nederlands akoestisch genootschap
NAG

journaal

nr. 178
maart 2006

Speech understanding in background noise with the two-microphone adaptive beamformer BEAM^{TM} in the Nucleus Freedom^{TM} cochlear implant system *

Lieselot Van Deun[1], Ann Spriet[1,2], Kyriaky Eftaxiadis[3], Johan Laneau[1], Marc Moonen[2], Bas van Dijk[4], Astrid van Wieringen[1], Jan Wouters[1]

[1] ExpORL, Dept. Neurosciences, K.U.Leuven, Herestraat 49, bus 721, 3000 Leuven, Belgium
[2] ESAT/SCD-SISTA, K.U.Leuven, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium
[3] Cochlear Ltd., 14 Mars Road, PO Box 629, Lane Cove NSW 2066, Australia
[4] Cochlear CTCE, Schaliënhoevedreef 20, Building i-B, 2800 Mechelen, Belgium

Abstract

A double-blind evaluation of the two-microphone adaptive beamformer BEAM^{TM} and a hardware directional microphone, in the Nucleus Freedom^{TM} cochlear implant (CI) system, has been carried out with five adult Nucleus CI users. The test procedure consisted of a pre- and post-test in the lab and a two-week trial at home. In the pre- and post-test, the speech reception threshold (SRT) with sentences and the percentage correct phoneme scores for CVC words were measured in quiet and background noise at different signal-to-noise ratios. Performance was assessed for two different noise configurations (with a single noise source and with three noise sources) and two different noise materials (stationary speech weighted noise and multi-talker babble). During the two-week trial at home, the CI users evaluated the noise reduction performance in different listening conditions by means of the SSQ questionnaire. In addition to the perceptual evaluation, the noise reduction performance of the beamformer was measured physically as a function of the direction of the noise source.
Significant improvements of both the SRT in noise (average improvement of 5-16 dB) and the percentage correct phoneme scores (average improvement of 10-41%) were observed with BEAM^{TM} compared to the standard hardware directional microphone. In addition, the SSQ questionnaire and subjective evaluation in controlled and real-life scenarios suggested a possible preference for the beamformer in noisy environments. The evaluation demonstrates that the adaptive noise reduction algorithm BEAM^{TM} in the Nucleus Freedom^{TM} CI-system may significantly increase the speech perception by cochlear implantees in noisy listening conditions. This is the first monolateral (adaptive) noise reduction strategy actually implemented in a mainstream commercial CI.

1. Introduction

Although most cochlear implant (CI) users today achieve remarkably good speech understandig in quiet, they generally experience severe performance degradation in noisy acoustical environments. While people with normal hearing still succeed to understand 50 % of the speech in a noisy environment with a signal-to-noise ratio (SNR) as low as -5 dB [1, 2, 3], cochlear implantees have a speech reception threshold (SRT) between 5 dB and 15 dB and hence, require an SNR that is 10 dB to 25 dB higher [4, 5, 6]. An effective way to improve the ability to understand speech in background noise is to incorporate a noise reduction algorithm as a pre-processor to the CI's speech processor. In the last decade a number of noise reduction strategies have been evaluated, based on bilateral [7, 8, 9, 10] or monaural devices [6]. Until now the application of noise reduction algorithms in commercial CIs was

mainly limited to the use of a hardware directional microphone. In 1997, a bilateral noise reduction system, called the Audallion BEAMformer, was marketed by Cochlear Ltd. as a preprocessing device (consisting of a digital signal processor) for the Nucleus-22 CI system [11]. In 2005, the monaural two-microphone adaptive beamformer of Wouters and Vanden Berghe [6], referred to as BEAM, was implemented in the BTE speech processor of Cochlear's Nucleus Freedom CI system. In this paper, we present the results of a double-blind evaluation of the two-microphone adaptive beamformer BEAM^{TM} and the standardly used hardware directional microphone with five adult Nucleus CI users. The speech reception performance of the subjects has been assessed in different noise scenarios. Furthermore, the noise reduction programs have been subjectively evaluated by the subjects in real-life situations.

