THE BENEFIT OF SPEECH ENHANCEMENT TO THE HEARING IMPAIRED

N. Fink[1] M. Furst[2] and C. Muchnik[3]

[1] Department of Bio Medical Engineering, Faculty of Engineering, Tel-Aviv University, ISRAEL

[2] School of Electrical Engineering-Systems, Faculty of Engineering, Tel-Aviv University, ISRAEL

[3] Department of Communication Disorders, Faculty of Medicine, Tel-Aviv University, ISRAEL

ABSTRACT

Decreased speech intelligibility in background noise is a common complaint of most hearing impaired in comparison to quiet conditions. Speech enhancement algorithms widely used in hearing aids are often used with some success to improve speech intelligibility for the hearing impaired. It is eminent from the literature that not all hearing impaired benefit from speech enhancement algorithms. From the literature it is not clear whether speech enhancement techniques are useful to either normal hearing or hearing impaired.

Our results have shown a trend that with the assistance of speech enhancement techniques and an ear model comprising outer-hair cell functionality, the speech intelligibility of normal hearing becomes worse as the signal-to-noise ratio is lowered. In a similar task, hearing impaired, particularly those who use cochlear implants have shown a significant improvement in speech intelligibility.

Index Terms— speech intelligibility, noise, hearing-impairment, speech enhancement, ear model

1. INTRODUCTION

Loss of hearing is a major health problem with serious social implications. More than 10% of the world's population has some degree of hearing loss (HL) [1], and the number is expected to grow rapidly as the population ages. Many who have suffered a hearing loss feel restricted socially and professionally.

One of the most common complaints among patients with cochlear hearing loss is difficulty in understanding speech in a noisy environment with or without their hearing assisting devices (HAD) (hearing aids-HA or cochlear implant-CI).

Current HAs work well in quiet environments and provide the hearing impaired (HI) with improved understanding of auditory signals. Yet HADs are less efficient in noisy environments. This is probably due to spectral and/or temporal abnormal processing in the cochlea and other parts of the auditory pathway.

Most speech enhancement techniques that have been designed to reduce background noise have been commonly implemented with spectral subtraction in conjunction with auditory masked threshold algorithms [e.g. 2, 3]. These algorithms are used, for example, in cell phones as part of the speech vocoders, or in other speech applications such as speech recognition. While normal hearing (NH) listeners gain some benefit from those algorithms, the advantages for the HI are controversial, with most studies reporting less improvement in speech intelligibility than in NH people. Only few studies demonstrated an improvement in intelligibility among HI even when it was not achieved by NH [4, 5].

The statistics of HL indicate that only twenty percent of adults who could be helped with a HA actually wear one [6]. Not all HA users experience improved speech audibility and intelligibility. Among HA users, a unique group is of interest - characterized by a long-term use of their HA. Can HI with a senso-neural HL (SNHL) improve their hearing following long-term use of a HA?

There is a need to take advantage of speech enhancement techniques to improve speech recognition for monaural and binaural HA users when competing noise is present. We deposit that speech enhancement can improve speech recognition for long-term HA users when competing noise is present.

2. METHODOLOGIES

2.1. Speech Enhancement Techniques

2.1.1. Cochlear Reconstruction Algorithm (CRA)

Acknowledgments: The study was partly supported by the Advanced Communication Center of Tel Aviv University.

A specific speech enhancement technique was developed – Cochlear Reconstruction Algorithm (CRA). It is based on a full computational model of the cochlea [7]. It integrated outer hair cells (OHC) activity in a one-dimensional cochlear model. The dynamic properties of the cochlear model output were used to reconstruct noisy speech signals with improved signal to noise ratio (SNR). The cochlear representation algorithm (CRA) represents the auditory signal by the basilar membrane's velocity along the cochlear partition and as a function of time. The algorithm distinguishes between noised and un-noised speech fragments in the input speech signal according to threshold considerations and noised/un-noised decisions (by the algorithm) of adjacent fragments of the speech signal [8].

