Research Article

Judgments of Emotion in Clear and Conversational Speech by Young Adults With Normal Hearing and Older Adults With Hearing Impairment

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

Purpose: In this study, we investigated the emotion perceived by young listeners with normal hearing (YNH listeners) and older adults with hearing impairment (OHI listeners) when listening to speech produced conversationally or in a clear speaking style.

Method: The first experiment included 18 YNH listeners, and the second included 10 additional YNH listeners along with 20 OHI listeners. Participants heard sentences spoken conversationally and clearly. Participants selected the emotion they heard in the talker's voice using a 6-alternative, forced-choice paradigm.

Results: Clear speech was judged as sounding angry and disgusted more often and happy, fearful, sad, and neutral less often than conversational speech. Talkers whose

clear speech was judged to be particularly clear were also judged as sounding angry more often and fearful less often than other talkers. OHI listeners reported hearing anger less often than YNH listeners; however, they still judged clear speech as angry more often than conversational speech.

Conclusions: Speech spoken clearly may sound angry more often than speech spoken conversationally. Although perceived emotion varied between YNH and OHI listeners, judgments of anger were higher for clear speech than conversational speech for both listener groups.

Supplemental Materials: <https://doi.org/10.23641/asha.5170717>

When speaking to individuals with hearing loss, talkers tend to modify their speech from a conversational style to a more “clear” style. The resulting clear speech can be more intelligible than conversational speech (e.g., Ferguson, 2004; Liu, Del Rio, Bradlow, & Zeng, 2004; Picheny, Durlach, & Braida, 1985), particularly when speech understanding is hindered by some barrier, such as hearing loss (e.g., Ferguson, 2012; Smiljanić & Bradlow, 2005). Communication partners of individuals with hearing loss are often counseled and trained to “speak clearly” to help their partners receive this intelligibility benefit and improve communication. However, audiology patients sometimes complain anecdotally of communication partners sounding angry when in fact

these communication partners are only trying to speak clearly. In the present study, we investigated the basis for these complaints by examining whether listeners perceive anger or other emotional content when they listen to speech that has been spoken clearly.

In the psychology literature, models of emotion differentiation suggest that emotions are either discrete (i.e., having distinct emotion categories; Ekman, 2003), dimensional (i.e., varying in levels of arousal or valence; e.g., Russell, 1980), or some combination of both (e.g., Mortillaro, Meuleman, & Scherer, 2012; Weninger, Eyben, Schuller, Mortillaro, & Scherer, 2013). Ekman and Cordaro (2011), in line with a more discrete model, have suggested that emotions can be characterized by identifiable markers (Ekman, 1992, 2003). Although a comprehensive list of emotions and their markers is quite extensive, Ekman (2003) suggests that complex emotions generally fall under the umbrella of a more basic emotion. For example, the basic emotion family of sadness encompasses emotions such as shame, depression, and others. Ekman and Cordaro list seven basic emotions—sadness, happiness, fear, disgust, anger, contempt, and surprise—and explain that these seven

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

Correspondence to Shae D. Morgan: shae.morgan@utah.edu

Editor: Nancy Tye-Murray

Associate Editor: Kathleen Cienkowski

Received June 20, 2016

Revision received October 17, 2016

Accepted December 6, 2016

https://doi.org/10.1044/2017_JSLHR-H-16-0264

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

emotions have distinctive universal signals that differentiate them (e.g., unique facial expressions).

When emotional content is presented through auditory information alone, humans are able to classify emotions on the basis of acoustic characteristics specific to that emotion. For example, sadness is characterized by decreases in fundamental frequency (f_0), articulation rate, and overall energy relative to neutral speech whereas happiness is characterized by increases in those three domains (Banse & Scherer, 1996). Listeners have been shown to be quite adept at identifying portrayals of emotion in speech and music even when visual cues (e.g., facial and body cues) are not available (Banse & Scherer, 1996; Juslin & Laukka, 2003; Juslin, Laukka, Davidson, & Scherer, 2001; Ruffman, Halberstadt, & Murray, 2009; Scherer, Johnstone, & Klammer, 2003).

The expression of anger in speech is associated with several acoustic changes from neutral speech, including raised f_0 , increased energy in the high frequencies, increased overall energy, and greater pitch variability (Banse & Scherer, 1996; Juslin & Laukka, 2003; Scherer, 1986; Scherer et al., 2003). It is interesting to note that clear speech, when compared with conversational speech, shares many acoustic modifications with angry speech, when compared with neutral speech. Similar to angry speech, clear speech is also characterized by raised f_0 , increased energy in the high frequencies, and greater pitch variability (e.g., Picheny, Durlach, & Braid, 1986; Smiljanić & Bradlow, 2009).

