

Proceedings of Meetings on Acoustics

Volume 9, 2010

<http://asa.aip.org>

159th Meeting
Acoustical Society of America/NOISE-CON 2010
Baltimore, Maryland
19 - 23 April 2010
Session 2aPP: Psychological and Physiological Acoustics

2aPP14. Performance of Phonemically-Targeted Processing in Conjunction with Compression Processing with Spectral Enhancement

Jeffrey J. DiGiovanni*, Janet C. Rutledge, Chessy S. Umble and Allison P. Mester

*Corresponding author's address: Communication Sciences and Disorders, Ohio University, W151a Grover Center, Athens, OH 45701, digiovan@ohio.edu

The goal of the present study is to test, individually and in combination, two signal-processing strategies designed to improve both consonant and vowel perception. In the first strategy, specific consonants were targeted for processing to increase amplitude and duration. For consonants with a duration increase, the adjacent vowel was decreased proportionately in order to maintain overall word and sentence duration. Second, Coliga, an adaptive compression-processing strategy incorporating spectral enhancement, was used to process stimuli. Hearing-in-noise-test sentences were presented monaurally to normal-hearing and hearing-impaired adults in the presence of speech-shaped noise. For normal-hearing listeners, it appears that Coliga is improving intelligibility in certain conditions. Conservative amounts of duration and amplitude processing need to be applied in combination rather than when applied independently in order to provide benefit. For the hearing-impaired listeners, Coliga with spectral enhancement resulted in poorer performance than without, for all conditions except for the Coliga-only conditions, where the intelligibility was the same. It appears that phonemically-targeted speech processing and Coliga are working antagonistically due to either 1) the processing method negating the benefits when the two are combined, or 2) combining the duration and amplitude components leading to no benefit despite documented benefit of each when employed individually.

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

INTRODUCTION

People with sensorineural hearing loss (SNHL) have difficulty understanding speech in noise due to inaudibility and lack of speech clarity¹¹. Current hearing aid processing uses amplification and compression to make speech audible; however this amplification of the speech signal does not necessarily result in increased clarity. Hearing aids often fail to solve the problem of poor speech intelligibility, especially in difficult listening situations such as background noise and reverberation⁷. With the introduction of digital signal processing (DSP) in hearing aids, researchers have strived to advance speech processing technology and to improve speech intelligibility for persons with hearing loss. DSP advancements have made it possible for hearing aid companies to implement processing strategies, such as spectral enhancement, into hearing aids. Spectral enhancement is a method of processing which attempts to improve speech intelligibility by enhancing the important features of speech by exaggerating spectral contrast (i.e. the dB difference between (formant) peaks and valleys in the spectrum.) By selectively amplifying the peaks, with or without attenuating the valleys, the spectral contrast becomes more prominent. This serves to improve speech intelligibility in two ways; partially ameliorating the reduced frequency selectivity commonly found in persons with SNHL and by improving the local signal-to-noise ratio (SNR). Results from psychophysical tasks have demonstrated the need for a greater peak-to-trough ratio by individuals with hearing impairment when compared to individuals with normal hearing^{6, 12} and have shown spectral enhancement to be a fundamentally sound technique^{1, 2, 3}. Studies investigating spectral enhancement algorithms have had a wide range of outcomes varying from reductions in speech intelligibility to modest improvements^{3, 4}. Even though the concept of spectral enhancement is well-documented, researchers are still challenged with the actual implementation of this processing into a hearing-aid algorithm as past attempts have failed to produce beneficial results on a consistent basis.

A mode of speaking known as clear speech, or the act of deliberately speaking clearly, has also demonstrated speech intelligibility to be able to be increased, 17-26% on average^{9, 10}. Several differences between clear and conversational speech have been documented⁵ of which two have been studied⁴. Two distinguishing characteristics between clear and conversational speech is the amplitude and duration of consonants. In running speech, certain consonants uttered in a “clear” manner have a greater amplitude and duration than the equivalently produced “conversational” consonants⁵. Researchers have post-processed conversationally-spoken sentences by increasing either the amplitude of those consonants or the duration of those consonants and found significant intelligibility improvements⁴. These were not studied in combination, however. Therefore, measuring improvements when both amplitude and duration are processed together was included in the current study.