2. Experiment

2.1 Adaptive noise reduction algorithm beam

Figure 1 depicts the two-microphone adaptive beamformer BEAM^{TM} that was evaluated in this study. The beamformer combines a directional microphone (i.e., the front microphone on the BTE) and an omnidirectional microphone (i.e., the rear microphone). The algorithm is based on the Generalized Side-lobe Canceller of Griffiths and Jim [12] and consists of a fixed spatial preprocessor and an adaptive noise cancellation (ANC) stage. The spatial preprocessor uses a fixed FIR filter to create a speech reference and a noise reference. The ANC attenuates the residual noise in the speech reference. To limit speech distortion, the ANC is adapted during periods of noise-only [13, 14, 15, 16]. The speech detection algorithm described in Vanden Berghe and Wouters [17] is used to determine periods of noise-only.

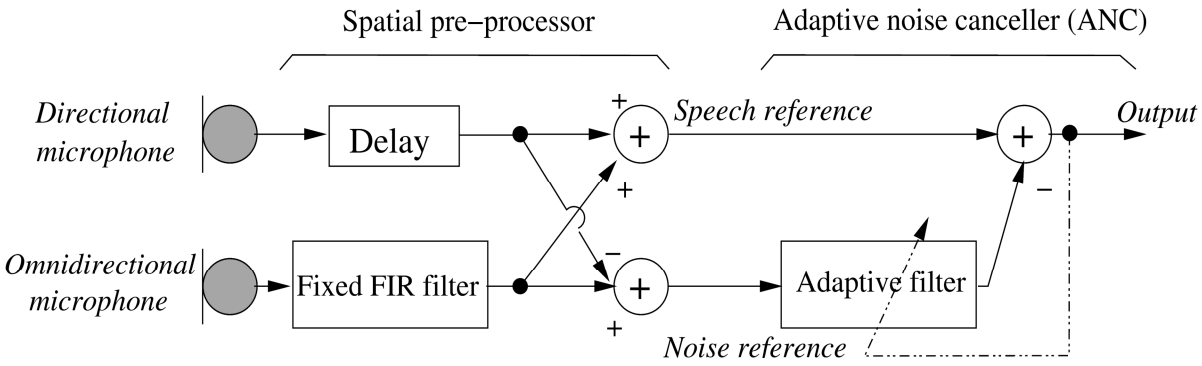


Figure 1: Two-microphone adaptive beamformer BEAM^{TM} .

2.2 Physical evaluation

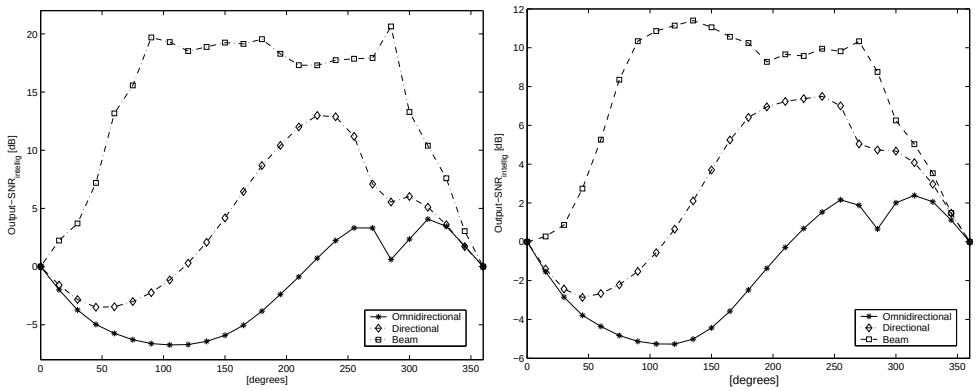
The physical performance of noise reduction algorithms is assessed in terms of the output intelligibility weighted SNR [18], defined as:

$$\text{output-SNR}_{\text{intellig}} = \sum_i I_i(output - SNR_i),$$

where output-SNR$_i$ is the output SNR (in dB) in the i-th one third octave band. The band importance for intelligibility function I_i expresses the importance of the i-th one-third octave

band with center frequency f_c^i. The center frequencies f_c^i and the values I_i are defined in ANSI S3.5-1997 [19].