Not all talkers produce clear speech that is drastically different from their conversational speech when instructed to speak clearly. Large individual talker differences have been demonstrated for vowel intelligibility in noise (Ferguson, 2004, 2012) and for perceived clarity (Ferguson & Morgan, 2015). Ferguson (2004) recorded 41 talkers who produced clear and conversational speech; these recordings are known as the Ferguson Clear Speech Database. The talkers showed considerable variability in their vowel intelligibility in clear and conversational speech as well as in the magnitude of their clear speech vowel intelligibility effect (i.e., the difference in vowel intelligibility between the two styles for a given talker).

Ferguson and Morgan (2015) showed significant individual differences between listeners judging the perceived clarity of clear and conversational speech using the talkers from the Ferguson Clear Speech Database. These results suggest that some individuals were more successful than others at producing clear speech that actually sounds clear. Talkers produce clear-sounding clear speech in part by making larger acoustic adjustments (e.g., raised f_0 and increased vowel perimeter) from their conversational speaking style than talkers who produce less clear-sounding clear speech; however, these alterations may also affect judgments that listeners make about the talker on the basis of their speech, such as their gender or their emotional state. For example, Booz (2016) found that young adults with normal hearing (YNH listeners) judged talkers as sounding more feminine in clear speech than in conversational speech.

Ferguson and Quené (2014) found that older adults with hearing impairment (OHI listeners) weigh acoustic

cues differently than YNH listeners do when identifying vowels. Some acoustic enhancements of clear speech that were beneficial to vowel intelligibility for YNH listeners (e.g., raising the second formant for front vowels) were detrimental to OHI listeners because the change moved a prominent acoustic cue to a frequency range with greater sensitivity loss. The limited access to high-frequency information in a signal due to hearing loss may affect not only intelligibility for some vowels, but also the identification of emotions that are characterized in part by high-frequency cues, such as anger and happiness.

In addition, age differences have been shown for both visual and auditory emotion perception (Ruffman et al., 2009). Ruffman et al. (2009) presented independent groups of young and older adults with pictures of emotional faces as well as acoustic stimuli produced in an emotional manner. The participants selected the facial expression that matched the emotion expressed in the auditory stimulus. Older adults were significantly less accurate than younger adults at matching some emotions (including anger) to the appropriate facial expression. These apparent age-related changes in emotion perception suggest that the degree to which listeners perceive anger or other emotions when listening to clear speech may differ for younger and older adults (see also Carstensen & Mikels, 2005).

In the current study, we examined if clear speech carries an emotional index. Although the acoustic similarities between clear speech and angry speech suggest that listeners may perceive anger when listening to clear speech, other emotion portrayals may surface as well. In addition, in the present study, we examined talker differences in listener judgments of emotion for two groups, “good” and “poor” clear speakers (as determined by clarity ratings from Ferguson & Morgan, 2015). Listeners may judge talkers with “good” clear speech (i.e., talkers with large perceived differences between styles) as sounding emotional more often than talkers with “poor” clear speech (i.e., talkers with small perceived differences between styles). Last, given that comments by individuals with hearing loss prompted this line of research, in the present study, we also examined the perception of emotion in clear and conversational speech in OHI listeners.

General Methods

Materials

Speech stimuli were selected from the Ferguson Clear Speech Database (Ferguson, 2004). The database consists of two lists of 188 sentences—one list recorded in a conversational speaking style and the other in a clear speaking style—spoken by 41 talkers. Conversational and clear sentences were recorded in separate sessions (with conversational speech always recorded first). Ferguson (2004) detailed additional information regarding the recording procedures and materials.

To minimize semantic priming of specific emotions, 14 emotionally neutral sentences were chosen as stimuli

(e.g., “He picked up the bead and put it away”; see Appendix for a complete list of the sentences used). The same set of 14 sentences was used for both speaking styles and for all eight talkers. The repeated use of a small set of sentences was intended to help listeners focus on the overall sound of the speech samples rather than on the words being said. The resulting 224 sentences (14 sentences $\times$ 2 styles $\times$ 8 talkers) were scaled to the same average root-mean-square amplitude (–25 dB) in Cool Edit 2000 to prevent listeners from using any amplitude differences between the two styles—a very common procedure in clear speech perception studies (e.g., Uchanski, 2005). This amplitude corresponds to approximately 85 dB SPL.

Talkers

Eight talkers (four men, four women) were selected for the present study on the basis of clear–conversational differences in perceived clarity as judged by YNH listeners (Ferguson & Morgan, 2015). Talker clarity ratings from the earlier study were assigned on a scale from 1 (*lowest possible clarity*) to 7 (*highest possible clarity*); average ratings for each talker in each style are summarized in Table 1. To explore potential effects of talker variability on perceived emotion, talkers were placed into two groups of four, each consisting of two men and two women. The talkers in Group A had some of the largest differences in perceived clarity between their clear and conversational speech and thus could be considered “good” clear speakers. In contrast, the talkers in Group B had some of the smallest differences in perceived clarity between the two styles and thus could be considered “poor” clear speakers.