Clear-speech processing has yet to be utilized in a hearing aid algorithm and therefore the effects of other hearing-aid processing (e.g., frequency-selective gain, compression) are unknown. By combining spectral enhancement and a clear speech processing scheme into a single algorithm speech intelligibility may be further enhanced. Coliga, a hearing aid algorithm that combines spectral enhancement with compression and linear gain, is a commonly-used processing strategy which allows all the sounds in the environment to be compressed into a hearing-impaired person’s reduced dynamic range. A draw back of compression is that it flattens the speech spectrum; Coliga preserves important spectral features of speech by maintaining or exaggerating spectral contrast, amplifying spectral peaks, and attenuating the valleys between results in a difference in the peak-to-valley ratio.

METHOD

Two experiments were performed in this study. In each experiment, background information was obtained from each participant and was used to identify any reasons for ineligibility such as vision disorders, chronic/serious illnesses, learning disabilities, psychological/behavioral disorders, fluctuating hearing, ear injuries requiring medical attention, ear surgeries, otalgia, or otorrhea. No participants were deemed as ineligible to partake in the current study due to the aforementioned criteria. However, subjects who scored less than 10.0% correct in any condition were excluded. Participants were recruited from the Ohio University student population as well as persons in the surrounding community of Athens, Ohio. This study was approved by the Ohio University Institutional Review Board and all participants provided informed consent.

Coliga with and without spectral enhancement as well as the amplitude and duration components of clear-speech processing were tested with normal and hearing-impaired individuals in two experiments. The first experiment included normal-hearing participants who listened to sentences that were either processed or unprocessed. The two types of processing under test were clear-speech processing (including variations of duration and amplitude processing) and Coliga with the spectral-enhancement feature enabled. The goal was to understand the benefits of the combined (i.e., both duration and amplitude) clear-speech processing and spectral enhancement. The only spectral enhancement algorithm available to us was the one incorporated into Coliga. Unfortunately, for the purpose of normal-hearing listeners, the gain processing normally reserved for those with hearing loss was unavailable. Therefore, in experiment 1 the use of Coliga was necessary to test the spectral enhancement. In this experiment, sentences were either unprocessed, processed either for clear speech or by Coliga, or processed for clear speech *and* by Coliga. In this manner, the processed sentences included variants of the clear-speech processing. The second experiment included participants with sensorineural hearing loss. Given the presence of hearing loss, all sentences were processed by Coliga. In this manner, the Coliga processing with the spectral-enhancement feature disabled and no other processing may be thought of as “unprocessed” except to provide frequency- and level-dependent gain commonly used in traditional hearing aids. In addition, the spectral enhancement feature was enabled and disabled in different conditions. Also, clear-speech processing was tested.

Experiment 1: The first experiment included 13 normal-hearing adults between the ages of 19 and 32 years. These participants underwent a pure-tone air conduction hearing screening with tones presented at 0.5, 1.0, 2.0, and 4.0 kHz at a level of 20 dB HL to each ear to ensure normal hearing sensitivity, bilaterally. Stimuli used for these participants included Hearing In Noise Test (HINT) sentences⁸ processed at +5 dB signal-to-noise ratio (SNR) and presented at -3 dB SNR. HINT sentences were selected as stimuli for this study based on the consistent level of difficulty and the phonetic equivalency across sentence lists⁸. For Experiment 1, duration and/or amplitude increases were made to the appropriate consonants to produce a combined duration and amplitude processing scheme. For the duration increases, 12 consonants were processed (/p b t d k g s ʃ m n ŋ l/). Similarly, 16 consonants were processed for amplitude increases (/p b t k d g l v a z ʃ θ ð m n h/). Consonants were processed with the use of Adobe Audition (Adobe, San Jose, CA) on a desktop computer (Dell, Round Rock, TX). Duration increases were 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, voice-onset time was increased in small segments (2-5 ms). For amplitude increases, the amplitude of the listed consonants was amplified while the vowels remained unchanged. The extent of amplification, however, was constrained by a 0 dB CV ratio. Four combinations were selected, based on the individual improvements for duration and amplitude found in literature⁴. These conditions were as follows, in the format of duration (ms)/amplitude (dB): 15/5, 15/10, 30/5, 30/10. For example, if a condition specified a duration increase of 15 ms and an amplitude increase of 5 dB, the conditions is referred to as 15/5. These conditions were run both with and without Coliga processing. For sentences processed by Coliga, the spectral-enhancement feature was always enabled in this experiment. The amount of increase in spectral contrast between peaks and troughs was a user-selectable parameter within Coliga, for which 12 dB was chosen¹. Different HINT sentences were used for each condition and the order of presentation was counterbalanced for all participants. All stimuli were recorded at 16-bit resolution at a sampling rate of 22.05 kHz. Throughout the HINT sentences, duration processing for 12 consonants were combined to create an overall clear-speech processing scheme⁴.

