Simplifying cochlear implant speech processor fitting

Christina Willeboer

Publication of this thesis was financially supported by:
Makker Hoortoestellen
EmiD, MED-EL, Oticon, Schoonenberg Hoorcomfort, Stichting Atze Spoor Fonds, and Veenhuis Medical Audio BV.

ISBN: 978-90-393-4794-2

Layout and cover design: Tom Willemyns
Cover photo: Krista Willeboer

Printed by: Universal Press, Veenendaal

Simplifying cochlear implant speech processor fitting

Vereenvoudiging van de afregeling van de spraakprocessor van het cochleair implantaat
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van
doctor aan de Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het
college voor promoties
in het openbaar te verdedigen op
dinsdag 27 mei 2008 des middags te 12.45 uur

door

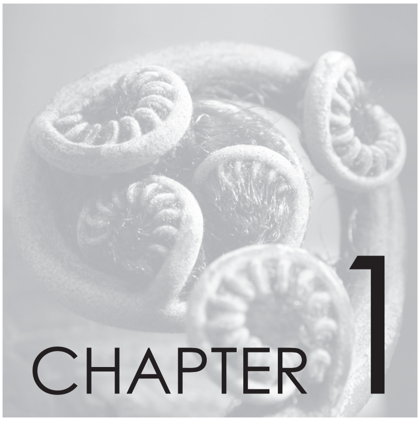
Christina Willeboer

geboren op 29 december 1977 te Arnhem

Promotor: Prof.dr. G.F. Smoorenburg
Co-promotor: Dr. G.A. van Zanten

Dit proefschrift werd mede mogelijk gemaakt met financiële steun van Cochlear.

Contents



Introduction

1.1 Cochlear implants

A cochlear implant (CI) can partially restore hearing by electrical stimulation of the auditory nerve in patients with severe hearing loss and deafness, who have limited or no benefit from conventional hearing aids. With over 150,000 users worldwide CIs are the second most applied neuroprostheses, after cardiac pacemakers.

The first demonstration that electrical stimulation, either using an intra- or extracochlear electrode, could result in an auditory sensation in a profoundly deaf person was given very early in the history of neuroprostheses (Gisselsson, 1950; Djourno and Eyriès, 1957). Initially, various designs and placements of auditory prostheses were evaluated. The designs were either single channel (House and Urban, 1973; House, 1976; Hochmair Desoyer et al., 1980) or multichannel (Eddington et al., 1978). Placements were extracochlear, e.g. in the mastoid or tympanic cavity with the electrode placed on the promontory (Douek et al., 1977), or intracochlear in the scala tympani (Clark et al., 1975).

Today's CI systems consist of a multichannel electrode array placed in the scala tympani of the cochlea. Recent designs with an electrode positioner or pre-curled electrode shape achieve a ‘modiolus-hugging’ placement, i.e. as close as possible to the modiolus, which incorporates the afferent endings of the spiral ganglion cells. The Nucleus 24 CI system (Cochlear, Lane Cove Australia) consists of 22 intracochlear electrodes and two extracochlear reference electrodes: a ball electrode (MP1) placed beneath the temporalis muscle and a plate electrode (MP2) on the metal housing of the implant. Due to the tonotopic organization of the cochlea, with high frequencies encoded at the basal end and low frequencies at the apical end of the cochlea, the place of stimulation determines the pitch perceived. Electrode 22 is the most apical, and electrode 1 is the most basal contact. However, the effect of the electrode arrangement on the perceived pitch depends on the amount of nerve endings that are intact. This varies strongly between CI recipients. The external parts of the CI system consist of a microphone to receive incoming sound, a speech processor that analyzes sound and determines the specific stimulus parameters, and a transmitting coil that sends the information via transcutaneous radio-frequency (RF) coupling to the implant (Fig. 1).

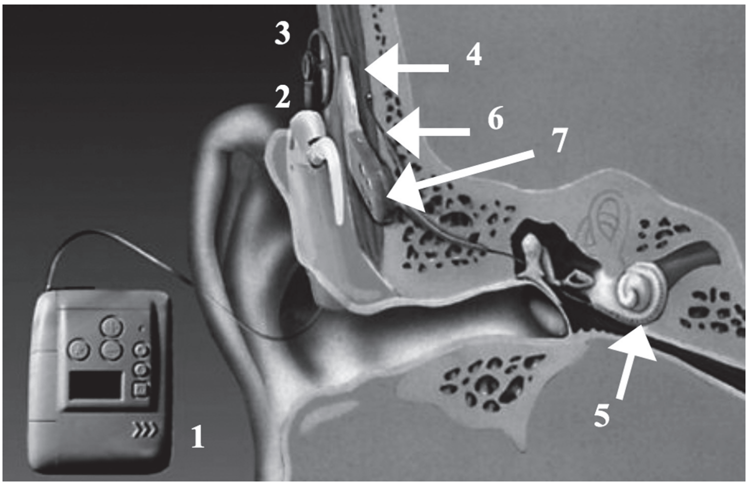


Figure 1: External and internal components of the cochlear implant (reprinted with permission from Cochlear Company). Externally are (1) the speech processor with (2) the microphone behind the ear and (3) the external coil. The implant consists of (4) an internal coil, (5) the electrode array placed in the cochlea, (6) the reference electrode MP1 placed beneath the temporal muscle, and (7) the reference electrode MP2 on the metal housing of the implant.

1.2 Fitting the Nucleus 24 multichannel cochlear implant system

The speech processor's procedure to analyze incoming sound and code it into stimuli for each of the electrodes is dependent on the speech coding strategy and other clinical fitting parameters chosen by the audiologist. As every cochlear implant recipient requires unique and specific electrical stimulation parameters to generate optimal auditory sensations, the CI system is programmed for each recipient individually.

1.2.1 Speech coding strategies

Three speech coding strategies are available in the Nucleus 24 CI system: spectral peak (SPEAK), continuous interleaved sampling (CIS), and advanced combination encoder (ACE).

