Published in final edited form as:

J Speech Lang Hear Res. 2012 June ; 55(3): 779–790. doi:10.1044/1092-4388(2011/10-0342).

Talker Differences in Clear and Conversational Speech: Vowel Intelligibility for Older Adults with Hearing Loss

Sarah Hargus Ferguson
University of Kansas

Abstract

Purpose—To establish the range of talker variability for vowel intelligibility in clear versus conversational speech for older adults with hearing loss and to determine whether talkers who produced a clear speech benefit for young normal-hearing listeners also do so for older adults with hearing loss.

Method—Clear and conversational vowels in /bVd/ context produced by 41 talkers were presented in noise for identification by 40 older adults with sloping sensorineural hearing loss.

Results—Vowel intelligibility within each speaking style and the size of the clear speech benefit varied widely among talkers. The clear speech benefit was equivalent to that enjoyed by young normal-hearing listeners in an earlier study. Most talkers who had produced a clear speech benefit for young listeners with normal hearing also did so for the current older listeners with hearing loss. However, effects of talker gender differed between listeners with normal hearing and listeners with hearing loss.

Conclusion—The clear speech vowel intelligibility benefit generated for listeners with hearing loss varies considerably among talkers. Most talkers who produce a clear speech benefit for normal-hearing listeners also produce a benefit for listeners with hearing loss.

Difficulty understanding speech is the primary complaint that leads individuals with hearing loss to seek hearing help, and thus the primary goal of audiologic rehabilitation is to improve everyday speech understanding. While hearing aids are a very important component of this rehabilitation, market surveys have consistently found that fewer than 25% of people with hearing impairments have hearing aids (Kochkin, 2009). Further, while over 80% of hearing aid users report satisfaction with the overall benefit they receive, only half say they are satisfied with the benefit received in noise (Kochkin, 2005). This suggests that to maximize speech understanding for the largest number of individuals with hearing loss, audiologists should offer other rehabilitative services. A recent survey of clinical practice showed that most audiologists do so (Prendergast & Kelley, 2002). For example, over 80% of audiologists surveyed reported offering information on assistive listening devices and training in communication strategies. Some audiologists (38%) also provide training for the frequent communication partners of their clients with hearing loss, including “speaking behaviors” intended to reduce the frequency of such breakdowns. A typical recommendation is that the communication partner “speak clearly and slowly” (Tye-Murray, 2004, p. 136).

A number of laboratory studies have demonstrated that this recommendation should lead to improved speech understanding. In these studies, talkers read printed speech materials, first in a conversational manner and later under instructions to speak as though talking to

someone who has difficulty understanding them. When presented to listeners with hearing loss for identification, “clear speech” is significantly more intelligible than conversational speech. This “clear speech benefit” has been found for sentence materials (e.g., Picheny, Durlach & Braid, 1985; Schum, 1996; Uchanski, Choi, Braid, & Durlach, 1996) and for single words (Uchanski et al., 1996). A clear speech benefit has also been found for individuals with normal hearing in degraded listening conditions identifying sentences (e.g., Liu, Del Rio, Bradlow, & Zeng, 2004; Payton, Uchanski, & Braid, 1994), words (e.g., Gagné, Masterson, Munhall, Bilida, & Querengesser, 1994), and consonant-vowel and vowel-consonant-vowel stimuli (Gagné, Rochette, & Charest, 2002). While the clear speech benefit observed for sentences and words has been similar for listeners with normal hearing and listeners with hearing loss (e.g., Payton et al., 1994), a study using vowel materials produced by a single talker (Ferguson & Kewley-Port, 2002) found divergent results for the two listener groups. Although vowels in /bVd/ context presented in a background of 12-talker babble were significantly more intelligible in clear speech than in conversational speech for listeners with normal hearing, no such clear speech vowel intelligibility benefit occurred for listeners with hearing loss.

Ferguson and Kewley-Port (2002) conceded that their surprising result “...begs the question of whether the strategies our talker employed are unique to him, or whether they represent a typical response to instructions to speak as though talking to a hearing-impaired listener.” (p. 270). This question has been difficult to answer, given that no previous study has directly assessed talker differences in the clear speech effect for listeners with hearing loss. However, differences among talkers are apparent in studies that included multiple talkers. In the largest such study, Schum (1996) presented sentences produced by 10 young and 10 elderly talkers to older adults with sloping sensorineural hearing loss. While no differences were observed between younger and older talkers, a table showing individual scores for each talker reveals a wide range of clear speech effects, from as low as 4 to as high as 45 rationalized arcsine units (RAU; Studebaker, 1985). Talker variability in the size of the clear speech benefit for listeners with hearing loss can also be seen in Picheny et al. (1985) and in Uchanski et al. (1996), which used the same three talkers.

Using a database of 41 talkers, Ferguson (2004) demonstrated that the clear speech vowel intelligibility benefit enjoyed by young listeners with normal hearing varies widely among talkers. The present study extended this work to one of the clinical populations for whom clear speech is intended: older adults with hearing loss. Specifically, vowel intelligibility in noise for the 41 talkers in the Ferguson Clear Speech Database was assessed for a group of 40 older adults with mild-to-moderately severe sloping sensorineural hearing loss. Talker differences in the magnitude of the clear speech vowel intelligibility effect were examined and compared to Ferguson's results for young listeners with normal hearing. The goals of the study were to establish the range of talker variability in the clear speech benefit among talkers in the 41-talker database for listeners with hearing loss, and to determine whether talkers who produce a clear speech vowel intelligibility benefit for listeners with normal hearing also do so for listeners with hearing loss. Based on the results of Ferguson (2004) and of Ferguson and Kewley-Port (2002), the hypotheses were that the clear speech benefit (a) would vary among talkers, and (b) would be smaller for listeners with hearing loss than for listeners with normal hearing.

