Research Article

Talker Differences in Clear and Conversational Speech: Perceived Sentence Clarity for Young Adults With Normal Hearing and Older Adults With Hearing Loss

Sarah Hargus Ferguson^a and Shae D. Morgan^a

Purpose: The purpose of this study is to examine talker differences for subjectively rated speech clarity in clear versus conversational speech, to determine whether ratings differ for young adults with normal hearing (YNH listeners) and older adults with hearing impairment (OHI listeners), and to explore effects of certain talker characteristics (e.g., gender) on perceived clarity. Relationships among clarity ratings and other speech perceptual and acoustic measures were also explored.

Method: Twenty-one YNH and 15 OHI listeners rated clear and conversational sentences produced by 41 talkers on a scale of 1 (*lowest possible clarity*) to 7 (*highest possible clarity*).

Results: While clarity ratings varied significantly among talkers, listeners rated clear speech significantly clearer than conversational speech for all but 1 talker. OHI and YNH listeners gave similar ratings for conversational speech, but ratings for clear speech were significantly higher for OHI listeners. Talker gender effects differed for YNH and OHI listeners. Ratings of clear speech varied among subgroups of talkers with different amounts of experience talking to people with hearing loss.

Conclusions: Perceived clarity varies widely among talkers, but nearly all produce clear speech that sounds significantly clearer than their conversational speech. Few differences were seen between OHI and YNH listeners except the effect of talker gender.

When individuals with hearing loss seek hearing help, their audiologist may offer them a list of communication strategies to help them address their everyday communication difficulties. These strategies can include turning off or moving away from sources of background noise when having conversation, making sure communication partners can see each other's faces, or asking communication partners to speak more clearly (Tye-Murray, 2014). Laboratory studies of such "clear speech" have established that it is generally more intelligible than "conversational" speech for a variety of speech materials and listener populations (Smiljanić & Bradlow, 2009) and differs from conversational speech on numerous acoustic

dimensions (Uchanski, 2005). However, individual talkers vary considerably in terms of the intelligibility benefit they achieve when they speak clearly (e.g., Ferguson, 2004, 2012), as well as the specific acoustic changes they make (e.g., Ferguson & Quené, 2014; Hazan & Baker, 2011). This study assessed talker variability in clear and conversational speech for another perceptual dimension: subjective ratings of speech clarity. It extended the work of Ferguson and Kerr (2009), in which young listeners with normal hearing rated clarity for sentences produced by eight talkers; this study uses materials produced by all 41 talkers from the Ferguson Clear Speech Database (Ferguson, 2004) and includes listeners with hearing loss.

The Ferguson Clear Speech Database consists of 21 female and 20 male talkers from the South Midland dialect region aged 18–45 years reading sentences under conversational and clear instructions. It was designed to explore talker variability in the acoustic and perceptual characteristics of clear speech—variability that could then be exploited in the search for the specific clear speech acoustic changes that underlie the superior intelligibility of clear speech. Most perceptual analyses of the database

^aDepartment of Communication Sciences & Disorders, University of Utah, Salt Lake City

Correspondence to Sarah Hargus Ferguson:
sarah.ferguson@hsc.utah.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Daniel Fogerty

Received March 1, 2017

Revision received June 13, 2017

Accepted July 11, 2017

https://doi.org/10.1044/2017_JSLHR-H-17-0082

Disclosure: The authors have declared that no competing interests existed at the time of publication.

have focused on the intelligibility of vowels in /bVd/ context that were extracted from various read carrier sentences (Ferguson, 2004, 2012). These studies revealed substantial variation among talkers not only in the size of the clear speech effect but also within each speaking style. In particular, when young adult listeners with normal hearing (YNH listeners) identified vowels in a background of 12-talker babble, talker scores ranged from 25% to 83% correct in conversational speech and 29% to 94% correct in clear speech, whereas the difference between clear and conversational speech ranged from –12 to 33 percentage points (Ferguson, 2004). When older adults with mild-to-moderately severe sloping sensorineural hearing loss (older adults with hearing impairment [OHI listeners]) performed the same task at a more favorable signal-to-noise ratio (–3 dB vs. –10 dB), variability was somewhat less dramatic than for the YNH listeners but still considerable (54% to 92% correct in conversational speech; 59% to 95% correct in clear speech; clear-minus-conversational differences from –9 to 24 percentage points; Ferguson, 2012). Ferguson (2012) also found a strong, positive relationship between clear-minus-conversational difference scores observed for YNH and for OHI listeners ($r = .75$). This indicated that, in general, when a talker's clear speech improved vowel identification for one group, it also did so for the other group.

Although it is encouraging that talkers seem to need to adopt only one set of clear speech strategies to improve vowel intelligibility for listeners with and without hearing loss (Ferguson, 2012), the proportion of talkers whose clear speech did not produce a clear speech vowel intelligibility benefit is disheartening. Using an arbitrary criterion of 5 percentage points as indicating a clinically significant improvement in vowel intelligibility, Ferguson (2004) found that just 23 of the 41 talkers (56%) produced a significant clear speech vowel intelligibility benefit for YNH listeners; Rogers, DeMasi, and Krause (2010) found a similar result (seven of 13, or 53%) for monolingual speakers. Ferguson (2012) tested the effect of speaking style statistically for each talker and found similar results for OHI listeners; 24 of the 41 talkers (59%) produced a significant clear speech vowel intelligibility benefit for OHI listeners. For both listener groups, then, approximately 40% of the talkers failed to produce a vowel intelligibility benefit when speaking clearly. The assumption of this work has been that vowel intelligibility is a reasonable index of overall intelligibility, an assumption supported by several studies showing the importance of vowel information to sentence intelligibility for listeners with and without hearing loss (Fogerty & Kewley-Port, 2009; Kewley-Port, Burkle, & Lee, 2007; Lee & Kewley-Port, 2009).

However, it is possible that talkers' success in producing clear speech that actually benefits listeners depends on the speech materials. While Picheny, Durlach, and Braidá (1985) reported both phonemes correct and words correct for their anomalous sentences, no study has measured perceptual characteristics of both vowels and other speech materials in clear and conversational speech for the

same set of talkers. Nonetheless, we can compare the range of clear speech effects measured for vowel intelligibility in the Ferguson Clear Speech Database to those observed for sentence intelligibility data in two clear speech studies that report individual data for several talkers. In Schum (1996), Figure 2 shows the magnitude of the clear-minus-conversational intelligibility difference for sentences produced by 10 young and 10 elderly talkers and presented in noise to OHI listeners. Rationalized arcsine unit (rau; Studebaker, 1985) difference scores, which can be compared roughly with percentage-point difference scores, ranged from 4 to 45 rau. Lam and Tjaden (2013), who examined the effects of different speaking style instructions on the intelligibility of clear versus “habitual” speech, also found sizeable variability among their 12 talkers. Of their three instruction conditions, their “clear” condition most closely resembled the instructions used in the Ferguson Clear Speech Database, and the clear speech benefit for sentences produced under these instructions and presented in noise to YNH listeners ranged from 2 to 32 percentage points. Wide talker variability in the magnitude of the clear speech benefit has thus been found both for vowels in /bVd/ context and for meaningful sentences. Note, however, that in both of these sentence studies, nearly all talkers produced a clear speech benefit. Using the 5-percentage-point criterion used by Ferguson (2004) and treating Schum's rau differences as percentage-point differences, 19 of the 20 talkers in Schum and 11 of the 12 talkers in Lam and Tjaden significantly improved their intelligibility when speaking clearly.