The SPEAK coding strategy is the oldest of the three and was introduced in 1994 (Skinner et al., 1994; Seligman and McDermott, 1995). It uses a spectral maxima selection process to stimulate an individualized fixed number (M), usually between 6 and 10, of the electrodes per analysis cycle. The acoustic signal from the microphone of the speech processor is band-pass filtered into a number of channels corresponding to the number of active electrodes. In case of a full insertion of the electrode array, SPEAK uses 20 active electrodes, and hence there are 20 filters. For each analyzed time frame of the audio signal, M electrodes coupled to the subset of filters with the largest envelope output amplitudes are selected. The selected electrodes are sequentially stimulated to minimize interference between electrodes and thus uncontrolled summation effects. Biphasic pulses are used to prevent accumulation of charge. The pulse repetition rate is 250 Hz, however, a jitter is applied to prevent the nerve fibers from firing exactly at 250 Hz and thereby introducing the percept of a pure tone of 250 Hz.

The CIS coding strategy for the Nucleus implant was introduced in 1997 (Wilson et al., 1995; Kiefer et al., 2001). It was developed to avoid channel interaction by activating all available electrodes continuously but interleaved with biphasic pulses. It makes use of higher stimulation rates, of up to 2400 Hz per channel, and a relatively small, fixed number (4 to 12) of channels, connected to the same number of electrodes. Stimulation amplitude of a specific electrode is dependent on the output in the specific filter, i.e. on the envelope of the signal in that frequency band.

The ACE coding strategy combines features of both SPEAK and CIS as it makes use of spectral maxima detection as well as high stimulation rates of up to 2400 Hz per channel. Whereas in the SPEAK strategy the number of active electrodes is limited to 20 of the 22 intracochlear electrodes, in the ACE strategy the maximum number of active electrodes is 22. The studies described in this thesis were all performed on recipients using the ACE strategy in the Nucleus 24 CI system.

1.2.2 Clinical fitting parameters in the ACE strategy

Fig. 2 displays a block diagram of the ACE strategy. At each stage in this diagram there are various clinical parameters that can be chosen during the fitting procedure.

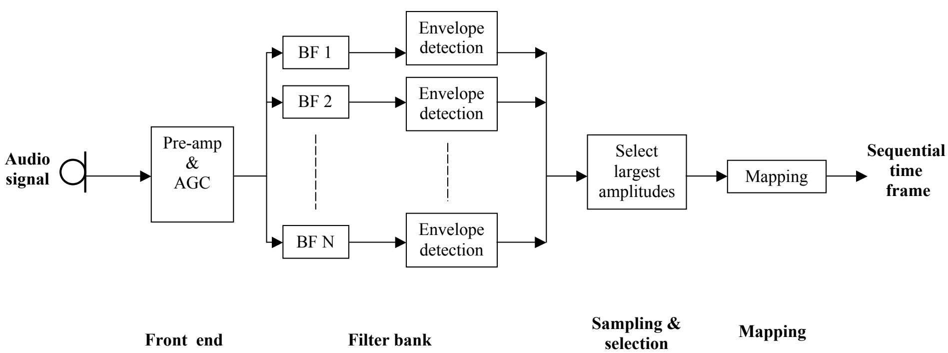


Figure 2: Block diagram illustrating the ACE strategy. Pre-amp is the pre-amplifier, AGC is the automatic gain control, BF are the bandpass filters.

A coding cycle starts at the front end with a directional microphone detecting the audio signal and transducing it into an electrical representation. The frequency response of the microphone provides a 6 dB/octave gain from 150 to 5000 Hz, and then drops off with 18 dB/octave. The microphone is pre-amplified, which is controlled by an automatic gain control (AGC) to produce a signal of which the peak value is fixed, and has an amplitude range of about 32 dB below this peak value. The sensitivity control on the speech processor determines from which sound input level the AGC starts operating. By default the AGC starts operating from input levels over 65 dB SPL.

The next processing stage is the filter bank. The signal is analyzed in a number of band pass filters corresponding to the number of available channels and hence electrodes. The frequency allocation table (FAT) is a clinical parameter that defines the width of the filters and their allocation to a specific electrode.

In the sampling and selection stage the channels with the highest amplitudes are selected for stimulation. The number of maxima (M) chosen in the clinical adjustment procedure determines the number of channels selected, and may vary between 4 and 10.

Finally, the selected electrodes are stimulated at amplitudes that are audible for the recipient. Stimulation amplitude is expressed in current level (CL), a quantity defined by the producer, Cochlear. The CL ranges from 1 to 255 current units (CU), which projects on the logarithm of the electrical current range from 10 μA to 1.75 mA; an increase by 34 CU corresponds to an increase in current by a factor of 2. For each electrode, the minimum current level that elicits an auditory percept (threshold level, T-level) as well as the maximum current level that produces a comfortable loudness level (comfort level, C-level) have to be determined. The range between T- and C-level is called the electrical dynamic range (EDR). A compression function is used to project acoustic envelope amplitudes onto the recipient's EDR. The shape of the acoustical-to-electrical mapping function, named loudness growth function (LGF) by Cochlear, is determined by the Q-parameter value and the base level (Fig. 3). The Q value is the upper percentage of the EDR for each electrode onto which the upper 10 dB of the input signal is mapped. The base level controls the acoustic input dynamic range and thereby the lowest input level that results in a stimulus at T-level. Due to the AGC, all input levels above a by the sensitivity control fixed level will be presented at C-level.