Procedures

All subject recruitment and other procedures were in accordance with a protocol approved by the University of Utah Institutional Review Board. Listeners sat in a sound-treated room in front of a computer monitor, mouse, and

Table 1. Subjective ratings of perceived clarity (Ferguson & Morgan, 2015) for conversational (CO) and clear (CL) speech along with the difference score (DIFF) between the two styles for two groups (Group A and Group B) of four talkers selected from the Ferguson Clear Speech Database.

Talker	CO	CL	DIFF
Group A			
M03	4.1	6.2	2.1
M12	3.8	5.8	2.0
F11	4.2	6.0	1.8
F08	4.4	6.1	1.8
Group B			
F21	4.3	4.9	0.5
F04	4.4	4.8	0.4
M15	3.8	4.2	0.4
M08	4.3	4.7	0.4

Note. Clarity was rated on a scale from 1 (*lowest possible clarity*) to 7 (*highest possible clarity*).

keyboard. A custom MATLAB graphical user interface (GUI) presented the instructions and guided the listeners through the experiment. Both on the screen and verbally, participants were told to “Judge the emotion you think you hear when listening to each sentence.” The category options “anger,” “fear,” “disgust,” “sadness,” “happiness,” and “neutral” were presented in a six-alternative, forced-choice task paradigm. Participants were instructed to select the “neutral” option if they heard no specific emotion. Recall that Ekman and Cordaro (2011) listed these emotions, along with contempt and surprise, as basic emotion categories. Contempt was excluded from this study because it is often difficult to conceptualize due to its lack of common use (Ekman, 2003). Surprise, on the other hand, was excluded because it is a brief emotion, quickly shifting to another emotion immediately after its expression (Ekman & Cordaro, 2011). Although we were particularly interested in if clear speech sounds angry, for this first study assessing the perception of emotion in different styles of speech, we chose to treat emotions categorically to see if listeners made judgments consistent with clinical observations and reports.

Stimuli were presented through a Tucker-Davis Technologies (TDT) RP2.1 real-time processor attenuated by a TDT programmable attenuator (PA-5) to approximately 70 dB SPL and routed via a headphone buffer (TDT HB7) to Sennheiser HD 515 headphones for monaural presentation. Participants were offered the opportunity to adjust the presentation level in 5-dB increments until they deemed the level comfortable (average YNH presentation level was approximately 71.5 dB SPL; average OHI presentation level was approximately 72.5 dB SPL). Monaural presentation was chosen to facilitate recruitment of OHI listeners for Experiment 2 as only four participants met the strict hearing loss criteria in both ears (see Listeners section of Experiment 2 for additional information). Listeners heard the 224 sentences in random order without replacement either in their preferred ear or, in the case of OHI listeners for whom only one ear matched selection criteria, in their qualifying ear. Participants responded by clicking on the emotion category they felt corresponded best to the sentence and then pressed “Enter” on the keyboard to advance to the next sentence. Listeners could request one additional presentation of the sentence by clicking a button in the GUI labeled “Listen again.” The GUI also offered listeners the option to take a break at regular intervals during the test block. Task familiarization consisted of 20 items that were presented in a practice block prior to testing. All participants confirmed that they felt sufficiently familiar with the task procedures after the practice block.

Experiment 1: Perceived Emotion in Clear and Conversational Speech

Methods

Listeners

Nineteen YNH listeners aged 18 to 33 years ($M = 22$, $SD = 3.5$) were recruited from the University of Utah

Psychology Participant Pool. Data from one participant were excluded due to a late disclosure of ineligibility (prior history of speech therapy). The remaining 18 participants had normal hearing by self-report, were native speakers of American English, and had no history of speech or language therapy or disorders. Participants received research course credit hours as compensation for their participation.

Statistical Analyses

Each listener ($n = 18$) provided a judgment of perceived emotion for each sentence ($N = 224$), resulting in 4,032 data points. For each sentence, responses were combined across all listeners to determine the percentage of listeners who chose each of the six emotion categories. For example, if half of the listeners labeled a given sentence as “neutral” and the other half chose “happiness,” that sentence received percentage scores of 50% “neutral,” 50% “happiness,” and 0% for all other emotion categories. The percentage scores for each category were then analyzed using separate linear mixed-effects models. Linear mixed-effects models account for intratalker correlations and random variance introduced by using multiple stimuli from the same talker and have been recommended for use in speech research (Quené & van den Bergh, 2004). For each emotion category, the fixed factors were speaking style and talker group; talker was also included as a random factor in each model. Fixed factors were dummy-coded with a binary such that the variable levels of interest (clear speech and talker Group A) were assigned a 1, and the others (conversational speech and talker Group B) were assigned a 0. Thus, the beta values resulting from the model output (summarized in Table 2) represent the average proportional difference between the two levels of a given fixed factor. All analyses were carried out using Stata 14 (StataCorp, 2015).