The study also tested several hypotheses regarding the effect of talker characteristics on the clear speech benefit for listeners with hearing loss. The first considered talker gender. When materials were presented to young listeners with normal hearing in Ferguson (2004), overall vowel intelligibility was higher for female talkers than for male talkers. Listeners with hearing loss, in contrast, often report difficulty understanding female talkers (Helfer, 1995). While the superior intelligibility of female talkers for young listeners with normal hearing

has been observed elsewhere (e.g., Bradlow, Torretta, & Pisoni, 1986), no previous study has assessed the effects of talker gender on speech intelligibility for listeners with hearing loss. The hypothesis was that for the present listeners, vowels produced by male talkers would be more intelligible than those produced by female talkers. With regard to other talker characteristics, the hypotheses were that older talkers and talkers with experience communicating with listeners with hearing loss would produce a larger clear speech benefit than younger talkers and talkers with no such experience.

Methods

Materials

Test stimuli were identical to those used in the perceptual study reported by Ferguson (2004). For each of the 41 talkers in the Ferguson Clear Speech Database, two tokens each of 10 vowels (/i/, /I/, /e/, /ɛ/, /æ/, /a/, /ʌ/, /o/, /ʊ/, /u/) in /bVd/ context were selected from each speaking style (clear and conversational). The /bVd/ words were excised from the carrier sentences in which they were elicited and then scaled to have the same peak RMS amplitude. The carrier sentences were meaningful but contained minimal contextual information regarding the identity of the /bVd/ word, which was centrally located within the sentence. Examples include “Vera put the _____ on the table” and “I think the word _____ is hard for kids to say” In the conversational speech condition, which was recorded first, talkers were instructed to read the sentences aloud, speaking 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. Additional details regarding recording procedures may be found in Ferguson (2004).

As in Ferguson (2004), test stimuli were presented in a background of 12-talker babble. A 15-s sample of babble was low-pass filtered at 8500 Hz and digitized from the noise channel of a recording of the Speech Perception in Noise Test (Kalikow, Stevens, & Elliott, 1977). On each test trial, a segment of babble selected from this 15-s sample was presented along with the /bVd/ word. The speech materials were recorded and the babble was sampled at 22050 Hz; they were resampled to 24414 Hz prior to the perceptual experiment for presentation via Tucker-Davis Technologies (TDT) System III audio hardware.

Listeners

The desired number of participants ($n = 40$) was determined via power analysis of pilot data. To achieve this target, 54 older adults (65 to 87 years, 28 female) were recruited from a subject pool maintained by the author. Subject pool members received a complete audiological evaluation upon joining the pool. To be invited to participate in the present study, pool members were required to be native speakers of American English with no history of speech or language disorders. They also were required to have mild-to-moderately severe sloping sensorineural hearing losses and good word recognition abilities ($> 80\%$ correct) for the Northwestern University Test No. 6 (NU-6; Tillman & Carhart, 1966) presented in quiet at 40 dB above the speech recognition threshold.

Upon enrollment in the present study, listeners were required to have normal cognitive status as determined by a score of 25 or higher on the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975), which was administered at the beginning of the experiment. All 54 participants met this criterion. Finally, listeners were required to achieve 88% correct performance on a vowel identification familiarization task (see below). Forty (22 female) of the 54 participants reached this criterion. Their mean audiogram is shown in Figure 1; their mean NU-6 score was 96%, they ranged in age from 65 to 87 years, and their mean MMSE score was 28.75. All listeners were paid for their participation.

Procedures

All test procedures were approved by the University of Kansas Human Subjects Committee. Listeners were tested individually in a double-wall sound-treated booth, seated in front of a computer monitor and mouse. On each trial, a test word and a segment of 12-talker babble were played from separate channels of a TDT RP2 real-time processor. The babble segment, which was 1 s longer than the test word, was selected from a random location within the stored 15-s babble sample; the test word and babble segment were centered temporally. The test word and babble segment were attenuated by separate TDT programmable attenuators (PA-5) to achieve the desired overall level and signal-to-babble (S/B) ratio. The speech and babble were then mixed (TDT SM5) and routed via a headphone buffer (TDT HB-7) to an insert earphone (E-A-RTONE 3A) for monaural presentation. To identify the vowel of the test word, the listener clicked on the response category corresponding to that vowel. The ten response alternatives were displayed on the computer monitor as 10 sets of three keywords: (1) feet, thief, bead; (2) sit, rib, bid; (3) tape, raid, bade; (4) head, said, bed; (5) back, mass, bad; (6) pot, sod, bod; (7) cup, rug, bud; (8) rode, own, bode; (9) good, should, book; (10) rude, news, boot. The same keywords were used in Ferguson (2004). After selecting their response, listeners could click “OK” to confirm it or “Cancel” if they wished to change their response due to uncertainty or a mouse-clicking error. Listeners were not permitted to replay the stimulus.

Listeners were tested in 4 sessions. These sessions were scheduled at the listeners' convenience; total test periods ranged from 5 days to 6 weeks. In the first session, which lasted 2 hours, listeners were given the MMSE and familiarized with the test procedures prior to beginning the experimental conditions. For familiarization, a single 41-trial block of clear vowel tokens created for Ferguson (2004) using one /bVd/ token from each talker was used. The order of the 41 stimuli was randomized each time the block was presented. After orientation to the vowel identification task and response alternatives by the experimenter, listeners completed the familiarization block, with response accuracy feedback indicating the correct answer for each trial, at a presentation level of 85 dB SPL in quiet. This block was repeated up to four times or until a criterion of 88% correct identification was achieved. On average, listeners required 1.83 presentations of the familiarization block to reach criterion. Once criterion had been achieved, listeners performed an additional familiarization block at 70 dB SPL in quiet without feedback¹, followed by three familiarization blocks in babble without feedback. The S/B ratios for these blocks were +3 dB, 0 dB, and -3 dB. After completing the familiarization, listeners completed two test blocks to finish out the first test session. The remaining 3 test sessions took approximately 90 minutes each. Listeners completed five test blocks in sessions 2 and 3, and four test blocks in session 4.

