

# Proceedings of Meetings on Acoustics

---

Volume 4, 2008

<http://asa.aip.org>

---

**155th Meeting  
Acoustical Society of America**  
Paris, France  
29 June - 4 July 2008

**Session 5aPPf: Psychological and Physiological Acoustics**

---

## **5aPPf10. The role of consonant duration and amplitude processing on speech intelligibility in noise**

**Jeffrey J. Digiovanni\* and Ashley K. Stover**

**\*Corresponding author's address: Hearing, Speech and Language Sciences, Ohio University, W151a Grover Center, Athens, OH 45701, [digiovan@ohio.edu](mailto:digiovan@ohio.edu)**

Hearing loss can have drastic effects on communication abilities leading researchers to strive to find solutions for the associated problems. Clear speech has been studied in the past as a possible avenue for improving speech understanding in those with hearing loss. In the present study, two characteristics of clear speech, increased consonant duration and consonant amplification, were applied to Hearing-in-Noise Test sentences. Stimuli were presented in sound field to hearing-impaired and normal-hearing individuals in the presence of speech-shaped noise, and percentage correct was measured. For HINT sentence presentations, both groups showed improvements for both amplitude and duration processing. Specifically, significant improvements were observed for the 10-dB condition for amplitude processing and the 15-ms condition for duration processing. Results show that increasing consonant duration and amplitude is a viable approach to improving speech understanding; nevertheless, this data suggest that there is a cutoff point where lengthening the duration of a consonant can result in a degraded speech signal.

---

Published by the Acoustical Society of America through the American Institute of Physics

## INTRODUCTION

Poor speech intelligibility is a common consequence of sensorineural hearing loss [1, 2, 3]. Hearing aids often fail to solve the problem and have especially fallen short in difficult listening situations such as background noise and reverberation [4]. Large improvements in speech understanding [5] have been demonstrated by a mode of speaking known as clear speech, but this concept has yet to be utilized in a hearing aid algorithm. This paper reports one of two experiments presented at the Acoustics'08 Paris conference.

Clear speech is the means of consciously altering your speaking style to a more enunciated and comprehensible manner and is often implemented when speaking to a hearing-impaired individual. It has been found that speech intelligibility increased by 17 to 26 percentage points on average when persons with hearing impairment listened to clearly spoken nonsense sentences versus the same material spoken in a conversational manner [5, 6].

Analyses of clear speech have revealed many characteristics that distinguish it from conversational speech [5, 6, 7, 8, 9, 10, 11, 12, 13]. Phonetic changes produced during clear speech include increased duration of consonants, vowels and voice onset time (VOT) as well as consonant-vowel (CV) ratio increases and articulated phonemes [7]. Picheny et al. [5, 7] found that, due to the introduction of more pauses and increases in the duration of individual sounds, the average clear sentence is twice as long as the same sentence spoken conversationally. Speaking rate was thought to be one of the key differences between clear and conversational speech. If speaking rate was truly the main difference between clear and conversational speech that is responsible for increases in speech intelligibility, it would be difficult to implement in an algorithm as the sequence of visual and auditory input would be disrupted. However, Krause and Braida [10] trained talkers to produce clear speech at a variety of speaking rates and found that it is possible for a clear speech intelligibility to be obtained at increased speaking rates up to approximately 200 words per minute. This finding provides researchers a basis on which to explore the other properties of clear speech that can be implemented while maintaining audio-visual synchrony.

Two characteristics of clearly enunciated speech that have been studied concomitantly are increased consonant duration and consonant amplification [14, 15, 16]. Increased consonant duration of individual speech sounds and greater root mean square (RMS) intensities for obstruent consonants are listed as some of the major acoustical modifications found in clear speech [5, 7]. Through this study, we aim to examine these two characteristics of clear speech (i.e., increased consonant duration and amplitude) through application of a duration and amplitude processing scheme to enhance these specific consonant attributes in order to understand the extent to which these augmentations improve speech intelligibility. Specifically, the investigators seek to evaluate the effectiveness of consonant duration trading and amplification to determine if a similar processing scheme could be successfully incorporated within a hearing aid algorithm in order to improve speech understanding in persons with hearing loss. It is hypothesized that when consonants are increased in duration and the adjacent vowel is decreased by the same amount, speech understanding will improve in individuals with normal hearing and hearing impairment in the presence of background noise. In addition, applying phonetically-targeted amplitude increases to consonants will lead to an improvement in speech understanding in noise.

