of the acoustic characteristics of tense and lax vowels in clear and conversational speech. To give relevant background information, the introduction will briefly explore the history of clear speech literature and then it will examine the previous findings on the characteristics of clear speech, focusing on those of tense and lax vowels. In doing so, it will also provide an overview of the acoustic theory of vowel production and vowel classifications. Finally, it will discuss how the investigation of the characteristics of tense and lax vowels in clear speech can contribute new acoustic research to the existing vowel literature.

1.1 Clear Speech History

The phenomenon of clear speech has been widely researched for nearly a century. Scholars have been interested in clear speech for a number of different reasons. Early studies focused on improved communications between normal hearing partners as a means to overcome a noisy environment, or in training talkers to improve telecommunications (Black, 1958; Snidecor, Malbry, & Hearsey, 1944; Tolhurst, 1955). These studies provided data supporting the fact that speakers could improve ratings of intelligibility by modifying speaking behaviors. In the 1980's clear speech research shifted in an attempt to investigate if the same speaking behaviors that provided improved intelligibility for normal hearing speakers extended to those with communication difficulty, namely, hearing loss (Chen, 1980; Picheny, Durlach, & Braida, 1985). In fact, clear speech was found to have significant intelligibility benefits for those with actual hearing loss, and those in conditions of a degraded signal. Perceptual studies have found that clear speech improves overall ratings of intelligibility in various populations including hard of hearing adults (Picheny et al., 1985), normal-hearing adult listeners (Bradlow, Torretta, & Pisoni, 1996; Krause & Braida, 2002), nonnative adult listeners (Bradlow & Bent, 2002), and school-aged children with and without learning disabilities (Bradlow, Krause, & Hayes, 2003). In order to determine if speaking modifications are in fact responsible for change, ratings of intelligibility are established by having speakers read words or sentences in both the clear and conversational style, and having listeners write what was said. The listener's transcriptions are then scored, and serve as the intelligibility ratings. More recent studies aimed to explore what aspects of clear speech production are responsible for the improved intelligibility associated with clear speech. The primary factors that have been investigated are speaking fundamental frequency (SFF), speaking rate, vowel lengthening, and vowel space expansion.

2. Background

As stated previously, many research efforts on clear speech have focused on it as a way of communicating with individuals who are hard of hearing. This body of literature can be divided into two categories, clear speech perception and clear speech production. Perception studies are driven on identifying the degree to which certain strategies in clear speech affect listener intelligibility ratings. Production studies are concerned with identifying salient changes in the execution of clear speech as compared to conversational speech. Often, researchers will conduct studies that incorporate both production and perception components. By doing so, investigators can understand which modifications translate into enhanced perception. For this reason the literature review for the perceptual aspects and the production aspects of clear speech are collapsed for some studies. The seminal work by Picheny et al. (1985) present findings strictly related to intelligibility and will be presented first.

Picheny et al. (1985) conducted a study in order to investigate if clear speech benefits found in normal hearing adults extended to adults with hearing impairments. In order to examine this effect, the investigators created 50 nonsense sentences, which included adjectives, nouns and verbs. In addition, articles, auxiliary verbs, prepositions, and article nouns were randomly selected and inserted in the appropriate places in order to create an English declarative sentence. Three college-aged male speakers were then recruited, and recordings of both clear and conversational styles of speech were made.

The speakers read each sentence twice in each of the styles, and sets of 50 conversational speech and 50 clear speech stimuli were dubbed into one recording. In order to ensure the sentences produced by the speakers were intelligible for normal listeners, the lists were then tested on normal hearing subjects before running the hard-of-hearing experimental group. To obtain intelligibility ratings for these sentences, the normal hearing pre-test listeners could choose to write the sentence they heard, or repeat it orally to the examiner. Intelligibility scores were based on the number of correctly identified adjectives, nouns and verbs. A word was marked as incorrect if the listener omitted or misidentified a sound, but if a listener omitted a past-tense suffix, or plural marking, the word was still considered correct. It was reported that the recordings were greater than 94% intelligible by the normal hearing listeners who participated in the pre-testing procedure.

After obtaining the preliminary data on intelligibility, hearing impaired listeners were tested. This group was comprised of five listeners who all had stable sensorineural hearing losses. Closer inspection of the group data available revealed that while the group all demonstrated sensorineural hearing loss, there was considerable variability in the ear in which hearing loss was present, the degree of hearing loss each person had (range = mild to severe), and the word recognition scores for each person. For the experiment, this group of listeners was presented the sentences at three different intensity levels; most-comfortable-level, maximum listening level, and 10 dB below the most-comfortable-level. For each listener, 36 conditions were presented, "(2 speaking modes × 2 frequency-gain characteristics × 3 talkers × 3 levels)" across 50 sentences containing 175 target words (pg. 98).

An ANOVA was completed on the responses of four out of the five individual's data, and revealed that 41% of the variance was attributable to the speaker. This was not surprising considering the variability in residual hearing ability across the listeners. The second highest source of variability was attributable to the mode of speaking (i.e. speaking style) at 19%. The authors interpreted this to mean that there are significant differences in intelligibility for hard of hearing individuals between the two modes (clear and conversational) with the clear mode more favorable for intelligibility gains.

Clear speech production has fascinated many researchers because of the perceptual benefits it affords listeners. The speaker-dependent factors that have been widely studied are SFF range, speaking rate, vowel lengthening, and vowel space expansion, which will be discussed below.

2.1 Speaking Fundamental Frequency Range

To start, several studies present evidence showing an overall increase in SFF range in clear speech. In 1986, Picheny et al. conducted a follow up study to their Picheny et al. (1985) study in order to investigate the acoustic differences between clear and conversational speech. They had three male speakers recite 50 nonsense sentences in both the clear and conversational style. No statistical analysis was presented in the paper. However, the authors did report that a wider SFF range was noted in the clear speech style.

Bradlow et al. (1996) conducted an experiment with the goal of identifying some aspects of clear speech that were directly related to increased intelligibility. Sentences were taken from the Indiana Multi-Talker Sentence Database (Karl & Pisoni, 1994). The database consists of 10 male speakers and 10 female speakers reciting 100 sentences. The experimenters recruited listeners (10 for each speaker) to transcribe the sentences produced by the speakers. The sentences were presented to the listener at 75 dB binaurally. The transcriptions were then scored in an all or none fashion. If a sentence was not transcribed correctly, one percentage point would be taken away from the speakers overall intelligibility score (100 % points as a maximum score for each speaker). The final intelligibility score was the result of an average across all 100 sentences. After gathering intelligibility data, the researchers then analyzed which factors of speech were related to increased intelligibility. An interesting finding was that female speakers received significantly higher ratings of intelligibility than their male counterparts did. Females were transcribed with 89.5% accuracy whereas males were transcribed with only 86.2% accuracy. Females were also found to have used a significantly wider range of speaking fundamental frequency than males as was shown by a 2-tailed unpaired *t*-test ($t(18) = 4.84$, $p < 0.001$). Although females had a wider SFF range and higher ratings of intelligibility, no positive correlation between increased speaking fundamental frequency range and intelligibility was found (Spearman $\rho = +0.384$, $p = 0.095$). Due to the correlational nature of the statistical analysis, it is impossible to ascertain whether increased intelligibility is the result of a speaker increasing SFF range. However, these findings suggest some aspect of female vocal quality positively influences intelligibly ratings.

A study conducted by Smiljanic and Bradlow (2005) found similar results. This study was conducted in an effort to determine if characteristics of the clear speech style extended cross-linguistically. Five English speakers and five Croatian speakers were recorded reading twenty analogous, nonsense sentences (in their respective languages) that each contained four target words. The average SFF range expansion for females, or the increased number of semitones in SFF range during clear speech, was 1.26 semitones and for males was 1.48 semitones. Interestingly, within the group who had the greatest SFF range expansion there was great individual variability (0.899-3.63 semitones) in the clear speech style. A significant main effect of style ($F[1,8] = 14.292$, $p < 0.005$) but not of language was found when analyzing the SFF range of all 10 speakers. In accordance with the previously mentioned studies, SFF range expansion was not definitively a characteristic of clear speech for all speakers.

2.2 Speaking Rate

Speaking rate is a speech modification that comes to mind when thinking about clear speech. However, only a few studies have published findings on rate changes associated with clear speech (Bradlow et al., 1996; Krause & Braida, 2002; Picheny et al., 1986; Smiljanic & Bradlow, 2005), and those findings, while informative, are not generally robust. First, differences in average speaking rate will be discussed and then pause changes will be presented. Although vowel lengthening impacts speaking rate in clear speech, there is extensive literature (Ferguson & Kewley-Port, 2007; Picheny et al., 1986; Smiljanic & Bradlow, 2008) to be discussed, and it will be presented in a separate section.

Picheny et al. (1986) found that when the three male subjects recited the 50 nonsense sentences in clear and conversational styles, the clear speech style had fewer words per minute (90-100) than the conversational style (160-200). Unfortunately, no statistical findings were present in the publication to determine if these differences between the styles were significant.

Bradlow et al. (1996) measured rate changes of the twenty speakers reciting the 100 Harvard sentences in Indiana Multi-Talker Database (Karl & Pisoni, 1994) by analyzing and comparing average sentence durations. The Harvard sentences are mono-clausal with five keywords and contain any number of function words. It was reported that the mean sentence duration for all the speakers was 2.115 seconds (standard deviation = 0.276 seconds). As the study was primarily interested in investigating what factors of clear speech are involved in enhanced intelligibility, the authors reported related findings. Specifically, there was not any obvious relationship between mean sentence duration and speech intelligibility scores.

Krause and Braida (2002) provide some insightful findings on the relationship between speaking rate and clear speech intelligibility ratings. The authors aimed to answer whether or not clear speech could also be achieved at fast rates. This challenges the general clear speech finding that a slow speaking rate is an inherent characteristic of clear speech. For this study, the experimenters screened and selected five speakers whom they thought would be able to produce fast speech without sacrificing intelligibility. The targeted speakers then participated in a training program where they practiced speaking quickly and maintaining intelligibility as was judged by a set of listeners (each listener heard a particular sentence only once). The listener and speaker were able to discuss potential changes that may increase intelligibility, and then the speaker would repeat the target sentence to the subsequent listeners until it was deemed clear. Final recordings of each speaker reciting 700 nonsense sentences were made. There were six speaking modes employed during the recitation (loud/normal, soft/normal, clear/normal, clear/slow, clear/quick) and six control modes which were done in conversational speech with varying intensity and rate.

The final step in the study was to present these sentence stimuli to a new group of listeners to obtain intelligibility ratings. The sentences were presented to eight listeners at a -1.8 dB signal to noise ratio (SNR). The listeners would repeat the sentence they heard aloud and the experimenter would judge the accuracy of the sentence. All listeners heard each of the six experimental and the six control styles. After analyzing the average percent correct across each listener in each style, the investigators found that the clear/slow speaking style received the highest intelligibility scores (63%). The intelligibility scores for the remaining styles are as follows: clear/normal (59%), loud/normal (53%), conversational/slow (51%), clear/quick (46%), conversational/normal (45%), conversational/quick (27%), soft/normal (26%) clear/slow (63%). The most interesting of the findings is that when comparing the clear speech styles, benefits are noted across speaking rate (i.e., clear speech was always found to be more intelligible than conversational speech of the same rate). After an ANOVA was complete, the authors interpreted the findings to suggest that clear speech was a statistically significant factor influencing ratings of intelligibility. They also noted that regardless of style, an increased rate would negatively impact intelligibility after a certain point (i.e., a person can only talk so fast before sacrificing intelligibility).

Smiljanic and Bradlow (2005) found that all 10 subjects recited the nonsense sentences more slowly in the clear speech style. The sentences in Croatian were reported to range from 10-16 syllables (average = 12.8) in length, and those in English ranged from 9-14 with an average length of 11.7. A significant main effect of style $F(1,8) = 94.713$, $p < 0.0001$, but not language was found. On average, speakers produced 1.44 syllables per second less in clear speech than in conversational speech. It was interpreted that speaking in fewer syllables per second was an attribute of both the English and Croatian clear speech styles.

Another critical component of rate change in clear speech is pause characteristics. It has been reported that when speaking clearly, there are more pauses inserted into a given unit of speech, as well as longer pause durations (Picheny et al., 1986; Smiljanic & Bradlow, 2005). Picheny et al. (1986) defined a pause as, "...any silent interval greater than 10 ms between words excluding silent intervals preceding word-initial plosives" (p. 435). After analyzing the speech of three males reciting nonsense sentences, it was found that both the number and duration of pauses significantly increased in clear speech. Unfortunately, no statistical evidence was presented so it is unclear if these findings were significant.

Smiljanic and Bradlow (2005) also present similar findings regarding number of pauses and pause duration. Nine out of ten subjects increased the total number of pauses in clear speech when compared to conversational speech. If fact, the authors state that the majority of the subjects did not use any pauses in conversational speech. The average increase in pause duration for the clear speech style was 0.12 seconds (range = 0.052s - 0.205s).

2.3 Vowel Lengthening

First, Picheny et al. (1986) found, for all three speakers, there is an overall increase in vowel duration in clear speech when compared to conversational speech. Interestingly, it was reported that lax vowels did not show as much lengthening as tense vowels. The authors speculated that lax vowels may be inherently short, and thus, the lesser degree of lengthening in clear speech is the result of that particular property of the vowel itself.

Next, Ferguson and Kewley-Port (2007) conducted a study to investigate which acoustic modifications were primarily responsible for improved intelligibility in clear speech. Utilizing 12 speakers from the Ferguson (2004) study, Ferguson and Kewley-Port obtained both conversational and clear recitations of the target words. The vowels used in the study were /i, ɪ, e, ɛ, æ, ɑ, ʌ, o, ʊ, u/ and were embedded in the syllable /bVd/. These 12 talkers were classified into two distinct groups based on previous ratings of intelligibility (low vs. high). Both groups' speech was then scaled to the same root-mean-square (RMS) intensity range, and presented to listeners at 70 dB SPL with a signal to noise (12-talker babble) of -10 dB. After analyzing the listener responses, it was determined that there was no main effect of group, suggesting that both groups received similar intelligibility ratings. More importantly to the investigation, however, the main effect of speaking style $F(1, 118) = 47.74$, $p < 0.01$, and the Style × Group interaction were significant $F(1, 118) = 50.27$, $p < 0.01$. It was hypothesized that the group who had received higher intelligibility scores would markedly differ acoustically in some salient aspects of their clear speech from the group with the lower scores, and thus, it was thought that the most influential clear speech attributes could be determined.

Each speaker produced two tokens for each vowel. In order to analyze vowel lengthening, the average was obtained for each speaker's production. A two-way repeated-measures ANOVAs revealed that in the clear speech style, all speakers produced significantly longer vowels than in conversational speech, $F(1, 98) = 172.94$, $p < 0.01$. An additional significant finding of importance concerning vowel changes was the Style × Group interaction, $F(1, 8) = 5.95$, $p < 0.05$. This was interpreted to suggest the group who had the higher intelligibility scores also had the most vowel lengthening. In particular, the group with better intelligibility scores was reported to have clear speech vowels that were, on average, 41% longer than conversational speech vowels as compared to the speakers with the low intelligibility score who had clear speech vowels that were only 25% longer than the vowels produced in conversational speech.

Smiljanic and Bradlow (2008) conducted a study in which they investigated the variability of vowel lengthening as a function of the surrounding speech sound contrasts. Similar to their 2005 study, five native speakers of English and five native speakers of Croatian (residing in the USA) were asked to read twenty anomalous sentences (in their respective native languages) in both the clear speech and conversational speech styles. These sentences contained target words comprised of vowels embedded in varying stop consonant combinations. For further analysis of prosody effects on lengthening, there were 20 additional sentences constructed for English speaking subjects to read. These sentences were manipulated so that the sentence-final words from the original set appeared in a non-final position.

Smiljanic and Bradlow (2008) found all vowels lengthened in clear speech. Specifically, they used a repeated measures ANOVA to investigate the effect of style (clear vs. conversational), length (tense vs. lax) and vowel pair on total vowel duration. There were five vowel pairs in the study: /e-ɛ/, /i-ɪ/, /o-ʌ/, /u-ʊ/, and /ɑ-æ/. The results showed both tense and lax vowels were longer in clear speech than in conversational speech. Despite overall vowel lengthening in clear speech, the style by length interaction was not significant, and Smiljanic and Bradlow (2008) interpreted this as meaning both tense and lax vowels lengthened to a similar degree in clear speech. However, they found a significant two-way interaction of length by vowel pair. When compared using paired *t*-tests for each vowel set, nearly all of the tense vowels were always longer than the lax counterpart, however, the lax vowel /æ/ did not follow this pattern and was found to be longer than /ɑ/. The investigators also considered sentence position (sentence-medial vs. sentence-final) and voicing (voiceless vs. voiced) effects on the vowels. It was found that vowels, regardless of type, were longer before voiced consonants than before voiceless, and the voicing effect was more prominent in sentence-final position than in sentence-medial position. Unfortunately, the use of nonsense sentences as stimuli directly calls into question the validity of results due to the unnatural way in which speakers may produce nonsense sentences.