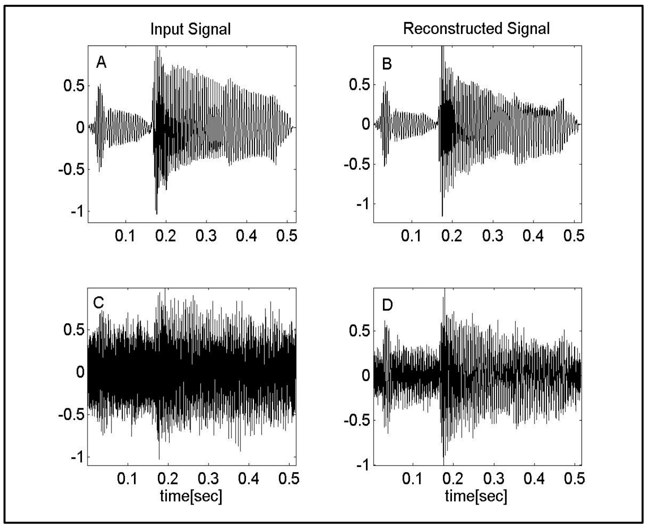

Fig. 1. Time domain representation of the word "BEN" (A) and its reconstructed representation (B). The word "BEN" + Gaussian white noise with SNR of -1.5 dB (C) and its reconstructed signal (D) with SNR of 6 dB

The reconstruction algorithm is demonstrated in Fig. 1 [9]. The original signals clean and noisy are presented in Figs 1A and 1C respectively, while the matched reconstructed signals are shown in figs. 1B and 1D respectively. Calculation of the mean SNR yielded -1.5 dB for the input signal (fig. 1C) and 6 dB for the reconstructed signal (fig. 1D), improvement of 7.5 dB.

Unique to the CRA (and advantageous over existing speech enhancement techniques) is the modeling of the OHCs dynamic behavior, thus providing some benefit in certain frequencies for the HI. The CRA can provide further effective speech enhancement for the HI by adjusting the CRA parameters according to the pathology of the HI.

2.2. Psycho-acoustical Experiments

2.2.1. Subjects

(1) 76 HI subjects (pre and post-lingual) participated in the experiment, 58 with HA (9 out of were tested with a different HA or re-tested) and 18 with CI (6 out of were re-tested). The average age of the HA users was 45 years, the average hearing-loss was 20 years and the average HA use was 8 years. Among the CI users, the values were 36, 34 and 15 respectively. (2) 43 NH subjects served as a control group.

2.2.2. Word Database and Noise Types

The database consists of the Hebrew adaptation to the AB List (HAB). The AB list, a set of monosyllabic real words, comprises consonant-vowel consonant (CVC) words. The list was designed so that the different phonemes in English shall be equally distributed throughout the entire list [10, 11]. The AB list is commonly used in hearing tests as it reduces the effect of word frequency and/or word familiarity on test scores. Corresponding lists are produced for other languages and accents [12]. The HAB list was designed for Hebrew natives, and it consists of 15 lists of 10 monosyllabic words such as "sir", "kir" [13]. A single female speaker with a sampling rate of 44.1 kHz recorded the HAB list.

Gaussian white noise was added to the database in various SNRs. The clean and noisy HAB lists were band-pass filtered between 500 Hz and 8 kHz. The filtered lists were applied to the CRA. Eventually, each input word had a corresponding reconstructed word. All together the complete database consisted of several HAB word lists (each list comprising of the same 150 words) in various treatments: noise levels.

Theoretical evaluation of CRA befit was obtained by computing the segmental SNR (SSNR). It was derived on the original database [SSNR(ORG)] and after application of the CRA algorithm [SSNR(CRA)]. A regression line was derived and found to be: SSNR(CRA)=7.2+0.74* SSNR(ORG). This result suggests that a maximum improvement of 7.2 dB (approximately) can be achieved for SSNR(ORG)=0 dB.