The contrast between the approximately 60% success rate observed in the two vowel intelligibility studies from the Ferguson Clear Speech Database (Ferguson, 2004, 2012) and the 90%–95% success rate seen in Lam and Tjaden (2013) and Schum (1996) suggests that whether or not a talker's clear speech is deemed “effective” may indeed depend on the materials. Unfortunately, the design of the database does not lend itself to comparing sentence intelligibility among individual talkers. While the database includes meaningful sentences, all 41 talkers recorded the same 28 sentences (a different set of 14 in each speaking style). The only way to evaluate intelligibility and the clear speech benefit without contamination by the strong learning effects associated with meaningful sentences is a between-subjects experiment comprising 41 different listener groups. Such a study might be feasible to carry out with YNH listeners in a university setting, but extending it to a population for whom clear speech is intended (i.e., OHI listeners) would be extremely difficult. Because of this, work using these meaningful sentences has been almost exclusively acoustic (e.g., Ferguson, Poore, & Shrivastav, 2010).

Ferguson and Kerr (2009) offered a solution to these methodological challenges: having listeners subjectively rate the materials rather than identify them. They presented sentences produced by eight talkers from the Ferguson Clear Speech Database to YNH listeners and instructed the listeners to rate “how clear the sentence sounds” on a 7-point scale. Sentences were presented via earphones in quiet at a slightly loud conversational level. Clarity ratings were

deemed more appropriate than intelligibility ratings given the high intelligibility of conversational speech in quiet for this population, particularly upon repeated presentation. Results demonstrated that clarity ratings not only reliably distinguished conversational and clear speech but also revealed differences among talkers. In addition, rated clarity was significantly correlated with other perceptual and acoustic measures, namely with vowel intelligibility and with vowel space perimeter ($r = .8, p < .001$, for both measures). Thus, Ferguson and Kerr concluded, “subjective ratings offer a viable alternative to objective measures of sentence intelligibility” (p. 13). A more recent study using a different set of eight talkers (Ferguson, Widener, & Mackey, 2011) and YNH listeners also found significant differences between speaking styles and among talkers, as well as relationships between clarity and other measures, further attesting to the viability of this approach.

This study was thus undertaken to determine the range of clear speech clarity differences that can be expected within a population of healthy talkers by obtaining subjective measures of clarity for all 41 talkers in the Ferguson Clear Speech Database. Furthermore, it did so not only for listeners with normal hearing but also for listeners with hearing loss, the population for whom clear speech is intended. Cox, Alexander, and Rivera (1991) and Cienkowski and Speaks (2000) showed that older adults with hearing loss are able to make reliable subjective ratings of intelligibility, and so, we anticipated that they would have no difficulty making subjective clarity measures. The perceived clarity data were also used to examine both the relationship between vowel intelligibility and sentence clarity and the acoustic features that underlie perceived sentence clarity for different listener groups. Although Ferguson (2012) found that talkers who produced a clear speech vowel intelligibility benefit for YNH listeners also produced a benefit for OHI listeners, sentences contain a much wider range of acoustic cues than vowels. In particular, sentences contain high-frequency stop and fricative consonants, the energy of which is often increased in clear speech, but the audibility of which is reduced for listeners with sloping hearing losses. We thus anticipated that the relationship between specific clear speech acoustic features and ratings of perceived clarity would differ for YNH and OHI listeners.

Ferguson (2012) found interesting interactions between talker gender and listener group and surprising effects of talkers’ previous experience communicating with individuals with hearing loss. For YNH listeners, vowel intelligibility was higher for female talkers than for male talkers (Ferguson, 2004), whereas no gender effect occurred for the OHI listeners (Ferguson, 2012). For both listener groups, however, female talkers produced a larger clear speech vowel intelligibility benefit than male talkers. This study therefore tested the effects of talker gender on perceived clarity. Given the complaints listeners with hearing loss often have understanding female speakers (Helfer, 1995), we anticipated that OHI listeners would judge male talkers as clearer than female talkers. In contrast, previous research showing superior intelligibility for female talkers for YNH listeners (e.g., Bradlow,

Torretta, & Pisoni, 1996) led us to expect YNH listeners to judge female talkers as clearer than male talkers. Finally, Ferguson (2012) found that for both OHI and YNH listeners, talkers with only a few previous experiences communicating with individuals with hearing loss produced a significantly larger clear speech vowel intelligibility benefit than talkers with either no prior experience or more experience. This counterintuitive result has both clinical and theoretical implications, and so, the same comparisons were included in this study.

Method

Materials

Test stimuli were 28 sentences selected from the Central Institute for the Deaf (CID) Everyday Sentence lists (Davis & Silverman, 1978) and produced by the 41 talkers (aged 18–45 years; 21 women, 20 men) comprising the Ferguson Clear Speech Database. Each talker read one list of 14 CID sentences in a conversational speaking style (i.e., they were instructed to speak as they would in everyday conversation) and a different list of 14 in a clear speaking style (i.e., they were instructed to speak as though they were talking to a person with a hearing loss). The sentences read in conversational and clear speech are listed in Appendixes A and B, respectively. The lists of CID sentences were just one of six blocks of sentences recorded in each style, and the order of these six blocks was randomized for each talker, although the order of the 14 sentences in each style was identical for all talkers. The experimenter monitored the talkers during the recording sessions and, if the talker was disfluent or made other production errors, instructed the talker to reread the sentences that contained errors. Otherwise, each talker produced each sentence just one time. Conversational and clear speech were recorded in two separate sessions, with conversational speech always recorded in the first session. Additional details about the recording procedures may be found in Ferguson (2004).

Research assistants excised the 1,148 individual sentence files (41 talkers $\times$ 28 sentences) from the raw recordings by hand and scaled them to a root-mean-square amplitude of -25 dB in Cool Edit 2000. Prior to the present experiment, the sentence files were resampled in MATLAB from their original recorded sampling rate of 22050 Hz to 24414 Hz for presentation via Tucker-Davis Technologies (TDT) System III audio hardware.

Listeners

Twenty-two YNH listeners were recruited from the University of Utah Department of Psychology Participant Pool. Normal hearing was confirmed by a screening test performed at 20 dB HL (American National Standards Institute, 2010) at octave frequencies between 250 Hz and 8000 Hz. One YNH listener completed the first test session but did not return for the second; their data

were discarded, yielding a final n of 21 YNH listeners (11 women, 10 men) aged 17 to 25 years ($\bar{x} = 20.95$ years). Twenty OHI listeners were recruited from Utah Senior Ears (USE), a participant pool maintained by the first author. Members of USE receive a complete audiologic evaluation when they join the pool. To be invited to participate in this study, USE members were required to have mild-to-moderately severe sloping sensorineural hearing losses and good word recognition abilities ($> 80\%$ correct) for the Northwestern University Auditory Test NU-6 (Tillman & Carhart, 1966) presented in quiet at 40 dB above the speech recognition threshold. Of these 20 OHI listeners, two did not return for the second test session, and another two did not follow test instructions regarding how to use the Likert scale (as described below). Finally, data were lost for one OHI listener, yielding a final n of 15 OHI listeners (eight women, seven men) aged 65 to 81 years ($\bar{x} = 70.8$ years). Their mean audiogram is shown in Figure 1; their mean NU-6 score was 93%.