## METHOD

Hearing-In-Noise Test (HINT) sentences [17] were used in the study. All stimuli were recorded in 16-bit resolution at a sampling rate of 44.1 kHz. Throughout the HINT sentences, twelve consonants (/p b t d k g s ʃ m n ŋ f/) were enhanced in the Consonant Duration Experiment and sixteen consonants (/p b t k d g f v s z ʃ θ dʒ m n/) in the Consonant Amplitude Experiment. Certain consonants were chosen for each experiment based on the findings of Picheny et al. [7]. In addition, other consonants (voiceless fricatives, stop/plosives, and nasals) were selected in order to analyze the effects of processing on a wider range of consonants. The processing did not change the overall RMS values of the stimuli substantially. The average RMS values between unprocessed and processed lists were equated to ensure the signal-to-noise ratio (SNR) remained constant across processed lists.

Consonants were processed with the use of Adobe Audition (Adobe, San Jose, CA) on a desktop computer (Dell, Round Rock, TX). The Consonant Duration Experiment was completed by increasing the duration of the consonant or part of the consonant and decreasing the duration of the adjacent vowel portion by the same amount. For stochastic (unvoiced) consonants or sections of consonants, small sections (3-10 ms) were duplicated. For deterministic (voiced) consonants or sections of consonants, the duration increase was achieved by duplicating

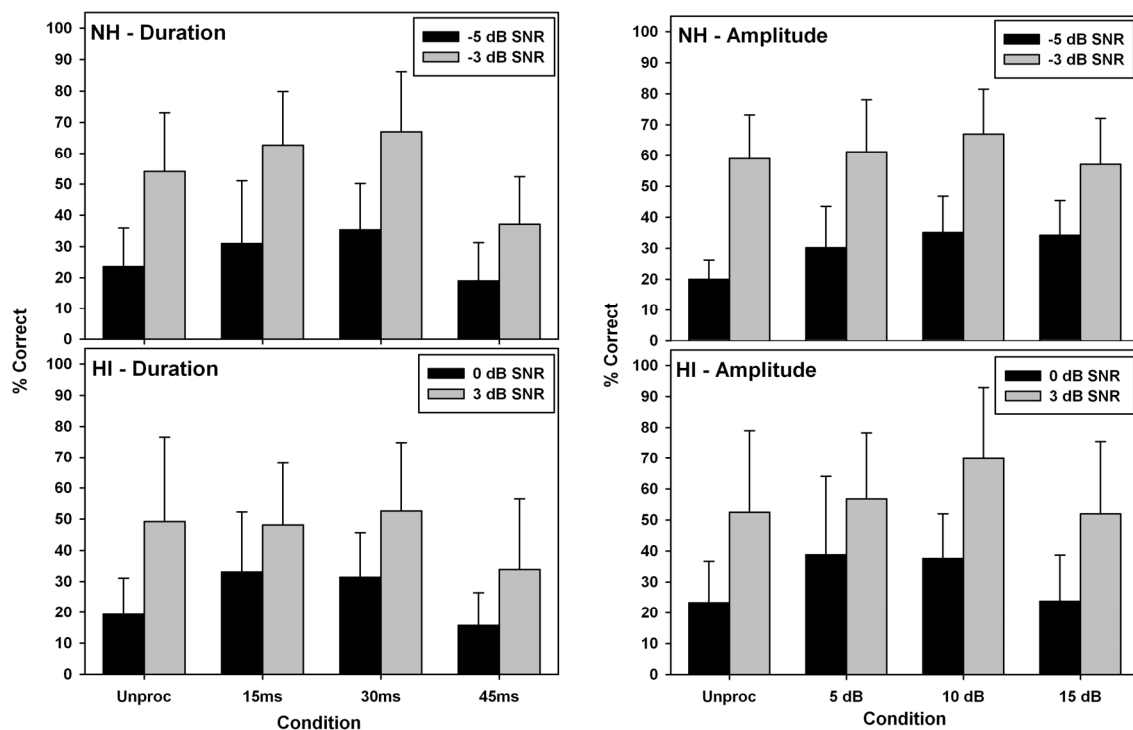
individual cycles of the waveform. In order to keep processing proportional throughout the entire consonant, VOT was increased in small segments (2-5 ms).