2.2.3. Experimental Procedure

Each subject passed a standard audiometric test procedure. The NH group consisted of subjects with 15 dB HL (pure tone average of 500, 1000, 2000 Hz). The hearing of the patients with HAs was assessed with and without their aid.

The psychoacoustical experiment was word recognition in an open-set. Subjects wearing their HA were introduced randomly with words chosen from the recorded database. The level of the tested words was adjusted for most comfortable level (MCL). The experiments took place in a sound proof room

Each subject was tested in 8 experimental sessions of 10-20 words each. Following each word, the subject was asked to repeat the word he/she heard.

Subjects were tested with SNR levels of 40, 30, 24, 18 dB. To avoid ceiling effect, NH subjects were also tested at SNR levels of 10, 5 and 0 dB.

3. RESULTS

3.1. Normal Hearing

The results of the NH group are depicted in Fig. 2. The average recognition score for various SNRs are indicated by the black and gray bars that represent the performance without the algorithm and with the algorithm respectively. A non-significant difference in performance with and without the CRA was achieved (-1%, -3%, -6%, 1%, 0% and -3% @ SNRs 30, 24, 18, 10, 5 and 0 respectively, where negative values depict poorer performance with the CRA rather than the original). A ceiling affect was derived at SNR40.

These findings are consistent with a comparative intelligibility study of single-microphone noise reduction algorithms on NH [14].

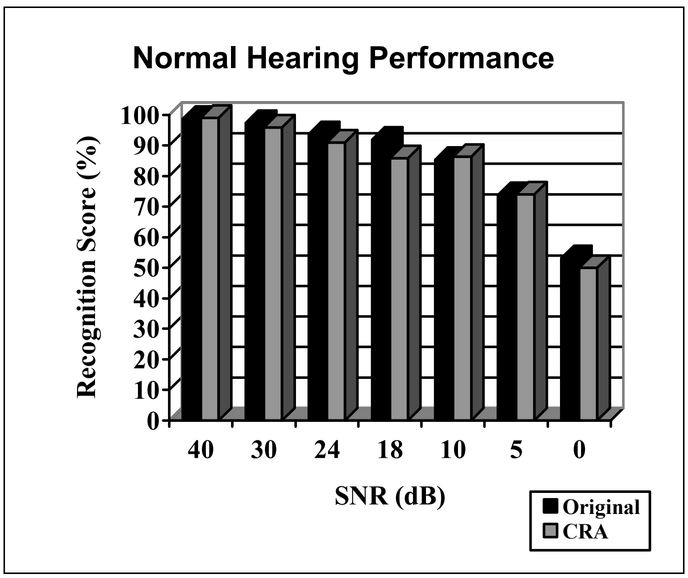

Fig. 2. Speech recognition score (percentage) of NH in an open-set word recognition task as a function of SNR.

3.2. Hearing-Aid Users

Speech recognition performance of the HI subjects was divided into two sub-groups. A group of 58 HI who use traditional HAs (analog or digital). Figure 3 represents the average performance of the first group of the hearing impaired. The use of the HA alone is characterized by a gradual decline in performance as SNR decreases, pointing to the difficulty of understanding speech as the energy of the competing noise increases. Applying the CRA to the HA users yielded improved performance of 1%, 1%, 4%, 15% @ SNRs 40, 30, 24, 18 respectively. Only for SNR18 the improvement was statistically significant (ANOVA, p<0.001).

