Acoustics of Clear Speech: Effect of Instruction

Jennifer Lam[a], **Kris Tjaden**[a], and **Greg Wilding**[a]

[a]University at Buffalo, NY

Abstract

Purpose—This study investigated how different instructions for eliciting clear speech affected selected acoustic measures of speech.

Method—Twelve speakers were audio-recorded reading 18 different sentences from the Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1984). Sentences were produced in habitual, clear, hearing impaired, and overenunciate conditions. A variety of acoustic measures were obtained.

Results—Relative to habitual, the clear, hearing impaired, and overenunciate conditions were associated with different magnitudes of acoustic change for measures of vowel production, speech timing, and vocal intensity. The overenunciate condition tended to yield the greatest magnitude of change in vowel spectral measures and speech timing, followed by the hearing impaired and clear conditions. SPL tended to be the greatest in the hearing impaired condition for half of the speakers studied.

Conclusions—Different instructions for eliciting clear speech yielded acoustic adjustments of varying magnitude. Results have implications for direct comparison of studies using different instructions for eliciting clear speech. Results also have implications for optimizing clear speech training programs.

Keywords

clear speech; instructions; acoustics

Speech clarity has been a topic of significant research interest over the past few decades. Following Smiljani and Bradlow (2009), the term *clear speech* refers to a speaking style wherein talkers voluntarily adjust their conversational or habitual speech to maximize intelligibility for a communication partner. Talkers may use clear speech in a variety of situations, such as when conversing in a noisy environment or when talking to someone who has a hearing impairment. Clear speech is further recommended for use as a therapy technique for speakers with dysarthria (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Dromey, 2000; Hustad & Weismer, 2007), for use in military aviation training programs (Huttunen, Keränen, Väyrynen, Pääkkönen, & Leino, 2011), and for use in aural rehabilitation training programs (Schum, 1997). Identifying variables responsible for the improved intelligibility typically associated with clear speech also may help to improve

Correspondence to: Jennifer Lam: jkyun@buffalo.edu.

signal processing technologies for auditory prostheses (i.e., Krause & Braida, 2009; Picheny, Durlach, & Braida, 1985, 1986; Uchanski, Choi, Braida, & Reed, 1996; Zeng & Liu, 2006). Thus, clear speech research has real-world implications for optimizing clear speech training programs and for developing technologies for use with populations who have difficulty understanding speech. In addition, the construct of clear speech figures prominently in certain speech production theories (Lindblom, 1990; Perkell, Zandipour, Matthies, & Lane, 2002).

The increase in intelligibility from conversational to clear speech has been termed the *clear speech benefit,* and the magnitude of this effect is known to vary widely across studies. That is, with the exception of a few investigations that used listeners with particular kinds of hearing impairment or simulated hearing impairment, the average clear speech benefit in a given study reportedly ranges from 3% to 34% (Smiljani & Bradlow, 2009; Uchanski, 2005). Variables contributing to cross-study variation in the overall clear speech benefit are not well understood. As noted in a following section, differences in instruction or cues for eliciting clear speech may be a contributing factor.

In a variety of studies, researchers have sought to determine the acoustic basis of the improved intelligibility typical of clear speech. The following review highlights findings from studies reporting acoustic measures that were of interest to the present investigation, including vowel formant frequency characteristics, speech durations, and vocal intensity. More comprehensive reviews may be found in Smiljani and Bradlow (2009) as well as Uchanski (2005).

Vowel space area, calculated from midpoint or steady-state values of F1 and F2, indexes the size of an individual's vowel articulatory–acoustic working space (Turner, Tjaden, & Weismer, 1995). The relationship between vowel space and intelligibility for neurologically normal talkers is suggested by studies reporting larger vowel space areas for talkers who are inherently more intelligible in conversational speech (Bradlow, Torretta, & Pisoni, 1996; Hazan & Markham, 2004). Some dysarthria studies have also suggested a relationship between vowel space area and intelligibility (e.g., Hustad & Lee, 2008; H. Liu, Tsao, & Kuhl, 2005; Tjaden & Wilding, 2004). An expanded vowel space area for clear versus conversational speech has been widely reported (Bradlow, Krause, & Hayes, 2003; Ferguson & Kewley-Port, 2002, 2007; Johnson, Flemming, & Wright, 1993; Moon & Lindblom, 1994; Picheny, Durlach, & Braida, 1986). There are exceptions, however. Goberman and Elmer (2005) did not find significant vowel space expansion for clear speech produced by speakers with dysarthria secondary to Parkinson's disease. Krause and Braida (2004) also reported minimal vowel space area expansion when neurologically normal talkers were trained to produce clear speech at normal or conversational speech rates. Picheny et al. (1986) further suggested a tendency for midpoint formant frequencies for lax vowels, and the derived lax vowel space area, to change more dramatically for clear speech as compared with tense vowels, although this trend has not been strongly supported in more recent studies (Krause & Braida, 2004).

Ferguson and Kewley-Port (2007) used a novel approach to investigate acoustic changes likely responsible for improved intelligibility of clearly produced /bVd/ words. Static and

dynamic spectral measures of vowel production for speakers exhibiting a large clear speech benefit (big-benefit talkers) were contrasted with measures for speakers exhibiting no clear speech benefit (no-benefit talkers). Clear speech for big-benefit talkers was characterized by greater vowel space area expansion, a greater overall increase in F1, and higher F2 values for front vowels when compared with clear speech produced by the no-benefit talkers. Vowel formant dynamics did not differ for the big-benefit and no-benefit talkers, nor did Tasko and Greilick (2010) find clarity-related differences in F1 or F2 diphthong transitions. However, other studies have reported that vowel formants are more dynamic in clear speech as compared with conversational speech (Ferguson & Kewley-Port, 2002; Moon & Lindblom, 1994; Wouters & Macon, 2002). Perceptual studies using speech resynthesis have also suggested the importance of dynamic spectral cues for accurate vowel identification (e.g., Assmann & Katz, 2005; Hillenbrand & Nearey, 1999).

Slowed speaking rate is probably the most widely reported characteristics of clear speech (Bradlow, Krause, & Hayes, 2003; Picheny et al., 1986; Smiljani & Bradlow, 2008). The slower than normal speaking rate typical of clear speech further tends to be associated with a reduced articulation rate, increased segment durations, and longer, more frequent pauses (Bradlow et al., 2003; Picheny et al., 1986; Uchanski et al., 1996). Although many studies have reported a slowed speech rate for clear speech, speakers can be trained to produce clear speech at a normal or conversational speech rate (Krause & Braida, 2004). This observation might be taken to suggest that a reduced rate, lengthened segment durations, and longer, more frequent pauses do not contribute to the improved intelligibility typically associated with clear speech. However, the fact that Krause and Braida (2004) found that the increase in intelligibility for clear speech produced at a normal speech rate was not as great as the increase in intelligibility for clear speech produced at a typically slower than normal rate suggests the contribution of temporal adjustments to the clear speech benefit. Other studies have also suggested that an increased number of pauses accompanying clear speech benefits intelligibility (S. Liu & Zeng, 2006).

Finally, clear speech studies reporting vocal intensity measures have found that clear speech is generally accompanied by at least some increase in vocal intensity (e.g., Dromey, 2000; Goberman & Elmer, 2005; Picheny et al., 1986). As noted by Uchanski (2005), the contribution of increased vocal intensity to the clear speech benefit is not attributable to improved audibility, as the relevant perceptual studies have equated clear and conversational speech for overall SPL. Rather, other speech production changes accompanying an increased vocal intensity, such as the use of more canonical vocal tract shapes, may help to explain improvements in intelligibility.

Although research conducted over the past few decades has made a great deal of progress in advancing our knowledge of clear speech, many important questions still need to be addressed. One is the extent to which the instruction or cue for eliciting clear speech impacts both production and perception (Smiljani & Bradlow, 2009). To date, clear speech has been elicited using a variety of instructions or cues, including "speak clearly" (Ferguson, 2004; Ferguson & Kewley-Port, 2007), "hyperarticulate" or "overenunciate" (Dromey, 2000; Moon & Lindblom, 1994), or "speak to someone with a hearing impairment or non-native speaker" (Bradlow et al., 2003; Smiljani & Bradlow, 2008). Other studies have used a

combination of these instructions (e.g., Picheny et al., 1985; Rogers, DeMasi, & Krause, 2010; Tasko & Greilick, 2010). Studies comparing different cues for increasing vocal loudness (e.g., Darling & Huber, 2011; Huber & Darling, 2011; Sadagopan & Huber, 2007) and studies investigating different cues or instructions for slowing speech rate (Van Nuffelen, De Bodt, Vanderwegen, & Wuyts, 2010) have suggested that speech output differs depending on the nature of the instruction or cueing provided to talkers. It seems reasonable to hypothesize that clear speech output also may differ depending on the cue or instruction used to elicit this speech style.

