J Am Acad Audiol 7: 212-218 (1996)

Intelligibility of Clear and Conversational Speech of Young and Elderly Talkers

Donald J. Schum*

Abstract

It has been documented that talkers can be trained to produce "clear" speech, which is significantly more intelligible for hearing-impaired listeners. In this study, the ability of both younger and older talkers to produce clear speech after a minimal amount of instruction and practice was investigated. Tape recordings were made with the talkers attempting to produce both conversational-style and also clear-style speech. Elderly patients with sensorineural hearing loss listened to and attempted to repeat these sentence productions. The results indicate that both groups of talkers could effect significant intelligibility improvements despite minimal instruction and practice. Further, the magnitudes of the intelligibility improvements were statistically similar between the two groups of talkers. Instruction in clear speaking techniques is thus advocated as a useful intervention technique for spouses and other family members of elderly patients with hearing loss.

Key Words: Aging, speech intelligibility, training

Elderly hearing-impaired individuals will often report difficulties understanding speech, especially for certain talkers (Gengel and Kupperman, 1980; Hood and Poole, 1980; Cox et al, 1987) and even after the provision of appropriate amplification (Schum, 1992). In order to improve speech understanding for elderly patients with hearing loss, a variety of intervention strategies are available. These intervention options include the use of, or improvement upon, amplification, reducing the noise or reverberation in the listening environment, the use of assistive listening devices designed to defeat the effects of competition in the listening environment, and the use of certain listening strategies on the part of the patient. With few notable exceptions (Erber and Lind, 1994; Tye-Murray and Schum, 1994), little focus has been placed on the role of the talker who provides the original speech message to the hearing-impaired listener.

Picheny et al (1985, 1986) have demonstrated that talkers can significantly improve the intelligibility of their speech for hearing-impaired listeners by attempting to speak clearly and accurately. Compared to normal, "conversational-style" speech, the "clear-style" speech is characterized by specific acoustic changes. These changes are different in nature and degree than the changes brought about when talkers attempt to speak more loudly or more slowly (Picheny et al, 1989; Moon and Lindblom, 1994) and include a decreased rate of speech, increased duration of phonemes, and fuller differentiation between phonemes. For listeners with sensorineural hearing loss, Picheny et al (1985) report a mean intelligibility increase of 17 percent for clear speech as compared to conversational speech.

Since the original work by Picheny and his colleagues, several laboratories have been evaluating the possibility of building a computer-based, hearing-aid-like device that can modify normal, conversational-style speech to have some of the properties of naturally produced clear speech (Gordon-Salant, 1986, 1987; Revoile et al, 1986, 1987; Guelke, 1987; Montgomery and Edge, 1988; Freyman and Nerbonne, 1989; Bunnell, 1990). The theoretical advantage of such devices is that any talker can be transformed into a "clear" talker without practice, instruction, or even effort. Unfortunately, a wearable, real-time version of such a device is probably years away. In the meantime, it would seem of value to evaluate the possibility that

*Oticon A/S, Hellerup, Denmark
Reprint requests: Donald J. Schum, Oticon A/S, 58 Strandvejen, DK-2900 Hellerup, Denmark

most any talker can quickly and easily learn to speak more clearly.

In the Picheny et al (1985, 1986) studies, three talkers were used. All were (presumably) young male adults and all had some experience producing highly intelligible speech, either as practiced experimental subjects in previous clear speech studies, as a relative of hearing-impaired individuals, or as an amateur public speaker. It is known that age-related physiologic changes in the speech-production mechanism may take place (Benjamin, 1981, 1988; Kahane, 1990), and that there are acoustic (Smith et al, 1987) and perceptual (Hartman and Danhauer, 1976; Ryan and Capadano, 1978; Linville and Fisher, 1985) differences between the speech produced by younger and older voices. Given these documented changes, elderly talkers may have limitations on their physical capacity to produce accurate and precise clear speech. Since the majority of hearing-impaired individuals are elderly, their most frequent home and social contacts might also be expected to be elderly. It is, therefore, relevant to examine the ability of these contacts to improve their speech intelligibility.