Table 2. Results from linear mixed-effects models for Experiment 1.

	Anger	Fear	Disgust	Sadness	Happiness	Neutral
Fixed effects						
Speaking style						
β	.15	-.05	.07	-.10	-.05	-.02
z	9.77	-6.12	5.49	-6.24	-2.49	-1.28
p	< .001	.005	< .001	< .001	.013	.201
Talker group						
β	.11	.03	-.05	-.05	-.05	-.01
z	2.42	2.81	-1.71	-0.48	-0.72	-0.26
p	.016	< .001	.087	.632	.47	.793
Two-way interaction						
Speaking style $\times$ talker group						
β	.20	-.06	.06	.00	-.14	-.05
z	6.84	-3.82	2.48	0.15	-4.10	-1.58
p	< .001	< .001	.013	.88	< .001	.114

Note. Separate models were tested for each emotion. Bolded values indicate significant p values at an α level of .05.

Results

Figure 1 illustrates, for each emotion category, talker group, and speaking style, the percentage of listeners who chose that emotion, averaged across sentences. The fixed effect of speaking style was significant for “anger,” “fear,” “disgust,” “sadness,” and “happiness” categories. To be specific, listeners chose “anger” and “disgust” more often and “fear,” “sadness,” and “happiness” less often when listening to clear speech compared with conversational speech. The fixed factor talker group was significant for the categories “anger” and “fear”—that is, for Group A talkers, listeners chose “anger” more often and “fear” less often than for Group B talkers. Significant interactions between speaking style and talker group were observed for “anger,” “fear,” “disgust,” and “happiness” categories. A stratified analysis revealed that the effect of speaking style (which was positive for “anger” and “disgust” and negative for “fear” and “happiness”) was greater for sentences produced by Group A talkers than for sentences produced by Group B talkers.

Discussion

In this experiment, we examined if speaking style had any effect on judgments of emotion. The results indicated that speaking clearly compared with conversationally did affect the emotion categories listeners chose. YNH listeners judged clear speech as sounding angry and disgusted more often and sad, fearful, and happy less often than conversational speech. In addition, the significant effect of talker group suggests that “good” clear speakers (Group A) were judged as sounding angry more often and fearful less often than “poor” clear speakers (Group B). Interactions between speaking style and talker group reveal minimal, nonsignificant differences between speaking styles for the categories “anger,” “fear,” “disgust,” and “happiness” for “poor” clear speakers and large, significant speaking style effects for the “good” clear speakers. To be specific, listeners selected “anger” and “disgust” more often in response to clear speech than in response to conversational speech for these “good” talkers, and they chose “fear” and “happiness” less often. Note that, although no significant main effects nor any interactions were found for the “neutral” category, a considerable number of listeners (around 50%) reported sentences as sounding neutral in each speaking style and talker group. This finding was expected, considering the sentence content was selected to be emotionally neutral and the talkers were not instructed to express any specific emotion at the time of recording.

The significant increase in the selection of the negative-valence emotion categories “anger” and “disgust” is consistent with audiology patient complaints of communication partners sounding angry when they use clear speech. Listeners also chose the available positive emotion (i.e., “happiness”) less often in clear speech than conversational speech. Furthermore, these clear speech effects (the increase in negative affect and decrease in positive affect in clear

Figure 1. The percentage of listeners who chose each emotion category averaged across all sentences spoken conversationally (CO) and clearly (CL) by two groups of talkers (Groups A and B).

speech) were much greater for the talkers who produced “good” clear speech. It is thus unsurprising that individuals with hearing loss feel they are being spoken to in an angry or disgusted manner when communication partners attempt to speak clearly. Such a perception of clear speech may create a barrier to communication instead of the desired improvement and may put unnecessary strain on the relationship between communication partners. However, the present results were obtained for YNH listeners. Experiment 2 was designed to extend these findings to the population for whom clear speech is intended: listeners with hearing loss.

Experiment 2: Perceived Anger in Clear Speech for Older Adults With Hearing Loss

Methods

Listeners

Thirty-three listeners participated in this experiment. Thirteen YNH listeners were recruited from the University of Utah Psychology Participant Pool. Three participants were excluded due to either experiment software error ($n = 1$) or late disclosure of ineligibility ($n = 2$; age > 35 years). The remaining 10 YNH listeners were ages 18 to 30 ($M = 22.0$, $SD = 4.0$). YNH listeners were included in the second experiment to verify the reproducibility of the results from Experiment 1.

