OBJECTIVE AND SUBJECTIVE EVALUATION OF NOISE REDUCTION ALGORITHMS FOR HEARING AIDS

Vijay Parsa[1,2], Gurjit Singh[1], Guo Chen[1] and Karthikeyan Umapathy[1]
[1]National Centre for Audiology, University of Western Ontario, London, Ontario, Canada N6G 1H1.
[2]Contact email address : parsa@nca.uwo.ca

1. INTRODUCTION

Speech is a highly redundant signal and because of this redundancy listeners with normal hearing are able to understand speech even in the presence of background noise. For listeners with sensorineural hearing-impairment however, there is a considerable loss of these redundant cues in the speech signal (Levitt, 2001). Thus, one of the most common complaints made by hearing-aid users is understanding speech in noise and consequently, one of the driving factors behind obtaining a hearing aid.

Recently, there has been an explosion in the number of digital hearing aids appearing on the market, with a number of these devices proffering noise reduction capabilities. Accurate and comprehensive evaluation of their noise reduction performance is important in order to quantify the relative benefits of these devices under a variety of listening conditions. In general, the noise reduction performance can be evaluated through objective electroacoustic measures and/or subjective listening tests. Subjective listening tests are preferred for their face validity, but they are often expensive, time-consuming and labour-intensive. Currently, there does not exist a validated or standardized electroacoustical procedure that allows clinical audiologists from assessing the relative benefits of various devices that offer similar, but not identical, noise reduction algorithms. The present study aims to address this issue through bridging electroacoustic and subjective measures of quality of speech processed through the noise reduction algorithms. The specific goals of this research were two-fold : 1) to evaluate the quality of speech processed through six different noise reduction algorithms with normal hearing and hearing impaired listeners, and 2) to identify instrumental measures of noise reduction performance that correlate best with the behavioural data.

2. METHOD

Six candidate noise reduction algorithms were evaluated in this study (Umapathy and Parsa, 2003). Algorithm 1 (EPH) is a technique based on minimizing the Mean Square Error (MMSE) of the Short Time Spectral Amplitude (STSA) estimator as proposed by Ephraim and Malah (1984). Algorithm 2 (WOL) is also based on the STSA estimator, however instead of the MMSE based amplitude estimator, the criterion used in this algorithm is the MMSE power spectrum estimator. Algorithm 3 (ES) is based on the subspace projection technique where noisy signals are decomposed into signal and noise subspaces, with the assumption that the signal is present only in the signal subspace whereas noise spans both subspaces (Klein and Kabal, 2002). Algorithm 4 (WV) uses wavelet packet decomposition and auditory masking properties (Lu and Wang, 2003). Both algorithms 5 (MP) and 6 (MP2) are based on the matching pursuit algorithm where time-frequency atoms from a Gaussian dictionary of time-frequency functions are adaptively moulded to fit the speech signal (Mallat and Zhang, 1993). The MP algorithm used a varying number of time-frequency functions to reconstruct the enhanced signal by applying a threshold on the slope of the rate of energy capture curve, while the MP2 algorithm used a fixed number of time-frequency atoms irrespective of the SNR values.

Ten adult participants with normal hearing (pure tone thresholds $\leq$ 20 dB HL at 1, 2, and 4 kHz) and 10 adult participants with hearing loss (mild to profound sensorineural hearing impairment) were paid to participate in this study. Participants rated the improvements in sound quality on 11-point scales, ranging from -5 to 5, on five dimensions: clarity, listening comfort, listening effort, background noise, and overall quality. Sound quality ratings were obtained using two sentences from the Hearing In Noise Test (HINT) database (Nilsson et al., 1994) which were corrupted by either speech-shaped noise (SSN) or multi-talker babble (MTB) at SNRs ranging between –4dB to +12 dB. In parallel, several instrumental measures of speech quality were computed from the speech stimuli processed by the noise reduction algorithms. These included the Perceptual Evaluation of Speech Quality (PESQ) (ITU, 2001), Perceptual Speech Quality Measure (PSQM) (ITU, 1998), Measuring Normalizing Blocks (MNB) (ITU, 1998), and Measuring Perceptual Spectral Density Distribution (MPSDD) (Chen and Parsa, 2004).