The question of whether different instructions or cues for eliciting clear speech result in equivalent articulatory–acoustic and perceptual changes is nontrivial. Differences in clear speech instruction could help to explain cross-study variation in production characteristics as well as cross-study differences in the magnitude of the clear speech benefit. Knowing the speech production adjustments associated with a particular clear speech instruction further has the potential to assist in optimizing clear speech training programs for dysarthria or aural rehabilitation. As an initial step in addressing this multifaceted issue, the current study examined the impact of three clear speech instructions on selected acoustic measures of speech for a group of healthy adult talkers. Instructions included "speak clearly" (clear condition), "overenunciate" (overenunciate condition), and "talk to someone with a hearing impairment" (hearing impaired condition). These cues or instructions were selected for several reasons. First, all instructions have been used in separate, previously published studies and are reasonably distinct from one another (Bradlow et al., 2003; Dromey, 2000; Ferguson, 2004; Johnson et al., 1993; Moon & Lindblom, 1994). In addition, these cues are probable candidates for use in clear speech training programs, and it was of interest to determine whether instructions likely to be used therapeutically would elicit similar types or magnitudes of acoustic change.

Method

Participants

A total of 12 neurologically healthy speakers (six men and six women) ranging from 18 to 36 years of age ($M = 24$, $SD = 6$ years) were recruited from the student population at the University at Buffalo to serve as participants. Participants were judged to speak Standard American English as a first language and reported no history of hearing, speech, or language pathology. All speakers passed a pure tone hearing screening, administered bilaterally at 20 dB HL at 500, 1000, 2000, 4000, and 8000 Hz (American National Standards Institute [ANSI], 1969). In addition, no speaker had received university training in linguistics or communicative disorders and sciences.

Speech Sample

For each of the 12 speakers, 18 different sentences, varying in length from five to 11 words, were selected from the Assessment of Intelligibility of Dysarthric Speech (AIDS; Yorkston & Beukelman, 1984). Sentences were audio-recorded as part of a larger speech sample, which also included /hVd/ words. Each sentence set was selected to include three to five occurrences of the tense vowels /ɑ, i, æ, u/ and the lax vowels /ɪ, ɛ, ʊ, ʌ/. Vowels also were

selected to occur in content words and in syllables receiving primary stress. In addition to measures of duration and intensity discussed below, spectral characteristics of vowels were of primary interest, following Ferguson and Kewley-Port (2007).

Speakers read the sentences in four conditions, including, habitual, clear, hearing impaired, and overenunciate. These conditions were selected for several reasons. First, other clear speech studies have elicited adjustments in clarity by instructing speakers to "overenunciate/ hyperarticulate" (overenunciate condition), "talk to someone with a hearing loss" (hearing impaired condition) or "speak clearly" (clear condition) (e.g., Bradlow et al., 2003; Ferguson, 2004; Johnson et al., 1993; Moon & Lindblom, 1994; Smiljani & Bradlow, 2008). When selecting instructions from among those utilized in previous studies, care also was taken to choose distinctive instructions. Thus, only "talk to someone with a hearing impairment" but not "talk to a non-native speaker" was included for study. In addition, these cues were deemed to be reasonable candidates for use in clear speech training programs, especially "overenunciate" and "speak clearly," and it would seem important to know whether training program instructions would elicit similar types or magnitudes of acoustic change.

For the habitual condition, speakers were simply instructed to read the sentences. Thus, the habitual condition is similar to "conversational" or "citation" speech in other clear speech studies. Speakers subsequently were asked to read the sentences "while speaking clearly" (clear), allowing each participant to determine what this phrase meant (see Ferguson, 2004). Speakers also were instructed to "speak as if speaking to someone who has a hearing impairment" (hearing impaired condition) and to "overenunciate each word" (overenunciate condition). The order of these latter two conditions was alternated among participants such that the order of conditions for six participants was habitual, clear, hearing impaired, and overenunciate, and the order of conditions for the remaining six participants was habitual, clear, over-enunciate, and hearing impaired. Participants were engaged in casual conversation for a few minutes between recordings for each condition to minimize carry-over effects. The clear condition was always recorded immediately following the habitual condition. This order of conditions was followed so that speakers could interpret the instructions to "speak clearly" without the prior influence of hearing the instructions to "overenunciate" or "talk to someone with a hearing impairment."

Procedure

Data collection took place in a sound-attenuated booth. A head-mounted CountryMan E6IOP5L2 Isomax condenser microphone was used to record the acoustic signal. A mouth-to-microphone distance of 6 cm was maintained throughout the recording for each speaker. The microphone signal was preamplified using a Professional Tube MIC Preamp, low-pass filtered at 9.8 kHz and digitized to a computer at a sampling rate of 22 kHz using TF32 (Milenkovic, 2002). Prior to recording each participant, a calibration tone of known intensity was also recorded. The calibration tone was later used to calculate vocal intensity from the acoustic speech signal (see Tjaden & Wilding, 2004). Stimuli were presented using Microsoft PowerPoint and were read from a computer screen. Written instructions for each condition were presented both visually on the computer screen and verbally at the beginning

of recording, as well as a quarter, half, and three quarters of the way throughout each condition. For each speaker, four random orderings of sentences were created such that sentences were produced in a different order for the four speaking conditions.

Prior to performing the acoustic analysis, audio files were recoded by a research assistant not involved in the study so that the investigators performing the acoustic measures were blinded to the identity of each speaking condition. The purpose of this procedure was to help control for possible experimenter bias when performing the acoustic analysis.

Table 1 summarizes acoustic measures of interest. These measures were selected from among those reported in previous clear speech studies (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002, 2007; Moon & Lindblom, 1994; Picheny et al., 1986). Because it was expected that the various clear speech instructions would elicit articulatory adjustments, as inferred from the acoustic signal, and possibly different magnitudes of adjustment, a measure of segmental articulation was included. The current focus on vowels is consistent with that of Ferguson and Kewley-Port (2007). Similarly, a measure of vocal intensity was obtained because it was speculated that the instruction to "speak to someone with a hearing impairment" might elicit a relatively greater adjustment in vocal intensity than other instructions. Finally, global speech timing was included for study because a reduced speaking rate has been so widely reported in clear speech research. Given that the instructions selected for study might be used in clear speech training programs for dysarthria, and that speaking rate reduction is often desirable in the treatment of dysarthria, it was of interest to determine whether the various clear speech instructions would elicit different magnitudes of speaking rate adjustment. Each of the acoustic measures is discussed in more detail below. All acoustic analyses were performed using TF32 (Milenkovic, 2002).

Segmental acoustic measures—Vowel segment durations were obtained using a combination of the waveform and wideband (300–400 Hz) digital spectrographic displays. Vowel duration was computed from the first glottal pulse of the vocalic nucleus to the last glottal pulse as indicated by energy in both F1 and F2 (Tjaden, Rivera, Wilding, & Turner, 2005; Turner et al., 1995). For each speaker and condition, vowel segment durations were averaged for tense (ɑ, i, æ, u) and lax (ɪ, ɛ, ʊ, ʌ) vowel categories. These averages were subsequently used in the statistical analyses for segment durations.

Linear predictive coding generated formant trajectories for F1 and F2 also were computed over the entire duration of each vowel of interest. Computer-generated tracking errors were inspected for errors and manually corrected as needed. Using Microsoft Excel, F1 and F2 values were extracted at three time points corresponding to 20%, 50%, and 80% of vowel duration (Hillenbrand, Getty, Clark, & Wheeler, 1995). Average midpoint formant frequencies (50% point) for each speaker and condition were used to calculate a variety of static measures of vowel production, including vowel space area, F1 range, F2 range, intravowel distance, and tense–lax spectral distance.

Using Heron's formula (see Goberman & Elmer, 2005; Turner et al., 1995), a separate vowel space area was calculated for tense vowels /ɑ, i, æ, u/ and lax vowels /ɡ, ɛ, ʊ, ʌ/, yielding a tense vowel space area and lax vowel space area for each speaker and condition for use in

the statistical analysis. Because tongue height and tongue advancement broadly correspond to adjustments in F1 and F2, respectively (Kent & Read, 2002), F1 range and F2 range also were calculated to supplement measures of vowel space area. Following procedures used by Ferguson and Kewley-Port (2007), F1 and F2 ranges were computed using midpoint formant frequencies from the four tense vowels that form the extremes or border of the articulatory–acoustic working space /ɑ, i, æ, u/. F1 range was determined by calculating the absolute difference between average F1 values for /i/ and /u/, and average F1 values for /æ/ and /ɑ/. Similarly, F2 range was determined by calculating the absolute difference between average F2 values for /i/ and /æ/, and average F2 values for /ɑ/ and /u/. Separate F1 and F2 ranges were calculated for each speaker and condition for use in the statistical analysis.