Further, in the Picheny et al (1985, 1986) studies, the talkers were provided with explicit instructions as to how to produce clear, intelligible speech, with significant feedback provided during recording sessions. If clear speech is truly significantly more intelligible, there would be benefit in having as many contacts of a hearing-impaired individual as possible producing this style of speech. However, for practical reasons, extensive training may not be possible.

The current study is an attempt to expand the original clear speech work of Picheny et al (1985, 1986) to include both young (n = 10) and elderly (n = 10) talkers, and to use talkers without formal experience in producing clear speech. Further, the talkers in the current study were provided with only a simple set of instructions and were given only a brief chance to practice.

This project was carried out via a series of specific experiments. In the first (reported in this paper), the intelligibility of conversational and clear speech productions of young and elderly talkers was evaluated. Elderly, hearing-impaired listeners listened to and tried to repeat verbatim these productions. The results are compared to those of Picheny et al (1985) to determine if unpracticed talkers from both age groups could achieve intelligibility gains similar to those of the highly practiced talkers studied by Picheny and his colleagues.

METHOD

Subjects

Talkers. Twenty persons served as talkers. Ten (6 female, 4 male) fell within the age range of 22 to 39 years and 10 (6 female, 4 male) fell within the age range of 62 to 70 years. All talkers demonstrated pure-tone air-conduction thresholds, measured at the octave frequencies from 250 through 4000 Hz, of 25 dB HL or better. All talkers reported being free of any neurologic or muscular condition that would be expected to affect their ability to produce speech. All talkers reported no formal speaking training. None of the talkers reported close relatives or significant others with longstanding, significant hearing loss, with the exception that some of the older talkers were spouses of persons with no worse than mild-to-moderate hearing loss. These hearing levels were able to be confirmed, as recent audiograms from these spouses were on file at the test facility.

Listeners. Sixty persons with sensorineural hearing loss served as listeners. The ages of these listeners ranged from 60 to 77 years, with a mean age of 67.4 years (SD = 4.1 years). Listening was performed monaurally under an earphone. All listeners were required to have auditory thresholds (averaged over 500, 1000, 2000, and 4000 Hz) in at least one ear falling in the range from 20 through 70 dB HL. If both ears fell within this range, the test ear was chosen at random. The mean audiogram (with standard deviations) for the test ears is presented in Figure 1.

Materials

The sentences were meaningful and varied from four to seven words in length. The materials were adapted from sentence lists included in the *Johns Hopkins Lipreading Corpus* (Bernstein and Eberhardt, 1986). Twenty sentences were selected to comprise 2 10-sentence practice lists. One hundred sentences were selected to comprise 10 10-sentence test lists. Across each 10-sentence list, the same number of words (N = 52) and syllables (N = 61) were included. Examples of these sentences would be:

She did a perfect cartwheel.
One chance is given to each man.
Their room was clean.

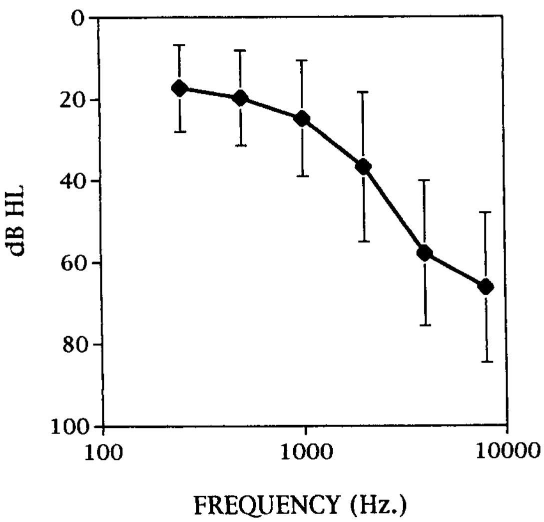

Figure 1 Mean and range of audiograms for the test ears of the 60 listeners.

