ISCA Archive
http://www.isca-speech.org/archive

Speech and Language Technology
for Disabled Persons
Stockholm, Sweden
May 31 - June 2, 1993

SPEECH INTELLIGIBILITY IMPROVEMENT FOR PEOPLE WITH COCHLEAR IMPLANTS

Wolfgang Helmut Döring
ENT-Department, University Hospital

Hans-Günter Hirsch
Institute for Communications Equipment and Data Processing

Aachen University of Technology, Templergraben 55, 5100 Aachen, Germany

ABSTRACT

With a cochlear implant deaf people are able to understand speech under good listening conditions. Problems occur in adverse conditions, e.g. in reverberant and/or noisy environments.

The speech intelligibility was examined in noisy situations for some persons using their implant system with and without the internal noise suppression. Experiments were carried out to improve the intelligibility of noisy speech using a single-channel noise suppression technique. This was realized by preprocessing i.e. by applying the resynthesized speech signal to the cochlear implant system.

COCHLEAR IMPLANT

Electronic hearing aids have been developed during the last twenty years for deaf people with a postlinguistic hearing loss. These are called cochlear implant systems. The persons involved in this study are using the intracochlear implant with 22 electrodes called "Nucleus 22" [1],[2]. Acoustic signals are recorded with a directional microphone placed directly above the ear. The signal is transmitted to a speech processing unit for analysis and coding. The extracted acoustic parameters are coded and inductively transmitted to a decoding and stimulation unit which is implanted above and behind the ear under the skin. The stimulation unit generates electrical pulses for an array of electrodes which is implanted in the inner ear. The electrical pulses lead to a stimulation of auditory nerve fibres. Selecting e.g. a specific pair of neighbouring electrodes, nerve fibers in the region of the electrodes position are stimulated corresponding to the well known frequency-to-place transformation in the inner ear.

The speech processing unit analyses the speech signal. Some spectral parameters and the pitch period are estimated. For voiced sounds two formant frequencies, the corresponding energy, the energy in the regions of 2000-2800 Hz and 2800-4000 Hz are calculated. The corresponding electrodes are stimulated with the rate of the estimated pitch period. For unvoiced sounds a random pitch and the energy above 4000 Hz instead of the first formant are used.

The speech processing unit offers an internal noise suppression mechanism. This is described as an adjustment of the threshold stimulation level dependent on the constant level of background noise [2]. It is comparable with a simple spectral subtraction in only two broad bands. These bands are used for the estimation of two formant frequencies in the frequency regions from 300 to 800 Hz and from 800 to 2300 Hz.

Because of the limited number of electrodes and because of the simple speech analysis technique these people have considerable problems to understand speech in adverse conditions, especially in the cases of reverberation and/or additive noise. The degradation of speech intelligibility was examined in noisy situations with and without the internal noise suppression of the implant system.

Furthermore we carried out some experiments with a noise suppression algorithm as a preprocessing of the noisy speech. The intelligibility was determined for a few people.

INTELLIGIBILITY IN NOISY SITUATIONS

The intelligibility of speech was examined for a couple of persons with cochlear implants in noisy situations. The intention was a determination of the intelligibility in a noisy background at all and a comparison of the results with the internal noise suppression turned on or off. Speech and noise were replayed in a nearly anechoic chamber via loudspeaker. Two-digit numbers were presented in these tests. The numbers are part of a German vocabulary which is often used for intelligibility tests. Some results are shown in figure 1 as an average of four persons with the actual version (MSP) of the speech processor.

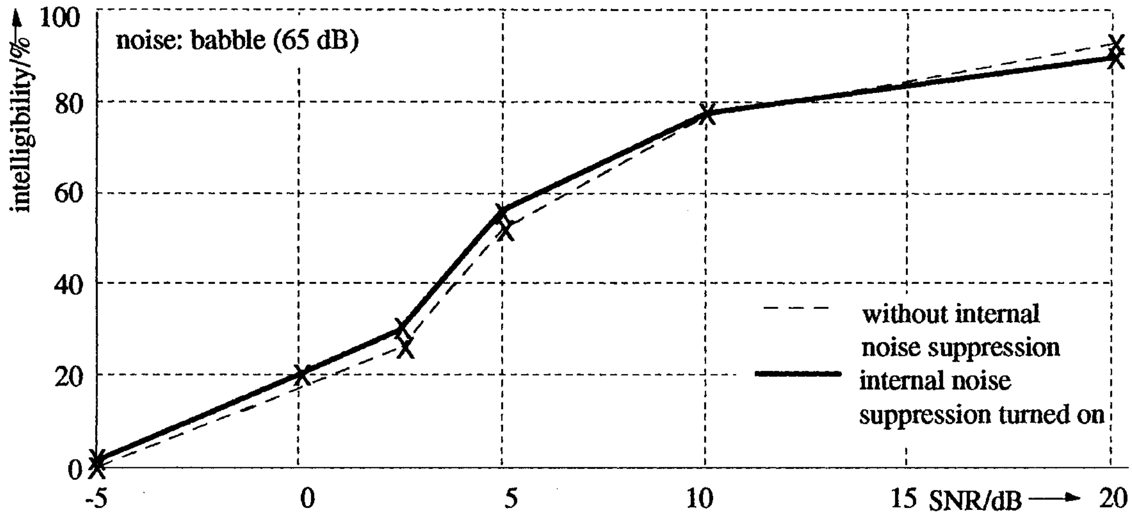

Figure 1. Intelligibility of noisy speech with and without the internal noise suppression