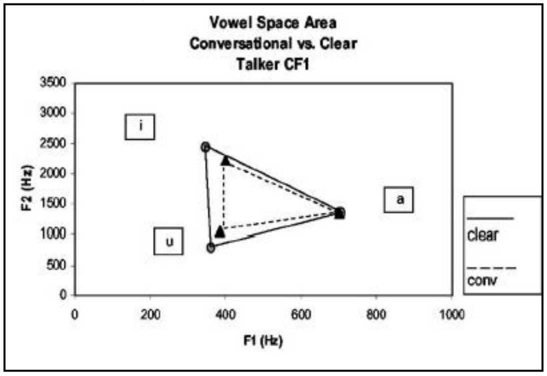
2.4 Vowel Space Expansion

In addition to vowel lengthening in clear speech, there are several other vowel transformations that have been documented in clear speech. Studies on the articulation of clear speech have shown that the tongue is located at more extreme positions in the oral tract when producing vowels in clear speech, resulting in vowels that are acoustically and perceptually more distinct from one another. For example, if the back vowel /u/, which has low F2, is produced in the more extreme back position, the F2 will be more low than its usual F2. Likewise, if the front vowel /i/, which has high F2, is produced in the more extreme front position, the F2 will be higher than its usual F2. This will make the distance in F2 between /u/ and /i/ greater. In order to analyze vowel changes from one speaking style to another, researchers commonly use several within subjects measurements of vowel space differences. Most commonly used are vowel space area, vowel space dispersion, and vowel peripheralization. The methodology of the measurements will be discussed before the literature review to provide the necessary background information.

2.4.1 Vowel Space Area

Most notably, vowel space area, which is based on the vowel triangle, has been used as a way to plot vowels based on second formant (F2) and first formant (F1) data (Fant, 1973; Picheny et al., 1986). Vowel space area is measured as the Euclidean area covered by the triangle, which is defined by the mean of each vowel category. Vowel space area is illustrated in Figure 1, where F1 (in Hertz) is labeled on the X-axis, and F2 (in Hertz) is labeled along the Y-axis. When the mean values of F1 and F2 for each vowel is plotted, the dots are connected by a line, and the area inside of the resulting triangle is mathematically calculated. In the figure, the conversational speech vowels are marked with triangles and connected by the dashed line, and the clear speech vowels are marked as circles and connected by a solid line. Because speaking clearly is known to effect the F1 and F2 of vowels, the data from vowels in clear speech will have greater distance between each other and thus, the overall area in clear speech will be greater when compared to the area in conversational speech. In fact, the vowel space area covered in clear speech should encompass the vowel space area plotted in conversational speech, and several researchers have found those results (Ferguson and Kewley-Port, 2007; Picheny et al., 1986; Smiljanic and Bradlow, 2005).

Figure 1. Vowel space area in conversational (conv.) and clear speech (Figure adapted from Smiljanic and Bradlow, 2005).



Picheny et al. (1986) found vowel space areas were larger in clear speech, but they observed that the vowel space area changes for tense and lax vowels in clear speech were different, with lax vowel expanding more than tense vowels. They gave an interpretation of the data stating, "The formant frequencies for lax vowels seem to be more sensitive to speaking mode than those for tense vowels, which change very little from conversational to clear speech" (p. 441). Thus, although the duration of lax vowels did not increase in clear speech, clarity of lax vowels was apparently achieved through the expansion of the vowel space. Unfortunately, no statistical data supporting this conclusion were presented.

A finding from the Bradlow et al. (1996) study of importance is the F1-F2 distance of the vowels /ɑ/ and /i/. The authors refer to these vowels as "point vowels" (pg. 266). It is widely accepted that each vowel has a characteristic acoustic representation based on the locations of formant frequencies. Based on this knowledge, the authors hypothesized that F1-F2 distance for the vowel /i/ would be positively correlated with increased intelligibility, since /i/ is characterized by widely spread F1 and F2 formants. Conversely the F1-F2 distance for the vowel /ɑ/ would be negatively correlated with intelligibility because there is relatively close spacing of F1 and F2. For example, if the F1-F2 distance of the vowel /ɑ/ spread out, where F1 became lower, and F2 became higher, the distance would become more like that of the vowel /i/ with the widely spread F1-F2 distance, thus the intelligibility of the vowel would decrease. This hypothesis was born out and analysis an all 20 speakers' productions revealed that this trend was associated with increased intelligibility (Spearman $\rho = +0.601$, $p = 0.009$ for the vowel /i/ and Spearman $\rho = -0.509$, $p = 0.027$ for the vowel /ɑ/). This suggests that clear speech elicits vowels that are produced with more extreme articulation, and the vowels with more extreme articulation can be understood more easily.

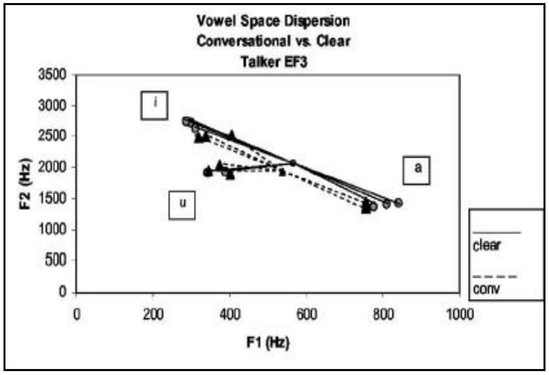
In the cross-linguistic study conducted by Smiljanic and Bradlow (2005) it was found that, subjects did have an increased vowel space area in clear speech when producing nonsense sentences. More specifically, the findings were significant for style ($F(1,8) = 48.691$, $p < 0.0001$), but not for language. Based on the significant finding for style, it was interpreted that the vowel space areas for both English (a language with a larger vowel inventory) and Croatian (a language with a smaller vowel inventory) changed to a similar degree for the tense vowels /ɑ, i, u/. Regardless of the mention of vowel inventory size, only three vowels were used in this study.

Most recently, Ferguson and Kewley-Port (2007) relate similar vowel space area findings. Recall that this study investigated the differences in speakers deemed to have a high degree of intelligibility and speakers with lesser ratings of intelligibility. The authors of this study chose to represent findings using the perceptually motivated Bark scale rather than the linear formant scales used in previously discussed studies. All speakers were found to expand their vowel space area in clear speech $F(1,8) = 65.57$, $p < 0.01$. More interestingly, those with higher ratings of intelligibility were reported to expand their vowel space area by 1.1 Barks (9%) compared to the lower intelligibility group expansion of .41 Barks (3%). The vowels of interest in this study were /i, ɪ, e, ɛ, æ, ɑ, ʌ, o, ʊ, u/, however, only the vowels /i, æ, ɑ, u/ were used in the analysis of vowel space area between clear and conversational speech.

2.4.2 Vowel Space Dispersion

Wishing to expand vowel space measures, Bradlow et al. (1996) devised the measurement known as vowel space dispersion. Vowel space dispersion uses the F1 by F2 plane to observe changes in clear speech by calculating how much a particular vowel moves from the center of a talker's F1 by F2 space. Rather than using the mean F1 and F2 for a given set of vowel data, all of the vowel productions are plotted, and the mean of the distances of each vowel from the central point in the talker's F1 by F2 space is considered the vowel space dispersion. When a talker produces a vowel in a more clear manner, the resulting vowel will subsequently move further away from the center point in the talker's F1 by F2 space. This phenomenon can be seen in Figure 2. Here, the vowels being represented by triangles and connected with the dashed line were produced in conversational speech, whereas the vowels represented by circles and connected with solid lines were produced in clear speech. It is clear that the lines connecting the vowels to the center point for clear speech are longer than those for conversational speech, suggesting that vowel formants for clear speech are dispersed further from the center of a talker's F1 by F2 space. While the change appears small, several authors present studies that established this method of change as reliable when comparing vowel changes across speaking styles (Bradlow et al., 1996; Smiljanic & Bradlow, 2005).

Figure 2. Vowel space dispersion in conversational and clear speech (Figure adapted from Smiljanic and Bradlow, 2005).



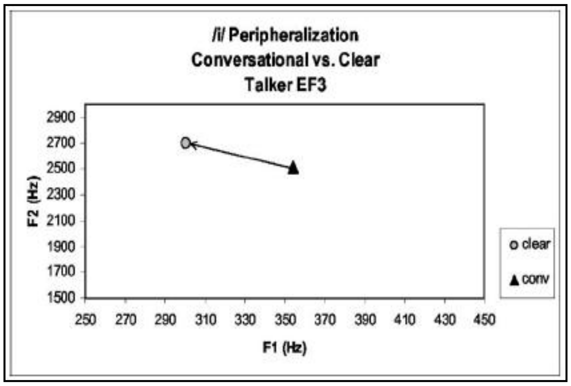
Bradlow et al. (1996) demonstrated the use of vowel space dispersion as a secondary method of measuring how tightly clustered or how widely spread from the center of an individual's vowel space each vowel token was in clear speech. It was reported that a moderate, positive rank order correlation (Spearman $\rho = + 0.431$, $p = 0.060$) was found for all 20 talkers when comparing ratings of intelligibility and vowel space dispersion. Further, Bradlow et al. (1996) reported that the measures of vowel space area and vowel space dispersion were highly correlated with each other (Spearman $\rho = + 0.782$, $p < 0.0001$). In fact, it was interpreted that due to the imperfect relationship between vowel space area and vowel space dispersion, "...each measure captures a slightly different aspect of the talkers' vowel production characteristics" (p. 265).

Next, it was found by Smiljanic and Bradlow (2005) that speakers of English and Croatian demonstrated greater vowel space dispersion in clear speech compared to conversational speech $F(1,8) = 27.577$, $p < 0.001$. However, despite the convincing evidence that vowel space dispersion is a characteristic of clear speech, the study mentioned above only included analysis from a small set of vowels, specifically, /i, u, ɑ/. Because all of these vowels are typically considered tense vowels, no data contributions were made to further the general understanding of how lax vowels behave in clear speech.

2.4.3 Vowel Peripheralization

The third method of measuring the difference in vowel articulation between clear and conversational speech is vowel peripheralization. According to Smiljanic and Bradlow (2005), vowel peripheralization differs from vowel space area and vowel space dispersion in that it calculates the Euclidean distance in the F1 by F2 space between the average token of a single vowel in conversational and clear speech styles. Using vowel peripheralization, it becomes evident how much a speaker changes an individual vowel from conversational to clear speech. The direction and magnitude of vowel changes can be easily observed in Figure 3. The vowel /i/ in conversational speech is designated by the triangle, and in clear speech by the circle. Notice that the F1 and F2 values of the high front vowel /i/ move into the upper left corner of the plane when switching from conversational speech to clear speech; this clearly demonstrates the decrease in F1 (suggesting an increase in tongue height) and the increase in F2 (suggesting an increase in tongue frontness) for the high front vowel /i/ in clear speech. The line drawn between the two points can be mathematically calculated, and can serve to show if certain vowels are more sensitive to the clear speech or conversational speech style.

Figure 3. Vowel peripheralization for /i/ (Figure adapted from Smiljanic and Bradlow, 2005).



Smiljanic and Bradlow (2005) found that the vowels used in their study peripheralized to similar degree in clear speech for both Croatian and English. Specifically ANOVA results were significant not significant for language and vowel. Vowel peripheralization has great potential to show differences in vowel category (i.e., to demonstrate differences in the magnitude of tense vs. lax vowel dispersion in clear speech), however, to date, no such studies have been done. Moreover, if articulatory strategies for achieving clear speech vary from person to person, statistical analysis of vowel space measures in combination with analysis of articulatory strategies may give researchers more information about which modifications are most commonly used in clear speech.

So far, this paper has presented the relevant background on clear speech research. Because this investigation is primarily motivated by determining the difference between tense and lax vowels in clear speech, the following section of the paper will provide background information on vowel production theory and phonemic classification of vowels. This will also provide insight into the stimuli development and methodological choices made for this study.

3. Vowel Production and Source Filter Theory

The shape of the vocal tract can be likened to a tube with superior openings at the lips and nostrils and the inferior opening at the vocal folds. Essentially, sound is generated at the vocal folds, and travels through the tube where it escapes out of the mouth or nostrils. All vowels are voiced, that is, they all require laryngeal vibration to be produced. Laryngeal vibration is caused by air from the lungs being passed through the glottis when the vocal folds are approximated to midline.. When set into vibration, the vocal folds move in a periodic manner which releases regular bursts (this is measured in cycles per second, or hertz [Hz]) of air into the pharynx. This is commonly referred to as the "source." Those bursts of air move up through the pharynx where they are modulated by its shape. Another critical component of the sound source that comprises vowels is that it is composed of more than one frequency component. These different frequencies are called harmonics. In fact, harmonics are multiples of the F_o of a tone. For example, if the F_o of a speaker is 150 Hz, the first harmonic frequency would be 300 Hz, the second would be 450, the third would be 600, and so on. Important to remember is that the harmonic frequencies of the sound moving through the vocal tract are distinct from the resonant frequencies of the vocal tract which will be discussed next (Kent & Read, 2002).

When acoustic energy travels, it does so in a longitudinal wave formation. These waves have certain regions where there is a high concentration of air particles (condensation) and areas where there is a low concentration of air particles (rarefaction). Depending on the rate at which the vocal folds are vibrating, a wave may have a long (slow vibratory rate) wavelength, or short (high vibratory rate) wavelength. Waves will fit differently in a vocal tract depending on size. For example, a particular vocal tract will accommodate more small wavelengths than it will long wavelengths. Curiously, given the way sound moves through a vocal tract, there is a certain relationship between vocal tract and wavelength that allows for the maximum vibratory amplitude; this is known as the function of the "filter" (Fant, 1973; Kent & Read, 2002).

The shape of the vocal tract and its resonant properties interact with the harmonic frequencies of a vowel resulting in the dampening of some of a vowel's harmonics. For example, the tongue is moved into various positions in the oral cavity in order to create the resonant property most characteristic of each vowel. Concentrations of acoustic energy centered around resonant frequencies of a vowel are called formant frequencies. Formant frequencies are important measures of vowel articulation, as they reflect the location of the tongue (i.e., the shape of the vocal tract or the resonator) when producing individual vowels. For example, to create the shape needed to produce the vowel /i/, the tongue is elevated and brought forward. This creates a relatively narrow passage for air to escape the oral cavity. This constriction in the superior-anterior aspect of the oral cavity translates acoustically to a lower first formant (F1), and a relatively high second formant (F2) which gives the overall spectrographic appearance of a widely spaced F1 and F2. Contrastively, the vowel /ɑ/ has a low tongue position, and relatively little constriction in the posterior dimension of the oral tract and is represented by a relatively high F1, and a relatively low F2. This is represented spectrographically as a relatively small space between F1 and F2.

4. Vowel Classifications

Vowels are classified in terms of height (high/central/low) and tongue frontness (front/central/back). Openness (close, close-mid, open-mid, open) is another term that is used interchangeably with height. Based on tongue movements, close vowels and open vowels were suggested in 1932 by Jones. Close vowels are said to be produced with the tongue in close approximation with the palate, whereas open vowels are produced further away from the palate. This becomes more evident when observing the IPA accepted vowel quadrilateral. This schematic representation of vowel placement is highly attributable to Jones (1932), and his work regarding cardinal vowels. A further description of the vowel triangle is necessary as all of the vowel space measurements previously discussed have a foundation in this early work.

Originally a triangle, the vowel quadrilateral has its underpinnings with the vowels /i, ɑ, u/ referred to as the cardinal vowels. The vowel triangle was conceptualized after studying X-ray data and plotting the corresponding tongue positions relative to the palate, and according to tongue retraction for each of the vowels (Ashby, 1989). The triangle formed by plotting these vowels then forms the perimeter for which the central vowels (/ɝ, ɚ, ə, ʌ/) fit inside.

Historically, linguists and phoneticians have been attempting to categorize vowels. One major debate of interest concerns the distinction between tense and lax vowels. In reviewing the literature, it becomes clear that there have been three main schools of thought on this matter that have persisted through time; tense vowels/lax vowels, long vowels/short vowels, and close vowels/open vowels (Bauman-Waengler, 2009). It is reported by Miller (1974) that the tense/lax distinction can be found in an early work of Sievers in 1901, where he described that certain vowels had more tension in both the tongue and larynx. After that, a number of early studies attempted to find a physical correlate to what Sievers called tension. Jones (1964) concluded that the tense/lax distinction could only reliably be noted for the high vowels. Moreover, he suggested a person could feel the tense-lax difference by palpating the throat. As cited by Miller (1974), Raphael conducted an electromyographic study of the genioglossus muscle in order to test the tense/lax hypothesis regarding muscular tension. The results of this study were, in order of descending muscular activity, /i, e, ɪ, ɛ/. Unfortunately, only front vowels were studied. Further, observing only the efforts of one muscle in order to achieve differing articulatory positions does not provide a clear picture of what is actually happening intraorally.

Next, according to Miller (1974), the long and short vowel classification system can be traced back to Jakobson, Fant and Halle in their seminal work published in 1952. They suggested that tense vowels have longer durations, and more distinction and pressure than their lax cognates. No reference data to support this hypothesis was presented. However, in a later paper in 1964, Jakobson and Halle did suggest that tense vowels are located at more extreme positions within the vowel triangle, and their lax cognates assume a more neutral position in the center of the triangle (Miller 1974).

For the remainder of this paper, tense vowels referred to in this study include /i, u, ɑ/ and lax vowels include /ɪ, ɛ, æ, ʊ, ʌ/. This is consistent with early work on tense and lax vowels that assumes an increased degree of overall tension is used to produce them (Bauman-Waengler, 2009). As mentioned previously, vowel space measurements use formant data as a way to observe changes in clear speech vowel production. Thus, it becomes clear that careful consideration must be taken when designing stimuli or selecting subjects for a speech production study aiming to compare vowels. One factor known to affect vowel production is dialectal variation.

