Talker Differences in Clear and Conversational Speech: Acoustic Characteristics of Vowels

Sarah Hargus Ferguson
University of Kansas, Lawrence

Diane Kewley-Port
Indiana University, Bloomington

Purpose: To determine the specific acoustic changes that underlie improved vowel intelligibility in clear speech.
Method: Seven acoustic metrics were measured for conversational and clear vowels produced by 12 talkers—6 who previously were found (S. H. Ferguson, 2004) to produce a large clear speech vowel intelligibility effect for listeners with normal hearing identifying vowels in background noise (the *big benefit talkers*), and 6 who produced no clear speech vowel intelligibility benefit (the *no benefit talkers*).
Results: For vowel duration and for certain measures of the overall acoustic vowel space, the change from conversational to clear speech was significantly greater for big benefit talkers than for no benefit talkers. For measures of formant dynamics, in contrast, the clear speech effect was similar for the 2 groups.
Conclusion: These results suggest that acoustic vowel space expansion and large vowel duration increases improve vowel intelligibility. In contrast, changing the dynamic characteristics of vowels seems not to contribute to improved clear speech vowel intelligibility. However, talker variability suggested that improved vowel intelligibility can be achieved using a variety of clear speech strategies, including some apparently not measured here.

KEY WORDS: speech perception, speech sound, acoustics

Most studies of clear speech, in which talkers are asked to speak as though talking to someone who has a hearing loss or is from a different language background, have shown significant improvements in intelligibility over ordinary conversational speech (e.g., Bradlow & Bent, 2002; Bradlow, Kraus, & Hayes, 2003; Gagné, Masterson, Munhall, Bilida, & Querengesser, 1994; Helfer, 1998; Krause & Braida, 2002; Picheny, Durlach, & Braida, 1985). A clear speech intelligibility benefit has been found for both sentences and words (Uchanski, Choi, Braida, & Durlach, 1996) and has been of similar magnitude for listeners with hearing impairment listening to linearly amplified materials as for listeners with normal hearing listening in noise (Payton, Uchanski, & Braida, 1994; Uchanski et al., 1996). A notable exception, however, is the data of Ferguson and Kewley-Port (2002). In contrast to listeners with normal hearing, who enjoyed a 15-percentage-point clear speech intelligibility benefit for vowels presented in noise, elderly listeners with hearing impairment listening to these vowels in noise without amplification showed no clear speech advantage. More recently, Ferguson (2004) showed that talkers vary in how much their vowel intelligibility improves when speaking clearly. For vowels in words presented in noise to listeners with normal hearing, the magnitude of the clear speech vowel

1092-4388/07/5005-1241

intelligibility benefit varied widely (from −12% to 33%) among a group of 41 normal talkers.

In the current study, acoustic analyses were performed on the vowels of 12 of these 41 talkers. The goal was to explore the extent to which various clear speech acoustic changes contribute to improved vowel intelligibility. Previous studies have shown that clear and conversational speech differ on many acoustic dimensions. For example, Picheny, Durlach, and Braida (1986) found that voice intensity was 5–8 dB greater and that talkers used a higher and more variable voice fundamental frequency in clear speech. Clear speech was also considerably slower than conversational speech because of longer and more frequent pauses as well as longer phoneme durations. Consonant power and the consonant–vowel ratio both increased in clear speech, and stop consonant bursts were deleted less frequently. Bradlow et al. (2003), who corroborated several of these results, also noted that alveolar "flapping" occurred less often in clear speech.

Vowel modifications have been reported in studies comparing clear and conversational speech (Bradlow et al., 2003; Ferguson & Kewley-Port, 2002; Picheny et al., 1986) as well as in studies comparing "hyperarticulated" and citation-style speech (Johnson, Flemming, & Wright, 1993; Moon & Lindblom, 1994; Wouters & Macon, 2002).[1] The most robust vowel change observed in these studies is an increase in vowel duration, as observed by Picheny et al. (1986) and corroborated by Ferguson and Kewley-Port (2002) and by Moon and Lindblom (1994). In addition, talkers increase the size of their acoustic vowel spaces in clear speech and in hyperarticulated speech (Bradlow et al., 2003; Johnson et al., 1993; Picheny et al., 1986). Finally, vowels seem to have greater dynamic formant movement when talkers speak clearly or hyperarticulate. Wouters and Macon (2002) found greater spectral rate of change during consonant–vowel transitions in hyperarticulated speech than in citation-style speech. Similarly, both Moon and Lindblom (1994) and Ferguson and Kewley-Port (2002) found that the amount of formant movement over the vowel nucleus was significantly greater in clear speech.

Although numerous acoustic differences between clear and conversational speech have been described, it is not yet clear which ones contribute to the superior intelligibility of clear speech. One reason is the sheer number of observed clear speech acoustic effects. To make the pool of potentially important changes more manageable, some investigators have used signal-processing techniques to acoustically modify conversational speech. For example, Gordon-Salant (1986, 1987) altered consonant intensity and duration in consonant–vowel syllables. Although increasing the consonant–vowel intensity ratio improved consonant intelligibility for several listening groups, increasing consonant duration provided no benefit. Picheny, Durlach, and Braida (1989) examined duration more globally using an algorithm that uniformly increased the duration of all speech sounds. Slowing conversational speech in this manner not only failed to improve intelligibility for listeners with hearing impairment but actually made conversational speech less intelligible. Uchanski et al. (1996) found similar results using a nonuniform time-scaling algorithm. Another attempt by Nejime and Moore (1998), using a different algorithm, was similarly unsuccessful.

Another way to reduce the number of potentially important clear speech acoustic effects is to use speech materials that vary over a relatively small number of dimensions. Krause and Braida (2004) noted the benefit of such an approach, speculating that acoustic variability due to phonetic context could disguise changes due to speaking style. This approach was taken by Ferguson and Kewley-Port (2002), who examined vowels in a fixed /bVd/ context. Furthermore, in contrast with earlier signal-processing studies, Ferguson and Kewley-Port (2002) investigated the acoustic features that underlie the clear speech intelligibility benefit for vowels by exploiting the natural variability in a large set of clear and conversational speech tokens (10 per vowel per style) produced by a single talker. Regression analyses on several acoustic measures revealed two interesting relationships. First, the enormous vowel duration increases produced by their talker explained very little of the variance in vowel intelligibility in noise. Second, another clear speech acoustic change, raising F2 for the front vowels (/i, ɪ, e, ɛ, æ/), seemed to negatively affect vowel intelligibility for listeners with hearing impairment. It appeared that raising F2 for these vowels made them less audible for these listeners, who had sloping hearing losses and for whom the speech was not amplified. The reduced F2 audibility, in turn, made the front vowels less intelligible in clear speech. This suggested that some of the clear speech acoustic changes this talker made were not well-suited to the needs of listeners with hearing impairment who are listening without amplification. This surprising outcome highlights the need to identify the acoustic characteristics that underlie the superior intelligibility of clear speech across a large group of talkers and for various types of listeners in a variety of listening situations.

Like Ferguson and Kewley-Port (2002), the current study takes advantage of naturally occurring differences within a single phoneme class in a fixed phonetic context to investigate how specific clear speech acoustic characteristics contribute to intelligibility. However, in contrast with that study, which exploited the variability found among multiple tokens produced by a single talker, the current study exploits the variability found within a large group

---

[1] Note that although clear and hyperarticulated speech are likely to be similar speaking styles, there has been no work directly comparing these two styles in terms of either intelligibility or acoustic characteristics.

of talkers. Specifically, it takes an approach similar to the one taken by Bradlow et al. (2003) when they discussed the clear speech acoustic changes made by the 2 talkers in their study. They observed that the female talker not only had a larger clear speech intelligibility benefit than the male talker but also reduced her speaking rate to a greater extent. This led them to infer that reduced speaking rate contributes to the superior intelligibility of clear speech. Conversely, a clear speech acoustic effect that was similar for the 2 talkers (e.g., increased pitch range) was interpreted as less important for enhancing intelligibility.