All listeners were required to be native speakers of American English with no history of speech or language disorders. They were also required to have normal cognitive status as determined by a score of 25 or higher on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), which was administered at the beginning of the experiment. Mean MMSE scores were 29.86 for YNH listeners and 29.05 for OHI listeners. Finally, each listener's test ear was examined otoscopically and deemed safe for use of an insert earphone.

Procedure

The University of Utah Institutional Review Board approved all test procedures. Listeners were tested individually in a quiet room, seated in front of a computer monitor, keyboard, and mouse. A custom MATLAB script with a graphical user interface presented the instructions and stimuli and recorded the listener responses. On each trial, a sentence was played from one channel of a TDT RP2 real-time processor, attenuated by a TDT programmable attenuator (PA-5) to the desired presentation level of 70 dB SPL, and routed via a headphone buffer (TDT HB7) to an insert earphone for monaural presentation. This presentation level is considered a slightly loud conversational level and was selected to be consistent with the level used in previous studies of vowel intelligibility (Ferguson, 2004, 2012). As in Ferguson (2012), the two listener groups heard the stimuli at the same presentation level to represent how OHI listeners might assess speech clarity in the real world, where most individuals with hearing loss do not use amplification (National Academies of Science, Engineering, and Medicine, 2016). Listeners heard the 1,148 sentence files in nine test blocks each containing four talkers (112 sentences) and one block containing five talkers (140 sentences). Each listener received a different random assignment of talkers to blocks, but the block containing five talkers was always administered last.

Listeners were instructed that their task was to listen to sentences and rate "how clearly the sentence was spoken"

Figure 1. Average audiometric thresholds for all older adult listeners with hearing impairment from 250 to 8000 Hz. Error bars indicate 1 SD . The reference line at 20 dB HL indicates the screening level criterion met by all young adult listeners with normal hearing.

on a scale from 1 to 7, where 1 represented *completely unclear (the lowest possible clarity)* and 7 represented *perfectly clear (the highest possible clarity)*. The labels attached to each of the categories are shown in Table 1. Listeners made their selection by clicking on the desired category and then pressing any key to continue. Listeners were familiarized with the graphical user interface and the task with a practice block of 30 stimuli not otherwise used in this study. These practice stimuli were chosen with a goal of eliciting ratings from the entire 7-point clarity scale and included both conversational and extremely hyperarticulated CID sentences spoken by the talker from Ferguson and Kewley-Port (2002), neutral conversational and clear sentences from two talkers recorded in a pilot project, and segments extracted from spontaneous speech produced by clinical nursing assistants in a residential healthcare setting wearing lapel microphones (Baldwin, Ferguson, & Williams, 2009). Listeners were told, “When you do your ratings, try to use the whole 1–7 scale, remembering to focus on how clearly the sentence was spoken with reference to the sentences you’re hearing today, not to speech you have heard in the past.” The experimenter watched the listeners’ responses during the practice block to make sure they were in fact using the whole range and reinstructed them and repeated the practice block as needed.

After the practice block, listeners were given a printed list of the 28 CID sentences they would be hearing and were asked to study it for 2 min “to help you focus on how clearly the speech was spoken, rather than how well you understand it.” They then completed five test blocks in the first test session. At the second session, listeners completed the practice block, studied the printed list, and completed the remaining five test blocks. In both sessions, listeners were offered short breaks after each test block. YNH listeners received course credit for their participation; OHI listeners were paid.

Data Analysis

Likert scale ratings are ordinal data, but often are treated as interval data and analyzed using statistical procedures such as analysis of variance (e.g., Ferguson & Kerr, 2009). Many recent authors in the areas of speech science, linguistics, and audiology have begun using mixed-effects regression models rather than analysis of variance to analyze their data, especially in studies involving multiple talkers, listeners, and/or items (e.g., Ferguson, 2012; McCloy, Wright, & Souza, 2015). Quené and van den Bergh (2004)

have written an excellent tutorial on mixed-effects models (which they call *multilevel models*), as well as other articles that both explore specific scientific questions and demonstrate the use of this powerful statistical analysis technique (Quené, 2008; Quené & van den Bergh, 2007). While linear mixed-effects models may be used for ordinal data, they assume that the data are normally distributed. A frequency plot of the present data (see Figure 2) suggests that they are considerably negatively skewed; a skewness test carried out in Stata 14 (StataCorp, 2015) confirmed this ($p < .0001$). Because this skewness violates the assumptions of linear regression, the raw ratings data in this study were analyzed using mixed-effects ordered logistic regression models. The fixed effects in each model varied depending on the specific question being tested and are specified below, but all models in which talker was not a fixed effect included talker and listener as nested random factors, whereas those that did include talker as a fixed effect had listener as a random variable. Models that also included item (i.e., the 28 individual sentences) as a random factor were attempted but would not converge. Stata 14 was used for all analyses.

Although having 41 talkers is helpful for characterizing the population of healthy talkers, it is statistically problematic to have a factor with 41 levels. This is especially the case when there are multiple factors of interest, because each interaction term will also have 41 levels. For this reason, we tested the effects of talker only for overall rated clarity and the magnitude of the speaking style effect. Effects of and interactions with the talker factor were not assessed for the two listener groups separately, thus avoiding a potential three-way interaction with 41 levels.

Results

Talker Variability and Speaking Style

Extensive talker variability is evident in Table 2, which shows average sentence clarity ratings for each talker in each style averaged across the listeners in each of the two listener groups. Among the 41 talkers, ratings (averaged over YNH and OHI listeners) ranged from 3.42 (M09) to 5.10 (F19) in conversational speech and from 4.35 (M10) to 6.09 (M03) in clear speech. One point of note is that, although listeners were instructed to rate the stimuli “with reference to the sentences you’re hearing today” and encouraged to use the whole scale, only 11 of the average conversational speech ratings and none of the clear speech ratings were less than 4, the midpoint of the rating scale. The skewness of the data distribution (Figure 2) is further evidence that, although listeners did use the whole rating scale, they tended to rate most of the speech they heard as “midway” or better. Indeed, the mode of the data distribution was a rating of 6.

To test the effect of talker statistically, three mixed-effects ordered logistic regression models were run with talker as a single fixed effect and listener as a single random effect. Each model had a different baseline talker to whom the other 40 talkers were compared. The three baseline

Table 1. Category labels used in rating task.

Category	Label
1	Minimum (lowest possible clarity)
2	Very unclear
3	Somewhat unclear
4	Midway
5	Somewhat clear
6	Very clear
7	Maximum (highest possible clarity)

Figure 2. Distribution of ratings data for all stimuli and listeners.

talkers were selected based on clarity ratings averaged across both speaking styles and both listener groups. To control the risk of Type I error, the numerous p values generated by each model were corrected in Stata 14 using the Holm-Šidák stepdown multiple comparison procedure (Ludbrook, 1998). In the first model, the baseline talker was M10, who had the lowest overall clarity (4.02). Clarity ratings were significantly higher for all other talkers ($z = 3.06$ to 23.1 , all corrected $p < .03$) except M09 ($z = 1.25$, corrected $p = .211$). In the second model, the baseline talker was F07, who had the highest overall clarity (5.38). Clarity ratings were significantly lower for all other talkers ($z = -23.1$ to -3.23 , all corrected $p < .03$). Finally, M20, whose average clarity rating (4.805) was closest to the mean across all talkers (4.802), was the baseline talker for the third model. Nine talkers had significantly lower clarity ratings than M20 ($z = -12.08$ to -3.06 , all corrected $p < .04$), 15 talkers had significantly higher clarity ratings than M20 ($z = 3.27$ to 11.67 , all corrected $p < .03$), and 16 had clarity ratings that did not differ significantly from M20 ($z = -2.53$ to 2.48 , corrected $p = .162$ to $.998$). The talker variability observed in Table 2 is thus statistically significant, despite nearly all ratings falling in the top half of the scale.