5. Dialectal Variation

It has been long known that regional variations exist in American English pronunciations. The University of Pennsylvania conducted an extensive study in an effort to investigate the differences more thoroughly. Of direct consequence to this study, Labov, Ash, and Boberg (2006) identified several main groups of speakers within the United States. Those were defined as the Inland North, The South, The West, The Midland, and the East. Each of these defined regions was found to have unique vowel characteristics.

Because Milwaukee, and subsequently the University of Wisconsin-Milwaukee (UWM), where the participants for this study were recruited, was identified as one of the cities that constituted the Inland North in the Labov et al. (2006) study, further attention will be given to the dialectal characteristics used by speakers in the Inland North. Of great importance is the trend for vowels to shift in a systematic way known as the Northern Cities Shift. This is known to occur mainly in the inland north, which includes Ohio, Indiana, Illinois, Michigan, and Wisconsin. This shift is characterized by the fronting of the vowel /ɑ/, the raising of the vowel /æ/, the backing of the vowels /ɪ, ɛ, ʌ/, and the lowering of the vowel /ɔ/ (Small, 2011). The consequence of this change is that those who speak with this dialect will use different articulatory patterns when speaking than those using a different dialect. As a direct result of articulatory movements, the acoustic output for these speakers would also be different than speakers from another region. After having completed a review of the literature on clear speech and establishing a background about vowel production, the rationale of this study will be presented next.

6. Goals

6.1 Rationale

One gap in the literature is that there has been no study that has completed a controlled investigation of how lax vowels change in clear speech as compared to tense vowels when grammatical sentences are read in both contexts. This is a fundamental area to investigate because there has long been controversy surrounding the definitions of tense and lax vowels (e.g., Stevens, 2000). Clear speech is known to provide a context in which speakers make exaggerated articulations to produce sounds (through lengthening and vowel space expansion), and it may provide a context to study the inherent properties of tense and lax vowels.

Moreover, there were only two studies that have investigated the degree of lengthening between tense and lax vowels (Picheny et al., 1986; Smiljanic & Bradlow, 2008), but they found contrastive findings. Picheny et al. (1986) found that the duration of tense vowels increased significantly more than lax vowels in clear speech, whereas Smiljanic and Bradlow (2008) found that tense and lax vowels were lengthened similarly in clear speech. To our knowledge, there is only one study, Picheny et al. (1986), that has investigated vowel space expansion between tense and lax vowels. They found a greater degree of expansion for lax vowels, however, as pointed out earlier in the Introduction, the total number of subjects was only three, and there was no statistical data to suggest if the observed differences were statistically significant. A final limitation of the study is that there was no mention of what vowels constituted the tense and lax series for the study.

Current research on vowel production in clear speech also lacks the systematic manipulation of the sentence position in which target vowels appear. Sentence position manipulation allows for the exploration of the effect of sentence level prosody on vowel production. It also provides an additional context for speakers to produce more exaggerated vowel forms thus, giving the opportunity to observe how vowels behave in their most canonical forms. It has been widely documented that the position of the target word in its utterance affects the acoustic characteristics of the target word (Oller, 1973). In particular, syllables and segments at prosodic boundaries, such as phrase-, clause- and sentence-final position, are found to be lengthened, a phenomenon known as final lengthening (e.g., Edwards, Beckman, & Fletcher, 1991). Thus, in sentence-final position, speakers tend to exaggerate words by increasing length which provides an acoustic contrast for the listener. For example, Crystal and House (1988) showed a reliable effect of voicing on vowel duration (i.e., longer vowel duration before voiced obstruents than before voiceless obstruents) only in sentence-final position. Similarly, Smiljanic and Bradlow (2008) found that the effect of voicing on vowel duration was greater in sentence-final position than in medial position, although the voicing effect was present in both positions.

6.2 Purpose

The purpose of this study was to provide a systematic investigation of the acoustic characteristics of tense and lax vowels across sentence position in clear and conversational speech. In particular, the following four acoustic characteristics were examined: duration, vowel space area, vowel space dispersion, and vowel peripheralization. Some of the basic predictions for the main effects that were addressed in the introduction are listed below.

- The clear speech style would elicit longer vowel durations and greater vowel space areas than vowels in the conversational style.
- The tense vowels would elicit longer duration and greater vowel space than the lax vowels would.
- The vowels in the sentence-final position would elicit longer vowel durations and greater vowel space areas than the vowels in sentence-medial position would.

7. Methods

In order to examine the acoustic characteristics of tense and lax vowels in clear speech, speaker subjects were invited to read sentences to two different listeners. One listener was used to elicit clear speech and the other was used to elicit conversational speech. The listeners will be discussed in greater detail below.

7.1 Listeners

In this study, two individuals were recorded for use in the experimental video that the speaker subjects watched, and are referred to as listeners. The role of the listeners was to prompt different speaking styles from the speaker subjects by acting as listeners for the productions. Adeline was an elderly woman who asked the subjects to read the sentences as clearly as possible because she said she had a hearing loss. Emma was a college student who asked the participants to read the sentences casually because she had no problem with hearing. It was expected that the speech directed to Adeline would prompt speakers to shift into the use of clear speech style, whereas the speech directed to Emma would prompt conversational speech style. Listeners for this study were recruited from the UWM campus by being personally contacted by the student investigator. The listeners were invited into the Phonetics Lab, and consent to obtain and use video and audio recordings was received in writing (Appendix A). Each speaker rehearsed and recorded the instructions that were to be provided to the speaker subjects. The full scripts for the hard of hearing listener and the normal hearing listener can be found in Appendix B. When satisfactory recordings were made of each listener, the subjects were paid $15.00 for their time, and were told more about the study. The recordings of each listener were then inserted into Superlab (Cedrus Cooperation, 2012) the experiment presentation software used in this study.

7.2 Speaker Subjects

To examine the characteristics of tense and lax vowels in clear speech, speech recordings were collected from 12 adults (6 males and 6 females), all of whom were Caucasian, monolingual speakers of American English between the ages of 18 and 28 (males 18-25; females 18-28) with a mean age of 21.36 years. Participants were recruited through flyers posted throughout the UWM campus (see Appendix C). Additionally, participants were all speakers of the Midwestern dialect, and had not resided outside of the Midwest for longer than a one-year period. All of the participants passed the hearing screening. In addition, none of them had a history of a speech, language, or hearing problem.

Six additional subjects were recruited but excluded from the final analysis for varying reasons. Three subjects were excluded because they missed numerous tokens. One subject failed to respond to a tone during the hearing screening and was discontinued from the study, and referred to Norris Health Center, at UW-Milwaukee. Two were excluded due to experimenter error in which no audio or video recordings were made.

7.3 Equipment

In order to examine the acoustic measures of tense and lax vowels in clear and conversational speech, sound recordings were made. Additionally, ultrasound recordings were also captured to examine the motions of the tongue during vowel production, but this portion of the study was not included in this thesis, and will not be further discussed. Audio data was collected using a Shure KSM137 unidirectional microphone. The video of the tongue movements was collected using a portable Sonosite 180 Plus ultrasound machine and a C11/7-4 MHz 11-mm broadband curved array transducer. The ultrasound transducer was held in place under each participant's chin with a specially designed ultrasound stabilization helmet (Articulate Instruments). The use of the stabilization helmet allows for the stabilization of the ultrasound transducer in relation to the head, independent of subjects' head movements. Both ultrasound and acoustic signals were recorded through a Sony GV-HD700 recorder in NTSC format (30 fps). This produced an ultrasound video with audio, which was downloaded to a computer after each recording session so the acoustic characteristics of speech could be analyzed.

7.4 Procedure

Consent was obtained in writing upon entering the Phonetics Lab in the Department of Communication Sciences and Disorders (see Appendix D). The consent form was approved by IRB on December 11, 2012 (see Appendix E). After consenting, subjects were verbally reminded of the procedures. Next, participants completed a brief questionnaire (see Appendix F) answering questions regarding their school level, area of study, and linguistic background. In this questionnaire, subject's names were replaced by unique code numbers. The consent form and questionnaire were then placed in a locked cabinet to ensure the safety of each subject's personal identification.

Next, participants were invited into the recording booth to have their hearing screened according to the ASHA (1997) standards (see Appendix G). If a participant did not pass the hearing screening, he or she was discontinued from the study, and referred to Norris Health Center. After the hearing screening, participants were asked to read a list of sentences (Appendix H) that would help the investigator determine if any dialectal differences were present in the subject's speech (Vaux, 2000-2005). This was done as an additional measure of control as it has been documented that regional dialect differences exist in American English. The responses were evaluated on a separate sheet (Appendix I). If in fact a subject was found to speak with a dialect different from that common in the Midwest, the subject was discontinued from the study. No one was excluded because of this criterion.

In order for subjects to know what to expect in the experimental video, and to minimize their potential speech errors during the recording, they were given the list of stimuli to look through for less than one minute. Subjects were then fitted with the specialized helmet used to stabilize the ultrasound transducer in relation to the head. Then subjects were brought into the recording booth and positioned in front of the computer monitor. The microphone was placed approximately 12 inches from the subject. The student investigator informed the subjects that further instructions would be given through the experimental video, started the experiment, and closed the doors to the recording booth. During the experimental video subjects were introduced to the two listeners who gave instructions asking participants to read sentences to them. The order of the presentation of the two listeners was randomized by the experimenter so that half of the subjects received the hard of hearing condition first, and the other half received normal hearing condition first. Within the video, the order of the sentences was randomized by SuperLab.

7.5 Stimuli

The tense and lax vowels studied in this investigation constituted of /i, u, ɑ/ and /ɪ, ɛ, æ, ʊ, ʌ/ respectively. These monophthong vowels were chosen because they are widely categorized into tense and lax distinctions (e.g., Small, 2011). Diphthongs were strictly excluded from this experiment due to the difficult nature of extracting steady-state formant data. Since some of the typical tense vowels are often considered to be diphthongs (e.g., /e/ and /o/), there are slightly more lax vowels than tense vowels in the stimuli set. The target vowels were embedded in words that were consonant-vowel-consonant (CVC) in construction. No words with approximants (liquids, nasals) were considered, as they are difficult to separate acoustically from the vowels. Also, all words had a voiceless coda (i.e., an unvoiced final consonant) (e.g., heat). This was done in an effort to control for the effect of the voicing status of the coda consonant on the vowels being analyzed. It is well known that the vowel duration is longer before voiced obstruents than before voiceless obstruents (House, 1961). The syllable shape and voicing control of the target words allowed for the vowels to appear in a similar context, thus providing a controlled comparison of the vowels. For each of the 8 vowels, there are four words containing the target vowel with a total of 32 words. The full list is in Table 1.

Table 1. Target words containing the tense and lax vowels.

Tense Vowels			*Lax Vowels*				
i	u	ɑ	ɪ	ɛ	æ	ʊ	ʌ
heat	suit	pop	sick	pet	fat	cook	chuck
sheet	boot	shot	tip	debt	pass	hook	cut
sheep	soup	cop	fit	deck	chat	book	shut
keep	duke	hot	kick	jet	tap	took	bus

The words containing the target vowel were also chosen by considering their word frequency and neighborhood density ratings. Word frequency refers to the number of times a word is used in a given word database. A word that has a high occurrence of use is considered a high frequency word, whereas, a word used infrequently is termed low frequency. Words with many similar sounding words are considered to have high/dense neighborhoods, whereas words with few similar sounding words are considered to have low/sparse neighborhoods (Luce & Pisoni, 1998). Wright (2004) found that speakers made articulatory adjustments for vowels depending on the target words' frequency and neighborhood density; that is, an increased vowel space area and greater vowel dispersion were found when producing words with low relative word frequency and high neighborhood density (i.e., lexically difficult words). Knowing that neighborhood density and relative word frequency affect speakers' productions of vowels, the words containing the target vowels in this study were controlled and had relative word frequency ratings of between 6,046 and 190,905 and neighborhood density ratings of 15-32. For these ratings, Washington University (2008) in St. Louis's English Lexicon

Project database and Neighborhood database was used. Specifically, the frequency ratings we used in the database were based on the Hyperspace Analogue to Language (HAL) speech corpus, which consists of approximately 131 million words gathered across 3,000 Usenet newsgroups (Lund & Burgess, 1996). The frequency ratings of the words ranged between 0 to 5,262,331, and the average frequency was around 5,600. The neighborhood density ratings we used in the database were determined from the Hoosier Mental Lexicon, a widely used database of almost 20,000 English words (Nusbaum, Pisoni, & Davis, 1984). To our best knowledge, the information on the exact range of neighborhood density is not available; however, neighborhood density of words can fall anywhere between 0 to 40, and the densities over or around 20 have been often considered as high in previous studies (e.g., Munson, Swenson, & Manthei, 2005).

The target stimuli for this experiment consisted of a set of 32 sentences containing two target words each (one in the medial position and one in the final position). To make all sentences a similar length, they were between 6 and 8 syllables, and were all simple sentences containing one independent clause. These considerations allowed us to maintain adequate control for sentence position. In the first sentence of example (1) below, there are two target words: *cop* (containing the tense vowel /ɑ/) and *shut* (containing the lax vowel /ʌ/). As well, the second sentence contains two target words: bus (containing the lax vowel /ʌ/) and cop (containing the tense vowel /ɑ/). Therefore, the word *cop* appears once in medial position and once in final position. Similarly, the word *shut* appears in final position, and it appears in medial position in a separate sentence (for a full list of the sentences, see Appendix J). Likewise, the word *bus* appears in medial position below, and it appears in final position in a separate sentence.

(1) The *cop* slammed the cell door *shut*.
He drove the *bus* past the *cop*.

Since there were 32 target words, the target stimuli for this experiment consisted of a set of 32 sentences. In addition to the target sentences, there were 8 filler sentences (one quarter of the target sentence number) (see Appendix J). The filler sentences contained words in medial and final position which had nasals (/n, m, ŋ/), voiced stops (/d/), and liquids (/l, r/) codas. These filler sentences provided variety to the sentence stimuli, and helped to reduce the likelihood of the subjects recognizing the target stimuli.

7.6 Data

All subjects produced 32 sentences. These 32 sentences were produced four times (two in the conversational style and two in the clear style), resulting in 128 total sentences. Each sentence production had two words containing a target vowel (one in the medial position and one in the final position) resulting in 256 words. Because there were 12 participants in total, the data set for all subjects should have consisted of 3,072 tokens (12 participants × 32 sentences × 4 repetitions × 2 target words). However, five tokens (across three participants) were excluded from the female subjects' token count resulting in the final count of 3,067. One token was omitted because the formant values were not clearly visible, thus no data could be extracted. One token was omitted because the speaker failed to produce the target. Three tokens said by the same speaker, which had the vowel /ʊ/, were excluded because the speaker produced an incorrect production the target word. The speech error changed the target vowel to /u/.

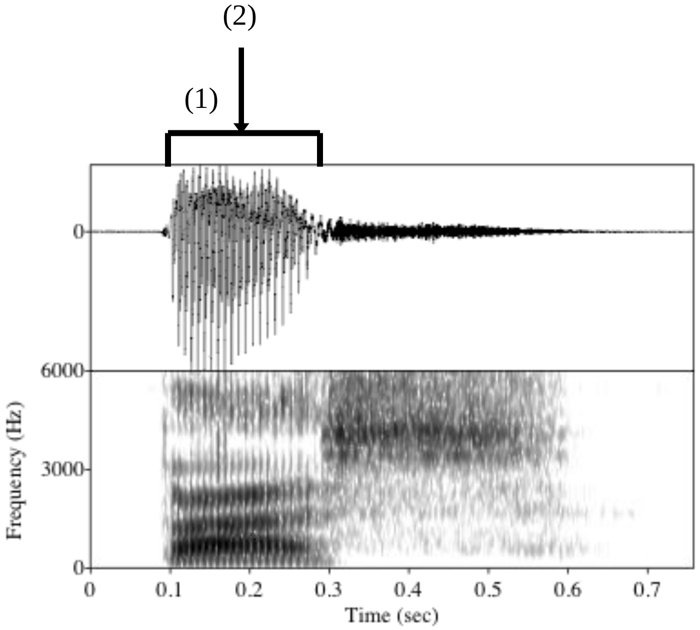
7.7 Acoustic Coding

The acoustic analyses of the audio data were carried out using Praat software (Boersma & Weenink, 2005). Four measures were used to determine the amount of vowel (tense and lax) change in clear speech: vowel duration, vowel space area, vowel space dispersion, and vowel peripheralization.

7.7.1 Durational Coding

As shown in the Figure 4, the vowel duration was defined as the interval between the onset and offset of the vowel. Vowel onset was determined by identifying the point in the vowel production where clear periodicity, F2 information, and glottal pulses were visually identifiable on the waveform and spectrogram. Vowel offset was defined as the point during the production where F2 information was no longer present, glottal pulses were absent, and constriction for the coda consonant started.

Figure 4. Representative waveform and spectrogram for the word *bus*, illustrating the (1) vowel duration and (2) mid-point of the vowel production.



Initial vowel boundaries were roughly determined by a lab assistant, and then finalized by the student investigator. Once the durational coding was confirmed, the durations for each production were extracted using a Praat script. The durational data were then organized by vowel, position (sentence-medial vs. sentence-final), and style (clear vs. conversational) in Microsoft Excel for further analysis.