As illustrated in Figure 1, a measure termed *intra-vowel distance* (ID) was calculated to index the amount of change in F1 × F2 space for a given vowel in the clear, hearing impaired, and overenunciate conditions relative to the habitual condition. Relatively greater IDs suggest that a given vowel is more amenable to condition effects. Thus, the ID measure is not unlike the Euclidean distance from habitual centroid measure reported by Turner et al. (1995). As shown schematically in Figure 1, ID was determined by calculating the length of the line between a vowel's habitual F1 × F2 coordinate and the corresponding F1 × F2 coordinate for that vowel in the clear, hearing impaired, and overenunciate conditions. In Figure 1, the habitual F1 × F2 coordinates for /u/ and /ʊ/ are indicated with open circles labeled with the appropriate phonetic symbol. Squares in Figure 1 correspond to F1 × F2 coordinates for each vowel in the clear, hearing impaired, and overenunciate conditions. Solid lines represent IDs. Thus, for each speaker and vowel, three ID measures were obtained. IDs were calculated to describe vowels contributing to adjustments in vowel space area and were not included in the parametric statistical analysis.

Previous studies have suggested that clear speech impacts dynamic characteristics of vowel production (Ferguson & Kewley-Port, 2002; Moon & Lindblom, 1994; Wouters & Macon, 2002). Thus, following procedures established by Ferguson and Kewley-Port (2007), formant frequency values at the 20% and 80% points were used to calculate the dynamic measure, lambda (λ), using the formula $|F1_{80} - F1_{20}| + |F2_{80} - F2_{20}|$. For each speaker and condition, lambda measures were calculated for each vowel and then averaged across tense and lax vowel categories, yielding a mean lambda measure for tense vowels and a mean lambda measure for lax vowels for use in the statistical analysis.

The measure tense–lax spectral distance between the four vowel pairs, /æ-ɛ/, /i-ɪ/, /u-ʊ/ and /ɑ-ʌ/, was calculated to provide an index of the relative spectral distinctiveness of tense–lax vowel pairs. This measure is illustrated in Figure 1 by the dashed line, which indicates the Euclidean distance in F1 × F2 space for the tense–lax vowel pair /u/ and /ʊ/. Several additional measures also were derived to further characterize the degree of acoustic contrast for the four tense–lax vowel pairs. These measures included duration ratios and lambda ratios. These ratios were calculated separately for each of the four tense–lax vowel pairs for use in the statistical analysis. Thus, average lax durations were divided by average tense durations, yielding four duration ratios per speaker and condition. Similarly, average lax lambda values were divided by average tense lambda values, yielding four lambda ratios per speaker and condition.

Global timing—Measures of total utterance duration, run duration, and syllable counts were obtained for use in calculating speaking rate and articulation rate in syllables per second. Pause durations and pause frequencies were also measured. A *pause* was defined as a silent period of 200 ms or greater between words (Turner & Weismer, 1993).

Standard acoustic criteria were used to identify onsets and offsets for total utterance durations and run durations. A *run* was defined as a stretch of speech separated by interword pauses (Turner & Weismer, 1993). Total utterance durations were calculated by summing all run durations and pauses within a sentence. Sentence or run onsets beginning with obstruents were defined as the left edge of the release burst or fricative noise (Klatt, 1975). The onset of a nasal was indicated by a reduction in intensity in the higher frequencies (Kent & Read, 2002). Onsets for vowels, liquid, and glides were defined as the left edge of the first glottal pulse, as indicated by energy in both F1 and F2. Similar criteria were applied when determining offsets.

Speaking rate for each sentence in syllables per second was calculated by dividing the number of syllables produced by total utterance duration (in milliseconds) and then multiplying by 1,000. For each sentence, articulatory rate (syllables per second) was calculated by dividing the number of syllables produced by the total duration of all runs (in milliseconds) composing a sentence and multiplying by 1,000 (Turner & Weismer, 1993). For each speaker and condition, measures of speaking rate, articulation rate, pause duration, and pause frequency were averaged across sentences for use in the statistical analysis.

Vocal intensity—Following procedures from previous studies, the average root-mean-square (RMS) voltage of each run was converted to dB SPL in reference to each speaker's calibration tone (Tjaden & Wilding, 2004). For each speaker and condition, a measure of average intensity was determined by averaging SPL across all runs. These averages were used in the statistical analysis.

Measurement Reliability

Acoustic measures were performed by the first author, with measurement questions resolved by conferral with the second author. The second author also performed ongoing accuracy checks for a randomly selected sample of approximately 10% of stimuli (i.e., $n = 2$ sentences) for each speaker and condition at the initial time of measurement. One set of sentences from each speaking condition (i.e., approximately 10% of stimuli) was selected from four different speakers for use in determining measurement reliability following the same procedures utilized for performing the initial measures. Absolute measurement errors and Pearson product–moment correlations were used to index reliability. The correlation between the first and second set of segment duration and run duration measures was 0.96 (M absolute difference measure = 0.008 s, $SD = 0.016$ s) and 0.99 (M absolute difference measure = 0.02 s, $SD = 0.08$ s), respectively. The correlation between the first and second set of spectral measures was 0.99 (M absolute difference measure = 20.0 Hz, $SD = 60.0$ Hz), and the correlation between the first and second set of SPLs was 0.95 (M absolute difference measure = 0.23 dB, $SD = 0.69$ dB).

Data Analysis

Descriptive statistics in the form of means and standard deviations were calculated for all acoustic measures. Parametric statistical analyses (i.e., analysis of variance) were performed for tense vowel space area, lax vowel space area, tense segment duration, lax segment duration, F1 range, F2 range, tense–lax spectral distance, SPL, lambda (λ), duration ratios, lambda ratios, speaking rate, and articulation rate. Pause frequency and pause duration data were characterized using descriptive statistics because pauses occurred fairly infrequently. Using SAS Version 9.1.3 statistical software, a multivariate linear model was fit to each dependent variable in this repeated measures design. For all dependent variables, models included the main effect of condition (habitual, clear, hearing impaired, overenunciate). The analyses for tense–lax spectral distances, lambda ratios, and duration ratios also included a vowel main effect as well as a Condition × Vowel interaction, as potential differences among the four tense–lax vowel pairs were of interest. To control for gender differences in dependent measures, a variable representing gender also was included in each analysis. To account for the within-subject dependence structure, each multivariate linear model assumed that the distribution of the error terms for each subject was multivariate normal with zero mean and an unstructured covariance structure (Brown & Prescott, 1999). Post hoc pairwise comparisons were performed using a Bonferroni correction for multiple comparisons. All tests were two-sided, and a significance level of $p < .05$ was used in all hypothesis testing. Finally, for convenience, the clear, hearing impaired, and overenunciate conditions are collectively referred to as the *nonhabitual conditions* throughout the Results and Discussion.

Results

Vowel Space Area

Figures 2 and 3 report average F1 × F2 coordinates and the associated vowel space areas for male and female speakers, respectively. The statistical analysis indicated a significant condition effect for both tense, $F(3, 10) = 28.65$, $p < .0001$, and lax, $F(3, 10) = 17.9$, $p = .0002$, vowel space areas. For tense vowels, post hoc comparisons indicated greater vowel space areas for the clear ($p = .0002$), hearing impaired ($p = .0007$), and overenunciate ($p < .0001$) conditions when compared with the habitual condition. Tense vowel space area in the overenunciate condition was also significantly greater than in the clear condition ($p = .001$). For lax vowels, post hoc comparisons indicated larger vowel space areas for clear ($p = .006$), hearing impaired ($p = .001$), and overenunciate ($p = .0003$) conditions compared with habitual.

For tense and lax vowels, the finding of greater vowel space area in the nonhabitual conditions as compared with habitual held for 12 and 10 speakers, respectively. In addition, the overenunciate condition was associated with the greatest tense and lax vowel space areas for 10 and nine speakers, respectively. The remaining speakers exhibited the greatest vowel space areas in the hearing impaired condition.

Because previous studies have suggested that lax vowels are more susceptible to the effects of speaking style than tense vowels (Picheny et al., 1985), the mean percent increase in vowel space area for the non-habitual conditions relative to the habitual condition was

calculated. As reported in Table 2, on average, lax vowels exhibited a greater percent change in vowel space area for the nonhabitual conditions when compared with the habitual condition, particularly for female speakers.

Figure 4 reports data for F1 and F2 range in the form of box and whiskers plots. Data for F1 are reported in the upper panel, and data for F2 are reported in the lower panel. For F1 range there was a significant effect of condition, $F(3, 10) = 15.97$, $p = .0004$. Post hoc analyses further indicated that F1 range was greater in all three non-habitual conditions as compared with habitual (p .008) but did not differ for pairs of nonhabitual conditions. This trend of a greater F1 range in the nonhabitual conditions relative to habitual held for 11 of 12 speakers. As illustrated in Figure 4, descriptive statistics further indicated that, on average, F1 range was greatest in the overenunciate condition, followed by the hearing impaired, clear, and habitual conditions. Results were similar for F2 range. That is, the statistical analysis indicated a main effect of condition, $F(3, 10) = 21.13$, $p = .0001$. Post hoc analyses further indicated that F2 range was greater in all three nonhabitual conditions as compared with habitual (p .001), and this trend held for 11 of 12 speakers. In addition, F2 range was significantly greater in the overenunciate condition versus the clear condition ($p = .001$). Finally, as illustrated in Figure 4, descriptive statistics indicated that on average, F2 range was greatest in the overenunciate condition, followed by the hearing impaired, clear, and habitual conditions.