The output-$SNR_{intellig}$ of the hardware directional microphone and the beamformer were measured with two different processors, mounted on the right ear of a Cortex MK2 artifical head. Figure 2 depicts the outcome for an anechoic room (fig 2a) and for an office room (fig 2b) as a function of the angle θ of the noise source. As a reference, also the output-$SNR_{intellig}$ of the omnidirectional microphone is depicted.



a) Anechoic room b) Office room

Figure 2: Average output-$SNR_{intellig}$ of the omnidirectional (*) and directional microphone (◊) and the two-microphone beamformer BEAM^{TM} (□) as a function of the angle θ of the noise source. a) Anechoic room; b) Office room with T60 = 0.76 s.

2.3 Perceptual and subjective evaluation

2.3.1 Subjects

Five adult Nucleus CI users, one female and four male subjects, participated in the evaluation. As an inclusion criterion subjects had to reach 50-60% of speech understanding in quiet with the LIST sentences [20, 21] in order to achieve an adaptive SRT in noise.

Three Freedom speech processors were used for the evaluation. The speech processors were programmed according to the clinical fitting data of each patient. All subjects use the ACE (Advanced Combination Encoders) speech strategy [22]. The volume of the speech processor (value between 0 and 9) was adjusted to a comfortable loudness. The sensitivity was set to its default, i.e., 12. Two noise reduction programs were provided, namely, the hardware directional microphone (i.e., the directional microphone signal in Figure 1) and the two-microphone adaptive beamformer BEAM^{TM}. The order of the programs was randomized over the different subjects and only known to the researchers carrying out the evaluations and the CI users after completion of the complete test procedure (double-blind evaluation).

2.3.2 Setup of the lab experiments

The speech understanding tests in the lab were performed in a standard office room. The average reverberation time T60 of the room, measured with a speech weighted noise signal, equals 0.52 sec. The estimated intelligibility-weighted direct-to-reverberant ratio was 7 dB at 1 meter. The desired speech signal and noise signal(s) were presented through identical loudspeakers (Yamaha CBX-S3) positioned at a distance of 1 meter from the patient's head:

the speech source in front of the head, the noise sources at an angle θ with respect to the speech source. The loudspeakers were calibrated separately such that they produced the same sound level for a speech weighted noise signal at a reference point corresponding to the center of the patient's head. Two noise configurations were considered: a single noise source at 90° (at the side of the patient's CI) and three noise sources at 90°, 180° and 270°. The noise signals in the tripple noise source scenario were uncorrelated and all had the same sound level at the reference point. Presentation of the speech and noise materials was controlled from an adjacent room by means of two CD players and two audiometers. The signals, produced by two CD players, were amplified through a MADSEN OB822 audiometer (the speech signal and noise signal at 90°) and an AMPLAID 309 audiometer (the noise signals at 180° and 270°) before being sent to the loudspeaker. The maximum amplification of the audiometers corresponds to a sound pressure level of 90 dB SPL at the reference point in the test room.

2.3.3 Test protocol

Each test trial consisted of a pre- and post-test in the lab and a two-week trial at home. Between the pre- and the post-test, both noise reduction programs were compared by the CI users in different scenarios during a two-week field test at home and evaluated by means of the SSQ questionnaire [23].

During the pre- and the post-test, SRTs of the CI users for sentences were measured adaptively and percentage correct phoneme scores for CVC-words were assessed. Both tasks were carried out in quiet and in background noise, with the two noise reduction programs. LIST sentences were used as speech material for the adaptive SRT test. The LIST consists of 35 sublists of 10 Dutch/Flemish sentences spoken by a female speaker [20, 21]. The CVC words were part of the Flemish recordings of the NVA list [3]. The NVA list consists of sublists of 12 Dutch/Flemish 3-phoneme monosyllables, produced by a male speaker. For each condition, i.e., program and noise scenario, one sublist was selected. Different sublists were used during the pre-and the post-test. Two different noise materials were used, i.e., stationary speech weighted noise and non-stationary multi-talker babble noise. The speech weighted noise had the same long-term spectrum as the presented speech material. For the adaptive measurement of the SRT, the speech weighted noise on the CD by van Wieringen, Wouters [21] was used. For the measurement of the percentage correct phoneme scores, the steady NVA speech-weighted noise was used. The multi-talker babble noise was taken from the CD 'Auditory Tests (Revised)' by Auditec [24].