The presentation level for the test blocks was 70 dB SPL, a level selected as a typical speech level in noisy environments (Pearsons, Bennett, & Fidell, 1977). The same level was used in Ferguson (2004). Because speech audibility differs for listeners with normal hearing and listeners with hearing loss in most real-world situations, the present study was not designed to equate audibility across listener groups. However, the S/B ratio for the listeners with hearing loss in the present study was -3 dB, in contrast with the -10 dB S/B ratio used for the listeners with normal hearing in Ferguson (2004). In Ferguson and Kewley-Port (2002) and in a pilot study, these two S/B ratios yielded roughly comparable identification performance for the two listener groups for vowels in conversational speech and prevented ceiling effects for the most intelligible vowels.

¹This was done to give listeners experience with the conditions of the experiment (i.e., 70 dB SPL presentation level and no feedback) prior to adding the challenge of background noise.

The 1640 test stimuli (41 talkers $\times$ 10 vowels $\times$ 2 tokens $\times$ 2 styles) were arranged into 16 blocks of 100 to 120 items. These test blocks were created by dividing the stimuli into a 2×2 factorial design of gender by speaking style. Each block contained 20 stimuli (10 vowels $\times$ 2 tokens) from 5 or 6 talkers of the same gender, from a single speaking style. That is, there were 4 blocks of clear stimuli from male talkers, 4 of conversational stimuli from male talkers, 4 of clear stimuli from female talkers, and 4 of conversational stimuli from female talkers. While stimuli were presented three times each in Ferguson (2004), data from that study and from a pilot study in which older adults with hearing loss identified vowels from a subset of the talkers revealed that identification was highly consistent among the three presentations. Listeners in the present study therefore heard each stimulus item just one time. However, to counteract the possibility that performance for a given talker might be affected by the specific combination of talkers in a given test block, the identical three different sets of 16 test blocks used in Ferguson (2004) were used here. Each of the three sets contained different random combinations of talkers across the 4 blocks within each gender and style. Roughly equal numbers of listeners heard each set. Within each set, each listener received the 16 blocks in random order, and stimuli were randomized within each block.

Data analysis

To test whether vowel intelligibility differed significantly between clear and conversational speech, and whether the magnitude of the clear speech vowel intelligibility benefit varied as a function of talker (gender, age, experience communicating with individuals with hearing loss) or listener (hearing status) characteristics, individual listener scores for each talker were converted to RAU and analyzed using mixed-effects models. Mixed-effects models have several advantages over the repeated-measures analyses of variance techniques that typically have been used in studies comparing clear and conversational speech (including Ferguson, 2004). These advantages include higher statistical power and robustness to both missing data and violations of sphericity (Quené & van den Bergh, 2004). The chief advantage of mixed-effect models for the present investigation is their ability to simultaneously account for multiple sources of intercorrelation. For example, in the present study, scores obtained by individual listeners for the 41 talkers may be correlated: there may be listeners who achieve high intelligibility scores for all talkers and listeners who have low intelligibility scores for all talkers. Similarly, to the extent that talkers differ from each other, intelligibility scores obtained from different listeners (or different listener groups) for vowels produced by a given talker will tend to be correlated. Mixed-effects analyses handle such non-independence by modeling its sources as random effects, keeping track of which observations are repeated measurements within the same subject. All listeners in the present study heard all of the talkers, and so talker and listener effects were modeled as crossed random effects in models that required both random factors. All analyses were performed using Stata 11 (StataCorp, 2009).

The design of the experiment includes a number of variables with several levels, most notably the talker variable with 41 levels. To reduce the possibility of error associated with multiple comparisons, all p values for contrasts tested within variables comprising more than two comparisons were corrected using a False Discovery Rate procedure (FDR; Benjamini & Hochberg, 1995). FDR is commonly used in behavioral genetics research, where strains may be compared on dozens of behavioral endpoints. Benjamini, Drai, Elmer, Kafkafi, and Golani (2001) argue that FDR strikes “a balance between the concern about making too many false discoveries and the concern about missing the discovery of a real difference that may arise from being too conservative” (p. 283).

Results

Talker and Speaking Style Effects

In the first set of mixed-effects models, which had speaking style and talker as fixed effects, a model that included listener as a random variable was found to fit better than a model with no random factors (likelihood ratio test, $\chi^2 = 1456.26$, $p < .001$). This suggests significant differences in vowel identification among the 40 listeners; these differences are considered further in the Discussion section. Returning to the fixed effects, both speaking style and talker were found to be significant. Vowels in clear speech were significantly more intelligible than vowels in conversational speech ($z = 18.97$, $p < .001$), with an average difference between the styles of 8.8 RAU.

Overall intelligibility also varied significantly among talkers. This variability is apparent in Table 1, which shows intelligibility scores (in percent correct) for each talker in each speaking style. Among the 41 talkers, vowel intelligibility scores spanned a range from 54% to 92% in conversational speech and from 59% to 95% in clear speech. Categorical variables with multiple levels are tested in mixed-effects models by comparing one level of the variable to all other levels. When a variable has just a few levels, it is reasonable to do this using each level of the variable as the baseline. Testing all 41 levels of the talker variable, in contrast, seemed unreasonable and was deemed unnecessary to show that the talkers differed significantly, and so just a few reference talkers were sampled for illustrative purposes. All comparisons used a criterion FDR-corrected p value of .05. The intelligibility of vowels produced by talker M17, who had the highest overall intelligibility, differed significantly from the intelligibility of vowels produced by all of the other talkers. Vowels produced by talker M07, who had the lowest overall intelligibility, had significantly lower intelligibility than those produced by all but one of the other talkers. Finally, F07's overall vowel intelligibility was very close to the mean for the group; her scores differed significantly from scores obtained for 25 other talkers.