Intravowel Distance

Table 3 summarizes findings for ID. This table reports the number of speakers exhibiting the greatest ID for a given vowel comprising the tense and lax vowel categories within each condition. Thus, the first row of Table 3 indicates that for the clear condition, five speakers exhibited the greatest ID for /i/, six speakers exhibited the greatest ID for /u/, one speaker exhibited the greatest ID for /æ/, and no speakers exhibited the greatest ID for /ɑ/. The implication is that in the clear condition, the tense vowels /i/ and /u/ were most amenable to adjustment in F1 × F2 space as compared with /æ/ and /ɑ/. Of the tense vowels, Table 3 further suggests that the majority of speakers produced the greatest IDs for /i, u/. For lax vowels, although not as robust, Table 3 suggests that speakers tended to produce the greatest IDs for /ɪ, ʊ/ within a given condition. Thus, these lax vowels tended to be more susceptible to condition effects than other lax vowels.

Vowel Spectral Change Measure (Lambda)

Descriptive statistics for lambda are reported in Table 4. The main effect of condition was significant for both tense, $F(3, 10) = 3.90$, $p = .04$, and lax, $F(3, 10) = 5.79$, $p = .01$, vowels. Post hoc comparisons for tense vowels were not significant. However, Table 4 indicates a trend for greater tense lambda values in the overenunciate and hearing impaired conditions as compared with the habitual and clear conditions. Moreover, six speakers had the greatest lambda values in the overenunciate condition, and four speakers had the greatest lambda values in the hearing impaired condition.

Post hoc tests for lax vowels indicated significantly greater lambda values, or more dynamic vowels, for the hearing impaired ($p = .02$) and overenunciate ($p = .04$) conditions when

compared with habitual, as well as for the overenunciate condition versus the clear condition ($p = .01$). Consistent with these post hoc analyses, eight speakers had the largest lax lambda values in the over-enunciate condition, and three speakers had the greatest lambda measures in the hearing impaired condition.

Segment Durations

Descriptive statistics for segment durations are reported in Table 4. The statistical analyses indicated a significant condition effect for both tense, $F(3, 10) = 41.84$, $p < .0001$, and lax, $F(3, 10) = 17.9$, $p < .0001$, vowel durations. For tense vowels, post hoc comparisons indicated that speakers significantly increased vowel durations for the hearing impaired ($p = .0006$) and over-enunciate ($p = .0002$) conditions when compared with the habitual condition. Tense vowels were also longer in the overenunciate condition when compared to the clear condition ($p < .0001$). Post hoc comparisons for lax vowels indicated longer vowels for the hearing impaired ($p = .0001$) and overenunciate ($p < .0001$) conditions as compared with the habitual condition. The over-enunciate ($p < .0001$) and hearing impaired ($p = .005$) conditions were also associated with longer vowels compared with the clear condition. Inspection of individual speaker data revealed that all but two speakers followed the trend of increased tense and lax vowel durations for the clear, hearing impaired, and overenunciate conditions relative to habitual.

Distinction of Tense–Lax Vowel Pairs

Figure 5 illustrates spectral distances for tense (filled circles) and lax (unfilled circles) vowel pairs for speaker M07 (upper panel) and M04 (lower panel). Greater distances or longer lines in Figure 5 indicate greater distinction of vowel pairs in F1 × F2 space. For each of the four vowel pairs, the statistical analysis indicated no significant difference among conditions. However, inspection of individual speaker data indicated that for /æ-ɛ/, /i-ɪ/, and /u-ʊ/, anywhere from six to eight speakers increased spectral distances for the nonhabitual conditions compared with habitual. For the /ɑ-ʌ/ tense–lax pair, however, nine of 12 speakers reduced tense–lax spectral distances in the clear condition, and five speakers increased tense–lax spectral distances in the hearing impaired and overenunciate conditions. Interestingly, for every vowel pair, M04—data for which are plotted in the lower panel of Figure 5—reduced all but one tense–lax spectral distance for the nonhabitual conditions compared with habitual.

Table 5 reports descriptive statistics for duration ratios and lambda ratios. Ratios approaching 1.0 indicate reduced distinction. For both measures, ratios did not differ significantly across conditions. Indeed, duration ratios in Table 5 for a given vowel pair vary only a small amount across conditions, although it is notable that average duration ratios for the tense–lax pair /ɑ-ʌ/ tend to approach 1.0 in the nonhabitual conditions, indicating relatively less temporal distinction, whereas duration ratios for the tense lax pair /i-ɪ/ suggest a trend toward increased temporal distinction in the nonhabitual conditions. Inspection of average lambda ratios in Table 5 further suggests a trend for all tense–lax vowel pairs to exhibit increased dynamic distinction in most of the non-habitual conditions, as lambda ratios tend to move away from 1.0 in the nonhabitual conditions relative to habitual. The main effect of vowel pair was significant for duration ratio, $F(3, 33) = 5.13$, $p = .005$, and

lambda ratio, $F(3, 33) = 13.67$, $p < .0001$. Although this main effect was not of particular interest, it serves to demonstrate that vowel pairs behaved differently with respect to ratio measures.

Articulation Rate and Speaking Rate

Table 6 reports descriptive statistics for global measures of speech timing. For articulation rate, the statistical analysis indicated a significant condition effect, $F(3, 10) = 78.46$, $p < .0001$. Post hoc comparisons indicated slower articulation rates for the clear ($p = .001$), hearing impaired ($p < .0001$), and overenunciate ($p < .0001$) conditions when compared with habitual. Articulation rates were also reduced for the overenunciate condition when compared with the clear ($p < .0001$) and hearing impaired ($p = .01$) conditions. Findings for speaking rate were identical to those for articulation rate. Inspection of individual speaker data further indicated that nine of 12 speakers followed the trend of having the fastest articulation and speaking rates in the habitual condition followed by the clear, hearing impaired, and overenunciate conditions. In addition, 11 of the 12 speakers had a slower articulation rate and speaking rate in the overenunciate condition when compared with the clear and hearing impaired conditions.

Pause Frequency and Duration

Pause frequency and pause duration data also are reported in Table 6, which indicates that speakers used the greatest number of pauses in the overenunciate condition. This trend held for nine of 12 speakers. Table 6 further indicates longer average pause durations in the nonhabitual conditions as compared with the habitual condition.

SPL

Figure 6 reports mean sentence-level SPL and *SD*s. The statistical analysis indicated a significant main effect of condition, $F(3, 10) = 3.80$, $p = .04$. No post hoc comparisons were significant. However, inspection of individual speaker data further indicated that six of 12 speakers increased mean vocal intensity by approximately 1–2 dB for the hearing impaired condition compared with the habitual condition, and three speakers increased mean vocal intensity by approximately 1–2 dB for the overenunciate condition relative to the habitual condition.

Discussion

Different clear speech instructions were associated with different magnitudes of acoustic change for measures of vowel production, speech timing, and vocal intensity. The instruction to "overenunciate" tended to elicit the greatest adjustments in vowel production and speech timing, followed by "speak to someone with a hearing impairment." For half of the speakers studied, this latter condition also elicited the greatest adjustments in vocal intensity. Finally, "speak clearly" elicited the smallest acoustic adjustments relative to the habitual condition. Results and their implications are considered below.

Static and Dynamic Measure of Vowel Production, Speech Timing, and SPL

In agreement with previous clear speech studies (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002, 2007; Moon & Lindblom, 1994; Picheny et al., 1985), speakers in the present study significantly increased tense and lax vowel space areas for all nonhabitual conditions relative to the habitual condition. Descriptive statistics further indicate that both tense and lax vowel space areas were consistently maximized in the overenunciate condition, with analyses for F1 and F2 ranges suggesting that speakers achieved adjustments in vowel space area by producing greater excursions in both tongue height and tongue advancement, with a trend toward the greatest excursions for the overenunciate condition.

To further explore whether particular vowels were more amenable to the effects of clear speech instructions than others, and by inference those vowels contributing to vowel space area change in the nonhabitual conditions, ID measures were calculated to measure each vowel's shift in F1 × F2 space across conditions. In general, more speakers tended to produce the greatest ID measures for the high vowels /i, ɪ, u, and ʊ/ in the clear, hearing impaired, and overenunciate conditions. Given the fact that all speakers increased tense and lax vowel space area in the nonhabitual conditions, findings for ID measures suggest that shifts in high tense and lax vowels contributed most to vowel space expansion.