In the SRT test the speech level was varied and the noise level was kept constant, at either 55 dB SPL or 65 dB SPL. Recognition of the NVA-words was evaluated at two SNRs, i.e., -5 dB and + 5dB, at a noise level of 60 dB SPL. In both procedures, the performance was assessed for two different noise configurations (i.e., a single noise source and three noise sources) and the two noise materials (i.e., speech weighted noise and multi-talker babble).

2.3.4 Perceptual evaluation: results and discussion

2.3.4.1 Reliability

The reliability of all test results has been assessed by comparing pre- and posttest results. A repeated measures analysis of variance (ANOVA) revealed that results of the two test moments did not differ significantly, with F-values of 0.358 and 0.04 and corresponding p-values of 0.582 and 0.856 for the SRTs and the percentage correct phoneme scores, respectively.

2.3.4.2 SRTs for LIST sentences

The mean SRT in quiet corresponded to 51.4 dB SPL and 50.5 dB SPL for the beamformer and the hardware directional microphone, respectively. Although the difference between the two programs was insignificant (ANOVA, F = 0.625, p = 0.473), for all but one subject the speech reception threshold seemed to be worse in the beamformer condition. An explanation for the above, unexpected observation could be a suboptimal first stage of the beamformer due to potential microphone mismatch and differences in the mechanics of the prototype speech processors and the Freedom speech processors used in the evaluation. In any event, the higher SRTs do not need to affect speech understanding in normal conditions as they can be compensated by slightly changing the sensitivity control.

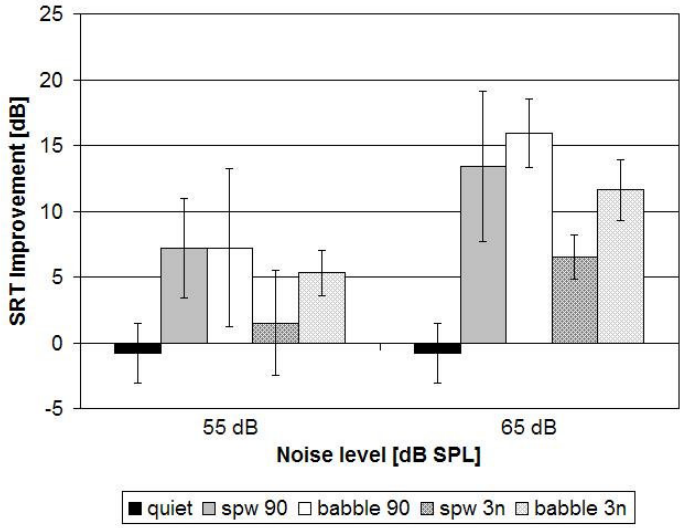


Figure 3: Average benefit in SRT [dB] of the two-microphone beamformer BEAM^{TM} compared to the hardware directional microphone in quiet and in the different noise scenarios, i.e., speech weighted noise at 90° (spw 90), babble noise at 90° (babble 90), speech weighted noise at 90°, 180°, 270° (spw 3n) and babble noise at 90°, 180°, 270° (babble 3n).

For all subjects, the SRT in noise was generally improved by the beamformer. The average benefit in SRT obtained with the two-microphone beamformer BEAM^{TM} compared to the hardware directional microphone is visualized in figure 3. The left part of the graph shows the results for the noise level of 55 dB SPL, while the results for 65 dB SPL can be found at the right-hand side of the figure. Each bar represents one condition (one noise material in one noise scenario). Error bars depict the standard deviation. As can be seen in figure 3 the beamformer improved the SRT in noise on average by 2 to 17 dB compared to the hardware directional microphone. A benefit is observed in all noise conditions, with the largest one in the condition with a single noise source at 90° and a level of 65 dB SPL (average improvement of 13.4 dB for speech weighted noise and 15.9 dB for babble noise). A repeated measures analysis of variance (ANOVA) proved the difference between two noise reduction algorithms to be significant ($F = 130.141$, $p < 0.001$).

The effect of other parameters on the speech recognition in general was also investigated. Babble noise is known to be more disturbing than stationary noise and indeed lead to significantly worse results ($F = 264.275$, $p < 0.001$). The level of the noise did not have a significant influence on the speech reception thresholds ($F = 1.011$, $p = 0.371$). Although one could expect the three noise sources at 90°/180°/270° to be less disturbing than the single

noise source at 90° because of head shadow, speech understanding did not seem to depend on the noise configuration (F = 0.006, p = 0.942).