Twenty OHI listeners aged 65–78 years ($M = 71.9$, $SD = 3.3$) were recruited from Utah Senior Ears, a participant pool maintained by the second author. Members of Utah Senior Ears receive complete audiological evaluations upon enrollment in the pool and are recruited into various research studies on the basis of their hearing status. The OHI listeners in this study were required to have, in at least one ear, a mild to moderately severe sloping sensorineural hearing loss with threshold limits of 35 dB HL at 500 Hz, 40 dB HL at 1000 Hz, 50 dB HL at 2000 Hz, and 65 dB HL at 4000 Hz. Although only four participants met the hearing loss criteria in both ears, 15 participants had no clinically significant asymmetry. The five listeners with asymmetries participated using the ear that met the stated criteria. The OHI listeners were also required to have good word recognition abilities, defined as scoring > 80% correct for the Northwestern University Test No. 6 (Tillman & Carhart, 1966) presented in quiet at 40 dB above the speech recognition threshold. Their average word recognition score was 93%. If the date of the participant’s last hearing evaluation was 3 years or more prior to the date of testing ($n = 1$), then a new evaluation was performed and candidacy was reassessed.

In addition to hearing status criteria, all listeners in this experiment were administered and passed a cognitive screening as indicated by a screening criterion of 23 points on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). Listeners were also native speakers of

American English and had no history of speech or language disorders or therapy. YNH listeners received research course credit hours for their participation, and OHI listeners were either paid ($n = 15$) or opted to participate as volunteers ($n = 5$).

Statistical Analyses

Thirty listeners judged each sentence, resulting in 2,240 data points from new YNH listeners and 4,480 data points from OHI listeners. As in Experiment 1, the responses were combined across listeners to determine the percentage of listeners who chose each emotion category for each sentence. These percentages were determined separately for the YNH and OHI listener groups and analyzed using linear mixed-effects models for each emotion category performed using Stata 14 (StataCorp, 2015). Fixed effects of speaking style, listener group, and talker group were analyzed, and talker was included in each model as a random factor. Preliminary analyses showed that the YNH results from Experiment 2 were consistent with those of Experiment 1; as a result, data for YNH listeners from both Experiments 1 and 2 were aggregated (total $n = 28$) for analyses involving listener group comparisons. Fixed factors were dummy-coded with a binary such that the variable levels of interest (clear speech, talker Group A, OHI listeners) were assigned a 1 and the others (conversational speech, talker Group B, and YNH listeners) were assigned a 0. Thus, the beta values in the model output (summarized in Table 3) represent the average proportional difference between the two levels of a given fixed factor. Any significant interactions were explored using stratified analyses, which examined the size and significance of the effect of each fixed factor for both levels of the other factor.

Results

Speaking Style and Talker Group Effects

Table 4 shows the group means for speaking style, talker group, and listener group for each of the emotion categories, and Table 3 summarizes the statistical results. Speaking style was a significant fixed effect for all emotion categories. To be specific, listeners chose the “anger” and “disgust” categories more often and “fear,” “sadness,” “happiness,” and “neutral” categories less often when listening to clear speech than when listening to conversational speech. Significant fixed effects of talker group revealed that talkers in Group A (i.e., “good” clear speakers) were judged as sounding angry and fearful more often than talkers in Group B (i.e., “poor” clear speakers); the effect of talker group was not significant for any other category. The two-way interaction between speaking style and talker group was significant for all categories except “sadness.” For all five of those categories, the effect of clear speech was larger for sentences spoken by talkers in Group A than it was for sentences spoken by talkers in Group B.

Listener Group Effects

The main effect of listener group was significant for all categories except “disgust.” OHI listeners chose the

Table 3. Results from linear mixed-effects models for Experiment 2.

	Anger	Fear	Disgust	Sadness	Happiness	Neutral
Fixed effects						
Style						
β	.17	-.03	.11	-.11	-.05	-.09
z	15.57	-5.45	11.72	-9.13	-3.87	-7.10
p	< .001	< .001	< .001	< .001	< .001	< .001
Talker group						
β	.11	.04	.01	-.06	-.03	-.06
z	3.20	3.20	0.27	-0.77	-0.67	-1.02
p	.001	.001	.791	.444	.503	.309
Listener group						
β	-.07	-.02	-.02	-.06	-.07	.24
z	-6.08	-4.63	-1.66	-4.82	-6.18	17.33
p	< .001	< .001	.096	< .001	.001	< .001
Two-way interactions						
Style × talker group						
β	.21	-.04	.15	-.01	-.14	-.17
z	11.32	-3.99	8.33	-0.44	-6.33	-6.73
p	< .001	< .001	< .001	.658	< .001	< .001
Style × listener group						
β	-.08	.03	.00	.10	.04	-.08
z	-4.47	2.93	-0.04	4.15	1.63	-3.17
p	< .001	.003	.972	< .001	.103	.002
Talker group × listener group						
β	-.03	-.03	.06	.03	.05	-.09
z	-1.45	-3.00	3.61	1.29	2.43	-3.70
p	.146	.003	< .001	.198	.015	< .001
Three-way interactions						
Style × talker group × listener group						
β	-.10	.05	.12	.00	.08	-.14
z	-2.75	2.60	3.35	-0.18	1.80	-2.69
p	.006	.009	.001	.856	.072	.007