The current study applies this approach on a much larger scale, applying an extreme groups design to the Ferguson clear speech database (2004). Two groups of 6 talkers were selected from this pool based on their clear-minus-conversational vowel intelligibility difference scores for listeners with normal hearing identifying vowels in noise. Talkers in the big benefit group showed large positive difference scores; the talkers in the no benefit group had difference scores close to 0 (see the second paragraph under *Talkers* for details). Vowels produced by each talker were analyzed, and between-style acoustic differences were compared for the two groups. Several potentially important clear speech acoustic changes were tested against the following premise: If a particular change is larger in talkers with a large clear speech intelligibility benefit than in those showing no benefit, then that change is likely to underlie the clear speech vowel intelligibility benefit. Acoustic features for which the two groups show similar clear speech acoustic changes, in contrast, are less likely to be responsible for enhanced vowel intelligibility in clear speech.

Method

Materials

The materials consisted of vowels (/i, ɪ, e, ɛ, æ, ɑ, ʌ, o, ʊ, u/) in /bVd/ context. These words were recorded in the context of meaningful sentences. Sixteen carrier sentences were used, each of which contained minimal contextual information regarding the identity of the /bVd/ word, which was centrally located within the sentence. Example sentences include "Vera put the _____ on the table" and "I think the word _______ is hard for kids to say." Seven tokens were recorded of each /bVd/ word, each in a different carrier sentence. In the conversational speech condition, which was recorded first, talkers were instructed to read the sentences as they would in everyday conversation. In the clear speech condition, talkers were instructed to say the sentences as they would if they were talking to a person with hearing loss. Complete details regarding recording procedures may be found in Ferguson (2004). For each talker, two tokens of each /bVd/ word were analyzed in each speaking style; these tokens were identical to those used in the intelligibility study in Ferguson (2004).

Talkers

Two groups of 6 talkers were selected from the pool of 41 talkers described in Ferguson (2004) based on the results of the perceptual experiment reported in that article. Briefly, the two /bVd/ test words for each talker in each style were presented for identification to listeners with normal hearing who were listening in a background of 12-talker babble. All test words were scaled to the same root-mean-square (RMS) amplitude prior to the experiment to eliminate intensity differences among the vowels and between the two speaking styles. The presentation level was 70 dB SPL, and the signal-to-babble ratio was −10 dB.

Talkers were chosen for the current experiment such that the two groups would (a) differ significantly in the size of the clear speech vowel intelligibility effect and (b) have equal numbers of male and female talkers. First, the 41 talkers were ordered by their clear speech vowel intelligibility difference scores in rationalized arcsine units (RAUs; Studebaker, 1985) as determined in Ferguson (2004). The ordered data were then scrutinized to identify two sets of 6 consecutive talkers that met the selection criteria. The selection was thus somewhat arbitrary, made with no regard for intelligibility within a given speaking style. The talkers with the fourth through ninth largest clear speech vowel intelligibility benefit in RAUs were labeled group BB (big benefit). The talkers with the third through eighth smallest clear speech effect were labeled group NB (no benefit). Percent correct vowel intelligibility data and clear minus conversational difference scores for the 12 talkers are shown in Table 1.

The clear speech vowel intelligibility effect for the two groups was analyzed statistically by converting percent correct scores for individual vowels to RAUs and submitting them to a two-way repeated measures analysis of variance (ANOVA). The main effect of group was not significant, $F(1, 118) = 0.61$, $p = .44$, indicating that overall intelligibility was similar for the two talker groups. In contrast, the main effect of speaking style and the Style × Group interaction were significant, $F(1, 118) = 47.74$, $p < .01$, and $F(1, 118) = 50.27$, $p < .01$, respectively. The significant interaction indicates that the groups differed in the extent to which speaking clearly improved vowel intelligibility. Planned comparisons revealed that the nearly 19-percentage-point clear speech vowel intelligibility benefit observed for group BB was significant, $F(1, 118) = 97.96$, $p < .01$, whereas the negligible effect for group NB was not, $F(1, 118) = 0.02$, $p = .90$. Planned comparisons also showed that group BB had significantly higher vowel intelligibility than group NB in clear speech, $F(1, 118) = 10.14$, $p < .01$. Although the group means

Table 1. Vowel intelligibility in clear (CL) and conversational (CON) speech, the percentage point difference between the two styles (DIFF), and age for BB and NB talkers.

Talker	Age	CL	CON	DIFF
		BB talkers		
F18	42	66.4	38.6	27.9
F17	37	90.2	72.1	18.1
M02	23	79.1	61.0	18.1
M03	20	76.7	58.6	18.1
F09[a]	29	87.9	72.4	15.5
M20	41	75.0	61.2	13.8
M	32.0	79.2	60.6	18.6
		NB talkers		
F08	22	50.0	49.1	1.0
M17	33	79.3	79.3	0.0
M08	45	69.3	70.0	−0.7
F20	35	57.6	57.9	−0.2
M16	26	68.3	71.7	−3.3
F15[a]	40	77.6	81.0	−3.3
M	33.5	67.0	68.1	−1.1

Note. CL and CON scores reflect percent correct vowel identification scores for listeners with normal hearing who were listening in a background of 12-talker babble (Ferguson, 2004). DIFF = CL score − CON score; BB = big benefit; NB = no benefit. Within each group, talkers are ordered by the clear speech vowel intelligibility effect.

[a] Atypical talker.

suggest that the NB talkers had more intelligible vowels in conversational speech (68.1%) than the BB talkers (60.8%), the group effect was not significant, $F(1, 118) = 3.58$, $p = .06$, and examination of individual talker means revealed substantial overlap among the two groups.

Acoustic Analyses

Seven acoustic metrics were selected to assess the three primary acoustic characteristics widely considered to be important for vowel identification in American English: duration (one metric), steady-state formant values (four metrics), and dynamic formant movement (two metrics). Vowel duration measures were made from the waveform, from the first point after the /b/ burst where both F1 and F2 were well specified to the sharp amplitude drop associated with the /d/ closure. Steady-state and dynamic metrics were derived from LPC formant tracking. After downsampling each stimulus file from 22,050 to 11,025 Hz, modified COLEA MATLAB code (Loizou, 1999) was used to accomplish this tracking using a 20-ms Hamming window and a 10-ms frame rate. LPC order was normally $M = 12$ but was adjusted as needed for each individual stimulus. Any tracking errors were corrected by hand editing.

Values of the first two formants (F1 and F2) were extracted from the formant tracks at three locations: at 20% and 80% of vowel duration (following Hillenbrand, Getty, Clark, & Wheeler, 1995) and at the vowel steady state. Although it is well known that steady-state values for formant transitions are not found simultaneously for F1 and F2 (Di Benedetto, 1989), if at all, various proposals have been made for measuring steady-state values for formants. Given that the vowels measured here in conversational speech were not short ($M = 176.4$ ms), the steady-state formant values were located by a simple rule of the 20% point plus 30 ms in order to position the measurements toward the beginning of the vowel but past the rapid part of the /b/ transitions.

Steady-state metrics. The four steady-state metrics were chosen based on the assumption, found in previous studies, that the F1 × F2 vowel space would be larger in clear speech. Prior to calculating the metrics, formant frequency data were converted to Barks (Traunmüller, 1990) to normalize psychological distance over F1 and F2 (Kewley-Port & Zheng, 1999). The first metric, perimeter, reflected the overall dimensions of the vowel space. It was calculated for each talker in each speaking style as the sum of four Euclidean distances between adjacent point vowels (/i/ to /æ/, /æ/ to /ɑ/, /ɑ/ to /u/, and /u/ to /i/). The values for each point vowel were obtained by averaging across the two tokens of the vowel produced in each speaking style. The other three metrics were designed to examine the vowel space expansion more closely. F1 range, which measured expansion in the F1 dimension, was calculated as the difference between the average steady-state F1 value for the low vowels /æ/ and /ɑ/, and the average steady-state F1 value for the high vowels /i/ and /u/. Finally, two metrics focused on F2: (a) F2 front, the average steady-state F2 value for the vowels /i/, /ɪ/, /e/, /ɛ/, and /æ/; and (b) F2 back, the average steady-state F2 value for the vowels /ɑ/, /ʌ/, /o/, /ʊ/, and /u/. Front and back vowels were treated separately because results from Ferguson and Kewley-Port (2002) suggested that the amount of F2 shift in clear speech is much greater for front vowels than for back vowels.