Great variability in the magnitude of the clear speech effect on perceived clarity can be seen in columns 4 and 7 of Table 2 and especially clearly in Figure 3. The data in Figure 3 are the average ratings shown in Table 2, but the talkers are ordered by the ratings given by YNH listeners in conversational speech. Among the 41 talkers, clear-minus-conversational differences between ratings averaged over

the YNH and OHI listeners ranged from 0.01 (M16) to 1.95 (M03). To demonstrate the talker variability for the speaking style effect statistically, mixed-effects ordered logistic regression models were run with speaking style and talker as fixed factors and listener as a random factor. The effects of speaking style and talker were significant ($z = 80.2$ and -18.63 , $p < .0001$), as were 27 of the 40 levels of the interaction between them when talker F01 and conversational speech were the (default) baselines ($|z| = 3.04$ to 13.73 , Holm-Šidák-corrected $p < .04$). To further explore this interaction, the effect of speaking style was tested for each individual talker; all but one talker (M16, $z = 0.25$, corrected $p = .8$) produced clear speech that was rated as being significantly clearer than their conversational speech (remaining 40 talkers, $z = 4.93$ to 20.58 , corrected $p < .03$). That is, when instructed to speak clearly, 40 of the 41 talkers produced speech that listeners rated as having been spoken significantly more clearly than their conversational speech.

Listener Group and Talker Subgroup Effects

The next analysis assessed whether the two listener groups differed in their ratings of speech clarity. Figure 3 shows a good deal of overlap between the lines for the YNH and OHI listeners, and the mean ratings for the two groups are practically identical ($\bar{x} = 5.55$ and 5.31 , respectively). In a mixed-effects ordered logistic regression model testing the fixed effects of speaking style and listener group (with talker and listener as nested random effects), the effect of speaking style was significant ($z = 87.82$, $p < .001$) while

Table 2. Average clarity ratings given by young adults with normal hearing (YNH listeners) and older adults with hearing impairment (OHI listeners) for sentences produced by 41 talkers in clear (CL) and conversational (CO) speech and the difference between the two styles (DIFF).

Talker ID	YNH listeners			OHI listeners		
	CL	CO	DIFF	CL	CO	DIFF
F01	5.4	4.6	0.8	5.3	4.9	0.4
F02	5.5	4.8	0.7	5.8	4.8	0.9
F03	5.2	3.9	1.3	5.1	3.9	1.2
F04	4.9	4.5	0.5	4.9	4.3	0.6
F05	6.0	4.3	1.7	6.0	4.1	1.8
F06	5.5	4.6	0.9	5.5	4.3	1.2
F07	6.2	4.9	1.3	5.9	4.5	1.4
F08	6.2	4.5	1.7	6.0	4.0	2.0
F09	4.9	4.7	0.3	5.2	4.6	0.6
F10	5.2	4.3	1.0	5.3	4.4	0.9
F11	6.0	4.3	1.7	6.0	4.3	1.6
F12	5.6	4.1	1.4	5.3	3.7	1.6
F13	5.2	3.3	2.0	5.3	3.9	1.3
F14	5.3	3.9	1.4	5.1	4.1	1.0
F15	5.6	4.5	1.1	5.3	4.0	1.3
F16	5.9	4.4	1.6	5.8	4.8	1.0
F17	5.7	4.1	1.6	5.7	4.2	1.5
F18	5.5	4.7	0.9	5.5	4.4	1.1
F19	5.6	5.3	0.3	5.6	4.9	0.7
F20	5.7	4.2	1.5	5.8	4.4	1.4
F21	4.9	4.4	0.5	4.9	4.4	0.6
M01	5.5	3.9	1.5	5.2	3.7	1.5
M02	4.3	3.6	0.8	4.7	4.3	0.5
M03	6.1	4.2	2.0	6.1	4.1	1.9
M04	5.2	4.2	1.0	5.4	4.5	1.0
M05	5.2	4.0	1.2	5.2	4.1	1.1
M06	5.4	3.6	1.8	5.7	3.9	1.7
M07	4.6	3.2	1.4	5.3	4.0	1.3
M08	4.7	4.3	0.3	5.3	4.9	0.4
M09	4.6	3.1	1.4	5.1	3.7	1.4
M10	4.3	3.4	0.9	4.4	4.0	0.4
M11	5.8	4.0	1.8	5.9	4.1	1.8
M12	5.8	3.9	1.9	5.7	4.1	1.7
M13	5.8	4.2	1.6	6.0	4.7	1.3
M14	5.0	3.6	1.5	5.1	3.9	1.2
M15	4.1	3.8	0.4	4.8	4.3	0.5
M16	5.0	5.0	0.0	5.1	5.1	0.0
M17	5.1	4.7	0.4	5.4	4.9	0.5
M18	5.5	4.0	1.4	5.4	4.2	1.2
M19	4.7	4.0	0.7	5.5	4.8	0.7
M20	5.0	4.3	0.7	5.4	4.6	0.8
M	5.3	4.2	1.1	5.4	4.3	1.1

Note. Mean values were calculated over all listeners.

the effect of listener group was not ($z = -0.98, p = .329$). The two-way interaction between speaking style and listener group was significant ($z = -4.58, p < .001$), and so, a stratified analysis was undertaken. It revealed that, in clear speech, ratings given by OHI listeners were significantly higher than those given by YNH listeners ($z = 2.1, p < .04$), whereas the group effect in conversational speech was not significant ($z = -0.35, p = .730$). Finally, Spearman rank correlations between the YNH and OHI ratings for the 41 talkers in the two speaking styles were strong, positive, and significant, $r_s(41) = .83$ in clear speech and $.73$ in conversational speech, both $p < .0001$. That is, YNH and

Figure 3. Average clarity ratings given by young adults with normal hearing (YNH) and older adults with hearing impairment (OHI) for sentences produced by 41 talkers in clear (CL) and conversational (CO) speech. Talkers are shown in order of ascending rated clarity for YNH listeners.

OHI listeners assigned very similar ratings to the individual talkers in each style. When conversational and clear ratings for the talkers were pooled into a set of 82 measures spanning a wider range of ratings, the Spearman rank correlation was even stronger, $r_s(82) = .92, p < .0001$.

The remaining two sets of mixed-effects ordered logistic regression analyses assessed the effects of two talker subgroup characteristics (gender and experience category) by adding them to the models as a third fixed factor (talker and listener remained as nested random effects).¹ The effects of talker gender for each speaking style and listener group can be seen in Figure 4. Listeners assigned higher clarity ratings to female talkers ($\bar{x} = 4.93$) than to male talkers ($\bar{x} = 4.66$); the mixed-effects logistic regression model that included talker gender as a fixed effect confirmed that this difference was significant ($z = 4.52, p < .0001$). However, the two-way interactions between speaking style and talker gender and between talker gender and listener group were also both significant ($z = 5.17$ and -8.86 , respectively, both $p < .0001$). Stratified analyses were then performed to determine the nature of these two-way interactions. The first focused on the speaking style by gender interaction, and it showed that women received significantly higher ratings than men in clear speech ($z = 12.89, p < .0001$), but not in conversational speech ($z = 1.61, p = .11$). The second stratified analysis assessed the listener group by gender interaction and found that, while both listener groups judged

¹In both of these analyses, the fixed effect of speaking style was significant, the effect of listener group was not, and the significant interaction between these two factors was attributable to a significant listener group effect that occurred in clear speech but not in conversational speech.