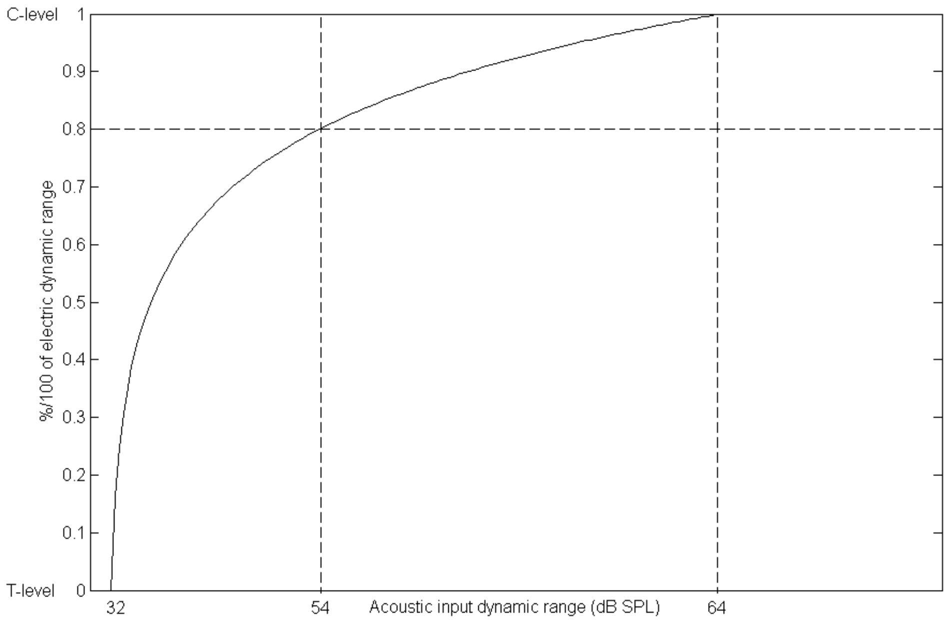


Figure 3: Acoustic-to-electric amplitude mapping: loudness growth function (LGF). This is the standard LGF setting, using a Q-parameter value of 20 and a base level of 4. The Q value of 20 implies that the upper 10 dB of the acoustic input range is mapped onto the upper 20 % of the electrical dynamic range. The base level of 4 results in an acoustic input dynamic range of about 32 dB SPL.

1.3 Conventional fitting procedure

The conventional fitting procedure requires determination of the optimal settings for all clinical parameters described above. In the present Custom Sound fitting software (Cochlear) many of the clinical parameters described above have a default value to be used as a starting point. For the ACE strategy the default for the sensitivity setting corresponds to an input dynamic range between 32 and 64 dB SPL. The number of maxima is 8. The default of the LGF has a Q value of 20 and a base level of 4 (Fig. 3).

Clinical parameters that cannot be chosen by default are the T- and C-levels. The conventional fitting requires a subjective estimation of the T-level and the C-level for each of the 22 intracochlear electrodes applying short pulse trains, using a psychophysical method similar to threshold and loudness discomfort level procedures in acoustic hearing.

In some recipients, adults as well as children, these behavioral measurements are difficult to perform and time-consuming. Adults, specifically those with a long duration of deafness or with disturbing tinnitus, may show inconsistent reactions to low stimulation levels, so it may be difficult to exactly determine their T-levels. In toddlers and infants it may also be difficult to get reliable responses within a restricted time, as the tone bursts may be meaningless to the child. Perception of these tone bursts is influenced by the child's cognitive maturation and the ongoing development of audition.

In the past years, attempts have been made to reduce the time required for fitting by determining T- and C-levels for a limited number of electrodes and setting these levels for the other electrodes by interpolation. However, as it would theoretically be best to try each combination of clinical parameters to optimize speech perception with the CI, this ideal fitting procedure is still too time-consuming, both for the clinician and the recipient.

1.4 Facilitating the fitting procedure: objective measures of the auditory system

Objective measures of the auditory system's response to electrical stimulation may be used to facilitate the speech processor fitting process. In the past two decades, several objective measures have been investigated to serve this purpose. These are the electrically evoked stapedius muscle reflex (ESR), the electrically evoked compound action potential (ECAP), the electrically evoked auditory brain stem response (EABR), and the electrically evoked middle latency response (EMLR).

1.4.1 ESR

The stapedius muscle reflex is normally elicited by strong acoustic activation of the auditory nerve. However, the reflex can also be elicited by electrical activation of the auditory nerve with a promontory electrode or through specific channels of a cochlear implant. The first publication on the feasibility of ESR measurements in CI recipients is from Jerger et al. (1986, 1988). They also demonstrated amplitude growth of the ESR with increasing stimulation levels. Later researchers reported a success rate of between 65 and 80% and a moderate to strong correlation of ESR thresholds with behavioral C-levels (Stephan et al., 1988; Spivak and Chute, 1994; Hodges et al., 1997). However, in most studies there was no one-to-one relationship with either T- or C-levels. Prediction of the behavioral levels from the ESR therefore is still a challenge. Some disadvantages of the ESR are the requirements of (1) specialized equipment, (2) an intact middle ear, and (3) a cooperative patient.

1.4.2 ECAP

The compound action potential (CAP) of the auditory nerve is the first action potential to arise after supra-threshold auditory stimulation. It represents the summed response of numerous fibers firing synchronously. The CAP can be evoked with an acoustical or electrical stimulus. In the latter case, the response is referred to as ECAP. Before the introduction of the neural response telemetry (NRT™) software in the Nucleus 24 CI system, it was only possible to measure ECAPs via direct access to the cochlea, i.e. intra-operatively (Gantz et al., 1994) or with the Ineraid CI system that incorporates a transcuteneous connector to directly control the electrodes (Brown et al., 1990; 1996). The NRT™ software, developed by Cochlear and the University of Zurich (Abbas et al., 1999; Dillier et al., 2002), provides an ECAP measuring system that uses one of the intracochlear electrodes and MP1 as the stimulating electrode pair, and a neighboring intracochlear electrode (usually positioned two electrodes more apically) and MP2 as the recording electrode pair. A bidirectional telemetry system allows the recorded ECAP to be sent back from the implant to the speech processor, from which it can be analyzed. Using this system, ECAP recordings can be performed quickly intra- or postoperatively, without additional equipment. Moreover, since the recording electrode is located inside the cochlea, muscle artifacts during ECAP recording are smaller than when using surface electrodes. Therefore, measurements can be performed without sedation, even in patients who are lively and making noises themselves.