In addition to clarity-related changes in midpoint vowel formant frequencies, tense and lax lambda measures—corresponding to the amount of spectral change in F1 and F2 between 20% and 80% of vowel duration and by inference within-vowel variation in tongue height and advancement—also differed significantly among conditions. Findings were similar to those for vowel space area in that lambda measures were significantly greater in the nonhabitual conditions compared with habitual. Moreover, both tense and lax lambda measures were greatest in the overenunciate condition (Table 4).

The different clear speech instructions also yielded different magnitudes of change in measures of speech timing, with the greatest change in the form of longer segment durations, longer pauses, greater numbers of pauses as well as slow articulation and speech rates for the overenunciate condition, followed by the hearing impaired, clear, and habitual conditions. Although previous clear speech studies have reported much greater numbers of pauses than reported in the current study (Krause & Braida, 2004; Picheny et al., 1985), the criterion for determining pauses used in these studies was much shorter (i.e., 10 ms) than that used in the current study, as well as in other studies (i.e., 200 ms or greater). Finally, on average, mean SPL for nonhabitual conditions only increased by 1–2 dB compared with that in the habitual condition. Half of the speakers produced a 3- to 6-dB increase in average SPL for the hearing impaired condition, however, and three speakers increased mean SPL by approximately 1–2 dB in the overenunciate condition relative to habitual. These SPL adjustments are less consistent than previously reported (Dromey, 2000; Moon & Lindblom, 1994; Picheny et al., 1985). However, as noted by Smiljaniċ and Bradlow (2009), an increased vocal intensity or increased vocal effort might not necessarily be expected for clear speech, as talkers presumably use clear speech to maximize intelligibility rather than vocal loudness.

The fact that the various clear speech instructions yielded different magnitudes of change in vowel space area, F1 and F2 range, lambda, and speech timing suggests caution in directly comparing these kinds of measures in studies that use different cues or instructions for eliciting adjustments in speech clarity. Results also have implications for clear speech training programs. That is, findings from the present study suggest that the cue "overenunciate" may prove most effective in a clear speech training program for increasing lingual excursion and thus vowel space area as well as vowel spectral dynamics. This cue also may be most effective for eliciting reductions in speech and articulation rate, at least relative to the other clear speech instructions studied. In contrast, the cue "speak to someone with a hearing impairment" would appear to be more effective than "overenunciate" or "speak clearly" in a clear speech training program that aims to increase average vocal intensity. Given the fact that only about half of the speakers increased vocal intensity in the hearing impaired condition, however, it is not necessarily expected that a clear speech training program would yield consistent changes in vocal intensity. Future studies are needed to investigate these suggestions. Additional studies also are needed to determine the extent to which the different magnitudes of acoustic change reported in the current study are perceptually relevant. A strong prediction is that clear speech instructions associated with relatively greater acoustic change would also be associated with a relatively larger clear speech benefit.

An explanation for why the various clear speech instructions or cues elicited different amounts of acoustic change may be related to the nature of the speech modifications suggested by each cue. Given that the clear condition consistently yielded the smallest acoustic adjustments, it would appear that the general instruction "speak clearly" may not adequately inform a speaker about how to adjust speech output. Ferguson (2004) also suggested that some speakers may require more specific instructions than "speak clearly" to maximize the benefits of clear speech. In contrast, the cues to "overenunciate" or "speak to someone with a hearing impairment" appear to have provided speakers with more focused information as to how to modify speech. The instruction to overenunciate appeared to direct speakers' attention to segmental articulation and speech durations, and the instruction "talk to someone with a hearing impairment" appeared to direct speakers' attention—for at least approximately half of the speakers—to vocal intensity as well as segmental articulation and speech duration, albeit to a lesser extent than "overenunciate."

Tense–Lax Vowel Distinctiveness

Chen (1980) reported that repeated productions of a given vowel (i.e., vowel clustering) become closer in F1 × F2 acoustic working space in clear speech, thus creating greater spectral distinction among vowel categories. More recent studies have shown that this type of spectral distinctiveness among neighboring vowels is important for intelligibility (H. Kim, Hasegawa-Johnson, & Perlman, 2011; Neel, 2008). Vowel clustering of the nature reported by Chen (1980) and H. Kim et al. (2011) was not computed in the present study due to the number of occurrences of each vowel ($N = 3$–5) per condition and the connected speech context. Instead, distinctiveness of tense–lax vowel pairs was examined via spectral distance in F1 × F2 space measures, duration ratios, and lambda ratios.

Although there was a trend for the four tense–lax vowel pairs to exhibit increased dynamic distinction in the non-habitual conditions, measures of spectral and temporal distinctiveness for tense–lax vowel pairs did not differ significantly among conditions. Apparently, speaking-condition related enhancement in distinctiveness for tense–lax vowel pairs is subtle —a finding also supported by Y. Kim (2011) in a recent study investigating the impact of increased vocal intensity on spectral distinctiveness of tense–lax vowel pairs. However, vowel distinctiveness in the present study was not wholly unaffected by clear speech instructions. That is, as noted by Smiljani and Bradlow (2009), results for vowel space area as well as F1 and F2 range suggest enhanced spectral distinctiveness among both tense vowels and lax vowels.

Interestingly, one speaker (M04) reduced almost all spectral distances for tense–lax vowel pairs in the non-habitual conditions. Inspection of M04's duration and lambda ratios suggested that this speaker did tend to enhance tense–lax vowel contrast in the nonhabitual conditions with duration or dynamic spectral cues, but the magnitude of the enhancement was modest. Informal listening to M04's sentences further suggested only subtle perceptual differences among conditions. More formal perceptual judgments are warranted, but M04's data support the hypothesis that spectral and temporal acoustic contrast in clear speech is perceptually important.

Clear Speech and Articulatory Effort

Perkell and colleagues (2002) reported that clear speech was associated with greater peak movement speeds, longer movement durations, and greater movement distances. Although there is no widely accepted metric of articulatory effort, Perkell et al. (2002) interpreted these adjustments as evidence of increased articulatory effort for clear speech much in the same way that an increased SPL has been interpreted to reflect increased respiratory– phonatory effort (see Fox et al., 2006). It has further been suggested that acoustic adjustments associated with clear speech, such as increased intensity, lengthened sound segments, expanded vowel space areas, increased formant transition extents, increased formant transition slopes, and greater spectral distances, are associated with changes in the duration and/or velocity of articulatory movements (Moon & Lindblom, 1994; Perkell et al., 2002).

In the present study, greater movement distances for the nonhabitual conditions are suggested by increased vowel space areas, F1 and F2 ranges, lambda values, and IDs. The nonhabitual conditions also yielded changes in segment durations, articulation rate, and speaking rate. To the extent that these types of acoustic changes for the nonhabitual conditions may be a reflection of or byproduct of articulatory effort, the implication is that speakers used the greatest articulatory effort when instructed to overenunciate. Relatedly, results for SPL suggest a trend for increased respiratory–phonatory effort in the hearing impaired condition, for approximately half of the speakers.

Caveats and Future Directions

Several factors should be kept in mind when interpreting results of the present investigation. First, speakers included for study were healthy, young adults who spoke Standard American

English with a dialect characteristic of the western New York region. Whether findings can be generalized to other ages, dialects, and languages is unknown. In addition, only neurologically healthy speakers were studied. Results may not translate in a straightforward manner to clinical populations, such as patients with dysarthria, for whom clear speech is used therapeutically. Future studies evaluating the effects of different clear speech instructions for a variety of populations, including dysarthria, are of importance.

As a result of the ordering of conditions, the cumulative effect of the clear condition on the following non-habitual conditions is unknown. However, in an effort to minimize carry-over effects across conditions, participants were engaged in a few minutes of conversation in between conditions. We also acknowledge that there was not a training component associated with the various clear speech instructions, nor was feedback provided regarding the adequacy with which speakers were able to produce clear speech. Thus, results should not be directly extrapolated to clear speech training programs or real-world clear speech situations in which listener feedback is available. Another factor to consider is the potential impact of clear speech instruction on measures of speech production that were not of interest in the present study. For example, clear speech has been shown to affect consonant production characteristics (Maniwa, Jongman, & Wade, 2009) as well as f0 (Bradlow et al., 2003), although Uchanski (2005) noted a fair amount of cross-speaker variability in the impact of clear speech on this latter measure. Future studies can now build on the present investigation to improve our understanding of how clear speech instruction affects a broader range of speech production measures. Studies also are needed to determine the perceptual consequences of different clear speech instructions. Finally, because many studies continue to use a hybrid of instructions to elicit clear speech (Bradlow et al., 2003; Picheny et al., 1985; Smiljani & Bradlow, 2008), additional studies are needed to determine how a combination of clear speech definitions affects both speech production and perception.

Acknowledgments

References

American National Standards Institute. Specifications for audiometers. New York, NY: Author; 1969. ANSI S3.6-1969

Assmann PF, Katz WF. Synthesis fidelity and time-varying spectral change in vowels. The Journal of the Acoustical Society of America. 2005; 117:886–895. [PubMed: 15759708]

Beukelman DR, Fager S, Ullman C, Hanson E, Logemann J. The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. Journal of Medical Speech-Language Pathology. 2002; 10:237–242.