Dynamic metrics. Two metrics were used to assess the amount of dynamic formant movement for individual vowel tokens. The assumption was that, as observed earlier (Ferguson & Kewley-Port, 2002; Moon & Lindblom, 1994; Wouters & Macon, 2002), vowels would be more dynamic in clear speech. The first metric, spectral change (λ), corresponded to the sum, in Barks, of the absolute formant frequency shift for F1 and F2. Thus, λ is calculated as

$$\lambda = |\, F1_{80} - F1_{20} \,| + |\, F2_{80} - F2_{20} \,|, \qquad (1)$$

where $F1_{20}$, $F1_{80}$, $F2_{20}$, and $F2_{80}$ are the F1 and F2 values in Barks at 20% and 80% of the vowel duration.

Increased formant movement in clear speech would be associated with higher λ values. Note that this two-point metric likely cannot capture details of complex formant movement or dynamics that might be observed over many consonantal contexts and would need validation before being applied more generally. However, for a single context, it appears from Hillenbrand and Nearey (1999) that two-point metrics capture formant dynamics as well as metrics with more temporal measurements.

The second dynamic metric, spectral angle (Ω), was inspired by the discrete cosine transform (DCT) analysis reported by Watson and Harrington (1999). Watson and Harrington (1999) showed that vowel formant trajectories, even those for diphthongs, could be classified with reasonable accuracy using just the first two DCT coefficients. The first DCT coefficient corresponds to the mean of the trajectory; the second corresponds to the tilt or angle of the trajectory. In the present study, to measure spectral angle for a single vowel, F1 and F2 were each compared to a theoretical steady formant with a frequency equal to that measured at the 20% point of the vowel. The angle θ for formant *n* was calculated as the arctangent of the difference between the formant frequencies at the 20% and 80% points, divided by the duration between these two points scaled to deciseconds. A formant showing no change from the 20% point to the 80% point would have a spectral angle of 0 radians. Individual formant angles were calculated as

$$\theta_n = \arctan\left(\frac{\mathrm{F}n_{80} - \mathrm{F}n_{20}}{d}\right), \qquad (2)$$

where

$$d = \frac{\mathrm{time}_{80} - \mathrm{time}_{20}}{100}. \qquad (3)$$

Spectral angle, Ω, was then calculated as the sum, in radians, of the absolute values of the F1 angle and the F2 angle, or

$$\Omega = |\,\theta_1\,| + |\,\theta_2\,|\,. \qquad (4)$$

Note that although increased formant movement in clear speech would be associated with higher Ω values, direction of formant movement is not captured by this metric.

Several studies have shown that the importance of dynamic formant information varies among individual vowels (Assmann & Katz, 2005; Hillenbrand & Nearey, 1999). For example, when Hillenbrand and Nearey (1999) synthesized vowels using original rather than flat formant contours, intelligibility improved by more than 22 percentage points for /e/, /æ/, /ʌ/, /o/, and /ʊ/ but by less than 10 percentage points for /i/, /ɪ/, /ɛ/, /ɑ/, and /u/. In a similar study by Assmann and Katz (2005), adding dynamic information most affected intelligibility for /o/, /e/, /ʊ/, /ɪ/, and /u/. It seems that whether a particular vowel is designated "dynamic" depends on the metric used and possibly on talker dialect. In Hillenbrand and Nearey (1999), the vowels for which formant dynamic information was most important tended to be those that were highly dynamic in natural speech. The current experiment, therefore, assumed that any speaking-style–related changes in formant movement would have the greatest intelligibility effect for vowels that were relatively more dynamic. Thus, statistical analyses on dynamic data were conducted for only the five vowels with the largest λ values (Equation 1) in conversational speech: /e/, /ɑ/, /ʌ/, /o/, and /ʊ/.

Statistical Analyses

As described above, the central premise of this study is that if a particular clear speech acoustic property underlies the superior intelligibility of clear speech, then the magnitude of this property will be greater for talkers showing a large clear speech intelligibility benefit (the BB talkers) than those showing no such benefit (the NB talkers). Each metric was tested against this premise using a two-way repeated measures ANOVA. Although the input data for these analyses varied, all used speaking style as the repeated measure and group as the categorical predictor. Acoustic measures that differed significantly between conversational and clear speech were considered potential contributors to the superior intelligibility of clear speech. The likelihood that a particular acoustic feature was responsible for the clear speech vowel intelligibility effect was much greater, however, when a significant Style × Group interaction was found, with a larger clear speech effect for group BB than for group NB.

Results

Atypical Talkers

The acoustic metrics were designed to assess characteristics generally considered important to vowel intelligibility in American English. It was expected, therefore, that at least one would be associated with improved vowel intelligibility in clear speech. That is, the Style × Group interaction was expected to be significant, with a larger style effect for group BB than for group NB, for at least one of the seven metrics. However, in initial analyses, none of the metrics showed a significant interaction. Inspection of individual talker data revealed that the clear speech effect for each metric varied greatly among the talkers. This is evident in Figures 1, 2, and 3, which show individual talker clear-minus-conversational differences for duration, the four steady-state metrics,

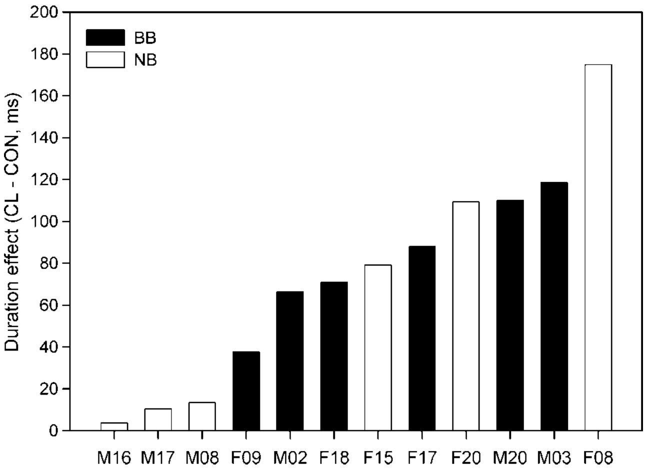


Figure 1. Difference between overall vowel durations in clear and conversational speech for individual talkers. Talkers are shown in ascending clear-minus-conversational difference order. BB = big benefit; NB = no benefit; CL = clear; CON = conversational.

and the two dynamic metrics, respectively. As a visual aid, talkers are ordered in each figure by the magnitude of the clear-minus-conversational difference for the metric shown. Talker group membership is indicated by filled (BB) and unfilled (NB) bars.

Group membership was not well separated for the seven metrics. Especially for the four steady-state metrics (see Figure 2), there was 1 talker in each group whose data differed markedly from the others. First, consider group NB. In all four panels of Figure 2, we see that Talker F15 had extreme values in relation to her fellow NB talkers. Panel D contains the strongest example. In contrast with the other NB talkers, who showed a shifted F2 for back vowels by −0.17 Barks or less, F15's shift was almost −0.50 Barks. In group BB, F09 had extreme values for F2 front, F1 range, and perimeter. This is most apparent in Panel A, which shows that F09's vowel space was actually smaller in clear speech. F15 and F09 resembled their respective groups somewhat more closely for duration and for the dynamic metrics, although F09 had the smallest clear speech vowel duration effect of group BB (see Figure 1), and F15 had a considerably larger clear speech effect for spectral change than the other NB talkers (see Figure 3, top).

This considerable talker variability, coupled with the absence of any significant Style × Group interactions in the initial analyses, suggested that the data from Talkers F09 and F15 might be obscuring trends present in the data from the other talkers. Indeed, when statistical analyses were repeated excluding these atypical talkers, several significant comparisons were obtained. These new analyses, excluding Talker F09 from group BB and Talker F15 from group NB, are thus reported henceforth. The rationale for this step is elaborated upon in the Discussion.

Duration

Average vowel durations in clear and conversational speech and the clear-minus-conversational difference for each talker group, excluding talkers F09 and F15, are shown in the top section of Table 2 . Note that columns 3 and 4 in Table 2 show the measured values for each metric in each speaking style; the last column shows the clear speech effect. Input data for the two-way repeated-measures ANOVA were average durations over two tokens of each vowel, for each talker and in each style. As expected, vowels were significantly longer in clear speech, $F(1, 98) = 172.94, p < .01$. In addition, talkers in group NB had significantly greater overall vowel duration than talkers in group BB, $F(1, 98) = 8.72, p < .01$, although the group effect was very small (<25 ms or 10%). Most importantly, the Style × Group interaction was significant, $F(1, 98) = 5.95, p < .05$. Talkers who showed a large clear speech vowel intelligibility benefit had a larger clear speech vowel duration effect (41% longer), as a group, than talkers who showed no benefit (25% longer).