To give a more detailed picture of the effect of the noise reduction algorithm, the influences of the noise parameters on the beamformer benefit are discussed. The benefit was significantly larger in babble noise than in speech weighted noise (F = 12.121, p = 0.025), which might be due to the larger disturbing effect of babble noise on speech understanding in general. Since the scores at a noise level of 55 dB SPL with the hardware directional microphone were already close to the scores in quiet, the amount of improvement that could be obtained at this noise level was smaller than at the noise level of 65 dB SPL (F = 33.188, p = 0.005). Finally, the beamformer benefit was significantly larger for a single noise source at 90° than for triple noise sources at 90°/180°/270° (F = 14.810, p = 0.018). This can be explained by the fact that because of head shadow, the total noise level reaching the microphones is lower in the tripple-noise scenario than in the single-noise scenario. Moreover, the two-microphone beamformer can better suppress a single noise source at 90° than multiple noise sources.

2.3.4.3 Percentage correct phoneme scores for NVA words (CVC)

With a mean percentage correct score of 71 % for the beamformer and 72 % for the directional microphone, the noise reduction processing again did not lead to a significant difference in speech understanding in quiet (F = 0.002, p = 0.962).

In noise, the two-microphone beamformer offered a large improvement compared to the hardware directional microphone (F = 192.243, p < 0.001). Figure 4 depicts the average benefit (+/- one standard deviation) of all subjects, of the two-microphone beamformer compared to the hardware directional microphone, in percentage correct phoneme scores in noise, for the different SNRs (-5 dB or +5 dB), noise materials (speech weighted noise or babble noise) and configurations (noise at 90° or at 90/180/270°). The beamformer resulted in an average improvement in percentage correct phoneme scores between 3.1% and 41% compared to the hardware directional microphone.

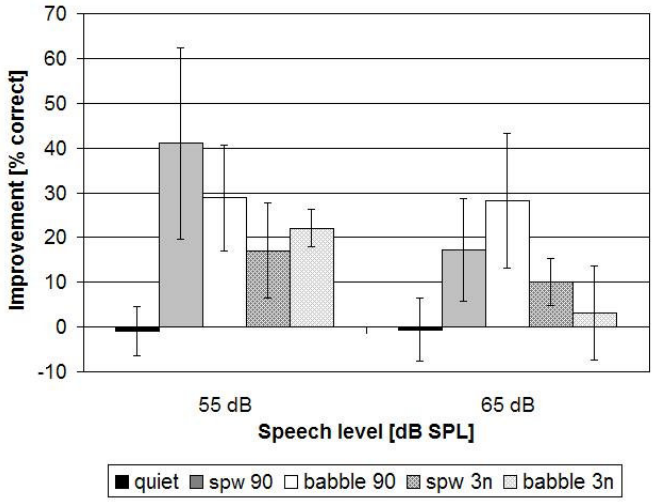


Figure 4: Average benefit in percentage correct phoneme scores of the beamformer relative to the directional microphone in quiet and in the different noise scenarios, i.e., speech weighted noise at 90° (spw 90) , babble noise at 90° (babble 90), speech weighted noise at 90°, 180°, 270° (spw 3n) and babble noise at 90°, 180°,270° (babble 3n).

Logically, speech recognition scores were better at a higher SNR ($F = 82.195$, $p = 0.001$). Furthermore, same observations were made as in the adaptive measurements: babble noise lead to significantly worse speech intelligibility ($F = 202.123$, $p < 0.001$) but a different noise configuration (90° or 90°/180°/270°) did not give rise to significantly different percentages ($F = 0.75$, $p = 0.435$).

The benefit of the beamformer, compared to the hardware directional microphone, was smaller at a signal-to-noise ratio of +5 dB than at a signal-to-noise ratio of -5 dB ($F = 9.012$, $p = 0.04$). The same explanation as for the adaptive measurements applies here: at +5 dB the scores with the hardware directional microphone were already close to those obtained in quiet, which left only little room for improvement by the beamformer. Following noise parameters did not significantly influence the size of the beamformer benefit measured with the CVC words: babble noise versus speech weighted noise ($F = 0.04$, $p = 0.851$) and single noise source at 90° versus triple noise source at 90°/180°/270° ($F = 7.316$, $p = 0.054$).