Note. Separate models were tested for each emotion. Bolded values indicate significant p values at an α level of .05.

categories “anger,” “fear,” “sadness,” or “happiness” less often and “neutral” more often than YNH listeners did. The two-way interaction between listener group and speaking style was significant for the categories “anger,” “fear,”

Table 4. Average percentage of young listeners with normal hearing (YNH) and older adults with hearing impairment (OHI) who assigned each emotion category to clear (CL) and conversational (CO) sentences spoken by members of two talker groups (Group A or Group B).

Emotion	Speaking style	YNH		OHI	
		Group A	Group B	Group A	Group B
Anger	CL	40	15	26	8
	CO	6	7	5	3
Fear	CL	5	3	4	3
	CO	12	4	6	3
Disgust	CL	23	21	28	13
	CO	7	14	6	12
Sadness	CL	4	11	4	9
	CO	20	28	11	15
Happiness	CL	5	20	4	10
	CO	20	17	12	8
Neutral	CL	23	29	34	57
	CO	34	30	60	59

“sadness,” and “neutral.” For these four categories, the effect of speaking style was larger for the YNH listeners than it was for the OHI listeners. Another two-way interaction, between listener group and talker group, was also significant for the categories “fear,” “disgust,” “happiness,” and “neutral.” For the categories “fear” and “happiness,” YNH listeners had larger differences between the Group A and Group B talkers than OHI listeners did. For the category “disgust,” there were no significant differences between listener groups for sentences spoken by Group A talkers, and for Group B talkers, OHI listeners chose “disgust” less often than YNH listeners. Last, YNH listeners assigned similar “neutral” ratings to the two talker groups, and OHI listeners chose “neutral” less often when listening to Group A talkers than when listening to Group B talkers.

Last, the three-way interaction between speaking style, talker group, and listener group was significant for the categories “anger,” “fear,” “disgust,” and “neutral.” For “anger” and “fear,” the magnitude of the clear speech effect was similar for both listener groups when listening to sentences spoken by talkers in Group B but was greater for YNH than OHI listeners when talkers in Group A produced the sentences. For “disgust” and “neutral,” the magnitude of the talker group effect was similar for YNH and OHI listeners when listening to conversational speech but was much larger for OHI listeners than YNH listeners for clear speech.

Discussion

Results of the linear mixed-effects regression analyses in this experiment were similar to those of Experiment 1—that is, listeners chose “anger” and “disgust” more often and “fear,” “sadness,” “happiness,” and “neutral” less often for clear speech than conversational speech. The results support the proposal that clear speech sounds angry more often than conversational speech, especially when the talker produces “good” clear speech. Unique to Experiment 2 was the finding that the selection of all emotion categories except “disgust” differed between YNH and OHI listener groups. These effects demonstrate that aging and/or reduced hearing sensitivity may affect the perception of emotion in clear and conversational speech.

General Discussion

Perceived Emotion in Clear Speech

The results of this study confirm that both YNH and OHI listeners judge clear speech as sounding angry more often than a typical conversational style of speech. Audiologists regularly counsel communication partners of individuals with hearing loss to “speak clearly” to improve understanding; however, the results from these experiments indicate that clear speech may sound angry to listeners. Thus, negative affect may accompany the improved clarity of clear speech. The frequent perception of negative emotion may exert strain on the relationships between individuals

with hearing loss and their loved ones and could cause these individuals to withdraw from social activities.

This concern about negative affect for different styles of speech extends to communication between members of other demographic groups as well. For example, talkers often adopt a clear speaking style when speaking to second language learners (Smiljanić & Bradlow, 2009). Clear speech intended for individuals who are nonnative speakers may also show patterns of increased negative and decreased positive valence. If so, nonnative speakers may make incorrect assumptions regarding the native speaker’s emotional state. Future investigations could address if nonnative listeners also perceive negative emotional valence in clear speech.

Talker Differences

Although six emotion categories were tested, the foregoing discussion focuses primarily on the category “anger” due to its relevance to the rationale for and research questions of the study. Figure 2 illustrates the percentage of listener judgments for the category “anger” for each speaking style, talker group, and listener group. Three-factor bar graphs for the other emotion categories are included in the online Supplemental Materials S1, S2, S3, S4, and S5. On average, listeners judged the clear speech produced by talkers in Group A as sounding angry more often than the clear speech produced by talkers in Group B. This suggests that producing clear speech that sounds clear also results in clear speech that sounds angry. Both YNH and OHI listeners showed this effect. However, it is important to note that listeners did judge some clear speech as sounding emotionally neutral, suggesting that some talkers use clear speech strategies that also make them sound angry and other talkers use different strategies that yield neutral-sounding clear speech. For example, both groups of listeners consistently rated one of the male talkers who produced “good” clear speech (M12) as sounding angry in his clear speech but judged the other male talker who produced “good” clear speech (M03) as sounding neutral in his clear speech.