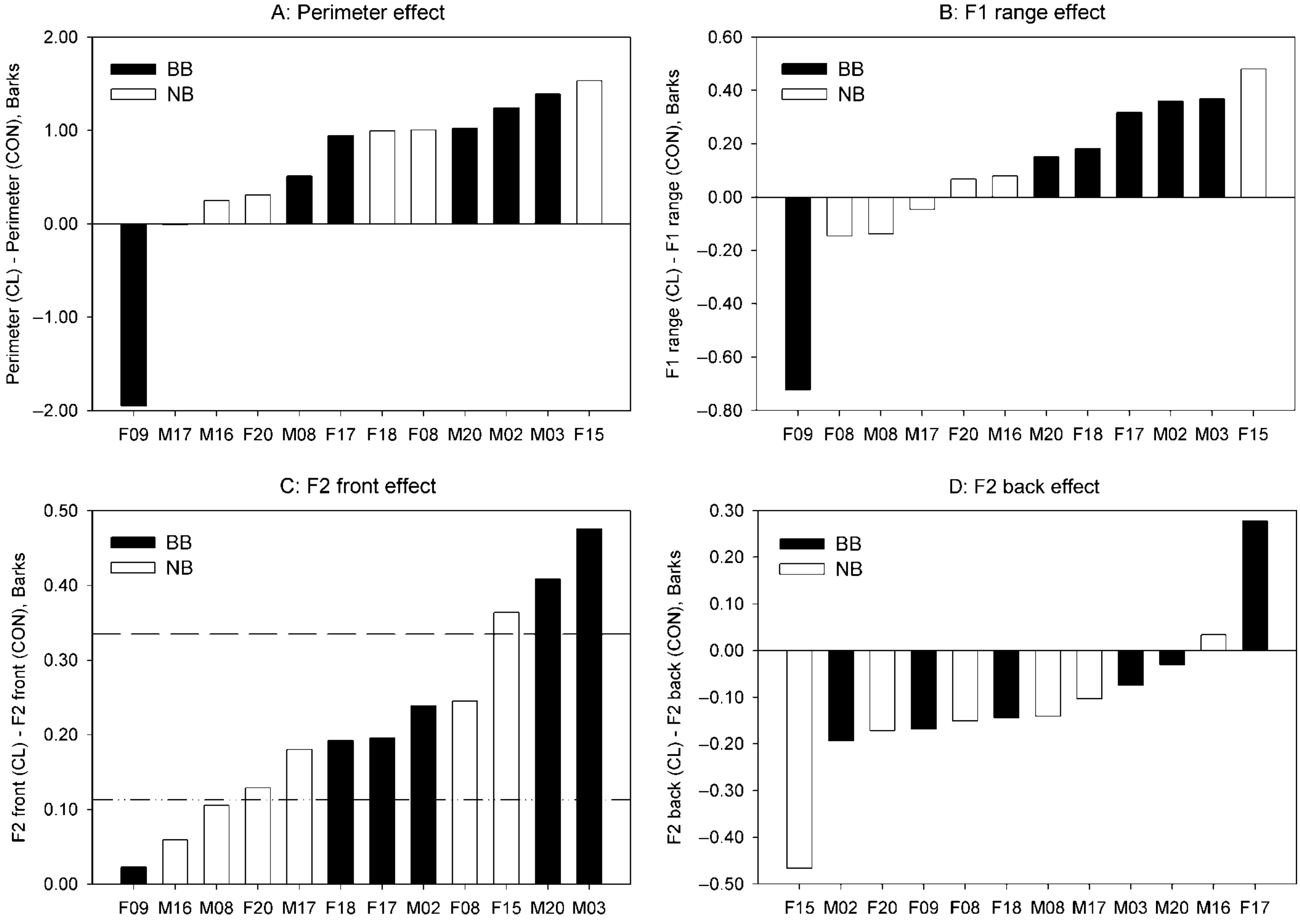
Steady-State Formant Measures

Perimeter. Four-point vowel spaces in clear and conversational speech for individual talkers are shown in Figures 4 and 5 for groups BB and NB, respectively. Input data for the two-way repeated measures ANOVA were perimeter values for each talker and in each style. Group means in each speaking style are shown in Table 2. On average, vowel spaces were significantly larger in clear speech, $F(1, 8) = 65.57, p < .01$. Vowel space dimensions did not differ significantly for the two talker groups, $F(1, 8) = 0.66, p = .44$. However, the Style × Group interaction was significant, $F(1, 8) = 13.89, p < .01$, indicating that the clear speech effect was significantly greater for group BB.

F1 range. Input data for this analysis were F1 range values for each talker in each style. Group means in each speaking style are shown in Table 2. Across groups, F1 range was significantly greater in clear speech, $F(1, 8) = 13.02, p < .01$. Across speaking styles, F1 range was not significantly different for the two groups, $F(1, 8) = 0.497, p = .5$. However, the Group × Style interaction was significant, $F(1, 8) = 22.23, p < .01$. Although F1 range was significantly larger in clear speech for group BB in planned comparisons, $F(1, 8) = 34.63, p < .01$, essentially no difference was observed for group NB, $F(1, 8) = 0.61, p = .46$.

F2 front. Input data for this analysis were average F2 values for the individual front vowels /i/, /ɪ/, /e/, /ɛ/,

Figure 2. Clear-minus-conversational differences for individual talkers for four steady-state acoustic metrics. All measures are in Barks. In each panel, talkers are ordered by ascending clear-minus-conversational difference for the metric shown. In panel C, dashed-dotted line indicates minimal detectable formant shifts for synthetic vowels under minimal stimulus uncertainty (Kewley-Port & Zheng, 1999), and dashed line indicates minimal detectable formant shifts for natural vowels under more ordinary quiet listening conditions (Liu & Kewley-Port, 2004).



and /æ/ (five values for each talker in each style); group averages in each style are shown in the middle row of Table 2. F2 values were significantly higher in clear speech, $F(1, 48) = 40.71, p < .01$, but did not differ for the two talker groups, $F(1, 48) = 0.29, p = .59$. Once again, the Group × Style interaction was significant, $F(1, 48) = 5.12, p < .05$, demonstrating that the clear speech effect was significantly larger for the BB talkers than for the NB talkers.

F2 back. F2 values for the five individual back vowels for each talker in each style (/ɑ/, /ʌ/, /o/, /ʊ/, and /u/) served as input data for this analysis. Group means in each speaking style are shown in Table 2. F2 back did not differ significantly between speaking styles, $F(1, 48) = 1.76$, $p = .19$, or between groups, $F(1, 48) = 2.1, p = .15$. Furthermore, the Style × Group interaction was not significant, $F(1, 48) = 0.49, p = .49$.

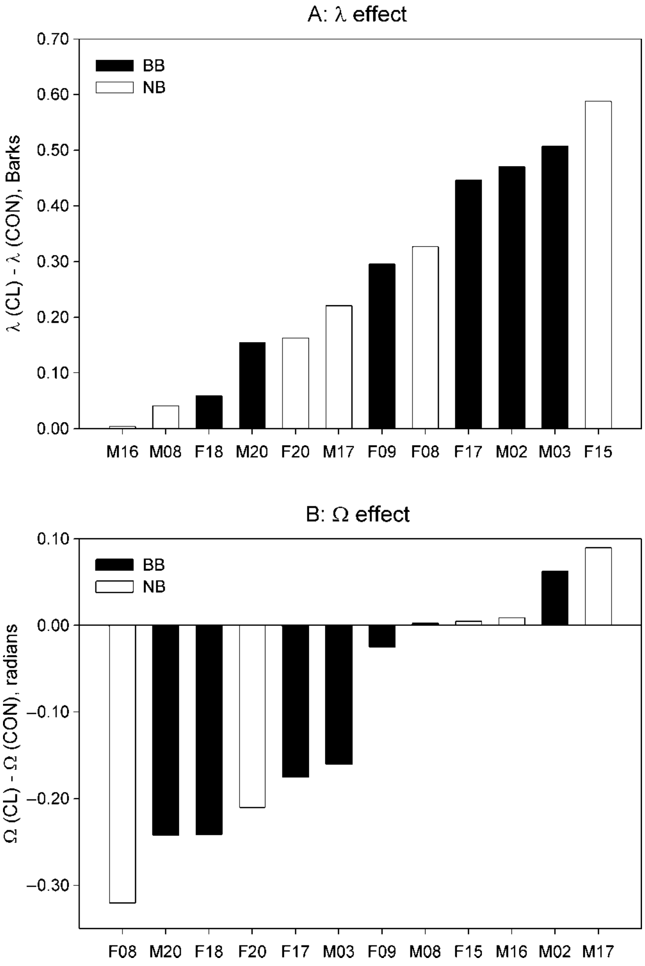
Summary of steady-state measures. When atypical talkers F09 and F15 were excluded from the statistical analyses, significant main effects of speaking style and significant Style × Talker group interactions were observed for the metrics F2 front, F1 range, and perimeter. Similar to vowel duration, each of these metrics was significantly greater in clear speech. Furthermore, the size of the clear speech effect on each of these metrics was significantly greater for talkers in group BB than for those in group NB. Neither the main effects nor the interaction were significant for the metric F2 back.