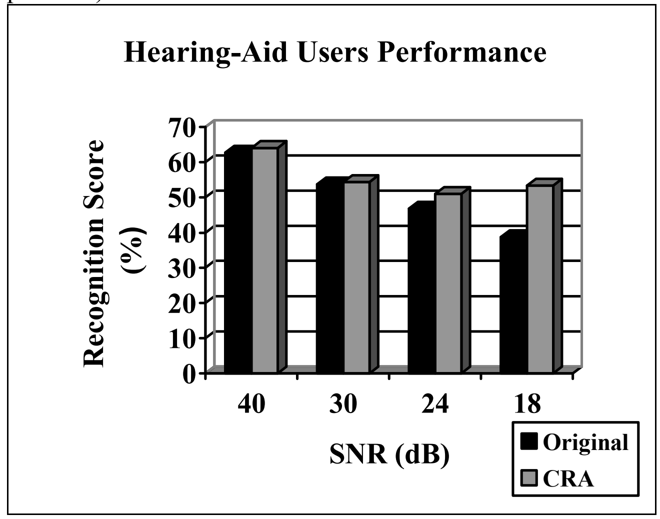

Fig. 3. Speech recognition score (percentage) of hearing-aid users in an open-set word recognition task as a function of SNR

3.3. Cochlear Implant Users

A second group of HI included eighteen CI users. Those subjects seem to benefit even more from the use of the CRA algorithm as can be seen in Fig. 4.

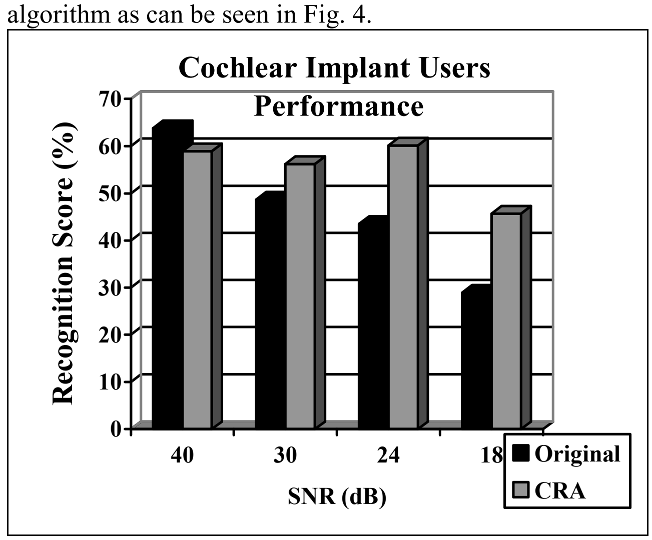

Fig. 4. Speech recognition score (percentage) of Cochlear Implant users in an open-set word recognition task as a function of SNR

Applying the CRA to the CI subjects yielded improved performance of 8%, 17%, 17% at SNR30, 24, 18 respectively) but also a decrease in intelligibility of 5% at SNR40. The improvement for SNR= 18, and 24 dB yielded a statistical difference (ANOVA, p<0.05). HA users exhibited a significant improvement at SNR18 only (fig 3).

4. DISCUSSION

The psychoacoustical experiments revealed an improvement in performance of HI subjects (with HAs or CI) using the CRA. On the other hand, such an improvement was not observed in NH listeners. It is, therefore, most likely that NH and HI subjects use different strategies in processing and recognizing noisy speech.

The CRA algorithm used in this study can be regarded as part of a family of speech enhancement techniques that use spectral subtraction. It is quite possible that other algorithms will yield similar results. However, a further research is required to explore this question.

We hypothesize that the different performance in recognition of the different groups is due to different strategies used by these groups, when processing the input from the cochlea. We assume that the strategy used by NH for speech recognition is optimal usage of speech representation redundancy. The speech can be recognized in different overlapping spectral ranges, and the decision is based on one or more frequency range. Long-term hearing-loss subjects using HAs have probably accustomed to the distorted signal produced by their HA which emphasized certain frequency ranges and ignored other frequencies where these subjects did not have residual hearing. CI users also have a hearing loss for a long term. The way the speech signal is presented to them perfectly fits the enhancement algorithm. Short-term hearing-loss subjects using HAs probably still try to use the strategies of NH which look at the whole frequency range of the speech signal. Regarding the CI users, as their ear is not functional; they are dependent on the strategy of their CI.