On the basis of their perceived clarity ratings in Ferguson and Morgan (2015), talkers who produced “good” clear speech (Group A) were assumed to be more effective clear speakers than the “poor” talkers in Group B. As a result, it was anticipated that the talkers in Group B would show little or no difference in how often their clear versus conversational speech would be judged as sounding angry. Nonetheless, talkers in Group B were indeed judged as sounding angry more often in clear speech than in conversational speech although the magnitude of the speaking style effect was smaller than that observed for Group A talkers. This suggests that factors in addition to perceived clarity may account for differences in the emotions perceived in the two styles.

Listener Group Differences

OHI listeners judged sentences spoken clearly as sounding angry less often than YNH listeners. Although

Figure 2. The percentage of young adults with normal hearing (YNH) and older adults with hearing impairment (OHI) who chose the emotion category “anger” for sentences spoken conversationally (CO) and clearly (CL) by two groups of talkers (Groups A and B).

this listener group difference was larger for clear speech, it occurred for conversational sentences as well (see Figure 2). Two possible mechanisms may be driving this reduced judgment of anger for the OHI listeners. First, reduced audibility (due to hearing loss) of the acoustic cues that characterize anger and other emotions may be responsible for the overall decreased perception of emotion. High-frequency energy is an acoustic cue important to not only anger, but other emotions as well, including happiness, fear, and sadness. Aging together with hearing loss can also result in reduced frequency selectivity (especially for high frequencies), altered loudness perception, and poorer processing of auditory information (Howarth & Shone, 2006), which may further reduce the perception of acoustic cues for anger and other emotions.

Second, recall that in some previous studies, older adults performed more poorly than younger adults on emotion recognition tasks (as in Ruffman et al., 2009). Despite decreases in general performance, however, Carstensen and Mikels (2005) reported that older adults tend to give greater memory and attention to positive stimuli than negative stimuli, a pattern referred to as *the positivity effect*. Socioemotional selectivity theory suggests that this positivity effect is due to shifts in motivation (Carstensen & Mikels, 2005) and posits that, as individuals age, they seek more meaningful experiences and interactions and shift their attention toward positive events rather than negative events. For example, a young adult may volunteer to undergo the experience of incurring debt and submitting to years of course work to gain an education and secure a job. According to the aforementioned theory, older adults focus instead on positive and enriching experiences in the present rather than on less desirable experiences that could result in future gain. The decreased judgments of anger, sadness, and fear by the OHI listeners in the present study

may thus be evidence of the positivity effect for auditory stimuli, which, to the authors' knowledge, only one previous study (Ruffman et al., 2009) has addressed. However, the present study also showed that OHI listeners reported hearing happiness less often than YNH listeners did, which argues against this explanation. Given that increased high-frequency energy is an acoustic cue for happiness as well as anger, perhaps both hearing status and age contributed to the listener group differences in perceived emotion.

Note that a limitation of Experiment 2 was that the listener group characteristics of age and hearing status were confounded. It was therefore impossible to glean from the present study which of these possible explanations—high-frequency hearing loss or the positivity effect—resulted in the listener group effects found here. Studies are underway investigating the effects of aging and hearing loss more closely by controlling for these factors. Despite the general decrease in the perception of anger for OHI versus YNH listeners, clear speech was still judged as sounding angry more often compared with conversational speech, consistent with patient reports and professional observation.

Clinical Implications

These results suggest that audiologists should exercise caution when counseling communication partners to use clear speech. Although intelligibility may improve, so may the perception of anger. However, some of the talkers in this study did produce “good” clear speech that listeners judged as sounding emotionally neutral. Future studies should examine the clear speech acoustic strategies used by these talkers with an ultimate goal of identifying speaking strategies that yield speech that is clearer but not perceived as angry. Because of the multifactorial nature of speech acoustic characteristics, it is likely that a combination of

strategies will be required to produce the desired result. Once this combination of strategies is identified, programs could be implemented to inform audiologists and others working with clear speech to modify their counsel to and training of frequent communication partners of older adults with hearing loss and other clients. Aural rehabilitation strategies that currently exist should continue to promote clear speaking strategies to retain the intelligibility benefit afforded by this speaking style, but clinicians should be aware that such a style may carry a negative tone and adjust their consultations accordingly.