Dynamic Metrics

Input data for each two-way ANOVA consisted of the value of the dynamic metric for each token of the five vowels /e/, /ɑ/, /ʌ/, /o/, and /ʊ/ in each speaking style for each talker.

Spectral change. Average F1 and F2 spectral change (λ) values for each talker group in each speaking style are

Figure 3. Clear-minus-conversational differences for individual talkers for dynamic formant metrics spectral change (λ) and spectral angle (Ω). In each panel, talkers are ordered by ascending clear-minus-conversational differences for the metric shown.



shown in Table 2. As expected, the five vowels showed significantly greater spectral change in clear speech, $F(1, 98) = 16.23, p < .01$. Surprisingly, the main effect of group was also significant, $F(1, 98) = 10.86, p < .01$, with larger λ values for the NB talkers than for the BB talkers. The Style × Group interaction, in contrast, was not significant, $F(1, 98) = 2.2, p = .14$.

Spectral angle. Average spectral angle (Ω) values for each group and style are shown in Table 2. In contrast with all other metrics, Ω was significantly smaller in clear speech, $F(1, 98) = 15.88, p < .01$. Neither the main effect of group nor the Style × Group interaction was significant, $F(1, 98) = 3.06, p = .08$, and $F(1, 98) = 1.21, p = .27$, respectively.

Summary of dynamic data. Although the two dynamic metrics differed significantly for the two speaking styles, the size of the clear speech effect on each metric was the same for the two groups of talkers. Thus, most talkers, both those who showed large clear speech vowel intelligibility effects and those who had no clear speech vowel intelligibility effect, changed the dynamic characteristics of their vowels in clear speech.

Discussion

This experiment measured acoustic characteristics of vowels in clear and conversational speech for two groups of six talkers. For the talkers in group BB, the average clear speech vowel intelligibility benefit for listeners with normal hearing listening in noise was approximately 19 percentage points; talkers in group NB showed essentially no difference between clear and conversational speech. Several clear speech characteristics considered potentially important to vowel intelligibility were measured. The acoustic measures included vowel duration, four metrics based on steady-state F1 and F2 frequencies, and two metrics intended to capture dynamic formant movement. Statistical analyses were conducted according to the premise that if a particular clear speech acoustic change was greater for the BB talkers than for the NB talkers, then that change contributed to improved vowel intelligibility in clear speech. Conversely, if a significant clear speech acoustic change was the same for the two groups, that particular clear speech strategy would be considered less likely to enhance intelligibility. Our results suggested that although vowel space expansion and increased vowel duration were associated with improved vowel intelligibility in clear speech, increased dynamic formant movement did not by itself improve intelligibility.

Atypical Talkers

As mentioned above, these results were obtained only when an atypical talker was excluded from each group. Figures 1, 2, and 3 show that Talkers F15 and F09 differed distinctly from the other talkers in groups NB and BB, respectively. The rationale for excluding these talkers from the statistical analyses is related to the multidimensional nature of speech. Given the large number of features that can vary during speech production, it seems likely that individual talkers might take very different approaches to speaking clearly. This scenario is particularly likely when talkers receive only very general clear speech instructions, as in the current study. In fact, considering both the limited constraints on the present talkers and the talker variability described in previous clear speech acoustic and kinematic analyses (Krause & Braida, 2004; Perkell, Zandipour, Matthies, & Lane, 2002; Smiljanić & Bradlow, 2005), it was certainly possible that no patterns differentiating successful and unsuccessful

Table 2. Average values in clear (CL) and conversational (CON) speech, and the clear-minus-conversational difference, for seven acoustic metrics.

Metric	Group	CL	CON	CL – CON difference
Duration (ms)	BB	267.14	176.38	90.76
	NB	277.19	214.83	62.36
Perimeter (Barks)	BB	13.77	12.65	1.12
	NB	14.26	13.85	0.41
F1 range (Barks)	BB	3.07	2.79	0.28
	NB	3.12	3.16	−0.04
F2 front (Barks)	BB	13.13	12.82	0.31
	NB	13.2	13.06	0.14
F2 back (Barks)	BB	10.13	10.26	−0.13
	NB	9.75	9.86	−0.11
Spectral change (λ, Barks)	BB	1.88	1.56	0.32
	NB	2.27	2.12	0.15
Spectral angle (Ω, radians)	BB	0.89	1.04	−0.15
	NB	1.03	1.12	−0.09

Note. Talker F09 from group BB (big benefit) and talker F15 from group NB (no benefit) were excluded in calculating these mean values.

clear speech strategies would be found. If anything, the consistency that emerged after removing the atypical talkers is more remarkable than the observation that two talkers differed from the others in their groups.

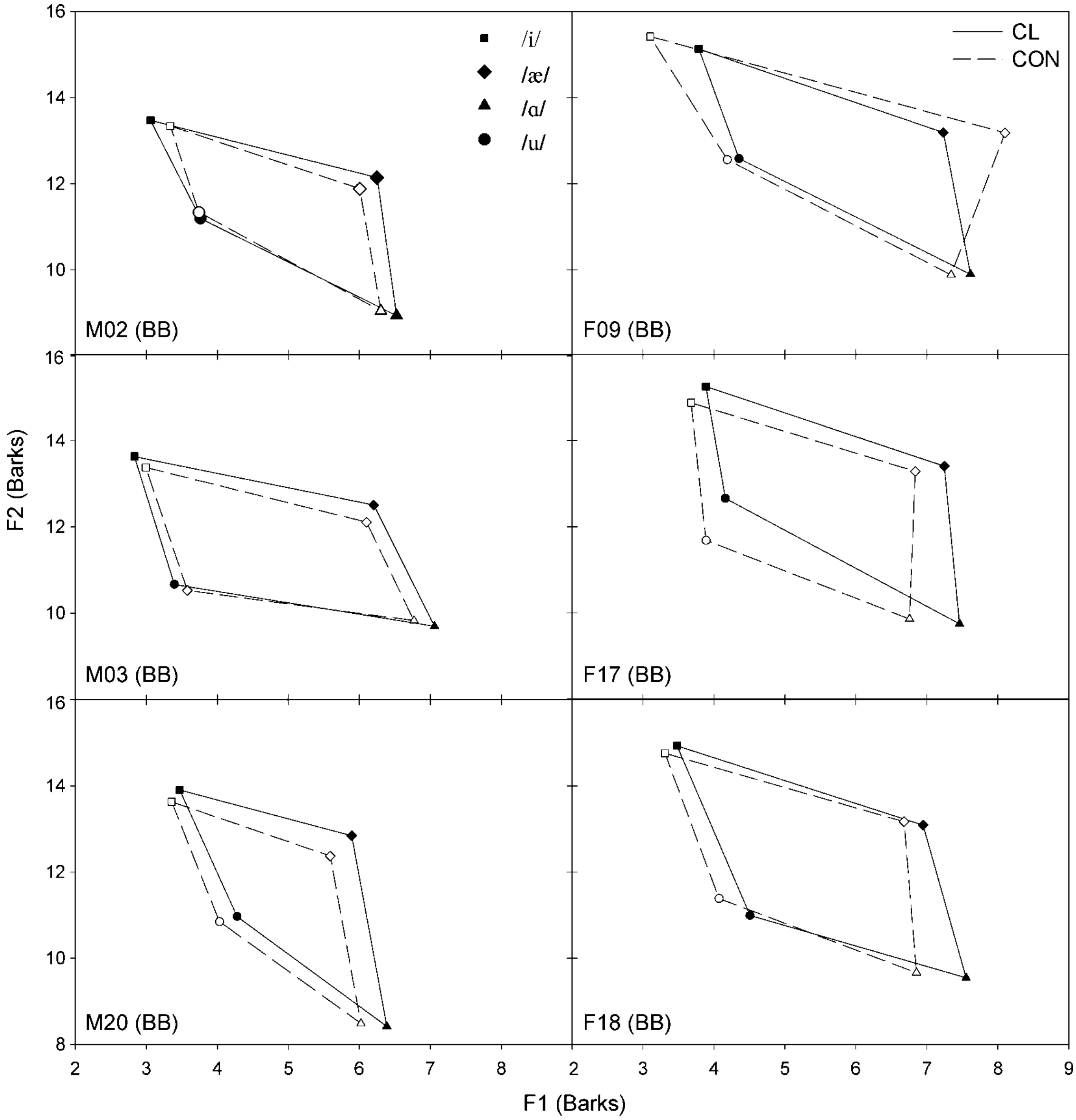
Demographic and other data for the atypical talkers were examined in search of an explanation for their unusual results. It was observed that Talker F15, who showed no clear speech benefit despite substantial vowel space expansion, had the highest conversational vowel intelligibility (81%) of the 12 talkers examined here. However, the Ferguson database (2004) included 2 talkers who had both higher conversational vowel intelligibility scores ($M = 83\%$) and a significant clear speech benefit ($M = 6$ percentage points). Talker F09, who showed a large clear speech benefit despite having a smaller vowel space in clear speech, differed from her group in a slightly more interesting way: She had grown up with two hearing-impaired siblings and thus had extensive experience communicating with people with hearing loss. Although Ferguson (2004) found no difference in the clear speech vowel intelligibility effect for talkers with varying degrees of this experience, it may be that certain talkers develop clear speech strategies that are particularly effective for these listeners. For example, F09 was the only talker in this experiment who did not raise F2 for front vowels, a strategy that Ferguson and Kewley-Port (2002) found negatively affected front vowel intelligibility for listeners with hearing impairment. Perhaps F09's previous experience taught her that raising F2 was ineffective and led her to develop other strategies to improve intelligibility for listeners with hearing impairment. Future acoustic analyses will seek to determine the exact nature of these strategies.