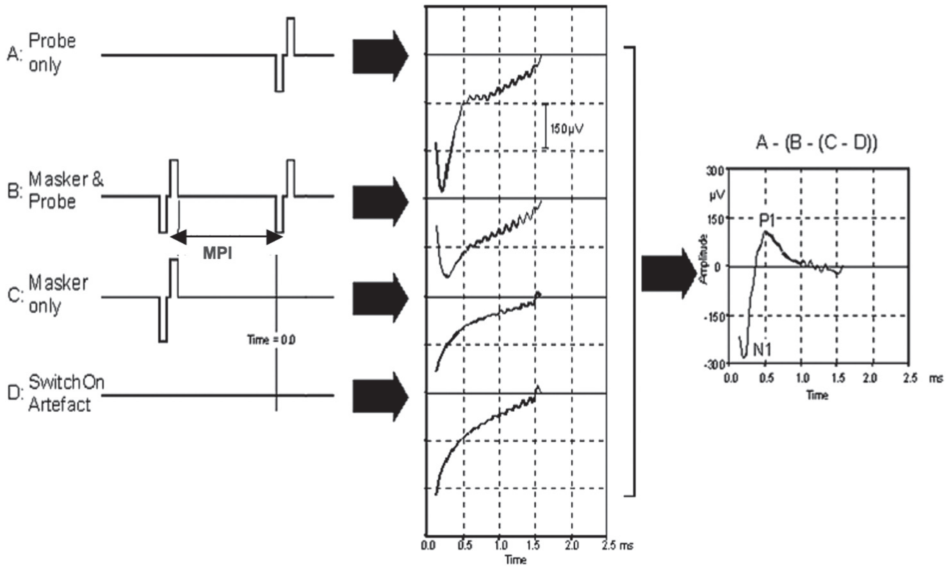


Figure 4: ECAP recording using the four measurements of the subtraction paradigm. (A) only the probe is presented: a neural response as well as a probe stimulus artifact is evoked. (B) masker-plus-probe: the masker evokes a large masker stimulus artifact and a neural response. Applying a very short masker-probe interval (MPI), the probe is presented during the refractory period of the nerve and hence does not evoke a neural response, but only a probe stimulus artifact. (C) only the masker is presented. (D) no stimulus: baseline correction for the switch-on artifact of the non-perfect amplifier.
The ECAP is derived from A – (B – (C – D)) (reprinted with permission from Cochlear Company).

A common problem in measuring electrically evoked auditory potentials is that the amplitude of the response is much smaller than the amplitude of the stimulus, resulting in a large stimulus artifact masking the response. The trick of the NRT™ software to work around this problem is to make use of the refractory state of the auditory neurons after stimulation. An ECAP recording consists of four measurements (Fig. 4). First, only the probe is presented, which evokes a neural response as well as a probe stimulus artifact. Secondly, a masker-plus-probe is presented. The masker evokes a large masker stimulus artifact and a neural response. Applying a very short masker-probe interval (MPI), the probe is presented during the refractory period of the nerve and hence does not evoke a neural response, but only a probe stimulus artifact. The second measurement is subtracted from the first, which separates the neural response evoked by the probe, i.e. the ECAP, from the probe stimulus artifact. To eliminate the stimulus artifact and neural response due to the masker, the masker is presented in the third measurement. The result of this measurement is added to the previous result. This eliminates the stimulus artifact and neural response evoked by the masker. The resulting response is then baseline corrected for the switch-on artifact of the non-perfect amplifier by measurement four, which does not contain any stimulus.

Initially, research was focused on the optimization of the measurement protocols, such as the optimal settings for the recording electrode, masker level, MPI, measurement delay, and amplifier gain. Detailed studies are performed by Lai (1999), who developed a standard measurement procedure. Further, the size and shape of the ECAP have been investigated and described (Brown et al., 1998 and 2000; Hughes et al., 2000; Cafarelli Dees et al., 2005). The ECAP consists of a negative peak N_1 with a latency of 0.2 to 0.6 ms, followed by a positive peak P_1 with a latency of 0.4 to 0.9 ms. Variations in this morphology between patients have been described, such as the absence of N_1, the absence of P_1, or a double P_1. The amplitude of the measured response varies with stimulus levels and across subjects, and ranges from approximately 50 µV to 1.5 mV. The amplitude growth function (AGF) describes the relation between stimulation level and amplitude of the response.

ECAP thresholds can be determined by visual inspection, which is rather subjective, or be defined as the intercept of the (extrapolated) linear part of the AGF and a certain amplitude. This can be the base line, i.e. an amplitude of 0 µV. However, at low stimulation levels there is a deviation from linear growth; the amplitude of the ECAP decreases only slightly with decreasing stimulation levels (Cafarelli Dees et al., 2005). Therefore, it may be better to define the ECAP threshold as the stimulus level at which a certain amplitude is reached. The set and shape of the ECAP thresholds across the whole electrode array, irrespective of overall level, is called the profile of the ECAP thresholds.

The intersubject differences in ECAP thresholds led to the suggestion that ECAPs could be useful in the programming of CIs (Abbas et al., 1999). Brown et al. (2000) and Hughes et al. (2000) investigated the correlations between ECAP thresholds and behavioral T- and C-levels in adults, and in children respectively. In adults, a moderate but significant correlation was found (r = 0.55 respectively 0.57). In children, the correlation was somewhat stronger (r = 0.70, respectively 0.71). These data were replicated in a number of studies. Thai-Van et al. (2001) found a good correlation between ECAP thresholds and behavioral levels for the apical electrodes (r = 0.70 – 0.90), whereas there was no significant correlation for the basal electrodes. Gordon et al. (2002) and Cafarelli Dees et al. (2005) found only a weak correlation between ECAP thresholds and behavioral T- and C-levels.