7.7.2 Formant Coding

Using Praat, the formants were measured at the mid-point of the vowel production (see Figure 4) where the influence of the surrounding consonants is minimal. Following the suggestions from the Praat Manual, the formant setting for male subjects was adjusted to 5 formants with a maximum frequency of 5000 Hz. For females it was 5 formants with a maximum frequency of 5500 Hz. If the formant tracking in Praat did not reflect the actual formant bands seen in the spectrogram, various adjustments were made to improve the formant tracking in Praat. For example, if poor tracking was seen, adjusting the number of formants counted by Praat to 4 or 6 often improved the formant tracking. An additional inspection was completed to ensure the accuracy of the vowel formant coding. This was done by creating F1 by F2 plots for each subject's individual vowel tokens (i.e., individual vowels spoken by each subject were plotted). Tightly clustered tokens represented reliable formant coding. Outliers were identified and then the formants were manually coded again. Once the coding was completed, a Praat script was used to extract the F1 and F2 data from the mid-point of each vowel token. The extracted formant results were then opened in Microsoft Excel and organized by vowel, position (sentence-medial vs. sentence-final), and style (clear vs. conversational) for further analyses.

Using the first two formants of each word, three measures of vowel space area were calculated: 1) vowel space area, 2) vowel space dispersion, and 3) vowel peripheralization (following Smiljanic & Bradlow, 2005). 1) Vowel space area was measured as the Euclidean area covered by the triangle (for the tense vowels /i, u, ɑ/ or the pentagon for the lax vowels /ɪ, ɛ, æ, ʊ, ʌ/ made from the average F1 and F2 coordinates of each vowel category. Specifically, for tense vowels, the F1 and F2 coordinate points for the three vowels were connected, resulting in a triangle. Then the area of the triangle was calculated using a general mathematical formula for calculating the area of a triangle when you know the coordinates of the three vertices of a triangle. The specific formula is as follows: Vowel space area = $\lvert \{F1_i*(F2_ɑ - F2_u) + F1_ɑ*(F2_u -$

$F2_i) + F1_u*(F2_i - F2_ɑ)\}/2|$ (adapted from Liu, Kuhl, & Tsao, 2003). In the formula, “$F1_i$” represents the F1 (*x* coordinate) for the vowel /i/ where “$F2_i$” represents the F2 (*y* coordinate) for the vowel /i/. This is the same for the remaining vowels so that $F1_ɑ$ and $F1_u$ are the F1 (*x* coordinates) for the vowels /ɑ/ and /u/ respectively, while $F2_ɑ$ and $F2_u$ are the F2 (*y* coordinates) for the vowels /ɑ/ and /u/. The vowel space areas for the lax vowels were also calculated by using a mathematical formula for a pentagon. The calculation of the areas of a triangle (for tense vowels) and a pentagon (for lax vowels) was performed using a Matlab (Garcia, 2007). For each speaker, there were vowel space areas for tense and lax vowels in four different conditions (2 levels of speech Style [clear vs. conversational] × 2 levels of word Position [sentence-medial vs. sentence-final]), and these served as dependent measures for the statistical analysis. Because the shapes of the areas were different for tense and lax vowels (one was a triangle and the other was a pentagon), the vowel space areas were not directly compared between tense and lax vowels.

2) Vowel space dispersion was calculated as the distance of each vowel token from the geometric center point (i.e., centroid) in the speaker’s F1 by F2 space. Again, a general mathematical formula to calculate the centroids of a polygon (i.e., a triangle for tense vowels and a pentagon for lax vowels) defined by vertices (i.e., F1-F2 coordinates) was used to get the central point in the vowel space. The calculation of the centroids was done using a Matlab program (Garcia, 2007). After calculating an average vowel space dispersion for each of the 8 vowel categories in each speaker, the averages of the three tense vowels and five lax vowels were obtained. Each speaker produced vowel space dispersion values for 8 different conditions (2 levels of speech Style [clear vs.

conversational] × 2 levels of vowel Tenseness [tense vs. lax] × 2 levels of word Position [sentence-medial vs. sentence-final], and these served as dependent measures for the statistical analysis.

3) The extent of peripheralization in clear speech relative to conversational speech for each vowel category was calculated as the Euclidian distance between two points (i.e., an average F1-F2 coordinate point in conversational speech and an average F1-F2 coordinate point in clear speech) in the F1 by F2 plane. To calculate the distances, a general mathematical formula to calculate the distance between two coordinates was used for each vowel: $\{(F1_{clear} - F1_{conversational})^2 + (F2_{clear} - F2_{conversational})^2\}^{1/2}$. This distance represents the amount of change each vowel underwent from its conversational speech form to its clear speech form. The average distance was calculated for the three tense vowels and the five lax vowels. Each subject had averages in 4 different conditions (2 levels of Tenseness [tense vs. lax] × 2 levels of word Position [sentence-medial vs. sentence-final], and these served as dependent measures for the statistical analysis.

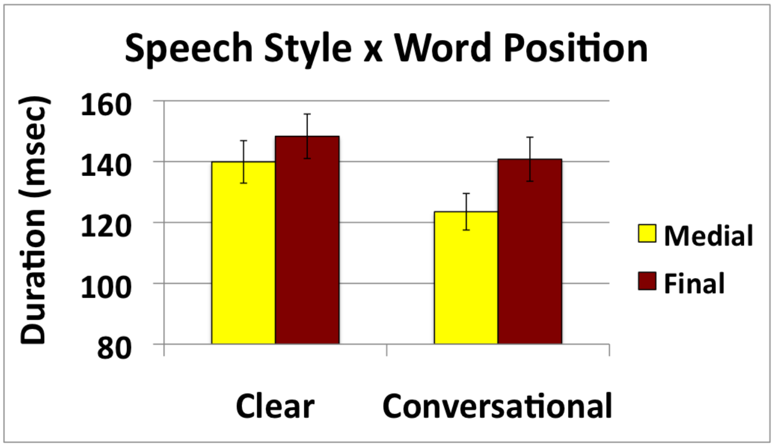
8. Results

Below the results for the four measures (vowel duration, vowel space area, vowel space dispersion, and peripheralization) will be presented.

8.1 Vowel Duration

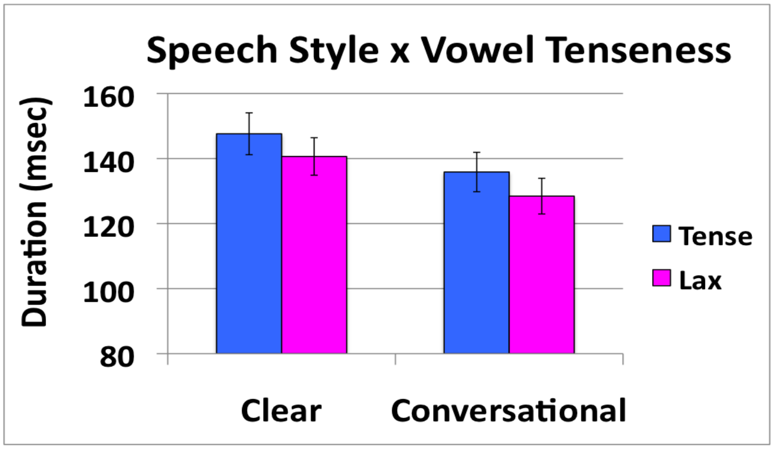
A 3-way repeated measures ANOVA was used to examine the effects of Style (conversational vs. clear), vowel Tenseness (tense vs. lax), and sentence Position (sentence medial vs. sentence-final) on vowel duration. The result showed that the main effect of Style was significant, *F*(1, 11) = .665, *p* = 0.001. As expected, vowels were longer in the clear speech style (*M* = 144.120 ms, SE = 6.203) than they were in the conversational speech style (*M* = 132.149 ms, SE = 5.582). Unexpectedly, the main effect of Position was not significant, *F* (1,11) = 3.344, *p* = 0.095, although the mean values for duration were overall longer in sentence-final position (*M* = 144.552 ms, SE = 7.102) than in sentence-medial position (*M* = 131.717 ms, SE = 6.372). However, the insignificant effect of Position might be due to the small differences between medial and final positions in clear speech (see Figure 5). In conversational speech, there appears to be a difference between medial and final positions. The interaction between Style × Position was significant, *F*(1,11) = 9.774, *p* = 0.010. From Figure 5, it is interesting to note that a lengthening strategy was utilized to make medial vowels more distinguishable in clear speech. That is, when switching from conversational to clear speech, there was a difference in the amount of durational increase between sentence-medial and sentence-final vowels. The duration of sentence-medial vowels increased dramatically in clear speech, while the duration of sentence-final vowels, which was already long presumably due to final lengthening, stayed more or less the same.

Figure 5. Effect of Speech Style × Word Position interaction on vowel duration. Error bars represent standard errors.



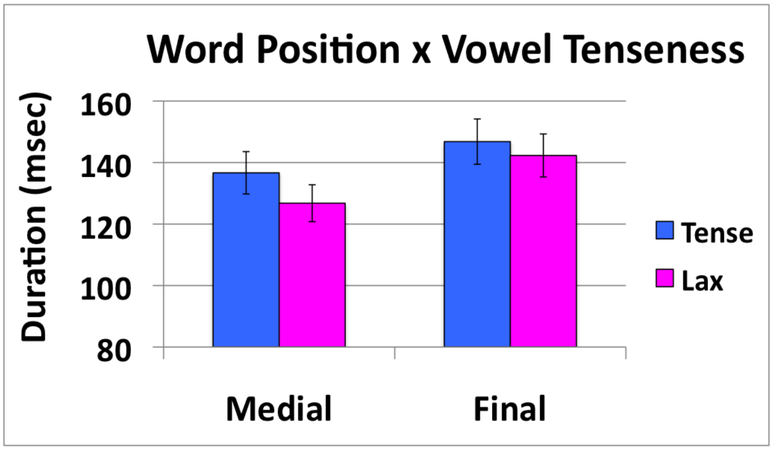
As expected, the main effect of vowel Tenseness was significant $F(1,11) = 29.176$, $p < 0.001$, with longer duration for tense vowels ($M = 141.732$ ms, SE = 5.952) than for lax vowels ($M = 134.536$ ms, SE = 5.645). This held true for both clear and conversational speech, as shown by no significant interaction between Style × Tenseness, $F(1,11) = 0.091$, $p = 0.768$ (Figure 6). That is, tense vowels were always longer than lax vowels to a similar degree in both clear and conversational speech. Interestingly, the lax vowel productions in the clear speech style lengthened to a point where they were longer than the tense vowel productions in the conversational style.

Figure 6. Effect of Speech Style × Vowel Tenseness interaction on vowel duration. Error bars represent standard errors.



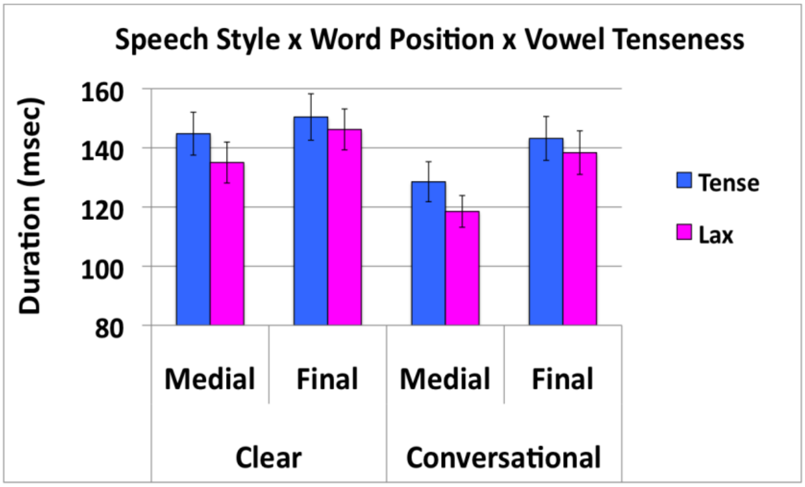
Furthermore, the tense-lax vowel distinction was equally present in the medial position and the final position, as shown by no significant interaction between Tenseness × Position, $F(1,11) = 2.442$, $p = 0.146$. The Tenseness × Position analysis suggests that both tense and lax vowels underwent final lengthening to a similar degree (Figure 7). Again, it is interesting to note that the lax vowel productions in the final position lengthened to a point where they were longer than the tense vowel productions in the medial position. This suggests that the short duration, a defining property of lax vowels, can be manipulated depending on the context.

Figure 7. Effect of Tenseness × Position interaction on vowel duration. Error bars represent standard errors.



Finally, three-way interaction of Style × Tenseness × Position was not found to be significant, $F(1,11) = 0.006$, $p = 0.940$. Again, it can be seen in Figure 8 that the distinction between tense and lax vowels was maintained regardless of speech style and sentence position because both tense and lax vowels underwent lengthening to a similar degree.

Figure 8. Effect of Style × Tenseness × Position interaction on vowel duration. Error bars represent standard errors.



The averaged values and standard deviations for each of the vowels can be seen in Table 2. Interestingly, the lax vowel /æ/ had the longest durations, and greatest variability of all of the vowels. Thus, although tense vowels were generally longer than lax vowels, there was some variability for individual vowels, especially for the lax vowel /æ/. An even more interesting observation is that generally, as the articulatory height of the vowels increased, the average durations became shorter. In other words, high vowels had shorter average durations than low vowels.

Table 2. Average duration (standard deviation) values (in msec) for individual vowels arranged by sentence position and style. Males and Females (N = 12).

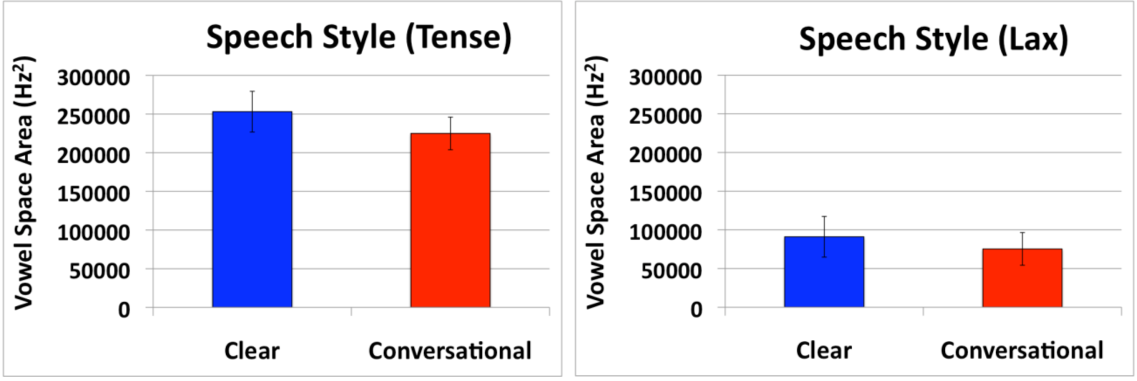
		Clear		*Conversational*		
		Medial	Final	Medial	Final	Average
Tense	i	123 (27)	129 (20)	111 (28)	125 (19)	122 (24)
	u	150 (30)	143 (25)	133 (34)	134 (26)	140 (29)
	ɑ	161 (23)	179 (28)	142 (21)	170 (32)	163 (26)
	Average	145 (27)	150 (24)	128 (28)	143 (25)	142 (26)
Lax	ɪ	102 (21)	115 (21)	90 (17)	108 (22)	104 (20)
	ɛ	148 (21)	157 (24)	134 (21)	154 (28)	148 (23)
	æ	195 (32)	202 (28)	169 (32)	188 (28)	189 (30)
	ʊ	102 (26)	119 (24)	88 (25)	110 (23)	105 (24)
	ʌ	127 (35)	139 (26)	111 (30)	131 (24)	127 (29)
	Average	135 (27)	146 (25)	118 (25)	138 (25)	134 (25)

8.2 Vowel Space Area

Two 2-way repeated measures ANOVAs were performed to examine the effects of speech Style (clear vs. conversational) and sentence Position (sentence-medial vs. sentence-final) on vowel space area for tense and lax vowels. Separate ANOVAs were performed for tense and lax vowels because of the imbalance between the tense vowels and lax vowels used for this study. That is, the three tense vowels (/i, u, ɑ/) formed a triangle, whereas the five lax vowels (/ɪ, ɛ, ʊ, æ, ʌ/) formed a pentagon; because the shapes of the areas were different to begin with, it was not meaningful to compare their sizes to examine the effect of vowel Tenseness.

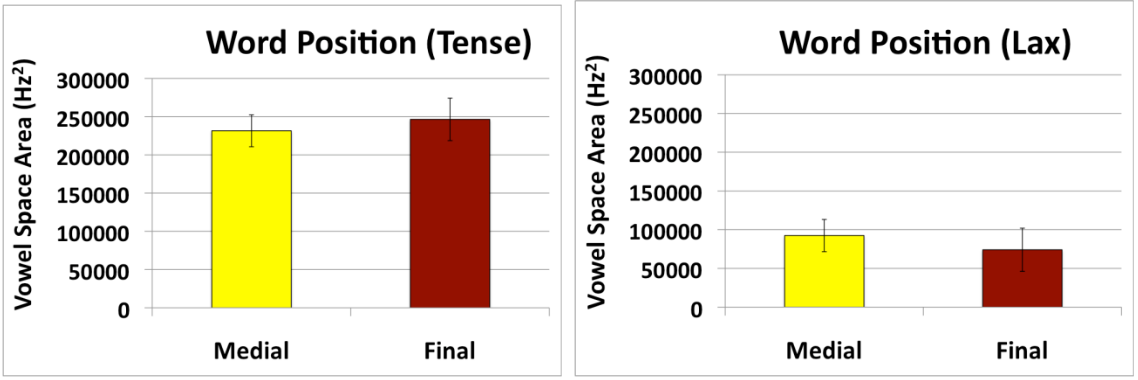
The results showed that the main effect of Style was significant. As expected, the clear speech style elicited a greater vowel space area for both tense [$F(1,11) = 10.993$, $p < 0.01$] and lax [$F(1, 11) = 14.096$, $p = 0.003$] vowels in clear speech than in conversational speech (Figure 9).