Assessment of interaction effects with categorical variables in mixed-effects models requires the creation of an interaction term for each level of the categorical variable. This was done, and the interaction between speaking style and talker was significant. That is, the magnitude of the clear speech effect differed significantly among the talkers. As with the talker effect above, the model containing the interaction terms was tested using three different baseline talkers and a criterion of FDR-corrected $p < .05$. F11 had the largest clear speech vowel intelligibility benefit (37 RAU); the magnitude of the benefit was significantly greater than that achieved by all 40 of the other talkers. M06, who produced the smallest positive clear speech effect (.5 RAU), differed significantly from 17 other talkers. Lastly, F21's clear speech vowel intelligibility benefit (8.8 RAU) was very close to the mean for the group; her clear speech effect differed significantly from that produced by 12 other talkers. The variability among talkers is apparent in the fourth column of Table 1, which shows the clear speech vowel intelligibility effect in percentage points produced by each talker (calculated by subtracting the conversational score from the clear score). The interaction between talker and speaking style can also be seen in Figure 2, which shows percent correct vowel intelligibility for all 41 talkers, with the talkers ordered by their scores in conversational speech. To further explore the interaction, the effect of speaking style was tested for each individual talker; it was significant for 24 of them (FDR-corrected $p < .05$).

Talker Gender Effects

The next analysis assessed whether talker gender affected overall vowel intelligibility or the magnitude of the clear speech vowel intelligibility benefit. To assess the effect of gender, a model was created that included speaking style and gender as fixed effects and talker and

listener as crossed random effects. Note that coefficients for individual fixed factors are unaffected by adding, subtracting, or changing other fixed factors, and so the effect of speaking style remained significant ($z = 18.97$, $p < .001$) in this and all subsequent models. Gender, on the other hand, was not significant ($z = -0.6$, $p = .55$), indicating that vowels produced by female and male talkers were equally intelligible (82 and 84 RAU, respectively, when averaged across speaking styles). To assess whether speaking style effects varied between male and female talkers, a gender $\times$ speaking style interaction term was added to the model. It was significant ($z = 4.43$, $p < .0001$), and so a stratified analysis was performed to test the effect of speaking style for each gender. Both male and female talkers produced significantly more intelligible vowels in clear speech than in conversational speech ($z = 10.55$ and 16.35 , respectively, both $p < .0001$), but the clear speech effect was greater for the female talkers than the male talkers (10.8 versus 6.7 RAU). The effect of talker gender was also assessed in each speaking style but was not significant in either case (conversational $z = -1.19$, $p = 0.234$; clear $z = 0.11$, $p = 0.915$). The interaction can be seen in Figure 3, which shows percent correct vowel intelligibility scores in each speaking style for female and male talkers both for the current older listeners with hearing loss and for the young normal-hearing listeners in Ferguson (2004).

Talker Age and Experience Effects

The talkers in the Ferguson database were recruited into four age brackets: (1) 18 to 25 years, (2) 25 to 31 years, (3) 32 to 38 years, and (4) 39 to 45 years. Each bracket contained 5 male and 5 female talkers, with an additional female in the 18-to-25 bracket. To determine whether talker age had any bearing on the magnitude of the clear speech effect, two mixed-effect models were tested. The first model contained age bracket and speaking style as fixed factors (with talker and listener as crossed random factors); the second added interaction terms. Age bracket had no effect on overall intelligibility ($|z| < 1.9$, $p > .06$ for all between-bracket comparisons) or on the magnitude of the clear speech effect ($|z| < 1.3$, $p > .2$ for all comparisons). Table 2 shows the means and standard deviations for the percent correct clear, conversational, and difference scores for the four age brackets.

Although talkers were recruited without regard to experience communicating with individuals with hearing loss, they were asked about such experience at the end of the clear speech recording session. Based on their responses, talkers were placed into one of four experience categories: none (no prior experience; $n = 14$), little (only a few experiences; $n = 8$), occasional (a relative or friend with hearing loss but less than one interaction per week; $n = 9$), or frequent (at least one weekly contact with one or more individuals with hearing loss; $n = 10$). To test whether such experience affected talkers' vowel intelligibility or their ability to produce effective clear speech, another two mixed-effect models were tested. One model included style and experience group (none, little, occasional, frequent) as fixed factors, while the other also included interaction terms. Both models included talker and listener as crossed random factors.

No differences were observed between the four experience categories in terms of overall intelligibility ($|z| < 1.4$, $p > .18$ for all comparisons). However, a significant interaction was observed between speaking style and experience when the group of talkers with "little" experience communicating with hearing loss was compared to any of the other three groups ($|z| > 4.3$, $p < .0001$ for all comparisons). A stratified analysis revealed that the speaking style effect was significant ($|z| \geq 8$, $p < .0001$) for each experience group, but was larger for those with "little" experience (14.7 RAU) than for talkers in the other three groups (none: 6.6 RAU; occasional: 8.4 RAU; frequent: 7.6 RAU). The other three groups did not differ significantly from each other. Average percent correct scores in each speaking style for the four experience brackets are shown in Figure 4.