The sentence lists were assigned to speakers in the following manner in order to minimize the effects of any intelligibility differences between lists. For the 10 10-sentence test lists, one talker from each group used lists 1 through 5 for the conversational productions and lists 6 through 10 for the clear productions. The next talker from each group used list 2 through 6 for the conversational productions and lists 7 through 10 and list 1 for the clear productions. This rotational pattern was continued such that, across the 10 younger talkers and across the 10 older talkers, each sentence list was used an equal number of times for clear and for conversational productions.

Recording Procedure

Each talker was recorded in a single-wall, sound-attenuating booth. The talker was seated approximately 6 inches from a Realistic 33-992B boom-mounted microphone. The output from the microphone was fed to a TEAC cassette tape recorder.

Procedure

Each talker produced 10 practice and 50 test sentences in what he/she considered a normal, conversational-style speaking mode. In addition, each talker was instructed on how to produce clear speech and then produced the remaining 10 practice and the remaining 50 test sentences in a clear-style speaking mode. The instruction set given to the talkers was as follows:

> Imagine that you are speaking to a person that you know is hearing impaired. I want you to speak as clearly and precisely as possible. Try to produce each word as accurately as you can.

If a talker inquired further as to whether or not they should speak louder, more dramatically, or slower, they were told to do whatever they felt was necessary in order to be better understood. The order of conversational versus clear speaking was counterbalanced across talkers. Again, a different combination of test lists for clear versus conversational productions was used for each talker.

A second generation cassette tape recording was made of the speech samples. Mis-starts, repeated sentences, unusually long between-sentence pauses, etc. were removed during the editing process. Also, the recording level on the second generation tape was adjusted for each talker and for each speaking mode such that the peak level for each sentence fell within a 3-dB range. In other words, even though some talkers produced more intense speech than others and although the clear speech productions were consistently more intense than the conversational productions, these intensity differences were purposely eliminated. This step was taken in order to minimize the effect of overall audibility on the intelligibility of clear versus conversational speech. None of the editing affected the spectral or temporal characteristics of any of the segments used during playback. Cafeteria noise was added to the second channel of this recording in order to avoid ceiling effects by increasing the difficulty of the listening task.

From the corps of 60 listeners, 3 were randomly assigned (without replacement) to each of the 20 talkers. Thus, each listener heard the productions of a single talker.

The speech was routed from the playback cassette tape deck (TEAC) through a clinical audiometer (Madsen OB-822). The cafeteria noise was mixed with the speech material at a +3-dB signal-to-noise ratio and the mixed signal was routed through a TDH-39 earphone. In order to further decrease the individualized effects of audibility, the mixed signal was presented at the highest comfortable listening level (chosen by the subject) within the range of 70 to 90 dB SPL.

Each listener heard 50 sentences produced in the conversational mode and 50 sentences in the clear speech mode. The first 10 sentences in each mode were considered practice items, with the remaining 40 sentences in each mode considered to be the test items. The order of presentation of the clear versus conversation productions was randomly varied from listener to listener.

The listener was instructed to repeat each sentence verbatim. A count was made of the number of words correctly repeated, including a requirement of appropriated tense and plurality word endings. There were 208 words scored across each 40 test sentences (52 words per each 10-sentence list).

RESULTS

Given that speech understanding ability can vary significantly within a group of subjects with sensorineural hearing loss, the percent correct scores were converted to rationalized arcsine transform units (RAUs) (Studebaker, 1985). This transformation stabilizes the variance across a range of percent correct scores. This allows easier comparison of differences in percent correct scores from across the entire range from 0 to 100 percent. Table 1 provides the difference in RAUs between the clear and conversational scores obtained by each listener. Positive scores indicate clear scores greater than conversational scores. Also indicated is whether or not the difference scores reach statistical significance ($p < .05$), based on the procedures specified by Studebaker (1985). For all 20 talkers, at least one of three listeners demonstrated significantly higher clear speech percent correct scores. For three talkers, one of three listeners demonstrated significantly higher clear speech percent correct scores. For two talkers, two of three listeners demonstrated significantly higher clear speech percent correct scores. For the remaining 15 talkers, all three of the listeners demonstrated significantly higher clear speech percent correct scores.