Experiment 2. The second experiment included seven hearing impaired adults between the ages of 23 to 84 years; these subjects had a sensorineural hearing loss ranging from mild to severe. Hearing sensitivity was established by a pure-tone air and bone conduction hearing screening with tones presented at 0.25, 0.5, 1.0, 2.0, 4.0, 8.0 kHz. Bone conduction was performed to ensure the participants had a sensorineural hearing loss. Stimuli presented to these subjects were also Hearing In Noise Test (HINT) sentences (Nilsson et al., 1994) presented at +5 dB signal-to-noise ratio (SNR). All stimuli were processed by Coliga, with either the spectral enhancement feature enabled or disabled. When enabled, +12 dB of spectral enhancement was chosen. Also, combined (i.e., both amplitude and duration) clear-speech conditions¹ were tested at the following duration (ms) and amplitude (dB) 15/5, 15/10, 30/5, 30/10. Clear-speech processing was performed in the same manner as described in the first experiment. The order of presentations was counterbalanced.

RESULTS

Experiment 1: Results of a two-factor, repeated measure ANOVA revealed that there was a main effect of Coliga processing, ($F(1, 15)=6.92, p<0.05$), and a main effect of condition, ($F(4, 15)=57.99, p<0.05$), while no interaction was revealed between processing and condition, ($F(1,4)=2.04, p=0.101$). A Fisher's LSD Multiple-Comparison post-hoc analyses revealed a significant improvement with Coliga processing in the 15ms/5dB and 15ms/10dB conditions ($p<0.05$). Analyses also revealed that the 15ms/10dB, 30ms/5dB, and 30ms/10dB conditions were significantly worse than the unprocessed and 15ms/5dB conditions for both Coliga and non-Coliga processing ($p<0.05$). The following graph shows the HINT sentence intelligibility at -3 dB SNR for normal hearing participants.

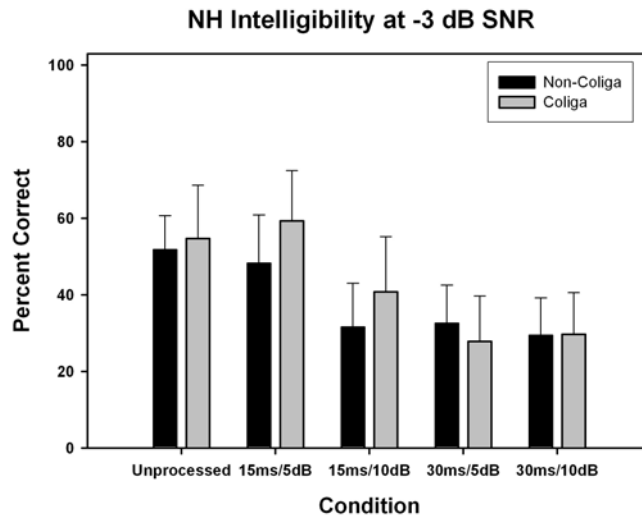


FIG 1. Shows HINT sentence intelligibility at -3 dB SNR for the normal-hearing listeners in Experiment 1.

Experiment 2: Results of a two-factor, repeated measure ANOVA revealed a main effect of spectral enhancement processing, ($F(1,6)=14.22, p<0.05$), and a significant effect of condition, ($F(4, 24)=5.86, p<0.05$). There was also an interaction between spectral enhancement processing and condition, ($F(4,24)=2.88, p<0.05$). The following graph shows HINT sentence intelligibility at +5 dB SNR for hearing impaired participants.

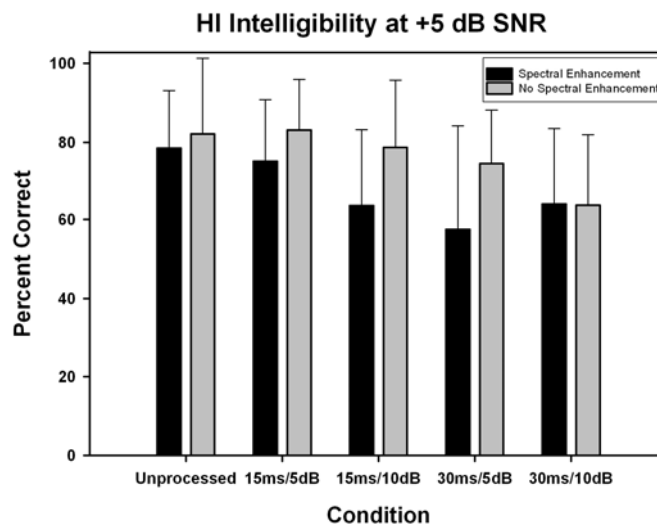


FIG 2. Shows HINT sentence intelligibility at +5 dB SNR for the hearing-impaired listeners in Experiment 2.