The results of these studies imply that the relation between ECAP thresholds and behavioral responses is not strong enough to allow for an accurate prediction of behavioral T- and C-levels in individual CI users. However, speech processor adjustments based on behavioral T- and C-levels do not necessarily yield the best speech performance. Therefore, later research focused on procedures for creating ECAP-based fittings and speech perception performance. Three ECAP-based fitting procedures have been incorporated in the Nucleus R126 2.0 fitting software (Cochlear, 2002) and are still present in the current Custom Sound fitting software (Cochlear). The first method is the 'Shift and tilt' approach, which is developed by our group, and described in Chapter 2. By principal component analysis it was shown that the profiles of ECAP thresholds as well as the conventional T- and C-levels across the full electrode array are governed by two factors. The first factor, overall level (termed shift), accounts for 90% of the variance. Inclusion of the second factor, roughly the slope of the profile (termed tilt), accounted for more than 95% of the variance. The second method is the 'T/C offset fitting', proposed by Seyle and Brown (2002), in which ECAP thresholds are combined with the behavioral T- and C-level measurement on one electrode in the centre of the array or a set of distributed electrodes. Speech perception scores showed no difference at a low speech level (55 dB SPL), but at 70 dB SPL the ECAP-based fitting was inferior to the conventional fitting. The third method is the 'Progressive pre-set fittings' by Almqvist (Reference Note 1). It makes use of four pre-set fittings based on the profile of the ECAP thresholds with increasing stimulus levels that are sequentially presented to the patient. Results were obtained only for young children. Sound field aided thresholds were around 25-35 dB HL. Children fit with the progressive pre-set fittings method scored high on the Meaningful Auditory Integration Scale (MAIS). A variation on this method was evaluated in ten adults by Seyle and Brown (2002). Instead of progressive fittings they used only one preset fitting, with T-level 20 CU below and C-levels 10 CU above the ECAP threshold for each electrode. Speech perception of sentences was significantly lower using this fitting compared to the conventional fitting.

1.4.3 EABR and EMLR

Several researchers focused on the EABR and EMLR as a tool for fitting the CI's speech processor (Shallop et al., 1991; Kileny, 1991; Hodges et al., 1994; Van den Borne et al., 1994; Brown et al., 1994 and 1999, Truy et al., 1998; Firszt et al., 1999). Most studies were performed before the introduction of the NRT™ software. In general, correlations between behavioral T-levels and EABR thresholds were highly significant. Further, EABR thresholds were obtained at approximately C-level, although occasionally they extremely exceeded C-levels. A main obstacle in EABR and EMLR recordings is stimulus artifact, which is large and can obscure the response. This problem is somewhat alleviated with the use of short biphasic pulses that are alternating in polarity. Because implant systems use radio frequency signals to transmit information across the skin, the radio frequency signal can be picked up by the recording electrodes and may contribute to the artifact problem. The use of a radio frequency filter can assist in successful recording of targeted responses. Nonauditory potentials, such as facial nerve stimulation and muscle artifact, can also interfere with recording the early latency potentials such as EABR. In small children EABR and EMLR recordings can therefore only be obtained with use of sedation.

1.5 Objectives of this thesis

The main objective of this thesis is to develop a CI fitting procedure that is easier for the recipient and less time-consuming compared to the conventional fitting procedure. These requirements are addressed by making use of the most promising of objective measures: the ECAP threshold. Using the ECAP threshold as a basis for speech processor adjustments minimizes the amount of subjective responses needed to do a fitting.

Whereas initial research on ECAPs, but also some of today's research, focuses specifically on the correlation between ECAPs and conventional behavioral T- and C-levels, implying that the conventional behavioral fitting is the golden standard and needs to be equaled, our research did not a priori assume that this is necessarily true. In our opinion it would be possible that a fitting based on ECAP thresholds would gain different T- and C-levels than the conventional fitting, but similar or better speech perception scores.

Another aspect in the search for the best fitting method is self-fitting by the patient, as a challenge of the current assumption in the conventional fitting procedure that the clinician knows what is best for the CI recipient. The influence of the recipient himself is conventionally restricted to giving oral feedback on the fitting made by the clinician. However, our spoken language is limited in words to define various aspects of sound. Vocabulary regarding hearing and sound is in the CI recipient population on average less well developed due to reduced auditory skills. This may result in miscommunication between the recipient and the clinician. Moreover, perception always depends on the individual subject, hence we are never able to hear or perceive what another person hears or perceives. We always make our own image of it. Therefore, it may be more effective to make the recipient himself adjust his fitting, instead of the clinician.

1.6 Outline of this thesis

The search for simplification of the fitting procedure started with an acute study, described in Chapter 2, in which 13 experienced CI recipients received an ECAP-based fitting for the period of two weeks. The results led to the 'Shift and tilt' fitting method.

Chapter 3 describes a prospective balanced cross-over trial in 18 new CI recipients comparing the ECAP-based fitting procedure to the conventional fitting procedure. Further optimization of the fitting procedure was performed in a study in which 18 recipients themselves gained the opportunity to adjust their fitting during daily situations (Chapter 4).

Chapter 5 describes a study performed in 25 experienced CI recipients to gain more insight in the effect of varying the clinical parameters T- and C-levels, EDR, and LGF on speech perception. In the study described in Chapter 6 we did not vary the clinical parameters, but we changed the stimulation pattern by eliminating stimuli in the lower part of the EDR in order to evaluate the importance of low-level speech stimuli on speech perception in CI recipients.

References

Abbas, P.J., Brown, C.J., Shallop, J.K. et al. (1999). Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear. 20(1): 45-59.

Battmer, R.D., Laszig, R., Lehnhardt, E. (1990). Electrically elicited stapedius reflex in cochlear implant patients. Ear Hear. 11(5): 370-74.