A possible strategy for NH may lie in the use of the redundancy in the speech signals along the frequency domain. A NH can understand different parts of the speech from various parts of the spectrum collected by his inner ear. Therefore, if one part of the spectrum is very noisy, he still can understand the speech from other parts of the speech that are less noisy. On the other hand, the HI perceives only part of the speech spectrum, and therefore cannot use the redundancy property. In order to understand what is being said, he concentrates on the partial available spectrum. A CRA algorithm may be of some benefit here by acting to diminish the noisy parts of the spectrum. The representation of the CRA (i.e, improved SNR) is "fed" to the HA, amplified and delivered to the cochlea where it is obstructed unfortunately by the residual hearing of the HA user, depending on the dead regions in his ear, not allowing the full bypass of the CRA representation further to the brain. Alternatively with the CI, the CRA representation is transferred as is to the nerve endings of the HI, bypassing the dead ear. This may explain why the CRA performs better for CI users than for HA users.

5. REFERENCES

[1] D.H. Wilson, P.G. Walsh, L. Sanchez, A.C. Davis, A.W. Taylor, G. Tucker, I. Meagher. "The epidemiology of hearing impairment in an Australian adult population". *Int J Epidemiol*. 1999; 28: 247-52

[2] B.H. Juang, C. Tsuhan. "The past, present, and future of speech processing". *IEEE Signal Processing Magazine*, 15: 24 - 48., 1998

[3] D.E. Tsoukalas, J.N. Mourjopoulos, G. Kokkinakis. "Speech enhancement based on audible noise suppression". *IEEE Trans. Speech Audio Process*. 5: 497–513, 1997.

[4] K.H. Arehart, J.H.L. Hansen, S. Gallant and L. Kalstein. "Evaluation of an auditory masked threshold noise suppression algorithm in normal-hearing and hearing-impaired listeners" *Speech communications* 40: 575-592, 2003.

[5] A.A. Montgomery, R.A. Edge, "Evaluation of two speech enhancement techniques to improve intelligibility for hearing-impaired adults". *J Speech Hear Res*. 31(3):386-393, 1988.

[6] M. Ross. "Why people won't wear hearing aids?" *International federation of hard of hearing people journal*, 1995.

[7] A. Cohen, M. Furst, "Integration of outer hair cell activity in a one-dimensional cochlear model". *J. Acoust. Soc. Am.*. 115, 5, 2185-2192, 2004.

[8] V. Weisz, A. Cohen, M. Furst. "Reconstruction of noisy speech signals based on a cochlear model". ARO 2003 Midwinter Meeting.

[9] N. Fink, M. Furst, T. Goshen, A. Cohen, V. Weiss. "Evaluation of a reconstruction algorithm for a comprehensive cochlear model – preliminary report". Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI), Granada, Spain. June 28 - July 3, 2004.

[10] A. Boothroyd. "Developments in speech audiometry". *British journal of audiometry* (formerly sound), 2, 3-10, 1968a.

[11] A. Boothroyd. "Statistical theories of the speech discrimination score". *J. Acoust. Soc. Am.*, 43, 362-367, 1968b.

[12] S.C. Purdy, B. Arlington, C. Johnstone, "Normative Data for the New Zealand Recording of the CVC (Revised AB) Word List". *New Zealand Audiological Society, Bulletin* 10(2):20-29, 2000.

[13] L. Kishon -Rabin, S. Patael M. Menahemi N. Amir. "Are the perceptual effects of spectral smearing influenced by speaker gender?" *Journal of Basic and Clinical Physiology and Pharmacology*, 15: 41-55, 2004.

[14] L.E. Humes, D.L. Wilson, N.N Barlow, C. Garner. "Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults". *Journal of Speech, Language, and Hearing Research*. Vol, 45, 772-782, 2002.