2.3.5 Subjective evaluation: results and discussion

A small preference for the beamformer was seen on some questions of the SSQ questionnaire. Most of the questions on which more than one subject preferred the beamformer above the hardware directional microphone were related to the ability to understand speech in noise (i.e., understanding speech in the presence of interfering speakers, car noise or a television). The SSQ results were less convincing than the perceptual scores, which is partly due to the fact that not all questions are adequate in revealing a difference between noise reduction programs: only a few questions focus on the ability to understand speech in noise. Besides that, the listening conditions that subjects encountered in daily life during the two-week trial at home were often not as noisy as the ones created in the lab. Some of the subjects even spent most of the time in quiet surroundings. This may lead to an underestimation of the beamformer's effect by the questionnaire.

When subjects had to judge their speech understanding on an SSQ scale in the lab setup in the different noise scenarios, the preference for the beamformer became more clear: the average difference between the hardware directional microphone and the beamformer was between 2 and 3.6 units for the different conditions.

A more controlled evaluation of the beamformer performance in real-life was carried out by taking three subjects outside to some noisy situations: bus, bus station, pub and street. The mean benefit in rated speech understanding obtained with the beamformer went from 0.25 to 1.17 on a score out of ten. The most important conclusion to be drawn from this test session is that also in real-life noisy conditions the subjective benefit from the beamformer, indicated on a score out of ten, was smaller than the benefit in controlled laboratory conditions. Possible explanations are that the real-life listening conditions were less noisy and/or more reverberant than the laboratory conditions and that the CI subjects used lip-reading. Note that the absolute scores for speech understanding with the directional microphone were lower in the lab (on average 3 out of 10) than in real-life (on average 7 out of 10). Hence, there was less room for improvement by the beamformer in the real-life test conditions.

The SSQ and subjective evaluation suggest that CI subjects might prefer the beamformer in real-life noisy listening conditions. Given the small group of subjects in this subjective evaluations, this preference may not be generalized to all CI users.

Conclusions

In this study, the two-microphone adaptive beamformer BEAM^{TM} in the new Nucleus Freedom^{TM} CI system has been evaluated through a double-blind evaluation with five CI

users. This is the first implementation of a monolateral (adaptive) noise reduction strategy in a mainstream commercial CI. Speech understanding tests in the lab demonstrated significant improvements of both the SRT in noise (average improvement of 5-16 dB) and the percentage correct phoneme scores (average improvement of 10-41%) with BEAM^{TM} compared to the standard hardware directional microphone in all tested conditions. The largest benefit is obtained for the noise level of 65 dB SPL and the single noise source at 90°. Self-report evaluations in controlled and real-life scenarios suggested a possible preference for the beamformer in noisy environments. This presentation is a summary of the research described in reference [25].

Acknowledgements

Lieselot Van Deun is a research assistant funded by the F.W.O.-Vlaanderen. Ann Spriet is a postdoctoral researcher funded by the F.W.O.-Vlaanderen. The authors would like to thank the CI subjects for their cooperation. We also thank Jean-Baptiste Maj for his help with the evaluation of BEAM^{TM} in the Nucleus Freedom^{TM} system. This research work was carried out at Lab. Exp. ORL of the Katholieke Universiteit Leuven, in the frame of IWT project 020540 ('Innovative Speech Processing Algorithms for Improved Performance of Cochlear Implants'), Research Project FWO nr. G.0233.01 ('Signal processing and automatic patient fitting for advanced auditory prostheses'), the Concerted Research Action GOA-AMBioRICS and was partially sponsored by Cochlear.