Figure 9. Effect of Style on Vowel space area for Tense and Lax vowels. Error bars represent standard errors.



Interestingly, the main effect of Position was found to be significant only for lax vowels, *F* (1,11) = 18.628, *p* = 0.001. Lax vowel space areas were greater in sentence medial position (*M* = 92430.000 Hz^2, *SE* = 9620.847 Hz^2) than in sentence final position (*M* = 74008.583 Hz^2, *SE* = 7588.063 Hz^2). In contrast to lax vowels, tense vowels did not show a difference in vowel space area between medial and final positions, *F*(1,11) = 1.044, *p* = 0.329. It can be seen in Figure 10 that the final position tense words had a slightly greater vowel space area (*M* = 246503.417 Hz^2, SE = 27850.186 Hz^2) than the tense vowels produced in medial position (*M* = 231508.667 Hz^2, SE = 20797.537 Hz^2). While not significant, it is interesting that tense vowels had a greater vowel space area in sentence final position unlike lax vowels, which had a greater vowel space area in sentence medial position.

Figure 10. Effect of Position on vowel space area for Tense and Lax Vowels. Error bars represent standard errors.



Word Position (Tense)

Word Position (Lax)

Vowel space area was found to be insignificant in the Style × Position interaction for both tense [*F*(1,11) = 1.262, *p* = 0.285] and lax [*F* (1,11) = 0.476, *p* = 0.504] vowels. That is, the effect of Position was independent of Style, suggesting that the amount of difference between medial and final vowels was equivalent in clear and conversational speech. As was mentioned above, for tense vowels, it can be seen that in both styles, the final position vowels had larger vowel space areas than the medial position vowels. For lax vowels, the medial position vowels had greater vowel space areas than the final position vowels in both styles. Also, in Figure 11, it can be seen for lax vowels, the average vowel space area for the Conversational/Medial average was slightly greater (*M* = 85163.250 Hz^2, SE = 8513.370 Hz^2) than the average vowel space area for the Clear/Final (*M* = 82430.083 Hz^2, SE = 8990.087 Hz^2). This suggests that while speech style is an important factor in eliciting larger vowel spaces, it seems for lax vowels, the position of the vowel is an equally important factor.

Figure 11. Effect of Style × Position on vowel space area for Tense and Lax Vowels. Error bars represent standard errors.

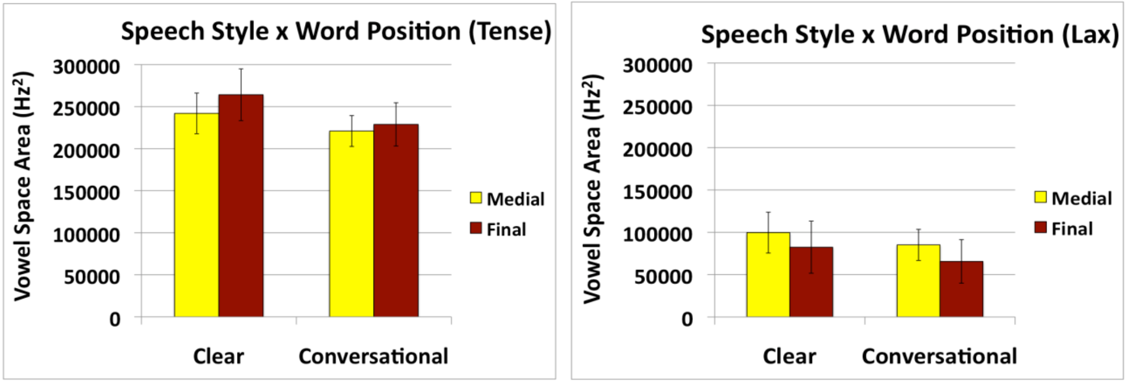
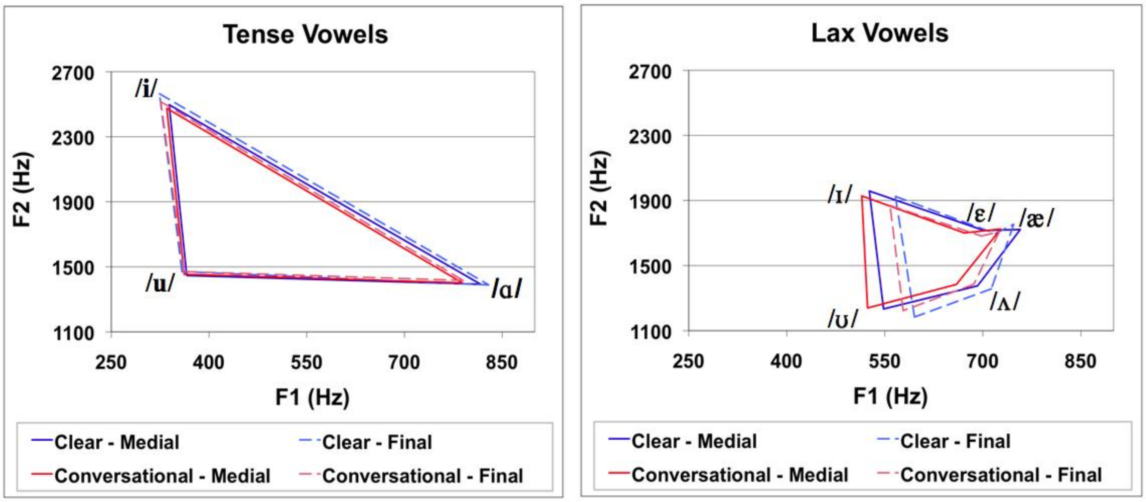


Figure 12 shows the F1-F2 coordinate plot of the tense and lax vowel area spaces. Recall, both tense and lax vowels had larger vowel space areas in clear speech than in conversational speech. However, for position, lax vowels underwent greater vowel space area expansion in sentence-medial position, whereas tense vowels were stable across sentence position. The vowel space expansion that lax vowels underwent in sentence-medial position can be clearly seen by looking at the solid lines in the F1-F2 plot for lax vowels below. In general, there seems to be a greater amount of F1-F2 change across both style and position for lax vowels than for tense vowels. Observationally, it seems as if the high lax vowels /ɪ, ʊ/ show the most variability.

Figure 12. Vowel space areas for tense and lax vowels arranged by style and sentence position.

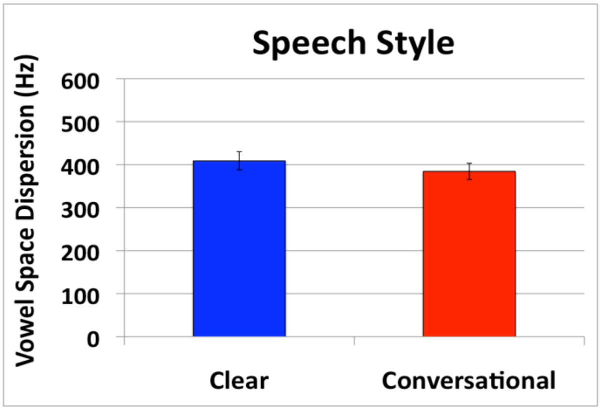


8.3 Vowel Space Dispersion

A 3-way repeated measures ANOVA was used to examine the effects of Style (conversational vs. clear), vowel Tenseness (tense vs. lax), and sentence Position (sentence-medial vs. sentence-final) on vowel space dispersion.

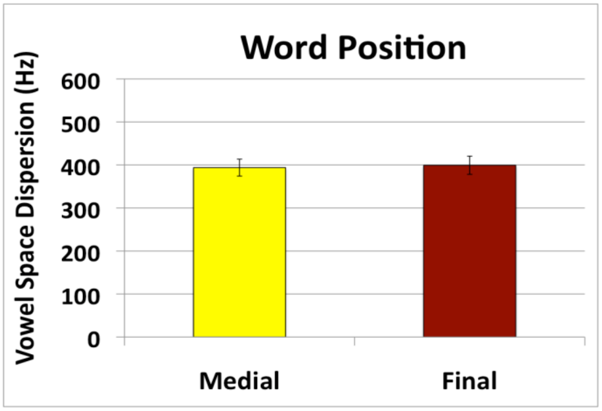
As expected, the main effect of Style was found to be highly significant for vowel space dispersion [F (1,11) = 25.437, $p < 0.0001$] (Figure 13) where vowels in clear speech had greater dispersion (M = 408 Hz, SE = 21.153 Hz) than those produced in conversational speech (M = 384.388 Hz, SE = 18.915 Hz).

Figure 13. Effect of Style on vowel space dispersion. Error bars represent standard errors.



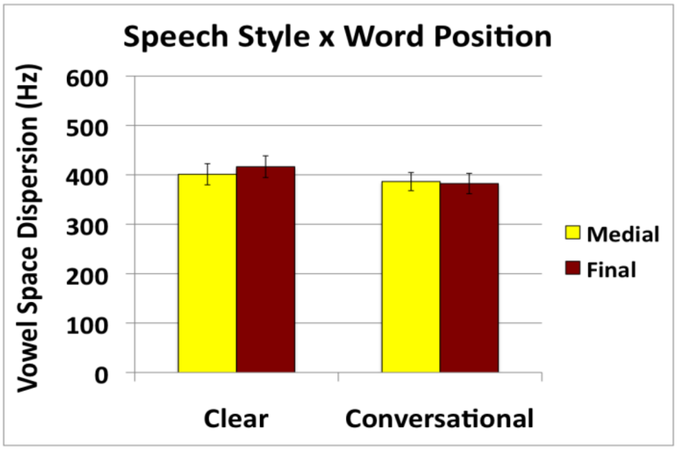
The main effect of position was not found to be significant *F* (1,11) = 0.384, $p = 0.548$ (Figure 14). The mean dispersion for sentence medial vowels was nearly equal to (*M* = 393.833 Hz, SE = 19.705 Hz) the mean dispersion for sentence final vowels (*M* = 399.387 Hz, SE = 21.104 Hz).

Figure 14. Effect of Position on vowel space dispersion. Error bars represent standard errors.



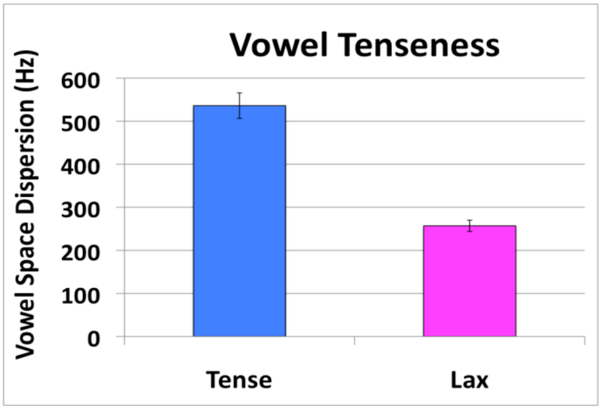
The interaction effect of Style × Position was significant, *F* (1,11) = 6.255, $p < 0.05$. It can be seen in Figure 15 that the amount of increase in vowel space dispersion from conversational to clear speech is greater for sentence-final vowels than for sentence-medial vowels. This observation was confirmed by a post-hoc analysis: Paired t-tests using a Bonferroni correction of the alpha level (0.05/4 = 0.0125) revealed that the vowel space dispersion of clear and conversational speech differed significantly in sentence-final position, *t*(11) = -5.53, *p* = 0.0001, but not in medial position, *t*(11) = -2.38, *p* = 0.03. In addition, there was no difference in vowel space dispersion between sentence-medial and sentence-final vowels in clear speech, *t*(11) = 1.509, *p* = 0.16, nor in conversational speech, *t*(11) = -0.43, *p* = 0.67.

Figure 15. Effect of Style × Position on vowel space dispersion. Error bars represent standard errors.



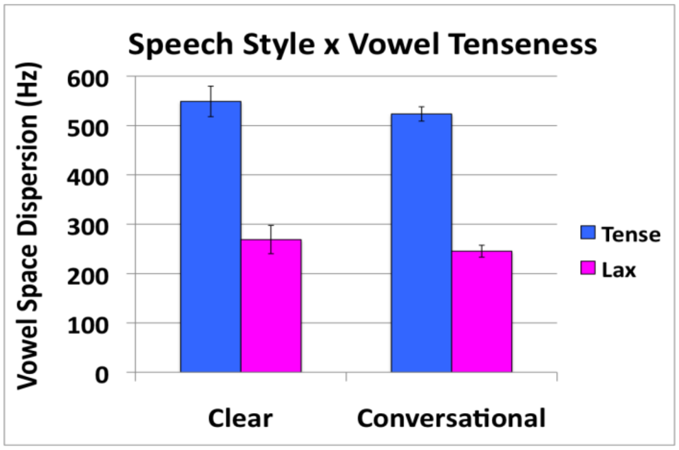
Vowel tenseness was highly significant *F* (1,11) = 155.373, *p* < 0.001 (Figure 16). Expectedly, tense vowels, which are produced at more extreme ends in the oral tract, had a greater degree of vowel space dispersion (*M* = 536.087 Hz, SE = 29.549 Hz) from the central point in the vowel space than lax vowels did (*M* = 257.133 Hz, SE = 13.066 Hz).

Figure 16. Effect of Tenseness on vowel space dispersion. Error bars represent standard errors.



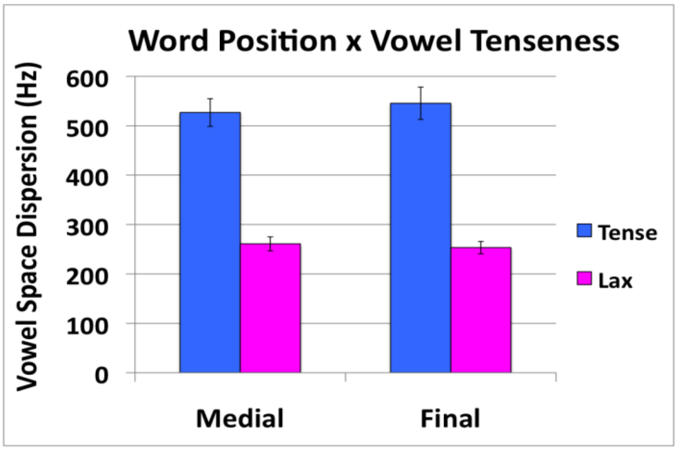
The Style × Tenseness interaction was not significant, *F* (1,11) = 0.039, *p* = 0.848. As it can be seen in Figure 17, the distinction between tense and lax vowels remained rather stable across the speech styles. The result also suggests that the amount of increase in vowel space dispersion from conversational to clear speech was equivalent for tense and lax vowels.

Figure 17. Effect of Style × Tenseness on vowel space dispersion. Error bars represent standard errors.



The Position × Tenseness interaction was also not significant, *F* (1,11) = 4.033, *p* = 0.070. It can be seen in Figure 18 that tense vowels always had greater vowel space dispersion than lax vowels did. This trend was equally present across sentence medial and sentence final positions, suggesting that speakers held vowel space dispersion for tense and lax vowels relatively stable across the sentence positions.

Figure 18. Effect of Position × Tenseness on vowel space dispersion. Error bars represent standard errors.



Finally, the three-way interaction between Style × Position × Tenseness was not significant *F* (1,11) = 0.169, *p* = 0.689. In Figure 19, notice how stable the distinction between tense and lax vowels is across sentence position and speech style.

Figure 19. Effect of Style × Position × Tenseness on vowel space dispersion. Error bars represent standard errors.

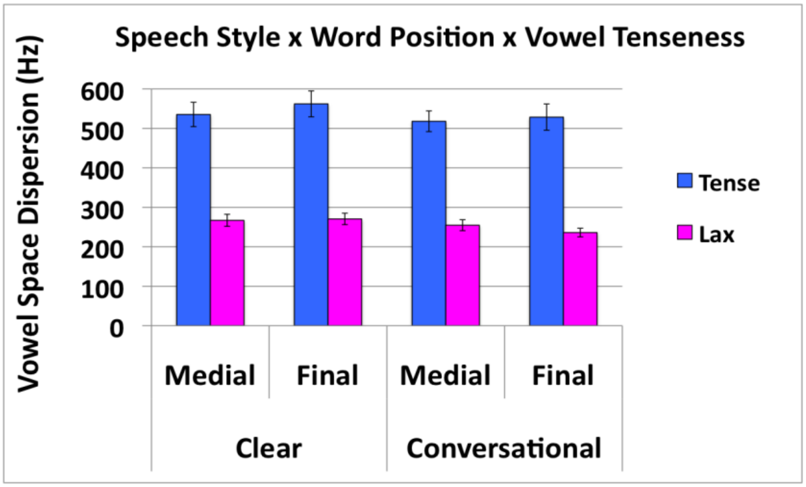


Table 3 below shows a comparison of the average vowel space dispersions in order from greatest to smallest for tense and lax vowel averages from the three-way interaction. Several observations can be made from the averages in the table. First, notice, that the clear speech style always elicited larger vowel space dispersion than conversational speech. Second, the largest dispersion values seem to affirm the finding of tense vowels and lax vowels differing across sentence position. That is, tense vowels seemed to disperse to a greater degree in sentence-final position (in both clear and conversational speech), whereas lax vowels showed a greater degree of dispersion in sentence-medial position when produced in conversational speech. Third, a closer look at the standard errors showed that tense vowels seemed to have had more variability across the sentence positions and speech styles than lax vowels had.