Duration

As expected, vowels were significantly longer in clear speech. Averaged across all vowels and talkers, the clear/conversational duration ratio was 1.4. The slowness of clear speech is perhaps the most robust finding in studies of clear speech acoustics. For example, in Picheny et al. (1986) and in Moon and Lindblom (1994), vowels were roughly 1.3 times longer in clear speech. In Ferguson and Kewley-Port (2002), the ratio was 2.1. The significant Style × Group interaction observed in the present study suggested that greater vowel duration increases were associated with greater vowel intelligibility improvements. Similarly, in Bradlow et al. (2003), the talker with the larger speaking rate reduction had the largest clear speech benefit. These results are also consistent with Bond and Moore (1994), who found that within a single speaking style, more intelligible talkers tended to use longer word and vowel durations than did less intelligible talkers.

These results conflict, however, with three studies in which slowing down conversational speech did not improve intelligibility: Picheny et al. (1989), Uchanski et al. (1996), and Nejime and Moore (1998). One reason

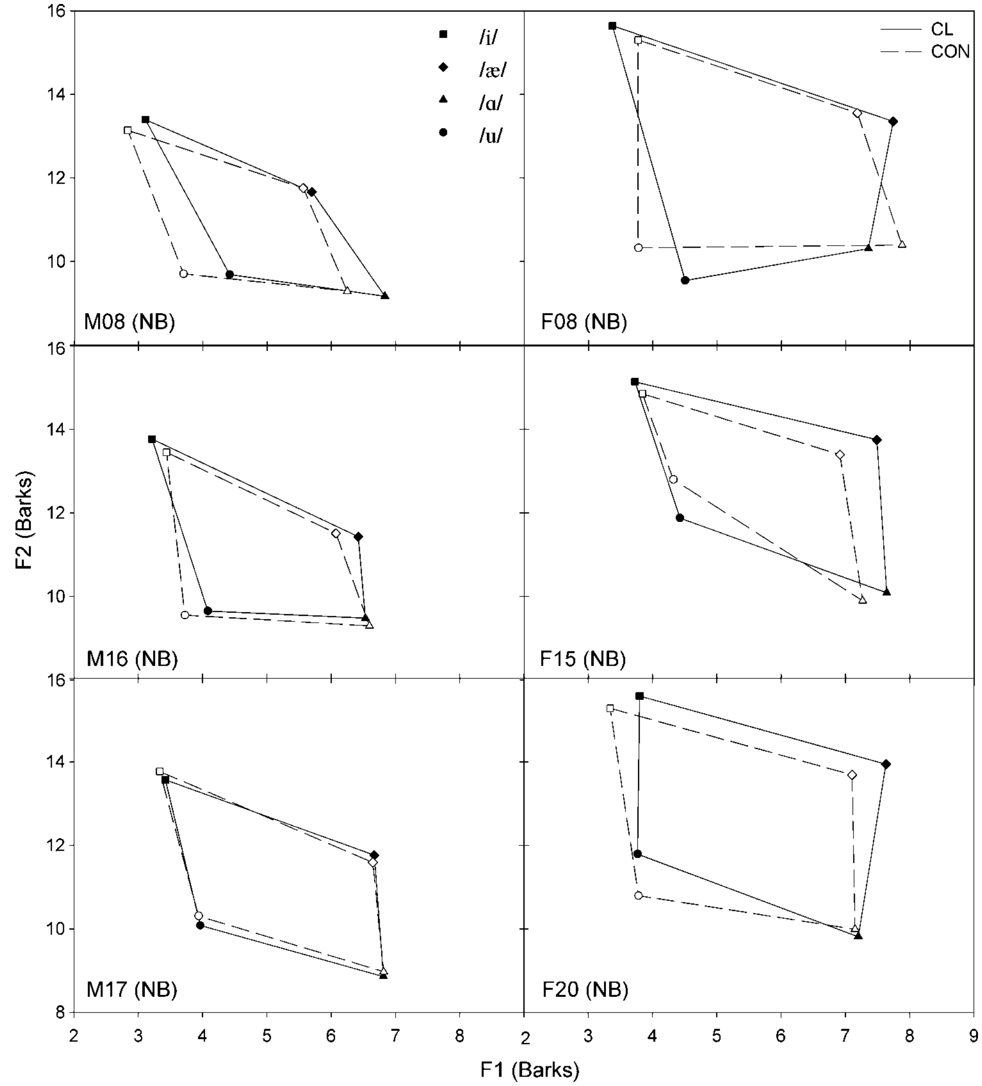
Figure 4. Four-point vowel spaces for talkers in group BB. CL = clear; CON = conversational.



for the varied results may be methodological. In contrast with these three studies, which used a signal processing approach to explore the relationship between speaking rate changes and intelligibility changes, the current study as well as Bond and Moore (1994) and Bradlow et al. (2003) did so by observing differences among talkers. It is possible that signal processing artifacts in these three studies cancelled any benefit associated with slowing, per se. However, because these studies usually included control conditions where expanded speech was compressed back to its original rate, this seems unlikely. The results of Bradlow, Torretta, and Pisoni (1996), a multitalker study that found no correlation between speaking rate and intelligibility within a single speaking style, further argue against this methodological explanation.

A close examination of Figure 1 further muddies the waters. Across groups, the clear speech vowel duration effect varied widely among the talkers. The significant Style × Group interaction seems to have resulted primarily from three NB talkers (M08, M16, and M17) who made virtually no vowel duration changes in clear speech. The other NB talkers actually showed larger vowel duration increases than half of the BB talkers. Thus, although statistical analyses suggested that larger increases in vowel duration were associated with greater vowel intelligibility, there were talkers who made substantial vowel duration increases without improving their

Figure 5. Four-point vowel spaces for talkers in group NB.



vowel intelligibility. Further research is needed to clarify the role of speaking rate in clear speech.