Listener Group Effects

For the final analysis, data from the seven young listeners with normal hearing (YNH listeners) reported in Ferguson (2004) were combined with the present data from older adults with hearing impairment (OHI listeners) so that effects of listener group could be assessed. As mentioned above, mixed-effects models are robust to missing data; they also permit comparisons between groups that differ in number of and variance among participants. Preliminary testing showed that including both talker and listener as random variables significantly improved the fit of the model, so the first model contained speaking style and listener group as fixed factors, and talker and listener as crossed random factors. Both fixed effects were significant. Averaged across the two listener groups, vowels remained significantly more intelligible in clear speech than in conversational speech ($z = 21.17$, $p < .0001$). Averaged across the two speaking styles, scores were significantly higher for the OHI listeners than for the YNH listeners ($z = 3.37$, $p < .01$). This difference can be seen in Figure 5, which shows overall percent correct scores in each speaking style for each listener group. The style $\times$ listener group interaction term was added to the model and found not to be significant ($z = -0.2$, $p = 0.84$). This suggests that the magnitude of the speaking style effect was the same for the two listener groups (8.8 and 9.0 RAU for OHI and YNH listeners, respectively).

Discussion

Speaking style, talker, and listener effects on vowel intelligibility

Averaged across all talkers, vowel identification scores for the current OHI listeners were significantly higher than those observed for the YNH listeners in Ferguson (2004) for both speaking styles. The most likely explanation for this difference is that the two groups were tested under different S/B ratios: -10 dB for the YNH listeners and -3 dB for OHI listeners. These S/B ratios had produced comparable conversational vowel intelligibility scores for OHI and YNH listeners in Ferguson and Kewley-Port (2002) and in a pilot study using just 12 of the database talkers. When the full set of 41 talkers was tested, however, the OHI listeners' overall score for conversational vowels was 77%, while YNH listeners achieved 65% correct vowel identification.

Despite the difference in overall intelligibility, the magnitude of the clear speech effect for OHI listeners (about 9 RAU) was essentially identical to that observed for the YNH listeners in Ferguson (2004). This result contrasts sharply with the results of Ferguson and Kewley-Port (2002), in which YNH listeners showed a clear speech vowel intelligibility benefit of 15 percentage points while OHI listeners showed no clear speech benefit. The listeners in Ferguson and Kewley-Port identified vowels produced by just one talker, however. For the present OHI listeners as well as for the YNH listeners in Ferguson (2004), the 41 talkers in the Ferguson Clear Speech Database varied considerably in the magnitude of the clear speech vowel intelligibility effect. The best talker for the OHI listeners produced a clear speech benefit of 24 percentage points while the poorest showed a clear speech decrement of -9 points. The range for YNH listeners in Ferguson (2004) was slightly larger, from a benefit of 33 percentage points to a decrement of -12 points.

To explore whether talkers who produced a clear speech benefit for YNH listeners also did so for OHI listeners, a correlational analysis was performed using the RAU difference scores for each talker obtained by each listener. The correlation between the two listener groups was strong, positive, and significant, $r(41) > .75$, $p < .001$, and is illustrated in Figure 6. In the scatterplot, 31 of the 41 data points fall within 5 percentage points of the diagonal. Of the 10 values that lie outside this range, two represent talkers for whom the clear speech vowel intelligibility benefit was larger for the OHI listeners than the YNH listeners. In the 8

remaining cases, a larger benefit was seen for the YNH listeners than for the OHI listeners, but for only three talkers did the OHI listeners have no clear speech vowel intelligibility benefit while YNH listeners had a significant benefit. That is, only three of the 41 talkers in the Ferguson Clear Speech Database showed results resembling those found for the talker in Ferguson and Kewley-Port (2002). To explore this further, a series of mixed-effects models were used to test the style $\times$ listener group interaction for each talker individually. When p -values for those 41 analyses were corrected using FDR, only one talker, F09, showed a significant interaction between speaking style and listener group. This talker will be considered again below, when effects of talker experience are considered.

As noted in the Results section, the fit of the various mixed-effects models improved significantly when listener was included as a random factor. This result implies that scores for individual OHI listeners tended to be correlated with each other across speaking styles, and that listeners differed from each other significantly in terms of their ability to identify the vowel stimuli. Wide variability among the listeners is evident in Table 3, which shows mean percent-correct vowel intelligibility scores for each OHI listener in each speaking style. Scores in each speaking style ranged from 49% to 89% in conversational speech and from 58% to 93% in clear speech, ranges comparable to those observed across the 41 talkers. In a set of mixed-effects models that included speaking style and listener as fixed effects and talker as a random effect, the effect of listener was significant. Both the highest-scoring listener (E07) and the listener (E03) whose performance was closest to the mean for the group performed significantly differently from all but 9 of the other listeners (FDR-corrected $p < .05$). When style $\times$ listener interaction terms were added to these models, however, none of them were significant (FDR-corrected $p > .05$). That is, despite differences in overall performance, the clear speech effect did not differ significantly among the listeners. Comparing Tables 1 and 2, we can see that the range of difference scores among the listeners (from .85 to 11.7 percentage points) is much smaller than the range of difference scores among the talkers (from -9 to 24 percentage points).

Possible explanations for the wide variability among the OHI listeners were assessed by computing Pearson correlations between vowel intelligibility scores for each listener and several listener characteristics: age, MMSE score, individual audiometric frequencies from 250-8000 Hz, and two puretone averages. PTA1 was the traditional pure tone average, calculated over thresholds recorded at 500, 1000 and 2000 Hz; PTA2 was calculated using 1000, 2000, and 4000 Hz. Of all those possible predictors, only audiometric thresholds at 2000-4000 Hz and the puretone averages were significantly correlated with overall performance, with PTA2 being the strongest predictor ($r = -.66$, FDR-corrected $p < .0001$). A t-test was used to assess whether listener gender affected vowel identification performance; the result was not significant ($p = 0.7$). Pearson correlations and the t-test for gender were also performed using difference scores for each listener, calculated by subtracting their conversational RAU score from their RAU score for clear speech. None of these tests were significant ($p > .19$). These results suggest that audibility has a significant impact on individual listeners' ability to identify vowels in general, but not on the magnitude of the clear speech vowel intelligibility benefit they enjoy.