Figure 4. Average clarity ratings given by young adults with normal hearing (YNH) and older adults with hearing impairment (OHI) for sentences produced by female and male talkers in clear (CL) and conversational (CO) speech.

women to speak significantly more clearly than men (YNH: $z = 4.41$, $p < .001$; OHI: $z = 6.79$, $p < .001$), the gender difference was appreciably greater for YNH listeners than for OHI listeners (as seen in Figure 4). The three-way interaction between speaking style, talker gender, and listener group was not significant ($z = 1.41$, $p = 1.60$).

The other way that talker subgroups were created was based on how much prior experience communicating with individuals with hearing loss the talkers reported having. Based on their answers to a small set of questions asked after the clear speech recording session, talkers were assigned to four experience categories: *none* (no prior experience; $n = 14$), *little* (only a few experiences; $n = 8$), *some* (a relative or friend with hearing loss but less than one interaction per week; $n = 9$), or *lots* (at least one weekly contact with one or more individuals with hearing loss; $n = 10$). The mixed-effects ordered logistic regression model that included experience category as a fixed effect, along with speaking style and listener group, showed that the effect of talker experience was not significant ($z = 0.01$ to 0.41 , Holm-Šidák-corrected $p > .15$). Furthermore, neither the two-way interaction between talker experience and listener group nor the three-way interaction between speaking style, listener group, and talker experience were significant ($|z| = 0.33$ to 1.17 , corrected $p > .5$; $z = 1.21$ to 2.07 , corrected $p > .10$).

The two-way interaction between talker experience and speaking style, however, was significant ($|z| = 4.54$ to 15.04 , all corrected $p < .005$); this interaction is illustrated in Figure 5, which shows the effect of speaking style for each talker experience group. A stratified analysis confirms the large, significant effect of speaking style apparent in Figure 5 for all four experience groups ($z = 32.31$, 39.03 , 46.42 , and 55.89 for *none*, *little*, *some*, and *lots*, respectively;

Figure 5. Average clarity ratings averaged over all listeners for clear (CL) and conversational (CO) sentences produced by talkers with varying amounts of experience communicating with individuals with hearing loss.

all corrected $p < .005$). The source of the interaction was revealed when the effect of talker experience was tested for each speaking style; this analysis showed that the effect of talker experience was not significant in conversational speech ($z = 0.12$ to 2.24 , corrected $p = .073$ to $.904$), but in clear speech, it was. In particular, perceived clarity ratings in clear speech for the *little* group were significantly higher than those for the other three experience groups ($z = 7.81$ to 8.91 , corrected $p < .005$), while ratings for the *none*, *some*, and *lots* groups did not differ significantly from each other ($z = 0.2$ to 1.47 , corrected $p = .226$ to $.84$). While this difference is not visible in the raw ratings data shown in Figure 5, it can be seen in Figure 6, which displays the logarithmic odds ratios generated by the mixed-effects ordered logistic regression models that tested the effect of experience group on perceived clarity within each style. In clear speech, the *little* group is clearly distinguished from the other three, whereas in conversational speech, the variability within each experience category is greater than any difference between the categories.

Correlations With Other Measures

Subjective ratings of clarity for each talker may be correlated with other measurements that have been reported in prior work. For example, Ferguson (2004) reported vowel intelligibility scores in clear and conversational speech for YNH listeners, and Ferguson (2012) reported vowel intelligibility scores in both speaking styles for OHI listeners. Additional work has obtained acoustic measures for these talkers, all of which can be averaged to create a single value for each talker and speaking style for acoustic measurements, such as speaking rate, vowel duration, and vowel perimeter (e.g., Ferguson & Quené, 2014). To assess the relationship between

Figure 6. Odds ratio plot showing effect of talker experience category on rated clarity. Each symbol shows the log odds of a 1-point difference in clarity ratings for each experience re: the lowest bracket (“none”). Error bars indicate 1 *SD*. CL = clear speech; CO = conversational speech.

subjective ratings of perceived clarity and these other metrics, Spearman rank order correlation coefficients were calculated and analyzed with a Holm-Šidák correction factor for multiple comparisons. To increase the data range for each of the measures, conversational and clear speech values for each talker were included as independent observations in the same analysis.

Ratings of clarity by YNH listeners were significantly, moderately, and positively correlated with YNH vowel intelligibility scores reported by Ferguson (2004), $r_s(82) = .40$, corrected $p < .01$. Ratings of clarity by OHI listeners were also significantly, moderately, and positively correlated with OHI vowel intelligibility scores reported by Ferguson (2012), $r_s(82) = .43$, corrected $p < .001$. Because no significant differences were found between clarity ratings assigned by YNH and OHI listeners, to assess the relationship between subjective clarity ratings and acoustic measures, each talker's clarity ratings were collapsed across all listeners, thus yielding a single average rating for each talker and speaking style. Spearman rank order correlations were then carried out on these average talker clarity ratings and measures of vowel perimeter, vowel duration, and speaking rate. Ratings of clarity were found to be strongly and positively correlated with vowel duration and vowel perimeter, $r_s(82) = .76$ and $.63$, respectively; corrected $p < .001$. A strong, negative

correlation was observed between ratings of clarity and speaking rate, $r_s(82) = -.70$, corrected $p < .001$.

Discussion

Talker Variability and Speaking Style

Consistent with vowel intelligibility in noise for YNH listeners in Ferguson (2004) and for OHI listeners in Ferguson (2012), subjective ratings of speech clarity showed wide-ranging, statistically significant differences among the 41 talkers of the Ferguson Clear Speech Database in both conversational and clear speech. The variability within the two styles (each spanning a range of 1.7 points on the 1–7 scale) and the highest and lowest ratings seen in conversational speech agree well with ratings given by university student listeners to sets of three monosyllabic words spoken by 12 talkers in Hazan and Markham (2004). The 12 talkers in that study were six “good” and six “poor” talkers selected from a group of 33 adult talkers on the basis of the intelligibility of monosyllabic words in noise for children and adults with normal hearing. The student listeners in Hazan and Markham used 7-point scales to rate the talkers on a number of perceptual dimensions, of which their dimension “mumbly–precise” most closely

resembles the scale used here. The average rating for the set of six poor talkers was 3.39, remarkably similar to the 3.42 average rating given to conversational speech produced by the lowest-rated talker in this study, whereas the average mumbly-precise rating for the six good talkers was 4.96, again quite close to the highest average conversational rating here, which was 5.1. The similarity between the present ratings and those reported in Hazan and Markham is even more noteworthy given how the studies differ in terms of speech material, speaking instructions, and talker dialect.