The speech as well as the noise were reproduced directly in front of the person at a distance of about 1,3 m. The noise consists of a babble which was artificially generated by a superposition of the whole test vocabulary of the same speaker. The noise level was fixed at 65 dB and the speech level was varied dependent on the desired signal-to-noise ratio.

Only a small gain can be observed for a certain range of the SNR comparing the results with and without internal noise suppression. Looking at the results more in detail persons show a very different behaviour. The intelligibility scores of two persons are shown in figure 2 where the speech was reproduced at an angle of 45 degree in relation to the front position on this head side where the implant's microphone is placed. The noise was generated at an angle of 45 degree on the opposite side at the same distance. An improvement can be observed for person WS using the internal noise suppression. Person MA exactly shows the opposite behaviour. One reason may be that some persons do not turn on the optional noise suppression at all so that they are not used to this kind of processing. Indeed it is a well known aspect that nearly everybody with an implant system needs at least some days to adapt to any modification of the processing algorithm.

This may be one of the reasons why person MA has a loss in intelligibility using the optional internal noise suppression and why the average results for the internal noise suppression in figure 1 are not as good as expected. But furthermore this result indicates

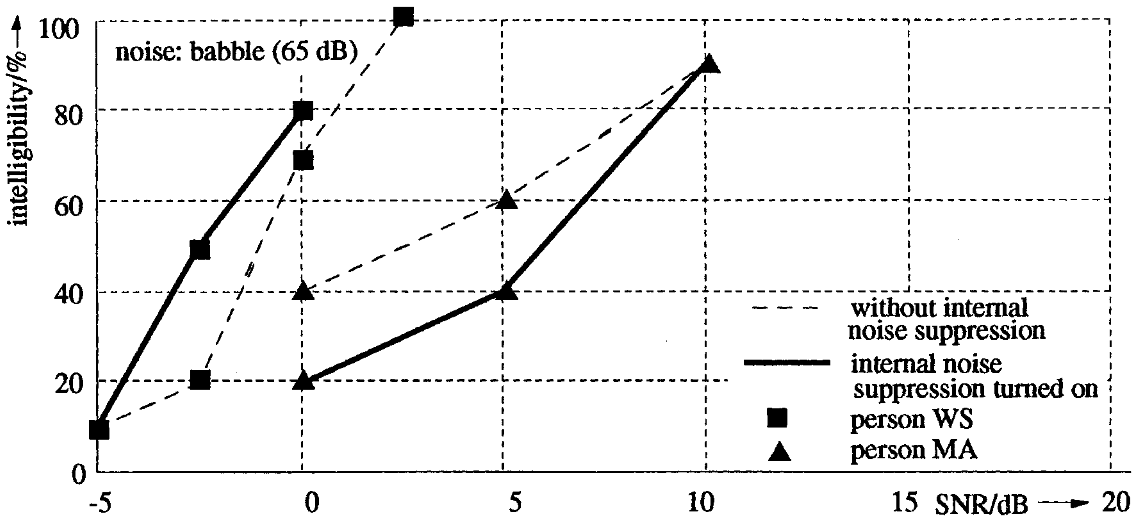

Figure 2. Intelligibility of noisy speech with and without the internal noise suppression

that the internal noise suppression algorithm does not only reduce a constant background noise. The processing also seems to modify the extracted features in segments of speech so that it is necessary for a person to adapt to this slightly modified processing. Because of this an external noise suppression was alternatively introduced applied as a preprocessing of the noisy speech.

EXTERNAL NOISE SUPPRESSION

Many algorithms exist for the preprocessing of speech to reduce the influence of noise and/or reverberation, e.g. [3]. A series of experiments was carried out using a slightly modified spectral subtraction technique[4], [5], [6]. The noise spectrum is estimated during speech pauses. An adaptive filtering is realized with the estimated noise spectrum during segments of speech. The Wiener filter is usually introduced to minimize the average squared error. It turned out that the best results were not obtained using a Wiener filter, i.e. an adaptive filtering based on the short-term power spectra, but using a weighting based on the magnitude of the spectral components [5].