Bradlow AR, Krause N, Hayes E. Speaking clearly for children with learning disabilities: Sentence perception in noise. Journal of Speech, Language, and Hearing Research. 2003; 46:80–97.

Bradlow AR, Torretta GM, Pisoni DB. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. Speech Communication. 1996; 20:255–272. [PubMed: 21461127]

Brown, H., Prescott, R. Applied mixed models in medicine. West Sussex, England: Wiley; 1999.

Chen, FR. Unpublished master's thesis. Massachusetts Institute of Technology; Cambridge: 1980. Acoustic characteristics and intelligibility of clear and conversational speech.

Darling M, Huber JE. Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. Journal of Speech, Language, and Hearing Research. 2011; 54:1247–1259.

Dromey C. Articulatory kinematics in patients with Parkinson disease using different speech treatment approaches. Journal of Medical Speech-Language Pathology. 2000; 8:155–161.

Ferguson SH. Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. The Journal of the Acoustical Society of America. 2004; 116:2365–2373. [PubMed: 15532667]

Ferguson SH, Kewley-Port D. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America. 2002; 112:259–271. [PubMed: 12141351]

Ferguson SH, Kewley-Port D. Talker differences in clear and conversational speech: Acoustic characteristics of vowels. Journal of Speech, Language, and Hearing Research. 2007; 50:1241–1255.

Fox CM, Ramig LO, Ciucci MR, Sapir S, McFarland DH, Farley BG. The science and practice of LSVT/LOUD: Neural plasticity-principled approach to treating individuals with Parkinson disease and other neurological disorders. Seminars in Speech and Language. 2006; 27:283–299. [PubMed: 17117354]

Goberman AM, Elmer LW. Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. Journal of Communication Disorders. 2005; 38:215–230. [PubMed: 15748725]

Hazan V, Markham D. Acoustic-phonetic correlates of talker intelligibility for adults and children. The Journal of the Acoustical Society of America. 2004; 116:3108–3118. [PubMed: 15603156]

Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America. 1995; 97:3099–3111. [PubMed: 7759650]

Hillenbrand JM, Nearey TM. Identification of resynthesized /hVd/ utterances: Effects of formant contour. The Journal of the Acoustical Society of America. 1999; 105:3509–3523. [PubMed: 10380673]

Huber JE, Darling M. Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: Respiratory physiologic and linguistic considerations. Journal of Speech, Language, and Hearing Research. 2011; 54:33–46.

Hustad KC, Lee J. Changes in speech production associated with alphabet supplementation. Journal of Speech, Language, and Hearing Research. 2008; 51:1438–1450.

Hustad, KC., Weismer, G. A continuum of interventions for individuals with dysarthria: Compensatory and rehabilitative treatment approaches. In: Weismer, G., editor. Motor speech disorders. San Diego, CA: Plural; 2007. p. 261-303.

Huttunen K, Keränen H, Väyrynen E, Pääkkönen R, Leino T. Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. Applied Ergonomics. 2011; 42:348–357. [PubMed: 20832770]

Johnson K, Flemming E, Wright R. The hyperspace effect: Phonetic targets are hyperarticulated. Linguistic Society of America. 1993; 69:505–528.

Kent, RK., Read, C. Acoustic analysis of speech. 2. Albany, NY: Singular/Thomson Learning; 2002.

Kim H, Hasegawa-Johnson M, Perlman A. Vowel contrast and speech intelligibility in dysarthria. Folia Phoniatrica et Logopaedica. 2011; 63(4):187–194. [PubMed: 20938200]

Kim, Y. Acoustic contrastivity in conversational and loud speech. Poster presented at the annual meeting of the Acoustical Society of America; San Diego, CA. 2011 Oct-Nov.

Klatt DH. Voice onset time, frication, and aspiration in word-initial consonant clusters. Journal of Speech and Hearing Research. 1975; 18:686–706. [PubMed: 1207100]

Krause J, Braida L. Acoustic properties of naturally produced clear speech at normal speaking rates. The Journal of the Acoustical Society of America. 2004; 115:362–378. [PubMed: 14759028]

Krause JC, Braida LD. Evaluating the role of spectral and envelope characteristics in the intelligibility advantage of clear speech. The Journal of the Acoustical Society of America. 2009; 125:3346–3357. [PubMed: 19425675]

Lindblom, B. Explaining phonetic variation: A sketch of the H & H theory. In: Hardcastle, WJ., Marchal, A., editors. Speech production and speech modelling. Dordrecht, the Netherlands: Kluwer Academic; 1990. p. 403-439.

Liu H, Tsao F, Kuhl PK. The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. The Journal of the Acoustical Society of America. 2005; 117:3879–3889. [PubMed: 16018490]

Liu S, Zeng F. Temporal properties in clear speech perception. The Journal of the Acoustical Society of America. 2006; 120:424–432. [PubMed: 16875238]

Maniwa K, Jongman A, Wade T. Acoustic characteristics of clearly spoken English fricatives. The Journal of the Acoustical Society of America. 2009; 125:3962–3973. [PubMed: 19507978]

Milenkovic, P. TF32 [Computer software]. Madison: University of Wisconsin; 2002.

Moon SJ, Lindblom B. Interaction between duration, context and speaking-style in English stressed vowels. The Journal of the Acoustical Society of America. 1994; 96:40–55.

Neel AT. Vowel space characteristics and vowel identification accuracy. Journal of Speech, Language, and Hearing Research. 2008; 51:574–585.

Perkell JS, Zandipour M, Matthies ML, Lane H. Economy of effort in different speaking conditions: I. A preliminary study of intersubject differences and modeling issues. The Journal of the Acoustical Society of America. 2002; 112:1627–1641. [PubMed: 12398468]

Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. Journal of Speech and Hearing Research. 1985; 28:96–103. [PubMed: 3982003]

Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. Journal of Speech and Hearing Research. 1986; 29:434–445. [PubMed: 3795886]

Rogers CL, DeMasi TM, Krause JC. Conversational and clear speech intelligibility of /bVd/ syllables produced by native and non-native English speakers. The Journal of the Acoustical Society of America. 2010; 128:410–423. [PubMed: 20649235]

Sadagopan N, Huber JE. Effects of loudness cues on respiration in individuals with Parkinson's disease. Movement Disorders. 2007; 22:651–659. [PubMed: 17266087]

Schum DJ. Beyond hearing aids: Clear speech training as an intervention strategy. Hearing Journal. 1997; 50:36.

Smiljani R, Bradlow AR. Temporal organization of English clear and conversational speech. The Journal of the Acoustical Society of America. 2008; 124:3171–3182. [PubMed: 19045801]

Smiljani R, Bradlow AR. Speaking and hearing clearly: Talker and listener factors in speaking style changes. Language and Linguistics Compass. 2009; 3:236–264. [PubMed: 20046964]

Tasko SM, Greilick K. Acoustic and articulatory features of diphthong production: A speech clarity study. Journal of Speech, Language, and Hearing Research. 2010; 53:84–99.

Tjaden K, Rivera D, Wilding G, Turner GS. Characteristics of the lax vowel space in dysarthria. Journal of Speech, Language, and Hearing Research. 2005; 48:554–566.

Tjaden K, Wilding GE. Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. Journal of Speech, Language, and Hearing Research. 2004; 47:766–783.

Turner GS, Tjaden K, Weismer G. The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. Journal of Speech and Hearing Research. 1995; 38:1001–1013. [PubMed: 8558870]

Turner GS, Weismer G. Characteristics of speaking rate in the dysarthria associated with amyotrophic lateral sclerosis. Journal of Speech and Hearing Research. 1993; 36:1134–1144. [PubMed: 8114480]

Uchanski, RM. Clear speech. In: Pisoni, DB., Remez, RE., editors. The handbook of speech perception. Malden, MA: Blackwell; 2005. p. 207-235.

Uchanski RM, Choi SS, Braida LD, Reed CM. Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. Journal of Speech and Hearing Research. 1996; 39:494–509. [PubMed: 8783129]

Van Nuffelen G, De Bodt M, Vanderwegen J, Wuyts F. Effect of rate control on speech production and intelligibility in dysarthria. Folia Phoniatrica et Logopaedica. 2010; 62:110–119. [PubMed: 20424466]

Weismer G, Laures JS, Jeng JY, Kent RD, Kent JF. Effect of speaking rate manipulationson acoustic and perceptual aspects of dysarthria in amyotrophic lateral sclerosis. Folia Phoniatrica et Logopaedica. 2000; 52:201–219. [PubMed: 10965174]

Wouters J, Macon MW. Effects of prosodic factors on spectral dynamics: I. Analysis. The Journal of the Acoustical Society of America. 2002; 111:417–427. [PubMed: 11831816]

Yorkston, KM., Beukelman, DR. Assessment of intelligibility of dysarthric speech. Austin, TX: Pro-Ed; 1984.