Brown, C.J., Hughes, M.L., Luk, B. et al. (2000). The relationship between EAP and EABR thresholds and levels used to program the Nucleus 24 speech processor: Data from adults. Ear Hear. 21(2): 151-63.

Brown, C.J., Hughes, M.L., Lopez, S.M. et al. (1999). Relationship between EABR thresholds and levels used to program the Clarion speech processor. Ann. Otol. Rhinol. Laryngol. 108: 50-7.

Brown, C.J., Abbas, P.J., Gantz, B.J. (1998). Preliminary experience with neural response telemetry in the Nucleus CI24M cochlear implant. Am. J. Otol. 19(3): 320-7.

Brown, C.J., Abbas, P.J, Borland, J. et al. (1996). Electrically evoked whole nerve action potentials in Ineraid cochlear implant users: Responses to different stimulating electrode configurations and comparison to psychophysical responses. J. Speech Hear. 39: 453-67.

Brown, C.J., Abbas, P.J., Fryauf-Bertschy, H. et al. (1994). Intraoperative and postoperative electrically evoked auditory brainstem response in Nucleus cochlear implant users: Implications for the fitting process. Ear and Hear. 15: 168-76.

Brown, C.J., Abbas, P.J., Gantz, B.J. (1990). Electrically evoked whole-nerve action potentials: data from human cochlear implant users. J. Acoust. Soc. Am. 88(3): 1385-91.

Cafarelli Dees, D., Dillier, N., Lai, W.K. et al. (2005). Normative findings of electrically evoked compound action potential (ECAP) measurements using the neural response telemetry (NRT™) of the Nucleus CI24M cochlear implant system. Audiol. Neuro-otol. 10(2):105-16.

Clark, G.M., Hallworth, R., Zdanius, K. (1975). A cochlear implant electrode. J. Laryngol Otol 89: 787–92.

Clarke, G.P., Sandfort, M.D. (1995). Interactive computer system for tuning the cochlear 22-channel implant in young children. Ann. Otol. Rhinol. Laryngol. Suppl. 166: 165-6.

Dillier, N., Lai, W.K., Almqvist, B. et al. (2002). Measurement of the electrically evoked compound action potential via a neural response telemetry system. Ann. Otol. Rhinol. Laryngol. 111(5 Pt 1):407-14.

Djourno, A., Eyriès, C. (1957). Prosthèse auditive par excitation électrique à distance du nerf sensorial à l'aide d'un bobinage inclus à demeure. Presse Médicale 35: 14–17.

Douek, E., Fourcin, A., Moore, B. et al. (1977). A new approach to the cochlear implant. Proc. R. Soc. Med. 70: 379-83.

Eddington, D.K., Dobelle, W.H., Brackmann, D.E. et al. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. Ann. Otol. Rhinol. Laryngol. 87:1-39.

Firszt, J.B., Rotz, L.A., Chambers, R.D. et al. (1999). Electrically evoked potentials recorded in adult and pediatric Clarion implant users. Ann. Otol. Rhinol. Laryngol. 108: 58-63.

Gantz, B.J., Brown, C.J., Abbas, P.J. (1994). Intraoperative measures of electrically evoked auditory nerve compound action potential. Am. J. Otol. 15(2): 137-44.

Gisselsson, L. (1950). Experimental investigation into the problem of humoral transmission in the cochlea. Acta Otolaryngol. 82(Suppl): 16.

Gordon, K.A., Ebinger, K.A., Gilden, J.E. et al. (2002). Neural response telemetry in 12- to 24-month-old children. Ann. Otol. Rhinol. Laryngol. 189(Suppl): 42-8.

Hochmair Desoyer, I.J., Hochmair, E.S., Fischer, R.E. et al. (1980). Cochlear prostheses in use: recent speech comprehension results. Arch. Otorhinolaryngol. 229(2): 81-98.

Hodges, A.V., Butts, S., Dolan-Ash, S. et al. (1999). Using electrically evoked auditory reflex thresholds to fit the Clarion cochlear implant. Ann. Otol. Rhinol. Laryngol. 108: 64-8.

Hodges, A.V., Balkany, T.J., Ruth, R.A. et al. (1997). Electrical middle ear muscle reflex: use in cochlear implant programming. Otolaryngol. Head Neck Surg. 111: 255-61.

Hodges, A.V., Ruth, R.A., Lambert, P.R. et al. (1994). Electric auditory brain-stem responses in Nucleus multichannel cochlear implant users. Arch. Otolaryngol. Head Neck Surg. 120: 1093-9.

House, W.F. (1976). Cochlear implants. Am. Otol. Rhinol. Laryngol. 85(Suppl 23): 1-93.

House, W.F., Urban, J. (1973). Long term results of electrode implantation and electronic stimulation of the cochlea in man. Ann. Otol. Rhinol. Laryngol. 82(4): 504-17.

Hughes, M.L., Brown, C.J., Abbas, P.J. et al. (2000). Comparison of EAP thresholds with MAP levels in the Nucleus 24 cochlear implant: Data from children. Ear Hear. 21(2): 164-74.

Jerger, J.F., Oliver, T.A., Chmiel, R.A. (1988). Prediction of dynamic range from stapedius reflex in cochlear implant patients. Ear Hear. 9(1): 4-8.

Jerger, J., Jenkins, H., Fifer, R. et al. (1986). Stapedius reflex to electrical stimulation in a patient with a cochlear implant. Ann. Otol. Rhinol. Laryngol. 95(2 Pt 1): 151-7.

Kiefer, J., Hohl, S., Stuerzebecher, E., et al. (2001). Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the nucleus CI 24M cochlear implant system. Audiol. 40(1): 32-42.

Kileny, P.R. (1991). Use of electrophysiologic measures in the management of children with cochlear implants: Brain stem, middle latency, and cognitive (P300) responses. Am. J. Otol. 12 (Suppl.): 37-42.

Lai, W.K. (1999). An NRT Cookbook. Guidelines for making NRT measurements. ISBN 3-9521853-0-2.