Taken together, these data suggest that, for clean laboratory recordings of conversational or citation-style speech, listeners judging how clearly this speech was spoken assign ratings around the middle of the 7-point scale and higher. Even the poorest laboratory conversational speech, such as that produced by the current talker M09 or by the six poor talkers in Hazan and Markham (2004), never seems to receive ratings lower than 3 on this scale. For this study, these skewed ratings could be due to the stimuli used for familiarization and the instructions given to our listeners. As noted above, the 30 practice items were selected to span a very large range of clarity, from extremely hyperarticulated read sentences to phrases extracted from spontaneous speech. Furthermore, listeners were instructed to rate the sentences “with reference to the sentences you’re hearing today, not to speech you have heard in the past.” Thus, if the practice set contained samples that had much poorer clarity than the test stimuli, those poor samples might have given the listeners an “anchor,” compared with which the test stimuli would all sound clearer. However, an examination of the data from the practice set revealed the ratings for those 30 samples to be nearly as skewed as the complete stimulus set. The average rating for the lowest-rated practice item was 2.97. This, along with the agreement between the present data and those of Hazan and Markham (2004), suggests that the skewness in our ratings data is not an artifact unique to this experiment.

Also consistent with the earlier vowel intelligibility studies, the difference between clear and conversational speech ratings varied widely among the talkers. In contrast with vowel intelligibility, however, where less than 60% of talkers produced a clinically (for YNH listeners in Ferguson, 2004) or statistically (for OHI listeners in Ferguson, 2012) significant clear speech vowel intelligibility benefit, statistically significant differences between subjective ratings of clear and conversational sentences were observed for 40 of the 41 talkers. This pattern is strikingly similar to the sentence intelligibility results seen in Schum (1996) and in Lam and Tjaden (2013), in which 19 out of 20 and 11 out of 12 talkers, respectively, produced a clear speech effect greater than 5 percentage points. If we consider a talker who produces a significant, positive clear speech effect to be a “successful” clear speaker, then it seems that the proportion of talkers who are successful depends on the materials used to measure success. When the materials are vowels for identification, the “clear speech success rate” is just 60%, but when the materials are sentences, we see success rates ranging from 92% in Lam and Tjaden to 98% in this study.

These high clear speech success rates for sentence materials are consistent with Lindblom’s hyperspeech versus hypospeech theory, which states that talkers spontaneously begin hyperarticulating when they encounter a listener who has difficulty understanding them (Lindblom, 1990). The talker differences found for the magnitude of the clear speech effect for vowel intelligibility, sentence intelligibility, and perceived sentence clarity all indicate that individual talkers adopt different strategies when instructed to read in a conversational and in a clear manner. While not all of these strategies or, more likely, sets of strategies apparently lead to improved vowel intelligibility, most talkers adopt enough strategies that their speech sounds clearer and may even be more intelligible. The relationship between vowel intelligibility and rated sentence clarity for the 41 Ferguson database talkers supports this idea; the Spearman correlation was only .4, indicating that vowel intelligibility only accounts for about 16% of the variance in perceived clarity. This suggests, in contrast with the assumption that underlies much of the previous work in this line of research, that vowel intelligibility is not a very strong index of overall intelligibility. A new database, the Utah Speaking Style Corpus (Hunter et al., 2016), includes vowels in /bVd/ context, as well as six lists of sentences from the Hearing in Noise Test (Nilsson, Soli, & Sullivan, 1994), and thus, will permit further investigation of the relationship between vowel intelligibility and overall intelligibility.

Listener Group and Talker Subgroup Effects

While YNH and OHI listeners gave similar overall clarity ratings, a significant difference between the listener groups was found for clear speech. In particular, OHI listeners assigned higher ratings to clear speech than YNH listeners did. One possible explanation for this is that certain clear speech strategies, such as producing greater stop burst and fricative energy, make consonants that are inaudible in conversational speech for OHI listeners, audible. While YNH listeners would perceive this increased energy as an improvement in clarity, OHI listeners might perceive clear speech as containing consonants that were not present in conversational speech. Thus, an acoustic change of a particular magnitude could be perceived quite differently by listeners with hearing loss than by listeners with normal hearing. If improved consonant audibility does explain OHI listeners’ higher ratings of clear speech, then we might expect a significant correlation between these ratings and the listeners’ high-frequency hearing sensitivity. To test this prediction, we carried out Spearman rank order correlations on the OHI listener clarity ratings for clear speech and two audiometric variables: the traditional pure-tone average (computed over thresholds measured at 500, 1000, and 2000 Hz) and a high-frequency pure-tone average (computed over thresholds measured at 1000, 2000, and 4000 Hz). Neither the pure-tone average nor the high-frequency pure-tone average was significantly related to the clarity ratings assigned by the OHI listeners ($r_s = .13$ and $.16$, respectively, $p > .9$).

Improved audibility of high-frequency phonemes thus does not appear to explain the listener group effect in clear speech. Another possible explanation involves age-related changes beyond the audiogram. Older adults are known to have poorer temporal processing than younger adults do, as well as slower cognitive processing. Either of these factors, or both in combination, could cause the slowed rate of clear speech to have a greater effect on OHI listeners' perception of clarity than it would for YNH listeners. A final, related possibility is that, given the difficulty that OHI listeners likely experience understanding speech outside the lab, they appreciate clear speech more and assign higher clarity ratings than YNH listeners. This study was designed to determine the range of clear speech clarity differences present in the Ferguson Clear Speech Database for two populations: a baseline group of listeners with normal hearing and a sample of the population for whom clear speech is intended, listeners with hearing loss. The OHI listeners were tested without any amplification because at least two thirds of this population do not use hearing aids (National Academies of Science, Engineering, and Medicine, 2016).

These design choices make it difficult to ascertain why the OHI listeners gave higher speech clarity ratings to the clear speech materials than the YNH listeners did. The effects of high-frequency sensitivity versus other age-related changes could be explored in future studies in several ways. One would be to include older adult listeners with normal hearing. Comparing their ratings to those given by the present OHI listeners would allow the examination of the effects of the pure-tone audiogram while controlling for the effects of age group differences. Another strategy would be to have YNH listeners rate sentences that have been processed in a manner simulating the sensorineural hearing loss of the OHI listeners. This would allow the effects of the age group to be uncontaminated by audibility differences.

Listener group was also found to interact significantly with talker gender, although not quite as we had anticipated. For the YNH listeners, previous speech intelligibility research had led us to expect that clarity ratings would be higher for female talkers than for male talkers, and indeed, this is what we found. However, for OHI listeners, the single previous study that has compared intelligibility for male talkers versus female talkers (Ferguson, 2012) found no effect of talker gender on vowel intelligibility in noise. Furthermore, listeners with hearing loss often report particular difficulty understanding female talkers (Helfer, 1995). We thus predicted that OHI listeners would give higher ratings to male talkers than to female talkers. This is not what happened, however: The effect of talker gender, while larger for YNH than for OHI listeners, was significant for both listener groups. This disconnect between audiology patient complaints and our OHI listener data might have its roots in the fact that stimuli in this study (as in all previous clear speech perceptual studies) were all scaled to the same root-mean-square amplitude. When communicating in the real world, it is possible that women produce speech at a lower intensity than men do, making women more difficult for listeners with hearing

loss to understand than men. Another possibility is that the complaints audiologists hear are not coming from all patients and directed at all talkers. It might be that our male patients are complaining about their most frequent communication partner, which for most men is their wife, and that these complaints have been inappropriately generalized to other patients and communication partners. A great deal more research is needed to determine the effects of talker gender on speech understanding by listeners with hearing loss, including patient surveys to pin down their specific complaints, as well as additional acoustic and perceptual studies examining the effect of talker gender for different listener groups.