Changes in perceived emotion in clear speech should also be investigated for other listener populations. For example, individuals with cochlear implants perform poorly on emotion recognition tasks relative to individuals with normal hearing (Chatterjee et al., 2015; Luo, Fu, & Galvin III, 2007). Although cochlear implants do a poor job of representing spectral information from the speech signal, they do an excellent job of coding temporal information. Temporal modifications in clear speech might thus help individuals with cochlear implants perceive emotion with greater accuracy. The present study successfully demonstrated that speaking clearly affects the perception of some emotions. A thorough investigation should be made into which clear speech acoustic changes enhance or hinder emotion perception.

Conclusions

In the present study, YNH and OHI listeners categorized sentences spoken in a clear style as sounding angry and disgusted more often and as sounding happy, sad, and fearful less often than conversational sentences. These findings support clinical observations of perceived negativity in clear speech. The increase in perceived anger in clear versus conversational speech was greater in talkers that produced “good” clear speech compared with talkers whose clear speech was similar to their conversational speech. Although increased judgments of anger in clear speech were apparent in a sample of YNH listeners, the results of the second experiment extended the findings to a clinical population—that is, OHI listeners. The OHI listener group (which differed from the YNH listener group in both age and hearing status) showed fewer anger judgments than YNH listeners did; however, they still identified clear speech as sounding angry more often than conversational speech.

Acknowledgments

This research was supported in part by National Institutes of Health Grant R01DC012315 to Eric Hunter. Development of the Ferguson Clear Speech Database was supported by National Institutes of Health Grant DC02229 to Diane Kewley-Port. Portions of these results were presented at the 169th meeting of the Acoustical Society of America and the 2015 annual meeting of the American Speech-Language-Hearing Association. A special thanks to Joan Morgan and Jaime Booz for their suggestions and comments on previous versions of this article.

References

- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 614–636.
- Booz, J. A. (2016). *Perceived gender in clear and conversational speech* (Unpublished master's thesis). University of Utah, Salt Lake City.
- Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition: Aging and the positivity effect. *Current Directions in Psychological Science*, 14, 117–121.
- Chatterjee, M., Zion, D., Deroche, M., Burianek, B., Limb, C., Goren, A., . . . Christensen, J. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, 322, 151–162.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169–200.
- Ekman, P. (2003). *Emotions revealed*. New York, NY: Holt Paperbacks.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3, 364–370.
- Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *The Journal of the Acoustical Society of America*, 116, 2365–2373.
- Ferguson, S. H. (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.
- Ferguson, S. H., & Morgan, S. D. (2015, November). *Acoustic and perceptual correlates of subjectively rated sentence clarity in clear and conversational speech*. Paper presented at the 170th Meeting of the Acoustical Society of America, Jacksonville, FL.
- 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.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Howarth, A., & Shone, G. R. (2006). Ageing and the auditory system. *Postgraduate Medical Journal*, 82, 166–171.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychology Bulletin*, 129, 770–814.
- Juslin, P. N., Laukka, P., Davidson, R. J., & Scherer, K. R. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1, 381–412.
- Liu, S., Del Rio, E., Bradlow, A. R., & Zeng, F.-G. (2004). Clear speech perception in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, 116, 2374–2383.
- Luo, X., Fu, Q. J., & Galvin III, J. J. (2007). Cochlear implants special issue article: Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in Amplification*, 11, 301–315.
- Mortillaro, M., Meuleman, B., & Scherer, K. R. (2012). Advocating a componential appraisal model to guide emotion recognition. *International Journal of Synthetic Emotion*, 3, 18–32.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, 28, 96–103.

- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434–446.
- Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103–121.
- Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions in older adults. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 64, P696–P703.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143–165.
- Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433–456). New York, NY: Oxford University Press.
- Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118, 1677–1688.
- 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.
- StataCorp. (2015). *Stata Statistical Software: Release 14* [Computer software]. College Station, TX: Author.
- Tillman, T. W., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6*. Brooks Air Force Base, TX: U.S. Air Force School of Aerospace Medicine.
- 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.
- Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M., & Scherer, K. R. (2013). On the acoustics of emotion in audio: What speech, music, and sound have in common. *Frontiers in Psychology*, 4, 1–12.

Appendix

Sentences used in the listening experiment

1. He picked up the bead and put it away.
2. He picked up the bid and put it away.
3. He picked up the bod and put it away.
4. He picked up the bud and put it away.
5. He picked up the bude and put it away.
6. Use the word *bad* in a sentence.
7. Use the word *bod* in a sentence.
8. Use the word *bode* in a sentence.
9. Use the word *bud* in a sentence.
10. Use the word *bude* in a sentence.
11. Look up the word *bad* on the Internet.
12. Look up the word *bade* on the Internet.
13. Look up the word *bud* on the Internet.
14. Look up the word *bude* on the Internet.