The intelligibility was examined for some persons under the following test conditions. The digitized speech and noise signals can be mixed to realize different signal-to-noise ratios. The noisy signal is processed with the noise reduction algorithm. The speech is resynthesized with the overlap add method from the processed spectra. The clean speech, the noisy speech, as well as the processed noisy speech can be fed to the external audio input of a person's cochlear implant system. Again the already mentioned two-digit numbers were used for the experiments. The numbers are randomly presented twice. Some results are given in figure 3.

Seven persons with an elder version (WSP) of the speech processor were involved in this experiment who were deaf for a period from 1 year to a couple of years. The intelligibility score of the clean speech was about 80 %.

If no noise suppression is applied, the degradation of the intelligibility is significant. Two noise signals were considered, a nearly stationary car noise and a speech babble in a cafeteria, both recorded by ourselves. The car noise was added at signal-to-noise ratios of 0 and 3 dB, the babble at a SNR of 0 dB.
As expected, the decrease of intelligibility was much higher for the speech like noise. The intelligibility can be improved by processing the noisy speech with the spectral subtraction technique.

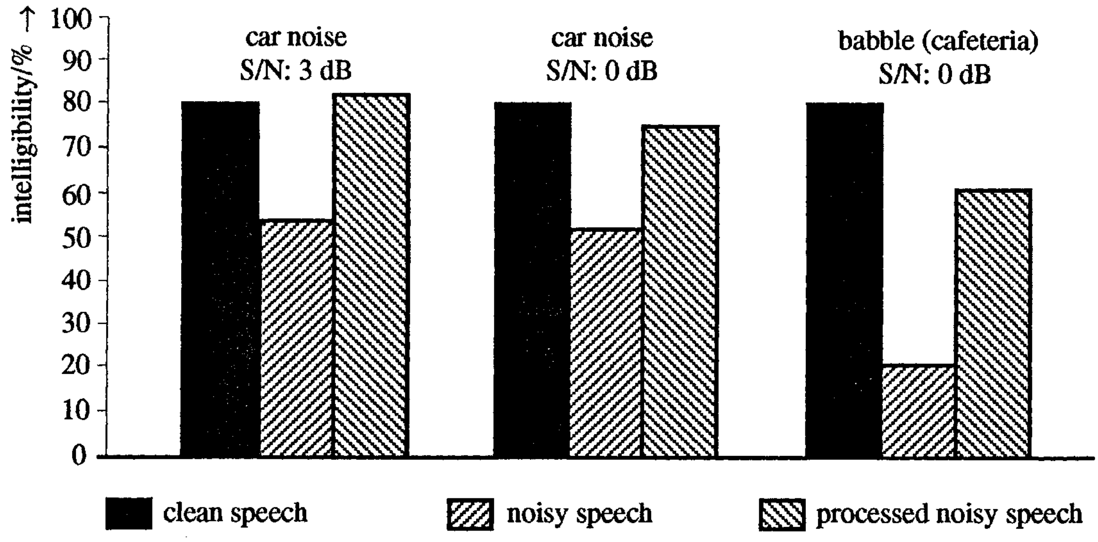

Figure 3. Speech intelligibility for different noise situations

Another small experiment was carried out with two persons with the actual version (MSP) of the speech processor. This allowed a comparison of the internal noise suppression function of the speech processor with the application of the external spectral subtraction method. Some intelligibility rates are presented in the following table.

	car noise (0 dB)	babble (0 dB)
internal noise suppression	65 %	48 %
spectral subtraction	92 %	75 %

It turns out that the improvement is much higher with the application of the external noise reduction scheme in comparison to the internal processing. This could be expected considering the simple internal processing in only two broad bands.

CONCLUSIONS

The loss of intelligibility in noisy situations was determined in some listening experiments with people with cochlear implants. No remarkable difference could be observed with or without the internal noise suppression of the implant system. Improvements could be shown applying an external preprocessing with a spectral subtraction technique. Further investigations will be carried out to improve the intelligibility for these persons in real environments.

REFERENCES

1. G.M. Clark et al.: Cochlear Prostheses, Curchill Livingstone, 1990
2. Cochlear Pty. Ltd.: Mini System 22 - Audiologist Handbook, 1989
3. J.S. Lim: Speech Enhancement, Prentice-Hall, 1983
4. S.F. Boll: Suppression of Acoustic Noise in Speech Using Spectral Subtraction, IEEE ASSP-27, No. 2, 1979, pp. 113-120
5. P. Vary: Noise Suppression by Spectral Magnitude Estimation - Mechanism and Theoretical Limits, Signal Processing, 1985, pp. 387-400
6. H.G. Hirsch: Intelligibility Improvement of Noisy Speech for People with Cochlear Implants, Proc. of the ESCA-workshop "Speech Processing in Adverse Conditions", Cannes, 1992, pp. 69-72