Steady-State Formant Measures

Perimeter. Consistent with previous studies (Bradlow et al., 2003; Johnson et al., 1993), overall vowel space dimensions were larger in clear speech. In addition, Figure 8 in Picheny et al. (1986) suggests that all three talkers increased their vowel space to at least some extent. Vowel space expansion, therefore, appears to be a fairly robust strategy when talkers are instructed to speak clearly. However, the current data suggest that the increase needs to be of a sufficient magnitude to increase vowel intelligibility for listeners with normal hearing. Although both groups expanded their vowel spaces in clear speech, group BB did this to a significantly greater degree than did group NB (see Figures 2, 4, and 5). On average, BB talkers increased their vowel space perimeters by 1.1 Barks, or almost 9%. For the NB talkers, the average perimeter increase was 0.41 Barks, or 3%.

New analyses of data from Ferguson and Kewley-Port (2002) offer additional support for this interpretation. Although their talker expanded his vowel space in clear speech, the percentage increase in the four-point vowel space perimeter (5%) was considerably less than that observed for group BB (9%). This talker also had a smaller clear speech vowel intelligibility benefit (only 18 RAUs versus 23.9 RAUs for group BB), perhaps due to his relatively smaller vowel space expansion.

F1 range. The F1 difference between high and low vowels was significantly greater in clear speech but only for the BB talkers. New analyses of data from Ferguson and Kewley-Port (2002) revealed a similar F1 range expansion. However, the principal F1 clear speech effect reported in that article was a significant F1 increase for 8 of 10 vowels. To test whether this F1 shift may be an important clear speech strategy, a two-way repeated measures ANOVA was performed on the F1 values for all vowels (without atypical talkers). Consistent with Ferguson and Kewley-Port (2002), F1 values were significantly higher in clear speech, $F(1, 118) = 39.63, p < .001$. However, the effect of group, $F(1, 118) = 1.251, p = .27$, and the Style × Group interaction, $F(1, 118) = 0.02$, $p = .88$, were not significant. This suggests that, in contrast with F1 range expansion, a simple increase in F1 in clear speech has a limited effect on vowel intelligibility.

F2 measures. As in Ferguson and Kewley-Port (2002), the average front vowel F2 value was significantly higher in clear speech. Although planned contrasts showed that the F2 increase was significant for both groups, the significant Style × Group interaction suggests that the shift the NB talkers produced was insufficient to improve vowel intelligibility. To support this interpretation, lines are included in panel C of Figure 2 that permit comparisons between individual talkers' front vowel F2 shifts and formant discrimination data from listeners with normal hearing tested under minimal stimulus uncertainty (0.11 Barks; Kewley-Port & Zheng, 1999) and for natural vowels in sentences (0.33 Barks; Liu & Kewley-Port, 2004). The dashed line shows that only the atypical NB talker, F15, produced an F2 shift that would be detectable in a sentence context. Of the five typical NB talkers, only three produced an F2 shift that could be detected even under optimal listening conditions (dash-dotted line). In contrast, all five of the typical BB talkers (excluding F09) raised front vowel F2 by an amount detectable in optimal listening, and two produced a shift that was detectable in more ordinary listening conditions.

In contrast with F2 for front vowels, F2 for back vowels did not differ significantly between the two speaking styles, nor was the Style × Group interaction significant. This result reflects a front/back contrast that was observed for group BB but not group NB. For the BB talkers, the clear speech F2 shift was more than twice as large for front vowels (0.3 Barks) as for back vowels (0.12 Barks). For the NB talkers, conversely, the F2 shift was only slightly larger for front vowels (0.14 Barks) than for back vowels (0.11). A front/back contrast similar to that observed for the BB talkers was reported by Ferguson and Kewley-Port (2002). For their talker, F2 for front vowels was an average of 0.6 Barks higher in clear speech; in contrast, the F2 shift for back vowels was only 0.16 Barks.

General discussion of steady-state metrics. Overall, these analyses suggest that when instructed to speak as though talking to a hearing-impaired person, most talkers expand their vowel space to at least some degree. Although this is consistent with the notion that clear speech is a hyperspeech form (Lindblom, 1990), the expansion is neither uniform nor simple (see Figures 2, 4, and 5). First, the expansion is greater on the F2 dimension than on F1, even after Bark transformation. Averaged across groups, the F2 space expanded by 0.34 Barks (the front and back vowel F2 shifts, summed), whereas the F1 space expanded by just 0.12 Barks. Second, the F2 expansion is due primarily to the front vowels. Across groups, the average F2 shift was 0.22 Barks for front vowels and 0.12 Barks for back vowels. Third, when the F1 expansion occurs, it is accomplished by an overall F1 increase that is greater for low vowels (e.g., /æ/ and /ɑ/) than for high vowels (e.g., /i/ and /u/).

Another important outcome of this experiment is that the amount of vowel space expansion varies among talkers. Specifically, talkers with varying amounts of clear speech vowel intelligibility benefit differed in several respects. First, the amount of overall vowel space expansion was greater in BB talkers (9%) than NB talkers (3%). Second, BB talkers had a larger F2 shift for front vowels than NB talkers (0.3 Barks vs. 0.12 Barks, respectively). Indeed, only the BB talkers showed the relatively larger F2 shift for front vowels versus back vowels described above. Finally, although all talkers had higher overall F1 values in clear speech, only the BB talkers expanded their F1 range (by 0.28 Barks versus just 0.05 Barks for the NB talkers). These results suggest that for listeners with normal hearing to enjoy improved vowel intelligibility in clear speech, two criteria must be met: (a) an expanded F1 range and (b) higher F2 values for front vowels.

A close examination of Table 2 suggests another interpretation of the significant Style × Group interactions observed for the perimeter, F1 range, and F2 front measures. For all three metrics, the mean values for the NB talkers in conversational speech are very similar to those of the BB talkers in clear speech. This raises the possibility that the NB talkers were already using clear speech values in their conversational speech, leaving no room to expand their vowel space when speaking clearly. However, examination of individual perimeter, F1 range, and F2 front values in each speaking style argues otherwise. For example, consider two talkers: M20 (a BB talker) and M16 (an NB talker). In conversational speech, M20 and M16 had very similar F2 front (12.3 Barks for both talkers) and perimeter values (12.4 and 12.3 Barks, respectively). However, although M20 increased his F2 front and perimeter values in clear speech (by .41 and 1.0 Barks, respectively), M16 made minimal increases (.06 and .25 Barks). For all three metrics, the fact that BB and NB talkers showed considerable overlap in conversational speech suggests that the NB talkers had as much room to expand their vowel spaces as the BB talkers. Vowel intelligibility data in conversational speech show similar overlap among the NB and BB talkers, further arguing against the idea that the NB talkers were already producing their best speech in this style.

Dynamic Metrics

The main effect of speaking style was significant for both dynamic metrics. Averaged across the five most dynamic vowels (/e/, /ɑ/, /ʌ/, /o/, and /ʊ/), spectral change (λ) was significantly greater in clear speech. This was consistent with an a priori expectation that talkers would produce more dynamic vowels when speaking clearly. Unexpectedly, however, spectral angle (Ω) was significantly smaller in clear speech. This was particularly surprising in light of results reported by Wouters and Macon (2002), who found increased spectral rate of change in hyperarticulated speech using a metric very similar to Ω. Moon and Lindblom (1994) also found a higher F2 rate of change in clear speech.

One important difference between the two dynamic metrics used here is that the spectral angle calculation includes duration. Spectral change, a purely spectral metric, demonstrates unequivocally that the amount of formant movement between the 20% and 80% points of the vowel was greater in clear speech. That an opposite effect was found for spectral angle suggests that this increased movement was offset by an increase in vowel duration. This explanation is supported by a significant negative correlation between clear-minus-conversational duration differences and clear-minus-conversational spectral angle differences for individual vowels ($r = -.48$, $p < .01$): Vowels with large positive duration differences (clear > conversational) tended to have large negative spectral angle differences (clear < conversational).

If the positive correlation between duration and spectral change yields decreases in spectral angle for clear speech, why did earlier studies (i.e., Moon & Lindblom, 1994; Wouters & Macon, 2002) find increases in spectral rate of change, a metric very similar to spectral angle, in hyperarticulated speech? One possible explanation is methodological differences between the current and previous studies. In contrast with the current experiment, which measured formant movement during the vowel nucleus, Wouters and Macon (2002) measured spectral rate of change for consonant–vowel formant transitions, diphthong transitions, and vowel–consonant transitions. Furthermore, Wouters and Macon (2002) included F3 transitions in their metric along with changes in F1 and F2. This is significant because all of the consonants in their study were liquids, including /r/. The distinct F3 characteristics of /r/ would necessarily create very large consonant–vowel F3 transitions. Similarly, Moon and Lindblom (1994) measured F2 rate of change only for front vowels preceded by /w/.