For the Consonant Amplitude Experiment, the same HINT sentences were used as in the Consonant Duration Experiment except duration was not manipulated. The amplitude of specific consonants was amplified while the vowels remained unchanged. Four conditions were examined; one included the unprocessed stimulus set while the remaining three conditions contained consonant amplifications of 5, 10, and 15 dB, respectively. The extent amplification, however, was constrained by a 0 dB CV ratio.

Ten normal-hearing and 10 hearing-impaired adults participated in the study. Hearing-impaired participants had a diagnosed bilateral, symmetrical sensorineural hearing loss  $\geq 25$  dB HL over at least 3 octave frequencies from 250 to 8000 Hz. Participants were tested in sound field in a two-room audiometric suite (Industrial Acoustics Company, New York). Stimulus presentation and data collection were controlled through the use of Matlab software (The Mathworks, Natick, MA) running on a desktop computer (Dell, Round Rock, TX). Stimuli were stored on the computer and presented via a 16-bit soundcard (Creative Labs, Singapore). The stimuli were amplified through a Crown CTs-4200 (Elkhart, IN) audio amplifier and presented over a loudspeaker (B&W DM601 S3, West Sussex, UK) at 0 degrees azimuth.

Speech stimuli were presented to both groups at 80 dB SPL for all conditions to ensure audibility. Hearing-impaired participants listened to stimuli without the use of their hearing aids in order to determine the effects of the speech processing free of hearing aid amplification and compression. Speech-shaped noise and speech stimuli were presented at two different SNRs for each subject group for each processing scheme. The SNRs used during the HINT presentation were -5 and -3 for the normal-hearing subjects, and 0 and 3 dB for the hearing-impaired group. During testing, each subject completed one sentence list for each condition in the SNR, resulting in four lists for each SNR. The Consonant Duration Experiment totaled eight sentence lists as well as the Consonant Amplitude Experiment. After each presentation, the participant was asked to repeat the sentence as it was heard. Scoring was completed using percentage of key words correct.

## RESULTS



**FIGURE 1.** Results of HINT sentence presentations for the Consonant Duration Experiment (left panels) and the Consonant Amplitude Experiment (right panels) across all conditions for normal-hearing subjects (top panels) and hearing-impaired subjects (bottom panels).

Consonant Duration Experiment results of HINT sentences across all conditions are represented in the left panels of Fig. 1 for both subject groups. In the hearing-impaired group, one subject was excluded due to the floor effect in the best SNR. Results did not reveal a significant difference between the normal-hearing and hearing-impaired groups across conditions, ( $F(1, 152)=0.88, p=0.36$ ). Conditions were significantly different from one another, ( $F(1,152)=30.8, p<0.05$ ), while no interaction was revealed between subject group and condition, ( $F(1, 152)=1.1, p=0.367$ ). However, the *post hoc* statistical analysis using Fisher's LSD revealed multiple findings. The 30 ms condition was significantly improved over most conditions, specifically, the unprocessed and 45 ms conditions at both SNRs, ( $p<0.05$ ). At the worse SNR, the 15 ms condition was improved over the unprocessed and 45 ms conditions while speech intelligibility at the 15 ms condition was only increased over the unprocessed condition when presented at the better SNR, ( $p<0.05$ ). The 45 ms condition was significantly worse than the 15 and 30 ms conditions at the worse SNR and all conditions at the better SNR at the level of 0.05.

