Acoustic correlates of clear speech vowel intelligibility for elderly hearing-impaired listeners

Sarah Hargus Ferguson
Department of Speech-Language-Hearing; Sciences and Disorders - University of Kansas

Introduction

• Clear speech is significantly more intelligible than typical conversational speech. It also differs from conversational speech on a wide array of acoustic dimensions. In many studies, normal-hearing and hearing-impaired listeners have enjoyed a similar clear speech intelligibility benefit.

• A striking exception to this was Ferguson & Kewley-Port (2002) in a study using words produced by a single talker. While younger listeners with normal hearing (YNH listeners) showed a large clear speech–vowel intelligibility benefit, elderly listeners with hearing impairment (EHI listeners) showed no benefit. These suggested that the clear speech acoustic changes that make speech more intelligible might be different for YNH and EHI listeners.

The Ferguson Clear Speech Database

• Ferguson (2004) recorded 41 talkers (20 M, 21 F aged 18 to 45 years) producing sentences in clear and conversational speaking styles. These included 188 sentences containing vowels in /tVd/ context placed in neutral sentence frames.

• For YNH listeners, the clear speech vowel intelligibility benefit was found to vary widely among the 41 talkers.

• Comparisons of acoustic data from talkers who produced either a large clear speech benefit or no clear speech benefit suggested that greater increases in vowel intelligibility in clear speech were associated with greater vowel space expansion, bigger differences between low and high vowels, greater increases in front vowel second formant frequency, and greater duration increases (Ferguson & Kewley-Port, 2007).

The Current Project

• The current project extended this line of research to the population for which clear speech is intended: listeners with hearing loss. In particular, elderly listeners with hearing loss (EHI listeners) were used, due to the high prevalence of hearing loss in this group.

Methods

Materials

• 10 vowels (/, /, /, /, /, /, /, /, /, / in /tVd/ context produced by the 41 talkers in the Ferguson database.

• 2 tokens of each /tVd/ word per talker in each speaking style (clear and conversational) excerpted from the recorded sentences (1640 stimuli total)

• All stimuli scaled to the same RMS amplitude

Acknowledgements

The project was supported by NIDCD R01-02229. Development of the Ferguson database was supported in NIDCD R01-02229 by Diane Kewley-Port, Ching Kieweg, Leela Parimisetty, Katie Beam, Jessica Stamey, Valrie Whitaker, and Clarisa Studebaker. This research was supported by NIDCD R01-02229 to Diane Kewley-Port. Doug Kieweg, Leela Parimisetty, Katie Beam, Jessica Stamey, and Clarisa Studebaker were responsible for data collection and analysis.

LPC formant tracks were generated for each token using Wavesurfer. The first and second formant frequency values (F1 and F2) were extracted from the tracks at 20% of the vowel duration. The vowel steady state (the 20% point plus 30 ms). Vowel space measures were calculated from steady state F1 and F2 values. Perimeter: the sum of Euclidean distances between adjacent point vowels (/i, /, /, /, /, /, /, /, /, /).

• Only three talkers showed the listener group differed (large benefits for YNH and no benefit for EHI) produced by the talker in Ferguson & Kewley-Port (2002).

• Three acoustic measures were used in the current study: F1 range, F2 front, and spectral change.

• Before and after talkers differed significantly on these three acoustic measures.

Results and Discussion

Vowel intelligibility

• Percent correct scores were converted to rationalized arcsine units (RAU; Snedecor & Cochran, 1980).

• As in Ferguson (2004), vowels were significantly more intelligible in clear speech than in conversational speech, and the clear speech vowel intelligibility benefit varied widely among the talkers.

Procedures

• Vowel intelligibility

• The /tVd/ words were presented via insert earphones at 70 dB SPL in a background of L2-talker babble (speech-to-babble ratio = 4:1). Each test block contained either clear or conversational tokens produced either by male talkers or female talkers. After each presentation, listeners selected the word they heard from a list of 10 keywords.

• Only three talkers showed the listener group differed (large benefits for YNH and no benefit for EHI) produced by the talker in Ferguson & Kewley-Port (2002).

• Acoustic analyses

• Two sets of 12 talkers were selected on the basis of the clear speech vowel intelligibility effect that the EHI listeners enjoyed. Big Benefit (BB) talkers showed a mean clear speech effect of 14 RAU while the No Benefit (NB) talkers showed a mean effect of 0.82 RAU.

• This extreme groups design assumed that if a given acoustic change were important to the clear speech vowel intelligibility benefit, then only the BB talkers would produce it. Any change produced equally by the two groups was deemed irrelevant to the clear speech benefit.

• BB and NB talkers differed significantly on three acoustic measures: BB talkers showed greater vowel space expansion (perimeter), bigger differences between low and high vowels (F1 range), and greater duration increases than the NB talkers.

• Overall performance was higher for these EHI listeners than for the earlier YNH listeners but was strongly correlated between the two groups, suggesting that talkers whose clear speech benefit prevailed for the YNH listeners also produced a clear speech benefit for the EHI listeners.

• Example vowel spaces are shown here. As a group, the BB talkers expanded their vowel space in clear speech by 1.8 Barks, while the NB talkers expanded by just 0.23 Barks.

• F1 range expansion was 46 Barks for the BB talkers and 04 Barks for the NB talkers.

• BB talkers increased vowel duration by an average of 100 ms, while NB talkers showed only a 30 ms increase.

• F2 front was significantly higher, F2 back was significantly lower, and spectral change was significantly greater in clear speech than in conversational speech, but no difference was observed between the BB and NB talkers for these measures.

• The only difference between these results and those found for YNH listeners in Ferguson & Kewley-Port (2007) is for F2 front, which was significant in the YNH analysis but not in the EHI analysis. This is not surprising given that front vowel F2 frequencies fell between approximately 1500 and 3000 Hz, a region in which these EHI listeners had decreasing hearing sensitivity. In contrast with Ferguson and Kewley-Port (2003), however, raising F2 for front vowels did not seem to be harmful for the EHI listeners.

Conclusions

• Results for EHI listeners were surprisingly similar to those found by Ferguson (2004) for YNH listeners.

• In general, vowels were more intelligible in clear speech than in conversational speech for these listeners.

• Listeners who produced a clear speech vowel intelligibility benefit for YNH listeners usually also did so for EHI listeners.

• Most of the clear speech acoustic changes that were associated with improved vowel intelligibility were also associated with EHI listeners, and none caused harm to the EHI listeners.

References