3. RESULTS

Figures 1a and 1b display the subjective quality ratings for different noise reduction algorithms from normal and hearing impaired listeners respectively. Algorithms based on spectral subtraction techniques (EPH and WOL)

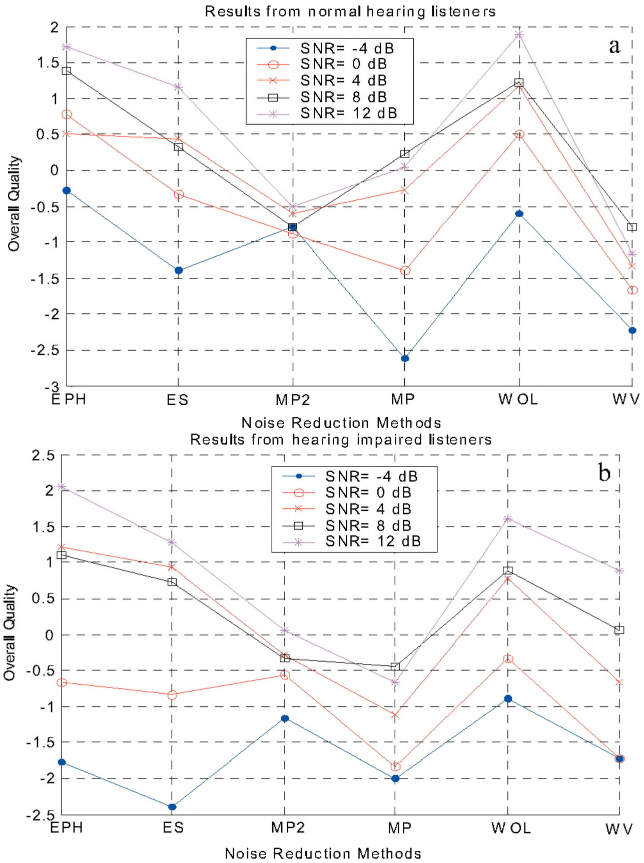

Fig. 1. Sound quality ratings from normal and hearing impaired listeners. Ratings were obtained from speech stimuli corrupted by the speech shaped noise at different SNRs and processed by different noise reduction algorithms.

can be seen to improve perceptual sound quality ratings by normal hearing listeners in conditions where the SNR is ≥ 0 dB. For hearing impaired listeners, the same algorithms were judged to improve the sound quality when the SNR is ≥ 4 dB. Table 1 displays the correlation between the instrumental measures of speech quality and the subjective ratings. These correlations were computed separately for perceptual data collected from normal hearing and hearing impaired (HI) subjects for the two noise conditions. It is evident from this table that the PESQ parameter correlated best with perceptual ratings under the speech-shaped noise condition, and its performance is similar to the PSQM parameter under the multi-talker babble condition.

Table 1: Correlations between instrumental and subjective measures of speech quality.

Measure	SSN		MTB	
	Normal	HI	Normal	HI
PESQ	0.79	0.83	0.71	0.85
PSQM	0.63	0.71	0.73	0.88
MNB	0.30	0.62	0.44	0.70
MPSDD	0.74	0.77	0.74	0.77

4. CONCLUSIONS

This paper investigated the performance of six noise reduction algorithms under a variety of listening conditions using instrumental measures and subjective ratings of speech quality. Ratings obtained from normal and hearing impaired listeners showed improvements in speech quality at positive SNRs for algorithms based on short-time spectral amplitude estimation. Instrumental measures such as the PESQ measure exhibited a good degree of correlation with quality ratings from normal and hearing impaired listeners. These results show that the PESQ measure can potentially be of use in the development and optimization of noise reduction algorithms for hearing impaired listeners.

REFERENCES

Chen, G. & Parsa, V. (2004). Output-based speech quality evaluation by measuring perceptual spectral density distribution, IEE Electronics Letters, 40, 783-785.

Ephraim, Y. & Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time amplitude spectral estimator. IEEE Transactions on Acoustics, Speech, and Signal Processing, 32, 1109-1121.

International Telecommunication Union (1998). Perceptual Speech Quality Measure , ITU-T recommendation P.861, Geneva, Switzerland.

International Telecommunication Union (2001). Perceptual Evaluation of Speech Quality , ITU-T recommendation P.862, Geneva, Switzerland.

Klein, M. & Kabal, P. (2002). Signal subspace speech enhancement with perceptual post-filtering. Proceedings of the IEEE International Conference on ICASSP, 1, 537-540.

Levitt, H. (2001). Noise reduction in hearing aids: a review. Journal of Rehabilitative Research and Development, 38, 111-121.

Lu, C.T. & Wang, H.C. (2003). Enhancement of single channel speech based on masking property and wavelet transform. Speech Communication, 41, 409-427.

Mallat, S.G. & Zhang, Z. (1993). Matching pursuits with time-frequency dictionaries. IEEE Transactions on Acoustics, Speech, and Signal Processing, 41, 3397-3415.

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.

Umapathy, K. & Parsa, V. (October, 2003). Objective evaluation of noise reduction algorithms in speech applications, Paper #Z6-2, 115th Audio Engineering Society Convention, New York, USA.

ACKNOWLEDGEMENTS

Financial support from the Oticon Foundation, Denmark, and the Ontario Rehabilitation Technology Consortium is gratefully acknowledged.