Speaking style also interacted with talker gender. In conversational speech, clarity ratings were similar for male and female talkers in conversational speech, and both male and female talkers showed significantly higher ratings in clear speech than in conversational speech. For clear speech, however, clarity ratings were significantly higher for female talkers than for male talkers. Ferguson (2012) also found that the clear speech vowel intelligibility benefit enjoyed by both YNH and OHI listeners was larger for female talkers than for male talkers. Female talkers thus appear to achieve a larger perceptual benefit when speaking clearly than male talkers do. Additional investigation should establish the different strategies or acoustic features that female talkers use to achieve this greater benefit.

Finally, some interesting group effects were observed when we compared groups of talkers with different amounts of reported experience communicating with individuals with hearing loss. While there was no significant difference among the four experience groups in conversational speech, the clear speech of talkers in the *little* group was rated significantly higher than the clear speech of talkers in the three other groups (*none*, *some*, and *frequent*). This same, counter-intuitive pattern was observed in Ferguson (2012) for vowel intelligibility, suggesting that having just a few experiences communicating with people with hearing loss leads to talkers producing both clearer and more intelligible speech than those talkers with more experience or no experience. Naturally, we would expect that frequent communication partners of individuals with hearing loss (individuals with more experience) would have developed the strategies requisite to ensure successful communication, presumably by acquiring strategies that they deem beneficial. Ferguson (2012) suggested that as talkers gain experience communicating with people with hearing loss, they discard or abandon various clear speech strategies, perhaps in favor of motor economy (e.g., Lindblom, 1990).

While the results from this study support this assertion, we note that the Ferguson Clear Speech Database was not specifically designed to test the effects of experience communicating with people with hearing loss on the ability to produce effective clear speech. Rather, talkers were assigned to the four experience categories on the basis of how they responded when asked whether they had experience talking to people with hearing loss. The 27 talkers who responded affirmatively were then asked about their relationship with

these individuals, as well as how often and over how long a period of time they conversed with them. The numbers of talkers in the three groups of talkers who had little, some, or frequent experience communicating with people with hearing loss were both relatively small ($n \leq 10$) and also unequal (though mixed-effects models are robust to unequal sample size). If talkers actually do become less effective clear speakers as they gain experience talking with people with impaired hearing, then clinicians working with the family members of such people need strategies to counteract this loss in effectiveness. To explore whether (and if so, how and why) experience communicating with listeners with hearing loss affects the ability to produce effective clear speech, future research should record much larger groups of talkers recruited specifically into different experience categories.

Correlations With Other Measures

Significant correlations were found between subjective ratings of clarity and other measures of speech. As ratings of talker clarity increased, vowel intelligibility increased irrespective of age or hearing status. For the 41 talkers, higher clarity ratings were associated with larger vowel space perimeters. Decreased speaking rate and increased vowel duration were also associated with higher ratings, suggesting that speaking more slowly can lead to improved clarity.

These results are consistent with previous data reported by Ferguson and Kerr (2009). In their study, they reported that sentence clarity was strongly and positively correlated with both vowel intelligibility and vowel perimeter, which was replicated in the current experiment. This study extended the findings to OHI listeners, showing consistent ratings and similar correlations between clarity ratings and acoustic or other perceptual measures in the population for whom clear speech is intended. Ferguson and Kerr also reported a moderate, negative correlation between speaking rate and ratings of clarity, but the correlation failed to reach significance. With the additional talkers included in this study, a strong, negative correlation was observed between rated clarity and speaking rate, suggesting that the earlier study may have been underpowered.

Limitations

A few weaknesses of this study have already been noted above. Our two listener groups differed in terms of both age and hearing status, and no steps were taken to equate audibility for the two groups, rendering us unable to determine the source of any group differences. Another limitation is that each listener rated each stimulus just one time. Although this permitted the experiment to be carried out in just two test sessions, it left us with no information about how reliable or stable the listeners' ratings were or whether intrasubject variability differed between the two listener groups. The use of a Likert scale for clarity ratings was also potentially problematic. First, as noted earlier, even though most studies treat them as interval data, Likert data are ordinal. In addition, our ratings data were positively

skewed, necessitating rather burdensome statistical analysis techniques. Using a visual analog scale would yield continuous data, solving the first problem, and might solve the second problem as well. Finally, the data here are ratings of how clearly the listeners judged the talkers to have produced each sentence. While we presume that improvements in clarity ratings when talkers speak clearly are a good proxy for improvements in intelligibility, we have no concrete evidence for this. A future study will address this issue by collecting both clarity and intelligibility data from the same listeners for the same material.

Conclusion

Talkers vary considerably in how clear YNH and OHI listeners judge their clear and conversational speech to be, but almost all produce perceptible and significant differences in rated clarity between the two speaking styles and for the two listener groups. The talker variability seen here, as well as a near-universal clear speech benefit, has also been reported for sentence intelligibility (Lam & Tjaden, 2013; Schum, 1996), whereas for vowel intelligibility, only 50%–60% of talkers have been found to produce a significant clear speech benefit (Ferguson, 2004, 2012; Rogers et al., 2010). These differences support our contention that the magnitude of the clear speech benefit depends on the speech materials. It also appears to depend on gender, with female talkers producing greater perceptual effects when they speak clearly than male talkers do, and may depend on the amount of experience talkers have communicating with listeners with hearing loss (though perhaps in a counterintuitive way).

The two materials for which perceptual data have thus far been measured in clear and conversational speech for all 41 talkers in the Ferguson Clear Speech Database, vowels (Ferguson, 2004, 2012) and sentences (this study), of course differ enormously from each other. The /bVd/ stimuli used in the vowel intelligibility studies are single words excised from neutral sentence frames that differed only in the identity of the vowel. In contrast, the CID sentences used here are meaningful, containing four to seven words and varying in prosody (the lists include questions, statements, and imperatives). Sentences thus provide much more information for listeners to judge than vowels in a fixed phonetic context. This additional information is both segmental (there are more phonemes) and suprasegmental (they contain temporal and pitch information within the sentence). It would be interesting to see how much talkers vary when the stimuli are meaningful and contain more phonemes but contain minimal suprasegmental information (e.g., individual words), or vice versa (e.g., anomalous sentences).

Experiments are underway to measure the perceptual characteristics of the first of these stimulus types, namely monosyllabic words. When these data have been collected and acoustic analyses of those words have been completed, we will be able to identify which clear speech acoustic changes not only make speech sound clearer but also make

words easier to understand. These strategies could then be taught to frequent communication partners of our patients with hearing loss. Having perceptual data from a wide range of materials will help this endeavor by revealing which materials, containing which balance of not just segmental and suprasegmental information but also lexical information, best index talkers' effectiveness in producing clear speech.

Acknowledgments

This research was supported by the National Institutes of Health Grant DC008886. The development of the Ferguson Clear Speech Database was supported by National Institutes of Health Grant DC02229 to Indiana University. Hearing evaluations for the older adults with hearing impairment were provided by the University of Utah Speech-Language-Hearing Clinic. Portions of this work were presented at the Academy of Rehabilitative Audiology Institute, Providence, RI, 2012. The authors are grateful to Craig Berg, Amanda Malone, and Patrick Pead for their assistance with data collection, Emily Kerr and Billy Speer for stimulus preparation, and Greg Stoddard, Hugo Quené, and Simone Graetzer for statistical guidance.