Literatuur

1. Plomp R. and Mimpen A. M., 1979, "Speech-reception threshold for sentences as a function of age and noise level", Journal of the Acoustical Society of America, 66(5),1333–1342.
2. Versfeld N. J., Daalder L., Festen J. M., and Houtgast T., 2000, "Method for the selection of sentence materials for efficient measurement of the speech reception threshold", Journal of the Acoustical Society of America, 107(3),1671–1684.
3. Wouters J., Damman W., and Bosman A. J., 1994, "Vlaamse opname van woordenlijsten van spraakaudiometrie", Logopedie, 7(6),28–33.
4. Hochberg I., Boothroyd A., Weiss M., and Hellman S., 1992, "Effects of noise and noise suppression on speech perception by cochlear implant users", Ear and Hearing, 13(4),263–271.
5. Parkinson A. J., Parkinson W., Tyler R. S., Lowder M. W., and Gantz B. J., 1998, "Speech perception performance in experienced cochlear-implant patients receiving the SPEAK processing strategy in the Nucleus Spectra-22 cochlear implant", J. of Speech, Language, and Hearing Research, 41(5),1073–1087.
6. Wouters J. and Vanden Berghe J., 2001, "Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system", Ear and Hearing, 22(5),420–430.
7. Weiss M., 1993, "Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor", J. Rehab. Res. Develop., 30(1),117–128.
8. Van Hoesel R. J. M. and Clark G. M., 1995, "Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients", Journal of the Acoustical Society of America, 97(4), 2498–2503.
9. Hamacher V., Doering W. H., Mauer G., Fleischmann H., and Hennecke J., 1997, "Evaluation of noise reduction systems for cochlear implant users in different acoustic environments", American Journal of Otology, 18(6 (Suppl)),46–49.
10. Kompis M., Feuz P., Valentini G., and Pelizzone M., 1999, "A four microphone noise reduction system for cochlear implants", In Conference on Implantable Auditory Prostheses, page 125, Asilomar, USA.
11. Cochlear, Inc., 1997, "Introducing the Audallion BEAMformer Digital Noise Reduction System. Cochlear", Clinical Bulletin, April, 1–5.
12. Griffiths L. J. and Jim C. W., 1982, "An alternative approach to linearly constrained adaptive beamforming", IEEE Trans. Antennas Propag., 30(1),27–34.
13. Greenberg J. E. and Zurek P. M., 1992, "Evaluation of an adaptive beamforming method for hearing aids.", Journal of the Acoustical Society of America, 91(3),1662–1676.

14. Hoshuyama O., Sugiyama A., and Hirano A., 1999, “A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters”, IEEE Trans. Signal Processing, 47,2677–2683.
15. Van Compernolle D., 1990, “Switching adaptive filters for enhancing noisy and reverberant speech from microphone array recordings”, In Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), volume 2, pages 833–836, Albuquerque.
16. Herbordt W. and Kellermann W., 2002, “Frequency-domain integration of acoustic echo cancellation and a Generalized Sidelobe Canceller with improved robustness”, European Transactions on Telecommunica-tions, 13(2),123–132.
17. Vanden Berghe J. and Wouters J., 1998, “An adaptive noise canceller for hearing aids using two nearby microphones”, Journal of the Acoustical Society of America, 103,3621–3626.
18. Greenberg J. E., Peterson P. M., and Zurek P. M., 1993, “Intelligibility-weighted measures of speech-to-interference ratio and speech system performance”, Journal of the Acoustical Society of America, 94(5), 3009–3010.
19. Acoustical Society of America, 1997, “ANSI S3.5-1997”, American National Standard Methods for calculation of the speech intelligibility index.
20. van Wieringen A. and Wouters J., 2006, “LIST and LINT: Dutch speech audiometry lists with sentences and numbers”, International Journal of Audiology in review.
21. van Wieringen A., and Wouters J., 2005, “LIST en LINT: Nederlandstalige spraakaudiometrielijsten met zinnen en getallen”, CD and booklet by Lab.Exp.ORL-NKO, K.U. Leuven.
22. Vandali A. E., Whitford L. A., Plant K. L., and Clark G. M., 2000, “Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 Cochlear implant system”, Ear and Hearing, 21(6), 608–624.
23. Gatehouse S. and Noble W., 2004, “The Speech, Spatial and Qualities of Hearing Scale (SSQ)”, International Journal of Audiology, 43,85–99.
24. Auditec, 1997, “Auditory Tests (Revised)”, Compact Disc, Auditec, St. Louis.
25. Spriet A., Van Deun L., Eftaxiadis K., Laneau J., Moonen M., van Dijk B., van Wieringen A., Wouters J., 2006, “Speech understanding in background noise with the two-microphone adaptive beamformer BEAM^{TM} in the Nucleus Freedom^{TM} cochlear implant system”, Ear and Hearing, in press.