DISCUSSION

Results from Experiment 1 indicate Coliga tends to improve speech intelligibility in noise, although this effect is not seen in all conditions. The significant improvement resulting from Coliga processing (i.e., 15ms/5dB and 15ms/10dB conditions) suggests that this algorithm can be used to improve speech intelligibility in noise. The degradation of the speech signal at the two extreme conditions (i.e., 30ms/5dB and 30ms/10dB) suggests a cutoff point in processing. Large amounts of processing may be distorting the signal beyond any benefit provided. Although not revealed during the statistical evaluation of the data, the best condition was Coliga 15ms/5dB. This suggests that this combined processing scheme may provide a viable approach to improving speech intelligibility. Previous data from DiGiovanni and Stover⁴ showed improvements when clear speech processing was used. Experiment 2 tested the Coliga algorithm with the spectral enhancement feature enabled and disabled as well as combining it with and without spectral enhancement as well as in combination with combined clear speech processing; this revealed that spectral enhancement may be working antagonistically with the combined clear speech processing. For hearing impaired participants, combined processing results in reduced intelligibility. These results show that either combining amplitude and duration is not beneficial, or clear speech principles are not interacting positively with Coliga. Future studies could include investigating the effects of adjusting the parameters of the Coliga algorithm. This is the first test of the new version of this algorithm. Parameters to be adjusted include the number of selected maximum peaks per frame, the spectral change rate per frame, the amount of spectral enhancement, and further developing the techniques used to combine the duration and amplitude processing. DiGiovanni & Stover (2008) showed benefit from duration and amplitude processing when applied individually. Therefore, we expect to see benefit when they are used in combination and perhaps the lack thereof is due to techniques used in the combination process. The duration and amplitude processing is done manually before noise is introduced. An algorithm is not yet available to apply this processing as it would be used in a hearing aid; therefore, algorithm development is essential in the application of these findings.

SUMMARY

Results from Experiment 1 suggest Coliga significantly improves speech intelligibility in noise for normal hearing listeners in 15ms/5dB and 15ms/10dB conditions; however tends to degrade the signal in 30ms/5dB and 30ms/10dB conditions. In experiment 2, combining Coliga and spectral enhancement processing appears to worsen speech intelligibility for hearing impaired listeners.

REFERENCES

1. DiGiovanni, J. J., and Nair, P. (2006). "Auditory filters and the benefit measured from spectral enhancement." *J. Acoust. Soc. Am.* 120, 1529-1538.
2. DiGiovanni, J. J., Nelson, P. B., and Schlauch, R. S. (2005). "A psychoacoustic evaluation of spectral enhancement," *J. Speech Lang. Hear. Res.* 48, 1-15.
3. DiGiovanni, J. J., Ries, D. T., Prewitt, J. M., and Seebohm, C. S. (in prep). "The relative contribution of low and high frequency decrements in increment detection related to spectral enhancement," *J. Speech Lang. Hear. Res.*
4. DiGiovanni, J. J., and Stover, A. K. (2008). "The role of consonant duration and amplitude on speech intelligibility in noise," *Proc. Mtg. Acoust.* <http://dx.doi.org/10.1121/1.3033930>.
5. Krause, J. C., and Braida, L. D. (2002). "Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility," *J. Acoust. Soc. Am.* 112, 2165-2172.
6. Leek, M. R., and Dorman, M. F. (1987). "Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* 81, 148-154.

7. Levitt, H. (2001). "Noise reduction in hearing aids: an overview," *J. Hear. Res. Dev.* 38, 111-121.
8. Nilsson, M., Soli, S. D., and 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.
9. Payton, K. L., Uchanski, R. M., and Braida, L. D. (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.
10. Picheny, M. A., Durlach, N. I., and Braida, L. D. (1985). "Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech," *J. Speech Hear. Res.* 28,96-103.
11. Plomp, R. (1978). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," *J. Acoust. Soc. Am.* 63, 533-549.
12. Summers, V., and Leek, M. R. (1994). "The internal representation of spectral contrast in hearing-impaired listeners," *J. Acoust. Soc. Am.* 95, 3518-3528.