Figure 2 provides the average (across three listeners) percent correct (converted to RAU) change for each of the 20 talkers. A positive score indicates a clear speech score greater than a conversational speech score. The average change for the younger talkers was 22.0 RAU and the average change for the older talkers was 16.9 RAU. A mixed model analysis of variance was performed on the difference scores, with talker as a nested (within-group) effect (Jennrich and Sampson, 1985). The results revealed no group effect ($p > .05$).

Table 1 Difference in RAUs between the Intelligibility of Clear and Conversational Speech Samples as Obtained by Each Listener

	Listener		
Talkers	*1*	*2*	*3*
Young			
Y1	24*	20*	20*
Y2	1	1	10*
Y3	12*	16*	13*
Y4	41*	40*	56*
Y5	14*	8	13*
Y6	9*	14*	13*
Y7	47*	36*	38*
Y8	42*	29*	35*
Y9	26*	32*	25*
Y10	14*	9*	8
Elderly			
E1	17*	16*	12*
E2	27*	25*	29*
E3	7	8	12*
E4	12*	13*	11*
E5	22*	22*	23*
E6	10*	0	7
E7	26*	14*	21*
E8	28*	16*	23*
E8	16*	13*	20*
E10	25*	14*	20*

A positive score indicates that the clear speech score was greater than the conversational speech score. An asterisk indicates that the difference was statistically significant ($p < .05$).

In order to confirm that the talkers were able to produce what has been termed "clear speech," as reported in a companion paper (Schum, 1995), the acoustic characteristics of the productions were analyzed in order to determine whether the unpracticed talkers of the current study could achieve the same type and magnitude of acoustic modifications as did the talkers of Picheny et al (1985). This analysis confirmed similar acoustic effects as reported by Picheny et al (1985).

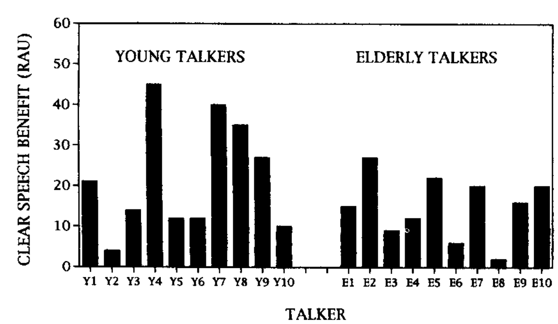

Figure 2 The average improvement afforded by clear speech for each of the 10 young and 10 elderly talkers.

DISCUSSION

In general, the lack of experience and the minimal instruction and practice period did not prevent the talkers in this investigation from being able to produce intelligibility improvements by adopting a clear speaking mode. The magnitude of these intelligibility improvements (22 RAU for the younger talkers and 17 RAU for the older talkers) was similar to the intelligibility improvements reported by Picheny et al (1985). Further, the elderly talkers were able to produce comparable intelligibility gains as did the younger talkers. In fact, certain elderly talkers (e.g., E2, E5, E7, E10) were better able than many of the younger talkers to improve the intelligibility of their speech.

In general, these results are encouraging. Without much instruction or practice, talkers were able to make fairly dramatic changes in their intelligibility. The results suggest that instruction in clear speech techniques is a viable intervention strategy for family members of elderly persons with hearing impairment. The speech production limitations of the elderly do not appear to be of a type or magnitude to preclude clear speech production. Therefore, elderly communication partners of persons with sensorineural hearing loss can be expected to be able to improve their intelligibility at will.

There are many opportunities within the typical audiologic practice to incorporate clear speech training. For example, many clinicians prefer to have spouses and other family members present at and involved in the hearing aid fitting and orientation session. Given that the instructional/practice requirements are minimal, clear speech can be discussed within the practical time constraints of the fitting and orientation process for new hearing aid users. Even in cases in which a patient receives an audiologic evaluation but chooses to defer amplification for whatever reason, the clinician can still take a few minutes to discuss this particular communication strategy with any family members present.