Table 3. Average vowel space dispersion and standard errors (in Hz) for tense and lax vowels ordered to show greatest to smallest magnitude of dispersion.

Tense Vowels			*Lax Vowels*	
Style × Position	Mean (SE)		Style × Position	Mean (SE)
Clear/Final	562.204 (32.864)		Clear/Final	270.635 (14.502)
Clear/Medial	535.373 (30.806)		Clear/Medial	267.116 (15.284)
Con/Final	528.659 (33.201)		Con/Medial	254.734 (13.863)
Con/Medial	518.112 (26.187)		Con/Final	236.049 (10.903)

Table 4 below shows the average vowel space dispersion for each vowel. Curiously, the vowels with the greatest amount of dispersion are the high front tense vowel /i/, and high front and back lax vowels /ɪ, ʊ/. Thus, except for the high back tense vowel /u/, high vowels generally had greater vowel space dispersion. These vowels were also greater in relative variability.

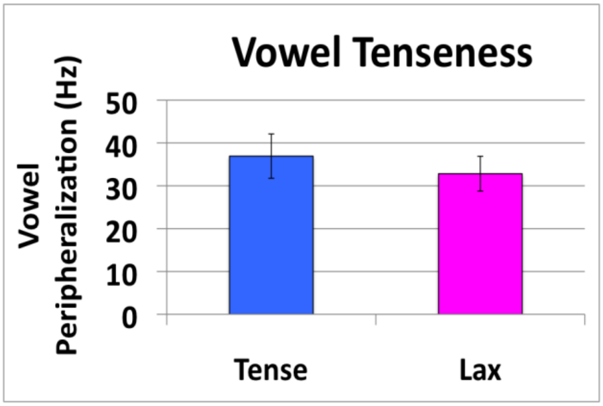
Table 4. Average vowel space dispersion (standard error) (in Hz) for individual vowels arranged by sentence position and style.

		Clear		*Conversational*		
		Medial	Final	Medial	Final	Average
Tense	i	738 (157)	778 (160)	717 (134)	734 (166)	742 (155)
	u	365 (124)	370 (140)	354 (120)	358 (142)	362 (132)
	ɑ	503 (126)	539 (143)	483 (129)	494 (139)	505 (134)
	Average	535 (136)	562 (148)	518 (128)	529 (149)	536 (140)
Lax	ɪ	376 (97)	370 (82)	353 (72)	316 (64)	354 (79)
	ɛ	193 (70)	198 (60)	179 (56)	181 (46)	188 (58)
	æ	168 (47)	182 (77)	177 (62)	153 (45)	170 (58)
	ʊ	367 (79)	384 (61)	353 (86)	342 (72)	362 (74)
	ʌ	231 (48)	220 (47)	212 (35)	188 (33)	213 (41)
	Average	267 (68)	271 (65)	255 (62)	236 (52)	257 (62)

8.4 Vowel Peripheralization

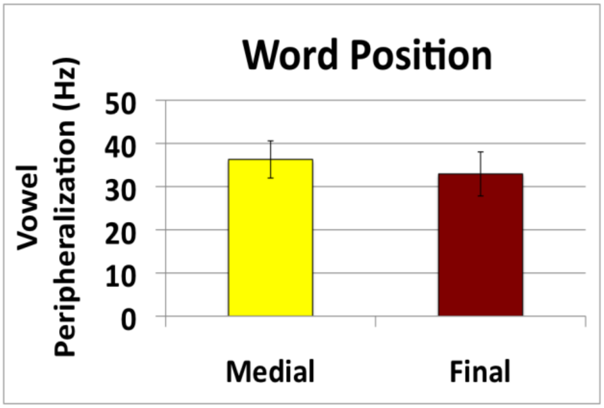
A 2-way repeated measures ANOVA was performed to examine the effects of vowel Tenseness (tense vs. lax) and sentence Position (sentence-medial vs. sentence-final) on the extent of peripheralization in clear speech relative to conversational speech. As for the main effect of Tenseness, there was no significant difference found for the amount of vowel peripheralization between tense and lax vowels [$F(1, 11) = .655$, $p = 0.435$] suggesting tense vowels and lax vowels peripheralized to the same degree in clear speech. It can be seen in Figure 20 that tense vowels peripheralized to a slightly greater degree (*M* = 36.934 Hz, SE = 5.170 Hz) than lax vowels did (*M* = 32.285 Hz, SE = 4.075 Hz), although the difference was not significant.

Figure 20. Effect Tenseness on vowel peripheralization. Error bars represent standard errors.



Also, the position of the vowel was not found to be significant, *F*(1, 11) = .317, *p* = 0.585. This was interpreted to mean both sentence-medial and sentence-final vowels peripheralized to the same degree in clear speech. However, in Figure 21 it can be seen that the amount of peripheralization for sentence-medial vowels (*M* = 36.290 Hz, SE = 4.312 Hz) was slightly greater than for sentence-final vowels (*M* = 32.929 Hz, SE = 5.107 Hz).

Figure 21. Effect of Position on vowel peripheralization. Error bars represent standard errors.



There was no significant interaction of Tenseness × Position, $F(1, 11) = 3.024$, $p = 0.110$. This supported the vowel space area and vowel space dispersion findings that, in clear speech, the acoustic changes made to tense and lax vowels are independent of the sentence position. An interesting observation can be made from Figure 22. Tense vowels appeared to have the greatest amount of peripheralization in sentence-final position (M = 38.657 Hz, SE = 7.847 Hz) whereas, lax vowels appeared to have the greatest amount of peripheralization in sentence-medial position (M = 37.638 Hz, SE = 6.429 Hz). Recall, a similar observation was found in the vowel space area Style × Position interaction where tense vowels had greater vowel space area in sentence-final position while lax vowels had greater vowel space area in sentence-medial position.

Figure 22. Effect of Position × Tenseness of vowel peripheralization. Error bars represent standard errors.

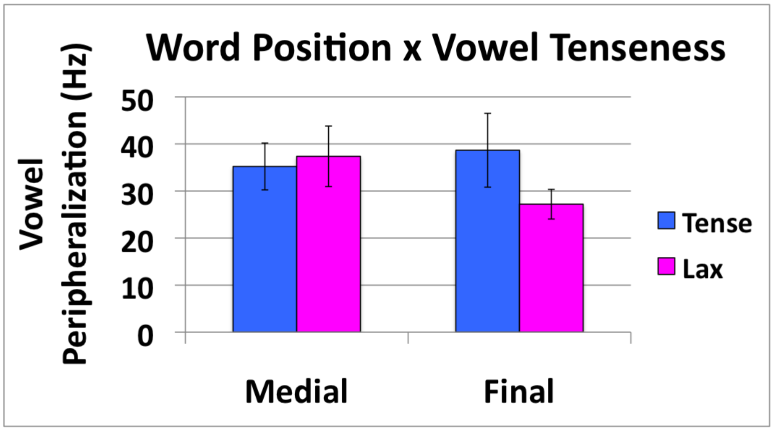


Table 5 shows the average vowel peripheralization for individual vowels across the varying sentence positions. Interestingly, sentence position averages for sentence-medial and sentence-final position are remarkably similar, highlighting the fact that sentence position does not seem like an important factor in the degree of vowel peripheralization. In addition, the average vowel peripheralization showed significant amount of variability suggesting that this measure varied greatly across participants.

Table 5. Average vowel peripheralization (standard error) (in Hz) for individual vowels arranged by sentence position.

		Medial	*Final*	*Average*
Tense	i	42 (38)	56 (34)	49 (36)
	u	79 (42)	64 (57)	72 (49)
	ɑ	48 (27)	59 (16)	54 (21)
	Average	56 (21)	60 (23)	58 (22)
Lax	ɪ	48 (37)	62 (35)	55 (36)
	ɛ	44 (20)	41 (26)	42 (23)
	æ	66 (36)	61 (38)	63 (37)
	ʊ	49 (40)	48 (28)	49 (34)
	ʌ	54 (30)	45 (27)	50 (28)
	Average	52 (24)	51 (21)	52 (23)

9. Discussion

9.1 Summary of Results and Comparison to Previous Research

The goal of the present study was to examine how various acoustic properties of tense and lax vowels (vowel duration, vowel space area, vowel space dispersion, and vowel peripheralization) vary across sentence position in clear and conversational speech. It was found that speakers used a wide variety of strategies in order to achieve clear speech. Table 6 is intended as a reference to provide an easy comparison of the results from this study to those of previous studies. Along the left column, the independent variables (both main and interaction effects) are listed. Along the top row, the dependent variables are listed. The results from the present study were compared against those from 4 previous studies, which are numbered chronologically from 1-4. The referent number 5, represents this study. An aspect of the summary table to mention is the use of the greater that (>) and less than (<) signs. These were used to show the relationship between the levels of the independent factors. For example, the results of Style on duration revealed that clear speech style had longer vowel durations than the conversational speech style did; this is represented as Cl > Co. In addition, bolded results indicate significant *p*-values for this study.

Table 6. Summary table of previous and current findings by dependent and independent variables.

Reference key:

	References	*Vowels used*
1)	Picheny et al. (1986)	tense vowels: /i, e, æ, ɑ, ɔ, o, u/, lax vowels: /ɪ, ɛ, ʌ/ (Note: no statistical results were published)
2)	Smiljanic & Bradlow (2005)	/i, u, ɑ/
3)	Ferguson & Kewley Port (2007)	/i, æ, ɑ, u /
4)	Smiljanic & Bradlow (2008)	tense vowels: /i, e, u, o, ɑ/ lax vowels: /ɪ, ɛ, ʊ, ʌ, æ/
5)	Current Study	tense vowels /i, u, ɑ/ lax vowels: /ɪ, ɛ, ʊ, æ, ʌ/

Factors	*Abbreviation key*
Style	Cl = Clear speech condition, Co = Conversational speech condition
Tenseness	T = Tense vowel, L = Lax vowel
Position	M = Sentence-medial, F = Sentence-final

Factors	*Duration*	*Vowel Space Area*		*Dispersion*	*Peripheralization*
Style	1) Cl > Co 3) $p < 0.01$ (Cl > Co) 4) $p < 0.05$ (Cl > Co) 5) **$p = 0.001$ (Cl > Co)**	1) Cl > Co 2) $p < 0.0001$ (Cl > Co) 3) $p < 0.01$ (Cl > Co) Vowel Space: Tense 5) **$p < 0.01$ (Cl > Co)**	Vowel Space: Lax 5) **$p = 0.003$ (Cl > Co)**	2) $p < 0.001$ (Cl > Co) 5) **$p < 0.001$ (Cl > Co)**	n/a
Tenseness	4) $p < 0.001$ (T > L) 5) **$p < 0.001$ (T > L)**	n/a	n/a	5) **$p < 0.001$ (T > L)**	5) $p = 0.435$
Position	5) $p = 0.095$	5) $p = 0.329$	5) **$p = 0.001$ (M > F)**	5) $p = 0.548$	5) $p = 0.585$
Style*Position	5) **$p = 0.01$ (F > M only in Cl)**	5) $p = 0.285$	5) $p = 0.504$	5) **$p < 0.05$ (Cl > Co only in F)**	n/a
Style*Tenseness	1) T > L 4) $p = 0.130$ 5) $p = 0.091$	n/a	n/a	5) $p = 0.848$	n/a
Tenseness*Position	5) $p = 0.146$	n/a	n/a	5) $p = 0.07$	5) $p = 0.11$
Style*Tenseness*Position	5) $p = 0.940$	n/a	n/a	5) $p = 0.689$	n/a

In order to summarize the findings from this study, and compare them with previous literature, each of the measures will be summarized separately beginning with vowel duration.

9.1.1 Vowel Duration

Picheny et al. (1986) published findings based on three speakers. For all three speakers, vowel duration increased in clear speech. Although no statistical values were presented, more recent studies have published similar findings. Ferguson and Kewley Port (2007) found the effect of style to elicit longer vowel durations in clear speech too ($p < 0.01$). Further, Smiljanic and Bradlow (2008) also found longer vowel durations for vowels in clear speech ($p < 0.05$). In this study, the effect of clear speech was found to be highly significant ($p < 0.0001$) with longer vowel durations observed in the clear speech style. Thus, it seems longer vowel durations are one characteristic of clear speech.

Regarding the effect of tenseness on duration, Smiljanic and Bradlow (2008) found tense vowels had longer durations than lax vowels ($p < 0.001$). This study also found the same result ($p < 0.001$). Regarding the interaction between Style × Tenseness, Picheny et al. (1986) mentioned that tense vowels had a greater increase in duration than did lax vowels in clear speech, but unfortunately did not provide any statistical results to support the claim. On the other hand, Smiljanic and Bradlow (2008) reported interaction effects of Style × Tenseness were insignificant ($p = 0.130$) and the present study provided the supporting evidence for the insignificant interaction ($p = 0.091$). This was interpreted to suggest that both tense and lax vowels were lengthened to a similar degree in clear speech.

So far, no previous studies have systematically examined the effect of position on the production of vowels in clear speech. The current study found no significant main effect of position on duration ($p = 0.095$). Interestingly, the Style × Position interaction was significant ($p = 0.01$) where sentence-final position elicited longer vowel durations than sentence-medial position did in conversational speech, but not in clear speech. The Tenseness × Position interaction effect was not significant ($p = 0.146$), suggesting that the tense and lax vowel distinction was maintained in both sentence-medial and sentence-final positions. Moreover, the three-way interaction of Style × Tenseness × Position was not significant ($p = 0.740$).

9.1.2 Vowel Space Area

Picheny et al. (1986) found that vowel space area increased when a speaker used clear speech. More recently, Smiljanic and Bradlow (2005) confirmed the findings of Picheny et al. (1986). In their study, the main effect of Style was highly significant ($p < 0.0001$) which meant that vowel space area did expand in clear speech. Ferguson and Kewley Port (2007) also found increased vowel space area to be a characteristic of clear speech ($p < 0.01$). Unfortunately, due to a limited number of vowels used in the previous studies, no data were available regarding how different types of vowels expanded in clear speech.

This study included a more comprehensive set of vowels, and analyzed tense and lax vowels separately in order to investigate whether there is a difference in vowel space expansion. Significant main effects were found for both tense and lax vowels ($p < 0.01$) ($p < 0.003$) respectively, which confirms both types of vowels do expand in the clear speech style.

This study also examined the effects of sentence position on vowel space expansion. It was hypothesized that vowels in sentence-final position would have greater vowel space areas than vowels in medial position. Interestingly, for tense vowels, no significant main effect was found for position ($p = 0.329$). More intriguing though, was the significant main effect of position for lax vowels ($p = 0.001$). In fact, lax vowels had greater vowel space area expansion in sentence-medial position. This is unexpected because vowels in sentence-medial position are generally more reduced than vowels produced in sentence final position. An observation revealed that it is primarily high lax vowels that were more vulnerable to change. The Style × Position interaction effect revealed no significant findings for tense or lax vowels ($p = 0.285$) and ($p = 0.504$) respectively.

9.1.3 Vowel Space Dispersion

Smiljanic and Bradlow (2005) published findings that vowels in clear speech underwent a greater degree of vowel space dispersion than the vowels produced in conversational speech ($p < 0.001$). It is obvious that if vowels are produced at more extreme points in the oral tract (as in clear speech), the vowels will show a greater distance from a talker's center point as measured in an F2 × F1 plot. Unfortunately, the only vowels analyzed by Smiljanic and Bradlow (2005) were / i, ɑ, u/ so no further comparisons between tense and lax vowels can be drawn from their paper.

This study confirms the findings that clear speech does elicit greater vowel space dispersion ($p < 0.001$). Additionally, a significant main effect was found for vowel Tenseness ($p < 0.001$). Expectedly, tense vowels underwent a greater degree of vowel space dispersion than lax vowels did. The Tenseness × Style interaction was not found to be significant ($p = 0.848$) which suggests that both tense and lax vowels have a similar degree of dispersion in clear speech.

Unexpectedly, the main effect of Position was not found to be significant ($p = 0.548$). However, the interaction effect of Style × Position was significant ($p < 0.05$). In fact, the amount of increase in vowel space dispersion from conversational to clear speech was greater for sentence-final vowels. The Tenseness × Position interaction effect was found to be insignificant ($p = 0.07$) suggesting that tense-lax distinction was stable across the sentence positions. The Style × Tenseness × Position interaction was also found to be insignificant ($p = 0.689$).