References

- American National Standards Institute.** (2010). *Specifications for audiometers (ANSI/ASA S.36-2010)*. New York, NY: American National Standards Institute.
- Baldwin, A. B., Ferguson, S. H., & Williams, K.** (2009). *Subjective ratings of foreign-accented speech by older and younger adults*. Poster presented at Aging and Speech Communication: An International and Interdisciplinary Conference, Bloomington, IN.
- 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. [https://doi.org/10.1016/S0167-6393\(96\)00063-5](https://doi.org/10.1016/S0167-6393(96)00063-5)
- Cienkowski, K. M., & Speaks, C.** (2000). Subjective vs. objective intelligibility of sentences in listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*, 43, 1205–1210. <https://doi.org/10.1044/jslhr.4305.1205>
- Cox, R. M., Alexander, G. C., & Rivera, I. M.** (1991). Comparison of objective and subjective measures of speech intelligibility in elderly hearing-impaired listeners. *Journal of Speech and Hearing Research*, 34, 904–915. <https://doi.org/10.1044/jslhr.3404.904>
- Davis, H., & Silverman, S. R.** (1978). *Hearing and deafness* (4th ed.). New York, NY: Holt, Rinehart & Winston.
- 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. <https://doi.org/10.1121/1.1788730>
- Ferguson, S. H.** (2012). Talker differences in clear and conversational speech: Vowel intelligibility for older adults with hearing loss. *Journal of Speech, Language, and Hearing Research*, 55, 779–790. https://doi.org/10.1092-4388_2011_10-0342
- Ferguson, S. H., & Kerr, E. E.** (2009). Subjective ratings of sentences in clear and conversational speech. *Journal of the Academy of Rehabilitative Audiology*, XLII, 51–66. <http://audrehab.org/jara/2009/Ferguson%20and%20Kerr.pdf>
- 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. <https://doi.org/10.1121/1.1482078>
- Ferguson, S. H., Poore, M. A., & Shrivastav, R.** (2010). Acoustic correlates of reported clear speech strategies. *Journal of the Academy of Rehabilitative Audiology*, XLIII, 45–64. Retrieved from <http://www.audrehab.org/jara/2010/Ferguson%20et%20al.pdf>
- Ferguson, S. H., & Quené, H.** (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 135, 3570–3584. <https://doi.org/10.1121/1.4874596>
- Ferguson, S. H., Widener, T. D., & Mackey, T.** (2011). Effects of talker experience on perceived clarity and acoustic features of clear versus conversational speech. *The Journal of the Acoustical Society of America*, 129, 2454(A). <https://doi.org/10.1121/1.3588061>
- Fogerty, D., & Kewley-Port, D.** (2009). Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *The Journal of the Acoustical Society of America*, 126, 847–857. <https://doi.org/10.1121/1.3159302>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R.** (1975). “Minimal mental state.” A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Hazan, V., & Baker, R.** (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, 130, 2139–2152. <https://doi.org/10.1121/1.3623753>
- Hazan, V., & Markham, D.** (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *The Journal of the Acoustical Society of America*, 116, 3108–3118. <https://doi.org/10.1121/1.1806826>
- Helfer, K. S.** (1995). Auditory perception by older adults. In R. A. Huntley & K. S. Helfer (Eds.), *Communication in later life* (pp. 41–84). Boston, MA: Butterworth-Heinemann.
- Hunter, E. J., Ferguson, S. H., Leishman, T., Maxfield, L. M., Graetzer, S., & Bottalico, P.** (2016). Gender differences and speech accommodation in occupational settings. *The Journal of the Acoustical Society of America*, 139, 2105. <https://doi.org/10.1121/1.4950254>
- Kewley-Port, D., Burkle, T. Z., & Lee, J. H.** (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 122, 2365–2375. <https://doi.org/10.1121/1.2773986>
- Lam, J., & Tjaden, K.** (2013). Intelligibility of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, 56, 1429–1440. [https://doi.org/10.1044/1092-4388\(2013\)12-0335](https://doi.org/10.1044/1092-4388(2013)12-0335)
- Lee, J. H., & Kewley-Port, D.** (2009). Intelligibility of interrupted sentences at subsegmental levels in young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 125, 1153–1163. <https://doi.org/10.1121/1.3021304>
- 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.
- Ludbrook, J.** (1998). Multiple comparison procedures updated. *Clinical and Experimental Pharmacology and Physiology*, 25, 1032–1037. <https://doi.org/10.1111/j.1440-1681.1998.tb02179.x>
- McCloy, D., Wright, R. A., & Souza, P. E.** (2015). Talker versus dialect effects on speech intelligibility: A symmetrical study.

- Language and Speech*, 58, 371–386. <https://doi.org/10.1177/0023830914559234>
- National Academies of Sciences, Engineering, and Medicine.** (2016). *Hearing health care for adults: Priorities for improving access and affordability*. Washington, DC: National Academies Press. <https://doi.org/10.17226/23446>
- Nilsson, M., Soli, S. D., & Sullivan, J. A.** (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, 95, 1085–1099. <https://doi.org/10.1121/1.408469>
- 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. <https://doi.org/10.1044/jshr.2801.96>
- Quené, H.** (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America*, 123, 1104–1113. <https://doi.org/10.1121/1.2821762>
- Quené, H., & van den Bergh, H.** (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103–121. <https://doi.org/10.1016/j.specom.2004.02.004>
- Quené, H., & van den Bergh, H.** (2007). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59, 413–425. <https://doi.org/10.1016/j.jml.2008.02.002>
- Rogers, C. L., DeMasi, T. M., & Krause, J. C.** (2010). Conversational and clear speech intelligibility of /bVd/ syllables produced by native and non-native English speakers. *The Journal of the Acoustical Society of America*, 128, 410–423. <https://doi.org/10.1121/1.3436523>
- Schum, D. J.** (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, 7, 212–218. http://www.audiologynow.com/sites/default/files/journal/JAAA_07_03_10.pdf
- Smiljanić, R., & Bradlow, A. R.** (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass*, 3, 236–264. <https://doi.org/10.1111/j.1749-818X.2008.00112.x>
- StataCorp.** (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462. <https://doi.org/10.1044/jshr.2803.455>
- Tillman, T., & Carhart, R.** (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words, Northwestern University Auditory Test No. 6*. Brooks AFB, TX: USAF School of Aerospace Medicine.
- Tye-Murray, N.** (2014). *Foundations of aural rehabilitation: Children, adults, and their family members* (4th ed.). Stamford, CT: Cengage Learning.
- Uchanski, R. M.** (2005). Clear speech. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 207–235). Malden, MA: Blackwell.

Appendix A

Central Institute for the Deaf Everyday Sentences (Davis & Silverman, 1978) Recorded in Conversational Speech

1. It's time to go.
2. Do you want to wash up?
3. The water's too cold for swimming.
4. It's no trouble at all.
5. Here are your shoes.
6. Have you been working hard lately?
7. How do you know?
8. Move out of the way.
9. Pass the bread and butter, please.
10. Weeds are spoiling the yard.
11. Breakfast is ready downstairs.
12. Let's get a cup of coffee.
13. I hate driving at night.
14. I'll carry the package for you.

Appendix B

Central Institute for the Deaf Everyday Sentences (Davis & Silverman, 1978) Recorded in Clear Speech

1. You'll get fat eating candy.
 2. I'll see you right after lunch.
 3. There's a good ballgame this afternoon.
 4. Music always cheers me up.
 5. How do you spell your name?
 6. What are you hiding under your coat?
 7. The phone call's for you.
 8. Come here when I call you.
 9. I'll think it over.
 10. Walking's my favorite exercise.
 11. Where are you going?
 12. Wait just a minute!
 13. I'll catch up with you later.
 14. Call me a little later.
-