The nature of the improvement afforded to elderly listeners by the use of clear speech may be twofold. First, it has been clearly documented by Picheny et al (1986) that, acoustically, the individual phonemes in clear speech are more completely formed and differentiated from each other. The vowel formant space is expanded in clear speech. Stop consonants are released more often. Consonant/vowel ratios are greater. Continuant consonants are longer. These acoustic enhancements are assumed to allow for more accurate identification of each individual phoneme, leading to improved word and sentence identification. In addition, recall that clear speech is significantly slower than conversational speech. The decreased speech rate is reflected in longer and more frequent between-word pauses and also longer phoneme durations. It has been observed that the time course of phonemic identification is longer in patients with sensorineural hearing loss (Dubno et al, 1987; Schum and Collins, 1990). Also, there is evidence that elderly listeners may have difficulty processing information at the same rate as younger listeners (Welford, 1985) and may benefit from decreases in speaking rate, under some conditions (Schmitt, 1983; Schmitt and Carroll, 1985). Therefore, a second benefit of the use of clear speech is that the rate at which the listener is required to process incoming speech is reduced. However, as indicated by Picheny et al (1989), artificial slowing of the rate of conversational speech does not yield the same magnitude of benefit as does clear speech. Schum (1996) reports that there are no strong correlations between any single acoustic measure and the intelligibility improvements afforded by clear speech. Rather, the intelligibility improvements are likely due to a complex interaction of durational, spectral, and intensity changes due to the adoption of the clear speaking style. Thus, the slower rate of clear speech in and of itself is not enough to explain the resultant intelligibility improvements.

It is reasonable to argue that the use of even the simple instructional set described above is not possible with all conversation partners that the elderly person with hearing loss is likely to encounter. For example, it is not practical to instruct every store clerk, telephone operator, physician, receptionist, etc. in the need to speak clearly. Therefore, it would be preferable if a hearing aid could transform any input speech into clear speech. Although digital-based implementations of such techniques are likely years from commercial availability, some currently available hearing aids can make changes in the input speech signal that replicate some clear speech effects. For example, consonant/vowel ratios can be increased by some analog hearing aid circuits (Preves et al, 1991). However, these types of spectral modifications alone do not allow for the important changes in phoneme durations and speaking rate inherent in clear speech. The benefits of any artificial attempts to decrease the speaking rate would have to be balanced against the disruption in the timing of the auditory versus visual cues. Also, Payton et al (1994) report that the advantage afforded by clear speech increases as background noise levels increase. Conversely, hearing aid benefit almost universally decreases as background noise levels increase. Therefore, the benefits of naturally produced clear speech are not likely to be replicated by hearing aid technology in the near future.

Finally, there is no reason to assume that the benefits afforded by clear speech cannot work in concert with the benefits afforded by hearing aid technology or, for that matter, any other available intervention strategy. For example, Tye-Murray and Schum (1994) describe an intervention program that incorporates clear speech instruction in a larger training program for frequent communication partners of persons with hearing impairment.

REFERENCES

Benjamin B. (1981). Frequency variability in the aged voice. *J Gerontol* 36:722–726.

Benjamin B. (1988). Changes in speech production and linguistic behaviors with aging. In: Shadden B, ed. *Communication Behavior and Aging: A Source Book for Clinicians.* Baltimore: Williams & Wilkins, 162–181.

Bernstein LE, Eberhardt SP. (1986). *Johns Hopkins Lipreading Corpus.* Baltimore: The Johns Hopkins University.

Bunnell HT. (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *J Acoust Soc Am* 88:2546–2556.

Cox RM, Alexander GC, Gilmore C. (1987). Intelligibility of average talkers in typical listening environments. *J Acoust Soc Am* 81:1598–1608.

Dubno JR, Dirks DD, Schaefer AB. (1987). Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *J Acoust Soc Am* 81:1940–1947.

Erber NP, Linda C. (1994). Communication therapy: theory and practice [monograph]. *J Acad Rehabil Audiol* 27:267–287.

Freyman RL, Nerbonne GP. (1989). The importance of consonant-vowel intensity ratio in the intelligibility of voiceless consonants. *J Speech Hear Res* 32:524–535.

Gengel R, Kupperman G. (1980). Word discrimination in noise: effect of different speakers. *Ear Hear* 1:156–160.

Gordon-Salant S. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *J Acoust Soc Am* 80:1599–1607.