9.1.4 Vowel Peripheralization

Smiljanic and Bradlow (2005) found significant results for vowel peripheralization ($p < 0.001$). They used only the tense vowels /i, ɑ, u/, and it was determined that all of the tense vowels peripheralized to a similar degree. For this study, vowel peripheralization was not found to be significant for Tenseness ($p = 0.435$). This was interpreted as tense and lax vowels peripheralizing to the same degree as each other. Further, the effect of position was not significant ($p = 0.585$), which suggested that vowels in sentence-medial and sentence-final positions also peripheralized to a similar degree for clear speech. Finally, the interaction effect of Tenseness × Position was insignificant (*p* = 0.11).

9.2 Theoretical Implications

Phonemes produced in clear speech are exaggerated forms of those produced in conversational speech with the resulting clear speech sounds thought to be the most ideal form of the sound. Thus, clear speech served as the vehicle to observe acoustic properties inherent to tense and lax vowels. By manipulating the sentence positions vowels appear in, it could be determined if speakers use different strategies to produce ideal forms of tense and lax vowels varying in sentence positions in clear speech. This is particularly useful because the existing vowel dichotomy may not fully capture the behavior of vowels in varying contexts. It was hypothesized that vowels would have longer durations, greater vowel space areas, and greater vowel space dispersions in clear speech. This hypothesis was found to be true. Clear speech did elicit longer durations, and greater measures of vowel space area.

It was also hypothesized that tense vowels would always have longer durations and greater vowel space dispersion than lax vowels. This hypothesis was found to be true for the measures of duration, where tense vowels were longer than lax vowels, and for dispersion where tense vowels dispersed more than lax vowels. However, by looking at average duration values, the lax vowel /æ/ was the vowel with the longest duration. Further, it was observed that vowels with low tongue positions (/æ, ɑ/) had longer durations than vowels with high tongue positions, such as /i, ʊ/. Also, interestingly, no interaction was found between Style × Tenseness suggesting that the amount of increase in vowel duration and dispersion was similar for tense and lax vowels. Thus, the defining properties of lax vowels (i.e., short duration and centralization) were also manipulated in clear speech. This suggests that these adjustments are part of a general process in clear speech.

As for position, it was hypothesized that vowels in sentence-final position would have longer durations, greater vowel space area, greater vowel space dispersion, and greater peripheralization than lax vowels. Unexpectedly, the effect of position alone was not enough to elicit longer durations, greater vowel space dispersion, or greater peripheralization; rather, it was the interaction between Style × Position that determined the properties of the vowels. That is, sentence-medial vowels differed from sentence-final vowels primarily in conversational speech. Interestingly, for vowel space area, position alone was sufficient to elicit greater vowel space area for lax vowels, but not for tense vowels. It was unexpected to find a greater degree of vowel space area expansion for lax vowels in medial position because vowels in sentence-medial position tend to be reduced. Based on this finding, it seems that speakers do, to some extent, use different strategies to achieve clear speech for tense and lax vowels.

9.3 Limitations

The group of speakers used in this study was homogeneous. Aside from gender differences, age, education, dialect and ethnicity were similar for all speakers. This was done as a way to control for dialectal variation. Because of these constraints, the findings from this study may not generalize to other American English speaking groups.

This study involved ultrasound recordings of speech to examine the articulatory aspects of speech production, although the data were not included in the present thesis. Ultrasound methods for analyzing tongue movements are rather new, and it was questioned whether the use of the ultrasound transducer would interfere with articulatory productions of phonemes during speech. It was hypothesized that acoustic measures of speech while a speaker was wearing the helmet and ultrasound transducer would be similar to the previously reported acoustic finding, thus, would provide evidence that the use of the helmet and transducer did not negatively impact a speakers articulatory movements during speech. The results of this study turned out to be remarkably similar to the findings from the previously published clear speech studies that have been presented in the discussion section of this thesis. The fact that the speakers in the present study were able to produce expanded vowel spaces in clear speech comparable to the previous studies suggests that they were able to position their tongues at extreme positions in the oral tract. For this reason, it seems the speakers in this study experienced no negative articulatory impact from wearing the specialized ultrasound helmet, or having the ultrasound transducer placed under their chins during connected speech.

9.4 Future Research

Future research in this area could further contribute to the understanding of tense and lax vowels by providing more in-depth analysis of the vowels. Analyzing individual vowels more systematically rather than vowel groups could provide vowel specific trends rather than group trends. Further, by investigating the segmental differences between tense and lax vowels based on the height or frontness of a vowel, it could be determined if there are durational/vowel space cues inherent to the articulatory placement. This could provide more insight regarding the categorization of vowels, and if the traditional dichotomy is valid. Also, more research is needed to identify if various groups, such as age, gender, ethnicity, or dialect, use different strategies when producing clear speech. Finally, the inclusion of tongue imaging by the use of ultrasound techniques could expand on the findings of vowel space area, vowel space dispersion, and vowel peripheralization by providing visual corroboration of the strategies used to achieve vowels spoken in clear speech.

References

American Speech-Language Hearing Association (1997). *Guidelines for audiologic screening*. Available from www/asha.org/policy.

Ashby, M. (1989). A note on the vowel quadrilateral. *Journal of Speech, hearing and Language Research, 3*, 3-10.

Bauman-Waengler, J. (2009). Introduction to phonetics and phonology from concepts to transcription. Boston: Pearson.

Black, J. W. (1958). Multiple-choice intelligibility tests. *Journal of Speech and Hearing Disorders*, *22*, 213.

Boersma, P. & Weenink, D. (2005). PRAAT: Doing phonetics by computer (Version 4.4.07) [Computer program]. Available at http://www.praat.org/.

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*, 255-272.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America.* 112, 272–284.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for learning-impaired children: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research, 4,* 80–97.

Cedrus Corporation. (2012). SuperLab [computer software]. San Pedro, CA. http://www.superlab.com/

Chen F. R. (1980). Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level. (Masters thesis). Retrieved from dspace @ mit.edu

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America, 83,* 1553-1573.

Edwards, J. E., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America, 89*, 369–382.

Fant, G. (1973). Speech sounds and features. Cambridge, MA: MIT Press.

Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal hearing listeners. *Journal of the Acoustical Society of America, 116,* 2365-2373.

Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research, 50*, 1241-1255.

Garcia, D. (2007) BiomeCardio: MATLAB (Version 7.5) [downloaded software]. Available from http://www.biomecardio.com/matlab/polycenter.html.

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America*, 33, 1174-1182.

Jones, D. (1932). *An outline of English phonetics.* Leipzig: B.G. Teubner.

Jones, D. (1964). *An outline of English phonetics.* Heffer: Cambridge.

Karl, J., & Pisoni, D. (1994). The role of talker-specific information in memory for spoken sentences. *Journal of the Acoustical Society of America, 95*, 2873.

Kent, R. D., & Read, C. (2002). Acoustic analysis of speech (2nd ed.). Clifton Park, NY: Delmar.

Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America, 112* (5), 2165-2172.

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: phonetics, phonology, and sound change.* New York: Mouton de Gruyter.

Li, M., Kambhamettu, C., & Stone, M. (2003). EdgeTrak: a program for band-edge extraction and its applications. Sixth IASTED International Conference on Computers, Graphics and Imaging, Honolulu, HI, August 13-15.

Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science 6* (3), F1-F10.

Luce, P. A. & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1-36.

Lund, K., & Burgess, C. (1996) Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28*, 203-208.

Miller, P. D. (1974). A Critical Bibliography on the tense/lax distinction in vowels. Working Papers In Linguistics (Ohio State University No. 17, 222-231).

Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America, 96* (1), 40-55.

Munson, B., Swenson, C.L., & Manthei, S.C. (2005). Lexical and phonological organization in children: evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research, 48*, 108-123.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America*, 54, 1235-1247.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words* (Research on Speech Perception Progress Report Number 10). Bloomington, IN: Speech Research Laboratory, Department of Psychology, Indiana University.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research, 28*, 96–103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29*, 434-446.

Small, L. H. (2011). *Fundamentals of phonetics: A practical guide for students (3rd Ed.).* Boston: Pearson Education.

Smiljanic, R., Bradlow, A. R., (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118* (3), 1677-1688.

Smiljanic, R., & Bradlow, A. R. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics, 36*, 91-113.

Smiljanic, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass, 3* (1), 236-264.

Snidecor, J. C., Malbry, L. A., & Hearsey, E. L. (1944). Methods of training talkers for increasing intelligibility in noise (ORSD Report 3178). New York: The Psychological Corporation.

Stevens, K. N. (2000). *Acoustic Phonetics*. Cambridge, Massachusetts: MIT press.

Tolhurst, G. C. (1955). The effect on intelligibility scores of specific instructions regarding talking (USAM Rep. No. NM 001 064 01 35). Pensacola, FL: Naval Air Station.

Vaux, B. (2000-2005). Harvard dialect study, Retrieved from http://dialect.redlog.net/index.html. April 25, 2012

Washington University. (2009). English lexicon project [Data File]. Retrieved from http://elexicon.wustl.edu

Wright, R. (2004). Factors of lexical competition in vowel articulation. In J. Local, R. Ogden, and R. Temple (Eds.), Papers in Laboratory Phonology VI, pp. 26–50. Cambridge: Cambridge University Press.

Appendices

Appendix A: Listener Consent Form

Consent Form

The purpose of this research is to investigate how vowels are produced when speaking using clear speech.

If you participate in this research you will be asked to be audio- and video-taped for research purposes.

There are no foreseeable risks or discomforts to you as the subject.

Participant Consent

I, (print full name)___________________________, have read and understand the foregoing information explaining the purpose of this research and my rights and responsibilities as a subject. My signature below designates my consent to participate in this research, according to the terms and conditions listed above.

Signature___________________________________

Date________________

[If audio- or video-taping], I, (print full name)___________________________, give the researcher permission to use, publish, and republish, in the context of this research, photographic, video, or audio reproductions of my likeness or voice made for this study.

Signature____

Date__________________

Appendix B: Listener Scripts

Hard of Hearing

Hi, my name is Adeline, I was wondering if you could read some sentences for me today. The sentences I would like you to say are going to show up on the computer screen one at a time. This part of the experiment will take about 20 minutes. I have difficulty hearing so please speak as clearly as possible. Are you ready? Let’s start with a few practice words.

Great job, now we are going to start the experiment.

Thank you very much for reading the sentences for me. The experiment is now over.

Normal Hearing

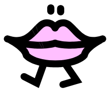
Hi my name is Emma. I was wondering if you could read some sentences for me today. The sentences I would like you to say are going to show up on the computer screen one at a time. This part of the experiment will take about 20 minutes. I don't have any trouble hearing so please speak as if you would in everyday conversation. Are you ready? Let's start with a few practice words.

Great job. Now we are going to start the experiment.

Thank you very much for reading the sentences for me. This part of the experiment is now over.

Appendix C: Recruitment Flyers

Paid Research Participants Needed



The Phonetics Lab in the Communication Sciences and Disorders Department is looking for people to participate in a speech production study.

- You will read a set of sentences on a computer monitor at a comfortable rate. We will make audio and ultrasound video recordings of your speech production. Ultrasound is a safe and non-invasive procedure that allows us to examine tongue movements during speech production.
- The experiment will take approximately 1 hour, and you will receive $15 for participation in this study.
- The research will take place at the Phonetics Lab located at Enderis Hall, Room B30.
- In order to for you to participate in the study, you must:
    - be at least 18 years old
    - be a monolingual, native speaker of English
    - have grown up in the Midwest(*especially*, Wisconsin, Minnesota, Iowa, Michigan, and the northern half of Illinois)
    - have normal hearing

If interested, please e-mail phonetic@uwm.edu to set up an appointment, and list a few times when you would be available.

If you have any questions or want to learn more about the study, please contact Lindsay Roesler at (412) 295-2665 or at phonetic@uwm.edu.

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Phonetics Lab speech production study
phonetic@uwm.edu

Appendix D: Consent

UNIVERSITY OF WISCONSIN – MILWAUKEE

CONSENT TO PARTICIPATE IN RESEARCH

THIS CONSENT FORM HAS BEEN APPROVED BY THE IRB FOR A ONE YEAR PERIOD

1. General Information

Study title:

Acoustic and articulatory characteristics of tense and lax vowels in clear speech

Person in Charge of Study (Principal Investigator):

Dr. Jae Yung Song

Assistant Professor in the Department of Communication Sciences and Disorders at UWM

2. Study Description

You are being asked to participate in a research study. Your participation is completely voluntary. You do not have to participate if you do not want to.

Study description:

The characteristics of speech production are affected by various factors such as the speaker's gender, age, and/or social status. The purpose of this study is to analyze the acoustic and articulatory characteristics of 'clear speech', a distinctive speaking style that speakers adopt when they are aware of a possible difficulty in the listeners' speech perception. In order to examine the acoustic characteristics of clear speech, audio recordings will be made and later analyzed using speech analysis software. The articulatory characteristics of clear speech will be investigated using ultrasound techniques, which are safe and non-invasive procedures that allow us to record the tongue motions during speech production. As there have been very few systematic investigations examining the actual tongue motions during the production of clear speech, this study is unique and will broaden our knowledge of how we produce clear speech. Furthermore, this type of research will lay foundation for building a more detailed and accurate model of speech production, which can be used in various settings such as language intervention programs.

All experiments take place in the Phonetics Lab, located in Enderis Hall, room B30 (2400, E. Hartford Ave, Milwaukee, WI 53211). Your participation in the study will last approximately 1 hour, and will require only one visit. Approximately 15 total individuals from the UWM community will participate in this study.

3. Study Procedures

What will I be asked to do if I participate in the study?

If you agree to participate you will be asked to complete the tasks outlined below:

1) Consent Form

You will be asked to complete the consent form. The purpose of the consent form is to inform you of your rights as a participant in the study, and of the procedures to be done during the experiment. If you would not wish to sign the consent form or would like to discontinue participating in the study at any time, you may withdraw from the experiment.

2) Short questionnaire

You will be asked to fill out a short questionnaire. The purpose of the questionnaire is to examine your gender, age, major, and in particular, your and your parent's language background. If you do not want to answer any of the questions or do not have the information to answer any of the questions, you may skip the questions and continue the study.

3) Hearing/dialect screening

1. Once you have completed the short questionnaire, you will be invited into the recording booth and have your hearing screened. During the screening, the experimenter will ask you a few basic questions about your hearing first. Then the experimenter will place headphones on you, and instruct you to raise your hand when you hear sounds in either your left or right ears. The purpose of the hearing screening it to ensure you meet the criteria to participate in the study. If you do not meet the hearing screening criteria, you will be discontinued from the study and referred to the Norris Health Center or to Community Audiology Services for further evaluation of your hearing. You will be given business cards with additional contact information so you can set up an appointment.

2. You will also be asked to participate in the dialect screening. The purpose of the dialect screening is to examine if you speak using a dialect from the Midwest or from elsewhere in the United States. You will be asked to read various words and sentences out loud to the experimenter. If you do not pass the dialect screening, you will be discontinued from the study.

4) Recording

After completing the short questionnaire and the hearing and dialect screening, you will be seated in front of a computer, and asked to produce a set of target sentences at a comfortable rate after receiving instructions on the computer. During the production of target sentences, both audio and ultrasound video recordings will be collected. The audio recording and ultrasound video are necessary to do analysis for this study. The ultrasound is a safe and non-invasive procedure that allows us to record the tongue motions during speech production. To collect ultrasound data, a small amount (less than a teaspoon) of ultrasound transmission gel will be applied to the ultrasound transducer, which will be placed under your chin. During the experiment, you will be asked to wear a helmet that is specifically designed to stabilize your head and the ultrasound transducer. If you are not comfortable being recorded and/or wearing the helmet, you can choose not to wear it and withdraw from the study**.**

5) Debriefing and compensation

After data collection is complete, you will leave the recording booth. Then, we will describe the experiment to you, outlining the main goals and hypothesis for data collection. You will be given the opportunity to ask questions. Additional information will be gladly provided at your request.

At the conclusion of the debriefing, you will receive $15 for full participation in this study. You will be asked to provide your name, and the amount paid on a separate form. All data will be kept confidential in a locked cabinet that only the primary investigator, Jae Yung Song, PhD has access to.

4. Risks and Minimizing Risks

What risks will I face by participating in this study?

There are no foreseeable risks (i.e., physical, psychological, and/or social) for this activity. You may experience slight discomfort or fatigue when applying ultrasound transmission gel and/or wearing a helmet during the speech production experiment. The ultrasound gel is non-toxic and non-fragrant, and only a small amount (less than a teaspoon) will be applied in each experiment. The weight to the helmet is 1.8 lbs. and it has been widely used in the field of ultrasound studies. There are no known risks in using either the gel or the helmet. This part of experiment will take only about 20 minutes, however, you will be allowed to take a rest or stop the experiment at any time if you want to.

5. Benefits

Will I receive any benefit from my participation in this study?

There are no benefits to you other than to further research.

6. Study Costs and Compensation

Will I be charged anything for participating in this study?

You will not be responsible for any of the costs from taking part in this research study.

Are subjects paid or given anything for being in the study?

You will be paid $15 after completing this study. If you do not pass the hearing or dialect screening, or withdraw before completing the study entirely, you will be paid $10 for your time.

7. Confidentiality

What happens to the information collected?

All information collected about you during the course of this study will be kept confidential to the extent permitted by law. After completing the experiment, each participant's name will be replaced with codes that do not contain any identifying information. The participants' codes will be used when analyzing and reporting data. The names and the matching codes will be kept on hard-copy papers and they will be stored in a locked cabinet that only the primary investigator, Jae Yung Song, PhD has access to. The audio and video data, which do not contain any identifying information, will be stored in the research network drive at the College of Heath Sciences for 3 years, maximum, for future use.