Zeng F, Liu S. Speech perception in individuals with auditory neuropathy. Journal of Speech, Language, and Hearing Research. 2006; 49:367–380.

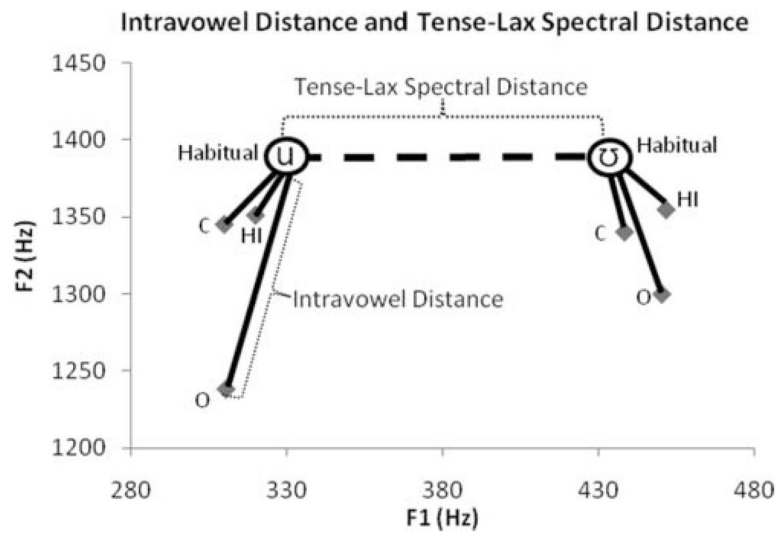**Intravowel Distance and Tense-Lax Spectral Distance**

Figure 1.
Examples of tense–lax spectral distance and intravowel distance are shown for the tense and lax vowels /u/ and /ʊ/. C = clear; HI = hearing impaired; O = overenunciate.
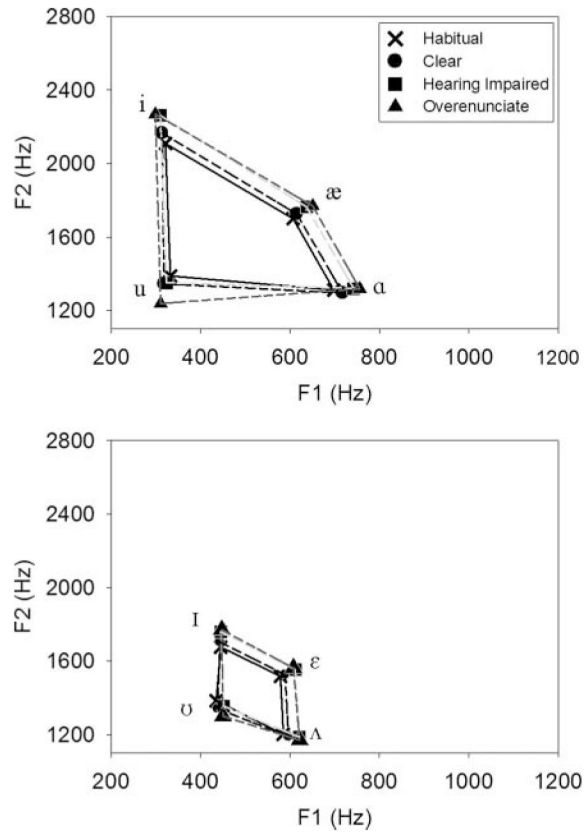
Figure 2.
Average tense (upper panel) and lax (lower panel) vowel space area for male speakers as a function of condition.
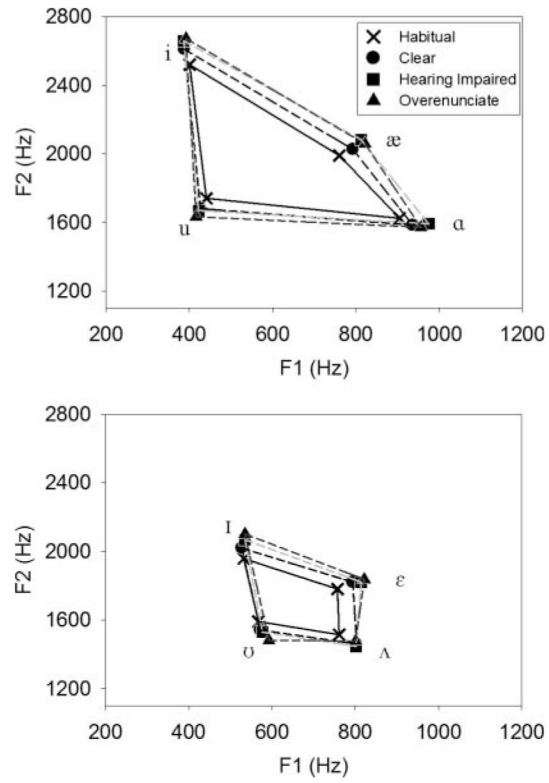
Figure 3.
Average tense (upper panel) and lax (lower panel) vowel space area for female speakers as a function of condition.
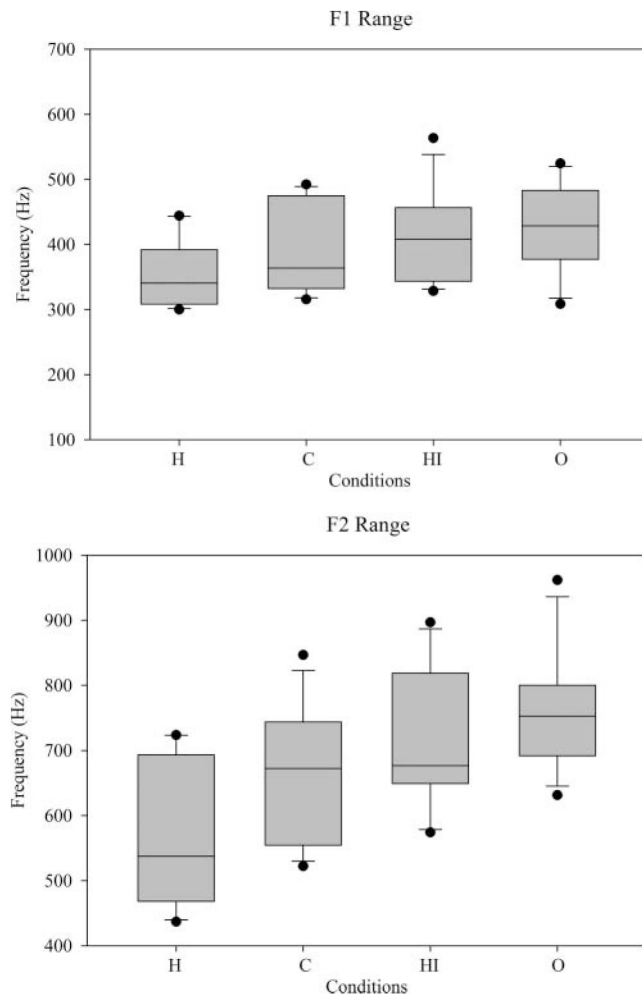
Figure 4.
Average F1 (upper panel) and F2 (lower panel) range as function of condition. Error bars represent the 90th and 10th percentiles. H = habitual.
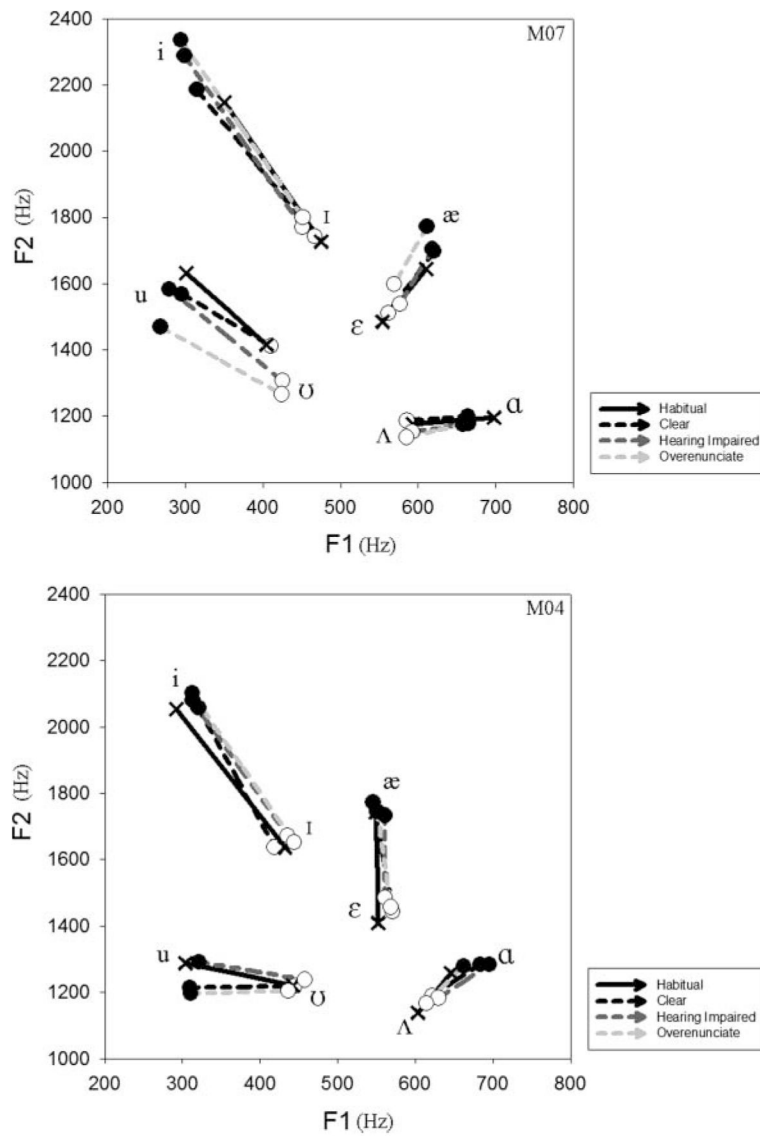
Figure 5.
Average tense–lax spectral distances for Speaker M07 (upper panel) and Speaker M04 (lower panel) as a function of condition.
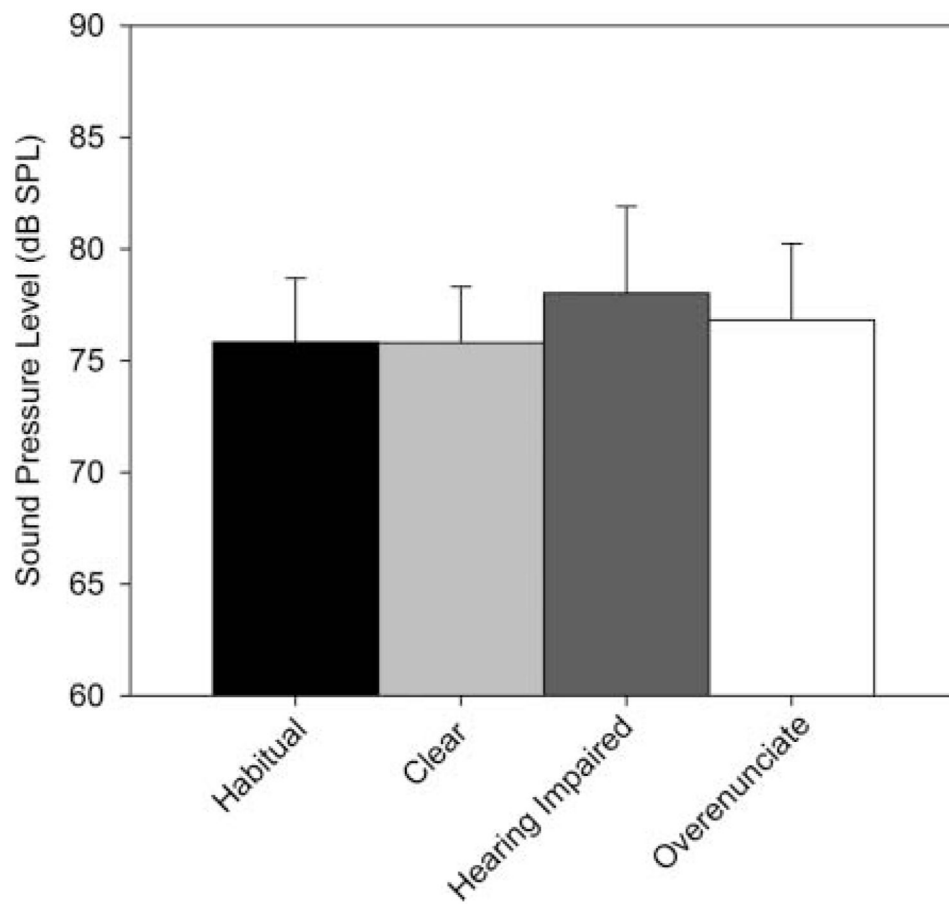
Figure 6.
Average SPL as a function of condition. Error bars represent standard deviations.