Talker characteristics and the clear speech effect

Talker gender—For the current OHI listeners, male and female talkers had similar overall vowel intelligibility. This result was contrary to two competing predicted outcomes. For the YNH listeners in Ferguson (2004), female talkers showed significantly higher overall vowel intelligibility scores (72.5%) than males (65%). Ferguson's results are consistent with other studies showing superior intelligibility for female talkers (e.g., Bradlow et al., 1996; Hazan & Markham, 2004) when materials were presented to YNH listeners. Given the absence of

any scholarly data on talker gender effects for OHI listeners, one might extrapolate from the YNH studies and predict that OHI listeners would also find female talkers to be more intelligible than male talkers. However, the experience of clinical audiologists would suggest the opposite outcome, at least for overall intelligibility. Individuals with hearing loss seeking hearing services frequently complain of particular difficulty understanding female voices (Helfer, 1995). On this basis, one would predict a finding of superior intelligibility for male talkers for the OHI listeners.

The present data fall between these two predictions, with no overall vowel intelligibility advantage for either gender. However, female talkers produced a larger clear speech benefit than male talkers for the OHI listeners, just as they did for the YNH listeners. In Figure 3, the interaction between speaking style and talker gender appears to be the same for the two listener groups, an impression that was confirmed with mixed-effects models of combined data from both groups. Figure 3 also suggests the possibility that gender differences might have occurred in conversational speech for the OHI listeners, but were offset by the absence of a gender difference in clear speech, but the gender effect was not statistically significant in either style. Taken together, these results suggest that whatever acoustic characteristics render female talkers more intelligible than male talkers for YNH listeners are unavailable for listeners with hearing loss. However, the acoustic changes that female talkers employ when speaking clearly that provide a larger clear speech advantage than that produced by male talkers seem to be equally beneficial to both listeners with and without hearing loss. Future research should examine the nature of these acoustic differences between male and female talkers and how they affect speech understanding by listeners with hearing loss.

Talker age and experience communicating with listeners with hearing loss—

Talkers were recruited into different age brackets under the hypothesis that those in the older brackets might have more contact with individuals with hearing loss than younger talkers, that this would make them better at producing effective clear speech. However, there was no difference between the age brackets in the size of the clear speech effect here or in Ferguson (2004). In addition, the correlation between talkers' actual age and the clear speech effect enjoyed by OHI listeners in the present study was close to zero, $r(41) = -.04$. In contrast, significant differences in the clear speech effect for OHI listeners were observed between talkers with varying amounts of experience communicating with listeners with hearing loss. Talkers who had had only a handful of such experiences produced a significantly larger clear speech vowel intelligibility benefit than talkers with no experience at all, with occasional experience, or with frequent experience. While Ferguson (2004) found no effect of experience on the clear speech benefit, this finding was obtained through a one-way analysis of variance carried out on clear-minus-conversational difference scores. Mixed-effects analyses of the YNH listener data, in contrast, did reveal a significant interaction between experience category and speaking style. Talkers with no experience and “occasional” experience showed a clear speech effect of 6 RAU. This was significantly lower than that produced by talkers with “frequent” experience (10 RAU), which in turn was significantly smaller than that produced by talkers with “little” experience (14 RAU).

It is counter-intuitive that talkers with only a few experiences communicating with listeners with hearing loss, rather than those with frequent experience, should produce the largest clear speech vowel intelligibility benefit. It certainly challenges the idea that the frequent communication partners of listeners with hearing loss should “just know” what to do to help their significant others. This idea is further challenged by the results for two specific talkers. The talker in Ferguson and Kewley-Port (2002) was a clinical audiologist with over 25 years of experience, but only YNH listeners benefited when he spoke “clearly”. The talker in the Ferguson database whose results most resembled his was talker F09. While her vowels were significantly more intelligible in clear speech than in conversational speech (by 22 RAU)

when YNH listeners identified them, the two styles yielded nearly identical performance (difference $< .1$ RAU) for the OHI listeners. She also resembled the earlier talker in having had extensive experience talking to individuals with hearing loss: she reported having grown up with two brothers who were hard-of-hearing. Interestingly, F09 was also one of the “atypical talkers” in Ferguson and Kewley-Port (2007), which examined acoustic differences between clear and conversational speech for subgroups of talkers who had produced either a large clear speech vowel intelligibility benefit or no clear speech benefit for YNH listeners in Ferguson (2004). In Ferguson and Kewley-Port (2007), F09 was one of six “big benefit” talkers, but her acoustic data bore little resemblance to those of the other five talkers. Most notably, F09 showed a substantially *reduced* vowel space in clear speech, in contrast with the vowel space expansion observed in the other five big benefit talkers.

Lindblom's H&H Theory (hyper- versus hypospeech; Lindblom, 1990) states that when a talker encounters a listener who has difficulty understanding the talker's speech, he/she will begin to hyperarticulate, sacrificing motor economy to achieve greater specificity in the speech signal. The results for talker F09 and the talker in Ferguson and Kewley-Port (2002) seem to negate this idea. However, H&H theory may help explain the pattern of results observed for talkers who had “little” versus more or no experience. Recall that clear speech was elicited in the Ferguson Clear Speech Database by having talkers read printed sentences aloud, imagining that they were speaking to someone with hearing loss. Talkers with no experience communicating with listeners with hearing loss would have no frame of reference for this, and so would just speak in a way that sounded clear to them. In contrast, talkers who had previously communicated with listeners with hearing loss would have had the experience of adjusting their speech in a way that was helpful to their communication partner. According to H&H theory, these adjustments would have been automatic. Later, when instructed to speak clearly in the lab, they would draw on that experience and use clear speech articulatory strategies they had used when actually talking to an interlocutor with hearing impairment.