We may decide to present what we find to others, or publish our results in scientific journals or at scientific conferences. Information that identifies you personally will not be released without your written permission. Only the PI will have access to the information. However, the College of Health Sciences at UWM, the Institutional Review Board at UW-Milwaukee or appropriate federal agencies like the Office for Human Research Protections may review this study's records.

8. Alternatives

Are there alternatives to participating in the study?

There are no known alternative treatments available to you other than not taking part in this study.

9. Voluntary Participation and Withdrawal

What happens if I decide not to be in this study?

Your participation in this study is entirely voluntary. You may choose not to take part in this study. If you decide to take part, you can change your mind later and withdraw from the study.

You are free to skip any questions you cannot provide an answer to without penalty. If you withdraw or are discontinued before completing the study, you will be paid $10. Your decision will not change any present or future relationships with the University of Wisconsin Milwaukee.

10. Questions

Who do I contact for questions about this study?

For more information about the study or the study procedures or treatments, or to withdraw from the study, contact:

Jae Yung Song
Department of Communication Sciences and Disorders
University of Wisconsin-Milwaukee
2400 E. Hartford Ave.
Milwaukee, WI 53211
(414) 229-2665

Who do I contact for questions about my rights or complaints towards my treatment as a research subject?

The Institutional Review Board may ask your name, but all complaints are kept in confidence.

Institutional Review Board
Human Research Protection Program
Department of University Safety and Assurances
University of Wisconsin – Milwaukee
P.O. Box 413
Milwaukee, WI 53201
(414) 229-3173

11. Signatures

Research Subject's Consent to Participate in Research:

To voluntarily agree to take part in this study, you must sign on the line below. If you choose to take part in this study, you may withdraw at any time. You are not giving up any of your legal rights by signing this form. Your signature below indicates that you have read or had read to you *this entire consent form, including the risks and benefits, and have had all of your questions answered, and that you are 18 years of age or older.*

____

Printed Name of Subject/ Legally Authorized Representative

____ ______________________

Signature of Subject/Legally Authorized Representative Date

Research Subject’s Consent to Audio/Video/Photo Recording:

It is okay to audiotape/videotape me while I am in this study and use my audiotaped/videotaped data in the research.

Please initial: ____Yes ____No

Principal Investigator (or Designee)

I have given this research subject information on the study that is accurate and sufficient for the subject to fully understand the nature, risks and benefits of the study.

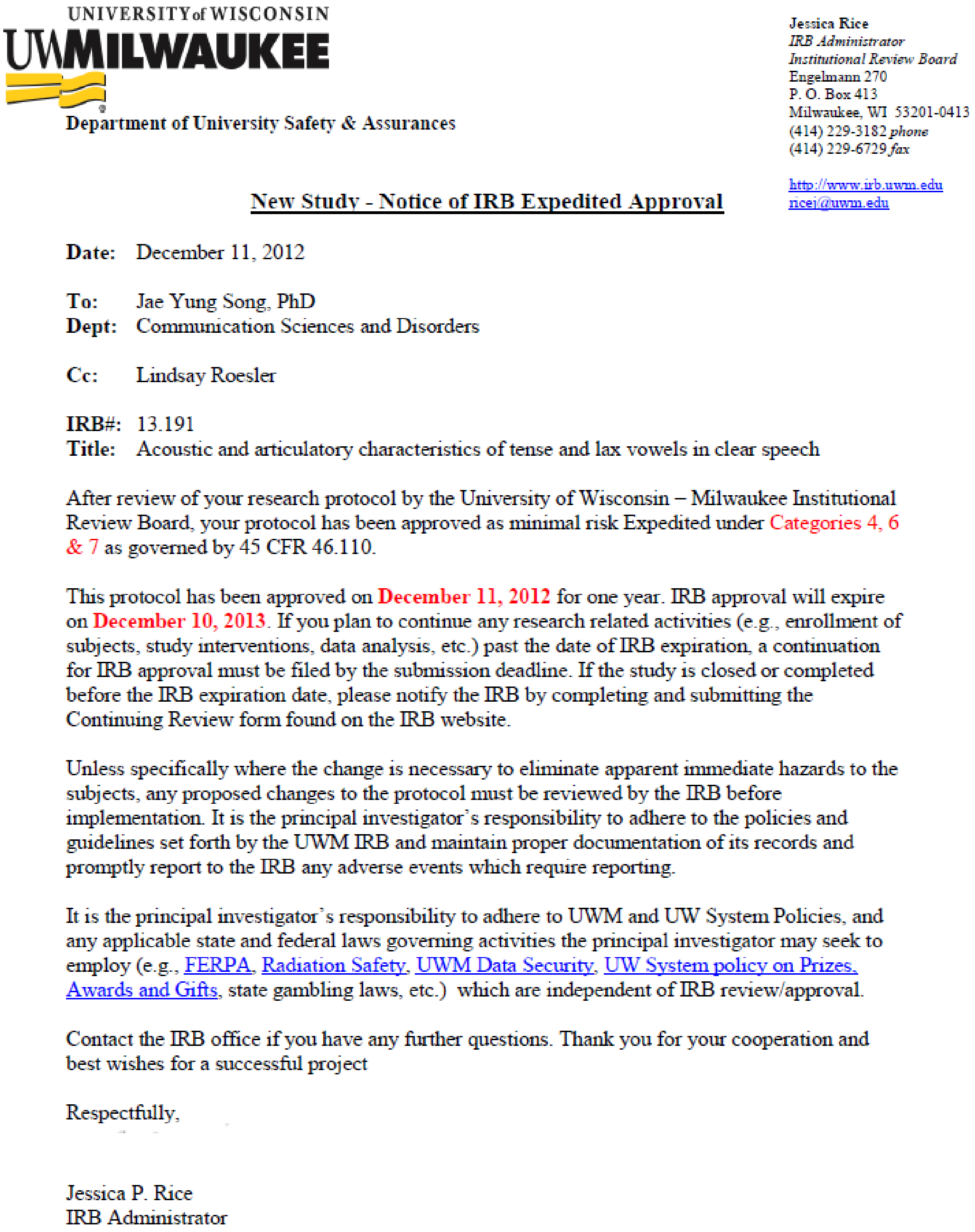
____ ______________________

Printed Name of Person Obtaining Consent Study Role

____ ______________________

Signature of Person Obtaining Consent Date

Appendix E: IRB approval



UNIVERSITY of WISCONSIN
UWMILWAUKEE
Department of University Safety & Assurances

Jessica Rice
IRB Administrator
Institutional Review Board
Engelmann 270
P. O. Box 413
Milwaukee, WI 53201-0413
(414) 229-3182 *phone*
(414) 229-6729 *fax*

http://www.irb.uwm.edu
ricej@uwm.edu

New Study - Notice of IRB Expedited Approval

Date: December 11, 2012

To: Jae Yung Song, PhD
Dept: Communication Sciences and Disorders

Cc: Lindsay Roesler

IRB#: 13.191
Title: Acoustic and articulatory characteristics of tense and lax vowels in clear speech

After review of your research protocol by the University of Wisconsin – Milwaukee Institutional Review Board, your protocol has been approved as minimal risk Expedited under Categories 4, 6 & 7 as governed by 45 CFR 46.110.

This protocol has been approved on **December 11, 2012** for one year. IRB approval will expire on **December 10, 2013**. If you plan to continue any research related activities (e.g., enrollment of subjects, study interventions, data analysis, etc.) past the date of IRB expiration, a continuation for IRB approval must be filed by the submission deadline. If the study is closed or completed before the IRB expiration date, please notify the IRB by completing and submitting the Continuing Review form found on the IRB website.

Unless specifically where the change is necessary to eliminate apparent immediate hazards to the subjects, any proposed changes to the protocol must be reviewed by the IRB before implementation. It is the principal investigator's responsibility to adhere to the policies and guidelines set forth by the UWM IRB and maintain proper documentation of its records and promptly report to the IRB any adverse events which require reporting.

It is the principal investigator's responsibility to adhere to UWM and UW System Policies, and any applicable state and federal laws governing activities the principal investigator may seek to employ (e.g., FERPA, Radiation Safety, UWM Data Security, UW System policy on Prizes, Awards and Gifts, state gambling laws, etc.) which are independent of IRB review/approval.

Contact the IRB office if you have any further questions. Thank you for your cooperation and best wishes for a successful project

Respectfully,

Jessica P. Rice
IRB Administrator

Appendix F: Short Questionnaire

Questionnaire

Part 1. (To be completed by participants prior to experiments)

Please answer the following questions.

1. Gender: Male Female

2. Age:

3. Major/Degree:

4. Where did you grow up (city, state)? List all places you have lived for at least 2 years.

5. Primarily, where did your guardian(s) grow up (city, state)?

6. What language(s) do you speak?

7. Do you have any history of speech, hearing, or language disorders? Y N
If yes, please describe: ___________________________________

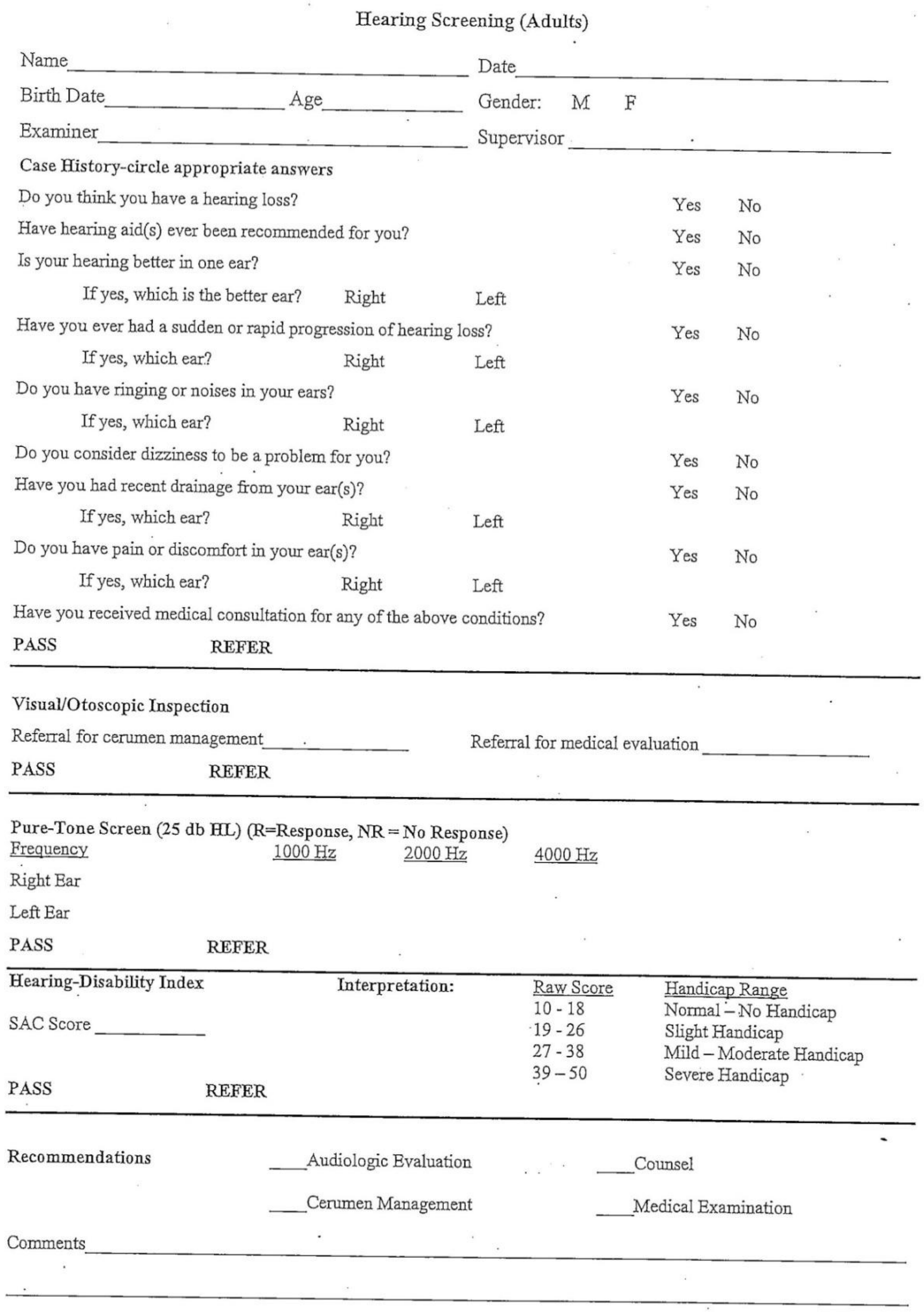
Part 2 (To be completed by the experimenter after experiments)

1. Date and time:

2. Experimenter's name:

3. Subject code:

Appendix G: Hearing Screening



Hearing Screening (Adults)

Name____________________________________ Date______________________________

Birth Date________________ Age____________ Gender: M F

Examiner________________________________ Supervisor __________________________

Case History-circle appropriate answers

Do you think you have a hearing loss?			Yes	No
Have hearing aid(s) ever been recommended for you?			Yes	No
Is your hearing better in one ear?			Yes	No
If yes, which is the better ear?	Right	Left		
Have you ever had a sudden or rapid progression of hearing loss?			Yes	No
If yes, which ear?	Right	Left		
Do you have ringing or noises in your ears?			Yes	No
If yes, which ear?	Right	Left		
Do you consider dizziness to be a problem for you?			Yes	No
Have you had recent drainage from your ear(s)?			Yes	No
If yes, which ear?	Right	Left		
Do you have pain or discomfort in your ear(s)?			Yes	No
If yes, which ear?	Right	Left		
Have you received medical consultation for any of the above conditions?			Yes	No

PASS **REFER**

Visual/Otoscopic Inspection

Referral for cerumen management______________ Referral for medical evaluation ________________

PASS **REFER**

Pure-Tone Screen (25 db HL) (R=Response, NR = No Response)

Frequency	1000 Hz	2000 Hz	4000 Hz
Right Ear			
Left Ear			

PASS **REFER**

Hearing-Disability Index

SAC Score __________

Interpretation:

Raw Score	Handicap Range
10 - 18	Normal – No Handicap
19 - 26	Slight Handicap
27 - 38	Mild – Moderate Handicap
39 – 50	Severe Handicap

PASS **REFER**

Recommendations

____Audiologic Evaluation
____Counsel
____Cerumen Management
____Medical Examination

Comments____

____

Appendix H: Dialect Screening

Dialect Screening

Please read the following lines of words out loud to the experimenter. When you read them, use your normal speaking voice. The experimenter will tell you when to read the next line. Do you have any questions?

1. father - brother
2. Mary - merry - marry
3. Florida
4. pajamas
5. cot - caught
6. log - fog - dog
7. high fives
8. park the car.
9. about town

Appendix I: Dialect Screening Data Sheet

Standard Midwestern American English Dialect Screening Protocol

Stimuli	**Target**	**Actual Performance**
Father-bother	Vowels sound the same.	
Mary – merry - marry	Vowels sound the same.	
Florida	First vowel is /ɔ/	
Pajamas	Second vowel is /æ/	
Cot - caught	Vowels sound different.	
Log – fog – dog	Vowels sound the same.	
High fives	Vowels are both /aI/	
Park the car.	/r/ is present in both words.	
About town.	Vowels are both /aʊ/	

Based on: Vaux, B. (2000-2005). Harvard dialect study. Retrieved from http://dialect.redlog.net/index.html. April 25, 2012.

Appendix J: Stimuli

Practice Sentences:

1. The boy are his dinner quickly.
2. His friend told him it was dangerous.
3. My teacher gave us our grade.
4. Jack filled his cup to the top.

Stimuli:

1. The *heat* made the corn *pop*.
2. The *sick* boy received a *shot*.
3. The poor *pet* was really *sick*.
4. Fried *fat* can help us *cook*.
5. The handsome *suit* was for *Chuck*.
6. My mom *took* the boarding *pass*.
7. His *cut* was the topic of the *chat*.
8. The *cop* slammed the cell door *shut*.
9. Her *sheet* had a picture of a *boot*.
10. My car seat *fit* in the *bus*.
11. That *debt* was hers to *keep*.
12. The *tap* is getting very *hot*.
13. I stuck my *hook* in the *soup.*
14. I *shut* the door with a swift *kick*.
15. The *pop* made me look at the *jet*.
16. He will *keep* his copy of the *book*.
17. A hard *kick* will hurt the *sheep*.
18. The *jet* pilot had a son named *Took*.
19. The *duke* does not drink from the *tap*.
20. Give the *pass* to the *duke*.
21. The brown *boot* was really *fat.*
22. I drove the *bus* past a *cop*.
23. The *hot* coal fell on the *deck*.
24. The *sheep* is the family *pet*.
25. Her useful *tip* let him off the *hook*.
26. The *deck* was very poorly *cut*.
27. Her *chat* was about the national *debt*.
28. The *soup* had lost all its *heat*.
29. My *book* had a signature *sheet*.
30. My uncle *Chuck* is very *fit*.
31. The robber *shot* a hole in the *suit*.
32. The *cook* received a handsome *tip*.

Filler sentences:

1. The doll came with a toy van.
2. The lease was up in June.
3. My 3 year-old niece can read.
4. She wore rouge in the scene.
5. The man made a bad deal.
6. The fool was really tan.
7. The lamb was painted teal.
8. I watched the news about the loon