Table 1

Dependent variables and their corresponding acoustic measure.

Dependent variable	Acoustic measure
Vowel space area (tense and lax)	50% formant frequencies for F1 and F2
F1 range	Average high F1 vowels (æ, ɑ)–Average low F1 vowels (i, u)
F2 range	Average high F2 vowels (i, æ)–Average low F2 vowels (ɑ, u)
Tense–lax spectral distance	50% formant frequencies for F1 and F2
Intravowel distance	50% formant frequencies for F1 and F2
Vowel spectral change measure (lambda)	20% & 80% formant frequencies for F1 and F2
Speaking rate	Total utterance duration
Articulation rate	Run duration
Vowel duration (tense and lax)	Vowel segment duration
Vocal intensity	dB SPL
Duration ratio	Average vowel durations: lax/tense
Lambda ratio	Average lambda values: lax/tense

Table 2

Mean (*SD*) percent change in vowel space area relative to habitual.

Measure	Gender	Vowel space (% increase relative to habitual)		
		Clear	**Hearing impaired**	**Overenunciate**
Tense vowels	F	46 (42)	61 (27)	76 (42)
	M	23 (11)	43 (34)	73 (48)
Lax vowels	F	85 (38)	88 (42)	101 (61)
	M	27 (23)	48 (41)	66 (55)

Table 3

Number of participants ($N = 12$) exhibiting the greatest intravowel distance measures as a function of vowel type.

Condition	Vowels			
Tense vowels	/i/	/u/	/æ/	/ɑ/
Clear	5	6	1	0
Hearing impaired	8	1	2	1
Overenunciate	8	4	0	0
Lax vowels	/ɪ/	/ʊ/	/ɛ/	/ʌ/
Clear	3	4	2	3
Hearing impaired	5	3	3	1
Overenunciate	4	6	1	1

Table 4

Mean (*SD*) lambda measures for tense and lax vowels and descriptive statistics for vowel duration, as a function of condition.

Measure	Gender	Habitual	Clear	Hearing impaired	Overenunciate
Tense lambda	F	288 (245)	308 (309)	333 (310)	335 (330)
	M	216 (178)	232 (201)	243 (199)	245 (206)
	All	253 (218)	271 (264)	289 (265)	291 (262)
Lax lambda	F	230 (153)	247 (170)	279 (188)	296 (269)
	M	229 (144)	225 (152)	238 (167)	251 (164)
	All	230 (149)	236 (162)	259 (179)	274 (225)
Tense vowel duration (ms)	All	116 (46)	136 (60)	152 (59)	169 (73)
Lax vowel duration (ms)	All	83 (33)	94 (40)	106 (46)	114 (46)

Table 5

Mean (*SD*) duration ratio and lambda ratio for each tense–lax vowel pair as a function of condition.

Measure	Vowel pair	Habitual	Clear	Hearing impaired	Overenunciate
Duration ratio	æ–ɛ	0.70 (0.08)	0.67 (0.07)	0.72 (0.13)	0.67 (0.11)
	ɑ–ʌ	0.65 (0.13)	0.68 (0.17)	0.70 (0.17)	0.72 (0.17)
	i–ɪ	0.70 (0.14)	0.63 (0.15)	0.65 (0.11)	0.60 (0.12)
	u–ʊ	0.76 (0.19)	0.76 (0.21)	0.73 (0.23)	0.78 (0.24)
	Average	0.71 (0.08)	0.69 (0.09)	0.69 (0.08)	0.67 (0.07)
Lambda ratio	æ–ɛ	1.03 (0.44)	0.97 (0.32)	0.95 (0.30)	0.88 (0.36)
	ɑ–ʌ	1.40 (0.57)	1.40 (0.50)	1.58 (0.73)	1.90 (1.03)
	i–ɪ	0.93 (0.38)	1.08 (0.62)	1.22 (0.70)	1.55 (1.02)
	u–ʊ	1.06 (0.96)	0.81 (0.41)	0.74 (0.38)	0.85 (0.46)
	Average	0.98 (0.19)	0.92 (0.22)	0.97 (0.25)	1.00 (0.18)

Table 6

Mean (*SD*) suprasegmental measures for each condition.

Measure	Habitual	Clear	Hearing impaired	Overenunciate
Articulation rate (syllables/s)	5.3 (0.9)	4.4 (1.0)	4.0 (0.7)	3.3 (1.0)
Speaking rate (syllables/s)	5.3 (0.8)	4.5 (1.0)	4.0 (0.7)	3.3 (0.9)
Tense vowel duration (ms)	116 (46)	136 (60)	152 (59)	169 (73)
Lax vowel duration (ms)	83 (33)	94 (40)	106 (46)	114 (46)
Pause duration (ms)	280 (127)	346 (112)	307 (91)	341 (138)
Pause frequency	2 (1)	4 (3)	3 (3)	11 (8)