Gordon-Salant S. (1987). Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects. *J Acoust Soc Am* 81:1199–1202.

Guelke RW. (1987). Consonant burst enhancement: a possible means to improve intelligibility for the hard of hearing. *J Rehabil Res Dev* 24:217–220.

Hartman D, Danhauer J. (1976). Perceptual features of speech for males in four perceived age decades. *J Acoust Soc Am* 59:713–715.

Hood JD, Poole JP. (1980). Influence of the speaker and other factors affecting speech intelligibility. *Audiology* 19:434–455.

Jennrich R, Sampson P. (1985). General mixed model analysis of variance. In: Dixon WJ, ed. *BMPD Statistical Software.* Berkeley, CA: University of California Press, 413–426.

Kahane J. (1990). Age-related changes in the peripheral speech mechanism: structural and physiological changes. In: Cherow E, ed. *Proceedings of the Research Symposium on Communication Sciences and Disorders and Aging.* Rockville, MD: American Speech-Language-Hearing Association, 75–86.

Linville SE, Fisher HB. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *J Acoust Soc Am* 78:40–48.

Montgomery AA, Edge RA. (1988). Evaluation of two speech enhancement techniques to improve intelligibility for hearing-impaired adults. *J Speech Hear Res* 31:386–393.

Moon S-Jae, Lindblom B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *J Acoust Soc Am* 96:40–55.

Payton KL, Uchanski RM, Braida LD. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *J Acoust Soc Am* 95:1581–1592.

Picheny M, Durlach N, Braida L. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *J Speech Hear Res* 28:96–103.

Picheny M, Durlach N, Braida L. (1986). Speaking clearly for the hard of hearing II: acoustic characteristics of clear and conversational speech. *J Speech Hear Res* 29:434–446.

Picheny M, Durlach N, Braida L. (1989). Speaking clearly for the hard of hearing III: an attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *J Speech Hear Res* 32:600–603.

Preves DA, Fortune TW, Woodruff B, Newton J. (1991). Strategies for enhancing the consonant to vowel intensity ratio with in the ear hearing aids. *Ear Hear* 12:139S–153S.

Revoile SG, Holden-Pitt LD, Edward DM, Pickett JM. (1986). Some rehabilitative considerations for future speech-processing hearing aids. *J Rehabil Res* 23:89–94.

Revoile SG, Holden-Pitt L, Edward D, Pickett JM, Brandt F. (1987). Speech-cue enhancement for the hearing impaired: amplification of burst/murmur cues for improved perception of final stop voicing. *J Rehabil Res* 24:207–216.

Ryan EB, Capadano III, HL. (1978). Age perceptions and evaluative reactions toward adult speakers. *J Gerontol* 33:98–102.

Schmitt JF. (1983). The effects of time compression and time expansion on passage comprehension by elderly listeners. *J Speech Hear Res* 26:373–377.

Schmitt JF, Carroll MR. (1985). Older listeners' ability to comprehend speaker-generated rate alteration of passages. *J Speech Hear Res* 28:309–312.

Schum DJ. (1992). Response of elderly hearing aid users on the Hearing Aid Performance Inventory. *J Am Acad Audiol* 3:308–314.

Schum DJ. (1996). Acoustic characteristics of clear and conversational speech of young and elderly talkers. Manuscript in preparation.

Schum DJ, Collins MJ. (1990). The time course of acoustic/phonemic cue integration in the sensorineurally hearing-impaired listener. *J Acoust Soc Am* 87: 2716–2728.

Smith BL, Wasowicz J, Preston J. (1987). Temporal characteristics of the speech of normal elderly adults. *J Speech Hear Res* 30:522–529.

Studebaker GA. (1985). A "rationalized" arcsine transform. *J Speech Hear Res* 28:455–462.

Tye-Murray N, Schum L. (1994). Conversation training for frequent communication partners [monograph]. *J Acad Rehabil Audiol* 27:209–222.

Welford AT. (1985). Changes of performance with age: an overview. In: Charness N, ed. *Aging and Human Performance*. New York: John Wiley & Sons, 333–369.