But if talkers learn from experience, why would talkers with more experience produce a smaller clear speech effect? It is proposed that when talkers encounter listeners with hearing loss for the first time, they (automatically) use a wide array of clear speech articulatory strategies. As talkers gain experience speaking with listeners with hearing loss, however, they gradually discard some of these strategies. This process would be driven by several factors, chiefly motor economy and auditory feedback. Having no access to how speech sounds to someone with hearing loss, talkers with normal hearing will naturally use strategies that make speech sound clear to them. It may be that the more time talkers spend speaking clearly, the further away they get from what a listener with hearing loss actually needs. There are several ways to test this explanation. Using the Ferguson database, acoustic differences between clear and conversational speech could be compared across the different experience groups. Perceptual data from other materials in the database, such as sentences (as in Ferguson & Kerr, 2009) or monosyllabic words, could also be compared across talker groups.

Conclusions

The current experiment demonstrates that the considerable talker variability that Ferguson (2004) found for vowel intelligibility and the clear speech vowel intelligibility benefit for YNH listeners also occurs for OHI listeners. When given general instructions to speak clearly, some talkers produce speech that is much more intelligible than ordinary conversational speech, while others do not. Across talkers as well as for most individual talkers, OHI listeners enjoyed a similar clear speech benefit to that achieved by YNH listeners. Thus, if a certain talker produced clear speech that was effective for YNH

listeners, it was also effective for OHI listeners. From a clinical perspective, this is a very encouraging result. The results from Ferguson and Kewley-Port (2002) for only one talker had suggested that talkers might need to adopt different clear speech strategies for different listener groups. The present results, however, suggest that, at least for vowels, the acoustic properties that make speech more intelligible for YNH listeners also make it more intelligible for OHI listeners.

Acknowledgments

This research was supported by the National Institutes of Health Grants DC008886, DC005803, and HD002528. Data analyses were performed under the guidance of Gregory J. Stoddard with of the University of Utah Study Design and Biostatistics Center, which is funded in part by Grants UL1-RR025764 and C06-RR11234 from the National Center for Research Resources. Douglas Kieweg, Leela Parimisetty, Katherine Beam, Jessica Stamey, and Kyung Ae Keum assisted with data collection and entry. Chiung-ju Liu and Susan Kemper assisted in participant recruiting, and the Schiefelbusch Speech-Language-Hearing Clinic provided hearing evaluations for the older adults. Diane Kewley-Port, Allard Jongman, Michael Blomgren, and Sean Redmond provided useful comments on an early draft of this manuscript. Development of the Ferguson Clear Speech Database was supported by National Institutes of Health Grant DC02229 to Indiana University.

References

- Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research*. 2001; 125(1-2):279–284. [PubMed: 11682119]
- Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995; 57(1): 289–300.
- 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]
- Ferguson SH. Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *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. *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.
- Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Gagné JP, Masterson V, Munhall K, Bilida N, Querengesser C. Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology*. 1994; 27:135–158.
- Gagné JP, Rochette AJ, Charest M. Auditory, visual, and audiovisual clear speech. *Speech Communication*. 2002; 37:213–230.
- Hazan V, Markham D. Acoustic-phonetic correlates of talker intelligibility for adults and children. *Journal of the Acoustical Society of America*. 2004; 116:3108–3118. [PubMed: 15603156]
- Helfer, KS. Auditory perception by older adults. In: Huntley, RA.; Helfer, KS., editors. *Communication in Later Life*. Boston: Butterworth-Heinemann; 1995. p. 41-84.
- Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*. 1977; 61:1337–1351. [PubMed: 881487]
- Kochkin S. MarkeTrak VIII: 25-year trends in the hearing health market. *Hearing Review*. 2009; 16(11):12–31.

- Kochkin S. MarkeTrak VII: Customer satisfaction with hearing instruments in the digital age. *The Hearing Journal*. 2005; 58(9):30–39.
- Lindblom, B. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, WJ.; Marchal, A., editors. *Speech production and speech modelling*. Netherlands: Kluwer Academic Publishers; 1990. p. 403–439.
- Liu S, Del Rio E, Bradlow AR, Zeng FG. Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America*. 2004; 116:2374–2383. [PubMed: 15532668]
- Payton KL, Uchanski RM, Braida LD. Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*. 1994; 95:1581–1592. [PubMed: 8176061]
- Pearsons, KS.; Bennett, RL.; Fidell, S. *Speech levels in various noise environments*. Washington, D.C.: U.S. Environmental Protection Agency; 1977. No. EPA-600/1-77-025
- 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]
- Prendergast SG, Kelley LA. Aural rehab services: Survey reports who offers which ones and how often. *The Hearing Journal*. 2002; 55(9):30, 34–35.
- Quené H, van den Bergh H. On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*. 2004; 43(1-2):103–121.
- Schum DJ. Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*. 1996; 7:212–218. [PubMed: 8780994]
- StataCorp. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP; 2009.
- Studebaker GA. A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*. 1985; 28:455–462. [PubMed: 4046587]
- Tillman, T.; Carhart, R. *An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words*, Northwestern University Auditory Test No 6. Brooks AFB, TX: USAF School of Aerospace Medicine; 1966.
- Tye-Murray, N. *Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members*. Clifton Park, NY: Thomson Delmar Learning; 2004.
- Uchanski RM, Choi SS, Braida LD, Durlach NI. 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]

Figure 1. Mean audiometric thresholds for listeners. Error bars indicate one standard deviation around the mean.