The current data suggest that increases in formant movement and increases in duration in clear speech are linked. It may be that talkers who make their vowels more dynamic in clear speech deliberately slow down to avoid overshooting the formant frequency targets for individual vowels. Alternatively, talkers may set out to speak more slowly and may find that this strategy permits them to also make vowels more dynamic. Whatever the reasons for this link, the absence of any significant Style × Group interactions for the dynamic metrics suggests that making vowels more dynamic actually does not differentially affect vowel intelligibility. This result was surprising because Ferguson and Kewley-Port (2002) found that formant dynamic changes in the clear speech of their single talker was important for some vowels. These conflicting results appear to comprise another instance where talkers vary considerably in how they alter the multidimensional properties of speech when speaking clearly. Further analyses of the Ferguson database might help clarify the relationship between dynamic formant movement and vowel intelligibility, both between and within the two speaking styles. Measures of dynamic formant movement different from those used here might also shed additional light on this relationship.

Conclusion

This experiment examined vowels in clear and conversational speech for 6 talkers who showed a large clear speech vowel intelligibility benefit (the BB talkers) and 6 who showed no vowel intelligibility difference between clear and conversational speech (the NB talkers). The strategy behind this extreme groups design was to determine which vowel acoustic changes from conversational to clear speech contributed to the improved vowel intelligibility observed in group BB. Two of the clear speech features studied here, increased vowel duration and vowel space expansion, were associated with increased vowel intelligibility in clear speech. Although both groups made their vowels longer, expanded their vowel spaces, and raised F2 for front vowels in clear speech, these increases were significantly greater for the BB talkers than for the NB talkers. In addition, only the BB talkers expanded their vowel spaces on the F1 dimension.

Note that these vowel space expansion patterns were not observed uniformly within the two groups. Group BB contained 1 talker (F09) whose vowel space was actually smaller in clear speech but who still produced a large clear speech vowel intelligibility benefit. Group NB contained 1 talker (F15) who showed significant vowel space expansion but no vowel intelligibility benefit in clear speech. In addition, the duration data showed overlap between the groups. These results illustrate the multidimensional nature of speech and suggest that a variety of acoustic factors can affect vowel intelligibility, probably not all of which were measured here. Considering the nonspecificity of the clear speech instructions and the wide array of acoustic characteristics that can vary during speech production, it is not surprising that talkers might accomplish a clear speech benefit in different ways. If anything, the robustness of the vowel space expansion results for the other 5 talkers in each group was the surprising outcome of this experiment.

In contrast with the duration and steady-state metrics, the two groups behaved very similarly on the formant dynamic metrics. Averaged across the talkers in each group, the amount of change in vowel formant movement in clear speech was similar for the NB and BB talkers. For both metrics, each group included a talker who made their vowels much more dynamic in clear speech as well as at least 1 talker who made only minimal changes to formant movement. These results suggest that altering the dynamic formant characteristics of vowels contributes little to improved vowel intelligibility in clear speech, at least when quantified using the present dynamic formant metrics.

The degree to which the vowel acoustic changes observed here are universal features of clear speech can be explored by measuring these characteristics for all 41 Ferguson database talkers (2004). In addition, correlational analyses could be used to examine the direct contribution of various acoustic properties to vowel intelligibility in clear speech, especially vowel duration. As in the current study, these analyses could focus on the acoustic and intelligibility changes observed (or not) in each talker's clear speech. Alternatively, relationships between the specific values of the acoustic characteristics and intelligibility (rather than between acoustic and intelligibility clear speech differences) could be examined without regard to speaking style, treating the intelligibility score for each talker in each style as a single data point. Such an analysis would not only provide insight about the nature of vowel intelligibility in clear speech but also might help explain why talkers differ so widely in vowel intelligibility, even within a single speaking style.

Another important application of the Ferguson database would be to extend these analyses to the group of listeners for whom clear speech is intended: individuals with hearing loss. The results of Ferguson and Kewley-Port (2002) suggest that some clear speech acoustic changes that benefit listeners with normal hearing, particularly raising F2 for front vowels, can negatively affect vowel intelligibility in listeners with hearing impairment who are listening without amplification. Both groups of talkers in the current experiment significantly increased F2 for these vowels, the BB talkers more so than the NB talkers. Any generalizations from listeners with normal hearing to listeners with hearing impairment about which acoustic features actually make clear speech more intelligible should be viewed with suspicion until further studies are conducted with listeners with hearing impairment.

Acknowledgments

This research was conducted as part of a doctoral dissertation completed at Indiana University and was supported by Grant NIHDCD-02229 from the National Institute on Deafness and Other Communication Disorders. We gratefully acknowledge Kenneth de Jong, Larry E. Humes, David B. Pisoni, and Charles S. Watson for their contributions to this research.

References

Assmann, P. F., & Katz, W. F. (2005). Synthesis fidelity and time-varying spectral change in vowels. *The Journal of the Acoustical Society of America, 117,* 886–895.

Bond, Z. S., & Moore, T. J. (1994). A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Communication, 14,* 325–337.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America, 112,* 271–283.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research, 46,* 80–97.

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.

Di Benedetto, M.-G. (1989). Vowel representation: Some observations on temporal and spectral properties of the first formant frequency. *The Journal of the Acoustical Society of America, 86,* 55–66.

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

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 112,* 259–271.

Gagné, J.-P., Masterson, V. M., Munhall, K. G., Bilida, N., & Querengesser, C. (1994). Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology, 27,* 135–158.

Gordon-Salant, S. (1986). Recognition of natural and time-altered CVs by young and elderly subjects with normal hearing. *The Journal of the Acoustical Society of America, 80,* 1599–1607.

Gordon-Salant, S. (1987). Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects. *The Journal of the Acoustical Society of America, 81,* 1199–1202.

Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology, 9,* 234–242.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America, 97,* 3099–3111.

Hillenbrand, J., & Nearey, T. M. (1999). Identification of resynthesized /hVd/ utterances: Effects of formant contour. *The Journal of the Acoustical Society of America, 105,* 3509–3523.

Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language, 69,* 505–528.

Kewley-Port, D., & Zheng, Y. (1999). Vowel formant discrimination: Towards more ordinary listening conditions. *The Journal of the Acoustical Society of America, 106,* 2945–2958.

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

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America, 115,* 362–378.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, the Netherlands: Kluwer Academic.

Liu, C., & Kewley-Port, D. (2004). Vowel formant discrimination for high-fidelity speech. *The Journal of the Acoustical Society of America, 116,* 1224–1233.

Loizou, P. C. (1999). *COLEA: A MATLAB software tool for speech analysis*. Retrieved February 26, 2002, from http://www.utdallas.edu/~loizou/speech/colea.htm

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

Nejime, Y., & Moore, B. C. J. (1998). Evaluation of the effect of speech-rate slowing on speech intelligibility in noise using a simulation of cochlear hearing loss. *The Journal of the Acoustical Society of America, 103,* 572–576.

Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America, 95,* 1581–1592.

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). 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, 112,* 1627–1641.

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.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing. III: An attempt to determine the contribution of speaking rate to difference in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research, 32,* 600–603.

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America, 118,* 1677–1688.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28,* 455–462.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America, 88,* 97–100.

Uchanski, R. M., Choi, S. S., Braida, L. D., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing. IV: Further studies of the role of speaking rate. *Journal of Speech and Hearing Research, 39,* 494–509.

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *The Journal of the Acoustical Society of America, 106,* 458–468.

Wouters, J., & Macon, M. W. (2002). Effects of prosodic factors on spectral dynamics. I: Analysis. *The Journal of the Acoustical Society of America, 111,* 417–427.

Received August 28, 2006

Accepted March 7, 2007

DOI: 10.1044/1092-4388(2007/087)

Contact author: Sarah Hargus Ferguson, Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas, Dole Center, 1000 Sunnyside Avenue, Room 3001, Lawrence, KS 66045-7555. E-mail: safergus@ku.edu.