Seligman, P.M, McDermott, H.J. (1995). Architecture of the Spectra 22 speech processor. Ann. Otol. Rhinol. Laryngol. 104(Suppl. 166): 172–5.

Seyle, K., Brown, C.J. (2002). Speech perception using maps based on neural response telemetry measures. Ear Hear. 23(1 Suppl): 72S-79S.

Shallop, J.K., VanDyke, L., Goin, D.W. et al. (1991). The prediction of behavioural T/C values for nucleus 22 channel implant patients from EABR results. Ann. Otol. Rhinol. Laryngol. 100: 896-8.

Shannon, R.V. (1985). Threshold and loudness functions for pulsatil stimulation of cochlear implants. Hear. Res. 18: 135-43.

Skinner, M.W., Holden, L.K., Demorest, M.E. et al. (1995). Use of test-retest measures to evaluate performance stability in adults with cochlear implants. Ear Hear. 16 (2): 187-97.

Skinner, M.W., Clark, G.M., Whitford, L.A. et al. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. Am. J. Otol. 15(Suppl. 2): 15–27.

Spivak, L.G., Chute, P.M. (1994). The relationship between electrical acoustic reflex thresholds and behavioral comfort levels in children and adult cochlear implant patients. Ear Hear. 1994 Apr;15(2): 184-92.

Stephan, K., Welzl-Müller, K. (2000). Post-operative stapedius reflex tests with simultaneous loudness scaling in patients supplied with cochlear implants. Audiology 39: 13-8.

Stephan, K., Welzl-Muller, K., Stiglbrunner, H. (1988). Stapedius reflex threshold in cochlear implant patients. Audiology. 27(4): 227-33.

Thai-Van, H., Chanal, J.M., Coudert, C. et al. (2001). Relationship between NRT measurements and behavioral levels in children with the Nucleus 24 cochlear implant may change over time: preliminary report. Int. J. Ped. Otorhinolaryngol., 58(2): 153-62.

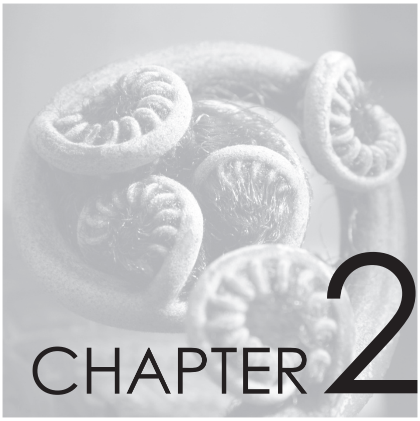
Truy, E., Gallego, S., Chanal, J. et al. (1998). Correlation between electrical auditory brainstem response and perceptual thresholds in digisonic cochlear implant users. Laryngoscope 108: 554-9.

Van den Borne, B., Mens, L.H., Snik, A.F. et al. (1994). Stapedius reflex and EABR thresholds in experienced users of the Nucleus cochlear implant. Acta Otolaryngol. 114(2): 141-3.

Wilson, B.S., Lawson, D.T., Finley, C.C. et al. (1995). New processing strategies in cochlear implantation. Am. J. Otology 16: 669-75.

Reference note

Almqvist, B. (2004). The Almqvist/Lund procedure for adjusting a speech processor for a congenitally deaf child using objective measurements. Cochlear Nucleus Report July/Agust, 3-4.



Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds

Guido F. Smoorenburg, Christina Willeboer and Johannes E. van Dijk
Audiology and Neuro-Otology 2002, 7: 335-47. http://www.karger.com/

Abstract

Adjusting the speech processor of a cochlear implant, per electrode, to the individual's response is a laborious task that may interfere with a user-friendly start of implant-mediated hearing, particularly in children. This research concerns the possibility of processor fitting based on a profile derived from measurements of the electrically evoked compound action potential (ECAP) thresholds across the electrode array, followed by adjustment of the overall level of the profile to the hearing

threshold and maximum comfortable loudness level using live voice. The results for CVC word lists show that speech perception is quite insensitive to the threshold setting of the speech processor. On average, the speech score does not decrease by more than 10% when, with the new method, the threshold setting comes out so much lower that the dynamic range has doubled. In contrast, the speech score appears to be sensitive to an increase of the maximum high-frequency stimulation settings for the basal electrodes, resulting in lower scores at these higher settings. The correlation between the overall ECAP thresholds and conventionally measured subjective thresholds is weak (r = 0.64). However, the correlation between the slopes of these threshold curves is satisfactory (r = 0.82). The correlation between the ECAP thresholds and the maximum stimulation levels is poor, both with respect to overall level and slope (r = 0.39 and 0.36, respectively). Applicability of the ECAP threshold in processor fitting could not be demonstrated in this study. Prediction of the most critical factor in speech perception, the slope of the maximum stimulation curve, from the ECAP thresholds is poor. However, considering habituation to the initial processor setting of at least 6 months, the small decrease in the CVC scores with the new setting suggests that a more user-friendly fitting procedure can be developed.

2.1 Introduction

Fitting the speech processor of cochlear implants is commonly based on subjective assessment of the detection threshold per electrode (the T-level) and the highest comfortable loudness level per electrode (the C-level). With 22 electrodes in the Nucleus CI24M implant this is a time-consuming procedure. When fitting the processor of children too young to give a reliable cooperative response this method becomes particularly cumbersome. Moreover, the first hearing experiences with cochlear implants may be difficult to interpret, may even be frightening to an implanted child. This may interfere with a rapid assessment of the correct fitting necessary to present the subject, as soon as possible, with the proper sounds. Frequent readjustments of the processor may seriously delay the process of learning to discriminate and recognize sounds because each processor readjustment may require adjustment of the subject to marked changes in the sounds. Therefore, other methods of fitting the speech processor, independent of the subject's response, would have significant benefit.