Figure 2. Overall percent correct vowel intelligibility in clear (CL) and conversational (CON) speech for individual talkers. Talkers are shown in order of ascending conversational vowel intelligibility.

Figure 3. Percent correct intelligibility for clear (CL) and conversational (CON) vowels produced by female and male talkers, as identified by OHI listeners in the present study and by YNH listeners in Ferguson (2004).

Figure 4.

Clear-minus-conversational (CL – CON) percent correct difference scores averaged across talkers in four categories of experience communicating with listeners with hearing loss. Error bars indicate 95% confidence intervals.

Figure 5. Overall percent correct vowel identification performance in clear (CL) and conversational (CON) speech by OHI listeners in the present study and by YNH listeners in Ferguson (2004). Error bars indicate 95% confidence intervals.

Figure 6. Clear-minus-conversational (CL – CON) percent correct vowel intelligibility difference scores for 41 talkers for OHI listeners in the present study versus YNH listeners in Ferguson (2004). Scores along the diagonal are identical for the two groups.

Table 1

Percent correct vowel intelligibility for older listeners with hearing loss for 41 talkers in clear (CL) and conversational (CON) speech, and the percentage point difference between the two styles (DIFF). Mean values were calculated over all listeners.

Talker ID	CL	CON	DIFF
F01	93.0	78.1	14.9
F02	87.1	82.8	4.4
F03	78.4	76.9	1.5
F04	76.9	79.9	-3.0
F05	79.6	72.3	7.4
F06	84.3	73.6	10.6
F07	88.9	71.3	17.6
F08	75.3	66.8	8.5
F09	84.6	84.5	0.1
F10	78.5	74.4	4.1
F11	94.5	70.6	23.9
F12	67.1	53.9	13.3
F13	84.9	67.5	17.4
F14	87.1	77.3	9.9
F15	79.5	83.9	-4.4
F16	89.9	86.6	3.3
F17	90.1	82.6	7.5
F18	82.5	63.9	18.6
F19	89.4	88.6	0.8
F20	68.8	66.9	1.9
F21	81.1	73.9	7.3
M01	70.3	79.6	-9.4
M02	90.9	78.6	12.3
M03	87.9	68.6	19.3
M04	76.8	71.5	5.3
M05	87.1	79.1	8.0
M06	87.5	86.4	1.1
M07	59.4	55.6	3.8
M08	86.9	83.3	3.6
M09	84.0	78.9	5.1
M10	81.3	80.3	1.0
M11	86.4	80.0	6.4
M12	81.9	77.8	4.1
M13	74.9	75.5	-0.6
M14	86.3	67.5	18.8
M15	81.9	77.6	4.3
M16	85.5	84.9	0.6
M17	93.3	91.5	1.8

Talker ID	CL	CON	DIFF
M18	89.6	86.1	3.5
M19	88.6	87.1	1.5
M20	88.0	80.1	7.9
MEAN	83.1	76.7	6.4

Table 2

Percent correct vowel intelligibility scores in clear (CL) and conversational (CON) speech, and the percentage point difference between the two styles (DIFF), for talkers in four different age brackets. Values were calculated over all talkers and vowels; standard deviations appear in parentheses.

Age bracket	CL	CON	DIFF
18 to 24 years	81.47 (6.7)	74.38 (4.7)	7.09 (8.4)
25 to 31 years	82.04 (6.8)	75.46 (10.2)	6.58 (7.2)
32 to 38 years	83.41 (11.2)	77.23 (10.5)	6.19 (6.9)
39 to 45 years	85.89 (4.4)	80.10 (7.2)	5.79 (6.8)

Table 3

Percent correct vowel intelligibility scores in clear (CL) and conversational (CON) speech, and the percentage point difference between the two styles (DIFF), for 40 individual older listeners with hearing loss. Mean values were calculated over all 41 talkers.

Listener ID	CL	CON	DIFF
OL01	91.6	85.9	5.7
OL02	89.6	84.3	5.4
OL03	83.9	78.4	5.5
OL04	90.0	89.1	0.9
OL05	92.8	86.8	6.0
OL06	92.3	87.6	4.8
OL07	91.7	88.8	2.9
OL08	83.5	76.2	7.3
OL09	76.5	68.4	8.0
OL10	79.1	73.3	5.9
OL11	92.2	86.7	5.5
OL12	87.7	81.6	6.1
OL13	76.3	69.4	7.0
OL14	92.3	88.0	4.3
OL15	92.6	84.8	7.8
OL16	89.5	82.9	6.6
OL17	85.2	82.8	2.4
OL18	80.7	72.3	8.4
OL19	75.9	66.6	9.3
OL20	75.2	64.9	10.4
OL21	84.9	76.0	8.9
OL22	86.8	81.8	5.0
OL23	87.3	80.0	7.3
OL24	73.0	66.2	6.8
OL25	86.5	81.6	4.9
OL26	75.9	64.1	11.7
OL27	83.8	74.8	9.0
OL28	86.8	82.1	4.8
OL29	58.3	49.4	8.9
OL30	83.8	76.1	7.7
OL31	81.1	74.1	7.0
OL32	93.4	87.0	6.5
OL33	76.3	72.0	4.4
OL34	88.3	79.4	8.9
OL35	78.3	71.3	7.0
OL36	72.4	66.2	6.2
OL37	76.7	69.6	7.1
OL38	89.9	86.2	3.7

Listener ID	CL	CON	DIFF
OL39	67.4	59.9	7.6
OL40	76.6	72.7	3.9