Consonant Amplification Experiment results of the HINT sentences are represented in the right panels of Fig. 1 for each subject group. The same subject was also excluded from this experiment due to a similar floor effect. Once again, results did not reveal a significant difference between the normal-hearing and hearing-impaired groups across conditions ( $F(1,152)=0.05, p=0.824$ ). Again, the effect of condition was significant, ( $F(1,152)=41.45, p<0.05$ ) and no condition by group interaction existed ( $F(1,152)=1.7, p=0.116$ ). Upon *post hoc* analysis of the conditions using Fisher's LSD, the 10 dB amplification condition was significantly improved over all conditions at the better SNR, ( $p<0.05$ ). Evaluation of the 10 dB condition at the worse SNR revealed the condition to be improved in comparison to the unprocessed and 15 dB conditions at the level of 0.05. The 5 dB amplification condition improved speech intelligibility over the unprocessed condition only at the worse SNR, ( $p<0.05$ ). Applying 15 dB of amplification also revealed intelligibility improvements at the worse SNR, ( $p<0.05$ ). The 15 dB condition was not significantly improved over any conditions in the better SNR.

## DISCUSSION

The results of both experiments suggest the consonant duration and amplitude characteristics of clear speech do play a role in the large speech intelligibility difference seen between clear and conversational speech. Consonant lengthening duration and phonetic-specific amplification are aspects of clear speech that resulted in a measurable improvement in speech recognition. However, a limit to the amount of processing was found since the extreme conditions of both experiments resulted in either diminished or unchanged speech intelligibility scores in relation to all the other conditions. Increasing consonant duration and amplifying consonants in limited amounts could be a viable approach to increasing speech intelligibility in applications such as hearing aids. For this to happen, however, advancements in hearing aid signal processing first need to take place because this precise processing is computationally intensive.

## REFERENCES

1. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Science of the United States of America*, 103, 18866-18869.
2. Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2005). Speech understanding in quiet and noise, with and without hearing aids. *International Journal of Audiology*, 44, 574-583.
3. Rowland, J. P., Dirks, D. D., Dubno, J. R., & Bell, T. S. (1985). Comparison of speech recognition-in-noise and subjective communication assessment. *Ear and Hearing*, 6, 291-296.
4. Levitt, H. (2001). Noise reduction in hearing aids: A review. *Journal of Rehabilitation Research and Development*, 38, 111-121.
5. 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.
6. Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, 95, 1581-1592.
7. 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.

8. Picheny, M. A., Durlach, N. I., & Braida, L. D. (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. *Journal of Speech and Hearing Research*, 32, 600-603.
9. Uchanski, R. M., Choi, S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech and Hearing Research*, 39, 494-509.
10. Krause, J. C. & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America*, 112, 2165-2172.
11. Krause, J. C. & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America*, 115, 362-378.
12. Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *Journal of the Acoustical Society of America*, 116, 2365-2373.
13. Gagné, J-P., Rochette, A., & Charest, M. (2002). Auditory, visual and audiovisual clear speech. *Speech Communication*, 37, 213-230.
14. Gordon-Salant, S. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *Journal of the Acoustical Society of America*, 80, 1599-1607.
15. Gordon-Salant, S. (1987). Effects of acoustic modifications on consonant recognition by elderly hearing-impaired subjects. *Journal of the Acoustical Society of America*, 81, 1199-1202.
16. Montgomery, A. A. & Edge, R. A. (1988). Evaluation of two speech enhancement techniques to improve intelligibility for hearing-impaired adults. *Journal of Speech and Hearing Research*, 31, 386-393.
17. 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. *Journal of the Acoustical Society of America*, 95, 1085-1099.