A potentially important tool for an alternative fitting procedure was introduced by Brown and Abbas (1990) and implemented by Cochlear Company in their Nucleus CI24M implant (FDA approval in June 1998). This tool is based on measurement of the electrically evoked compound action potential (ECAP) using one of the electrodes of the implant itself as a recording electrode. After analogue-digital conversion in the implant, the sampled ECAP is transmitted back from the implant to the speech processor. This objective (physical) method of assessing the response of the ear to electrical stimulation seems to be promising in comparison with stapedius muscle reflex measurements, which are difficult to quantify, and responses from the brainstem and higher brain centers, which require sedation in children and relatively long recording sessions.

Large stimulus artefacts impede ECAP measurements. The artefact due to the electrical stimulus is orders of magnitude larger than the response to this stimulus. In addition, the ECAP appears within 1 ms after the stimulus, a time interval within which an amplifier may not have recovered from overstimulation by the electrical stimulus. Brown et al. (1990) introduced a subtraction method that effectively copes with this problem. In experiments with cats Brown and Abbas (1990) showed that uncorrupted ECAPs could be obtained by taking advantage of the refractory period of neurons. Comparing the response to an electrical impulse P (the probe) and the response to the probe P, immediately preceded by another impulse M (the masker), the response to P alone will consist of the artefact plus the response to P while the response to the combination of probe and masker M + P will contain the artefact only. In the latter case, P does not elicit an ECAP since it was presented within the refractory period of the auditory neurones following their response to M. Thus, the difference between the two responses will reveal the ECAP. The implementation in the Nucleus CI24M includes a second artefact due to switching the amplifier on and off. This artefact is cancelled in a similar fashion. The method above is used in the present research. We applied the software developed by Dillier and co-workers in Zürich, Switzerland (Cochlear NRT™, v2.04).

So far, research on ECAPs in human subjects has been focused on optimization of the ECAP measurement itself and on relations between ECAP parameters and psychophysical measures (Brown et al., 1996, 1998, 2000; Abbas et al., 1999; Hughes et al., 2000). In particular the relation between ECAP thresholds and the T- and C-levels has been a point of focus in view of the question of whether or not ECAP measurements could replace the subjective assessment of the T- and C-levels used in speech processor fitting. In adults, Brown et al. (2000) found moderate correlation between the ECAP thresholds and the T-levels ($r = 0.55$) and C-levels ($r = 0.57$). In children, Hughes et al. (2000) found somewhat higher coefficients: $r = 0.70$ and 0.71, respectively. The authors of the two studies stated that, on an individual basis, these coefficients are too low to estimate the T- and C-levels confidently from ECAP measurements. They concluded that the application of ECAP thresholds in predicting T- and C-levels requires a contribution from subjective measurements.

In contrast to the studies summarized above, this research presents the evaluation of an ECAP-based fitting procedure without focusing on the prediction of T- and C-level per electrode from the ECAP threshold. Processor fitting is based on the ECAP threshold measured across the full electrode array, the ECAP threshold profile. The measurement is complemented by a simple adjustment of the overall level of the individual's ECAP profile to redefined T- and C-levels, T-NEW and C-NEW. This adjustment is based on the subject's response to live voice with the speech processor in SPEAK mode. T-NEW corresponds to detection of the speech signal, C-NEW to a comfortable speech level for a given T-NEW. As compared to the standard fitting method based on tone pips this procedure has the advantage, especially for children, that one starts the procedure with speech in direct interaction with the implanted subject. In the present study speech perception is measured immediately after the new, ECAP-based, fitting has been installed and about two weeks later. In these two weeks, subjects were asked to use the new fitting as much as possible. However, they were allowed to use their initial, conventional, fitting whenever they wanted. Both fittings were available to them via the program switch on the speech processor. The speech results for the ECAP-based fitting are compared to those for the initial fitting. The ECAP measurements were conducted post-operatively. Subjects were adults only.

2.2 Materials and methods

2.2.1 Subjects and initial processor fitting

Twenty-seven postlingually deafened adults entered the study. Thirteen subjects provided us with ECAP profiles across the full electrode array of 20 electrodes used in SPEAK mode (electrodes 3-22). The present study was based on these thirteen subjects. They all received the implant in our University Medical Centre Utrecht and had used their implant for at least 6 months.

All subjects used the Nucleus CI24M cochlear implant of Cochlear Company with either the Sprint™ or the ESPrit™ speech processor. The Clinical Programming System (CPS) and the Windows Diagnostic and Programming System (WinDPS, R116) software were used to measure the standard T- and C-levels and electrode impedance values. The T-level is defined as the lowest stimulus level that elicits a very soft, but consistent hearing sensation for each electrode separately. The C-level is defined as the maximum stimulus level per electrode that produces a comfortable loudness sensation. In addition to the adjustment of the C-levels per electrode they were balanced across electrodes for equal loudness. The T- and C-levels were measured with stimuli consisting of biphasic impulse trains with an impulse duration of 25 µs/phase and at a rate of 250 impulses/s. The duration of the impulse trains was 500 ms, separated by silent intervals of about 500 ms in which a response could be given. Stimulation mode was monopolar using both extracochlear reference (indifferent) electrodes, the ball electrode, usually inserted into the temporal muscle, and the plate electrode attached to the implant housing (MP1+2). Stimulation amplitude is expressed in the current level (CL), a quantity defined by Cochlear Company. The CL ranges from 1 to 255 current units, which corresponds to electrical currents from 10 µA to 1.75 mA. The relation between current units and electrical current is approximately logarithmic with 34 current units corresponding to a factor of two in electrical current or 6 dB. After three months of implant use, T- and C-levels had stabilized. Thus, at the commencement of the study, after at least six months of implant use, subjects had well adapted to the conventional fitting of the speech processor. For all subjects, the base level of the Cochlear speech processor was set at its default value of 4. With respect to the maximum value of 150 this implies that the dynamic range of the acoustic input signal amounts to 32 dB. Also, for all subjects the Q-value of the compressive non-linear transfer from acoustic input to electrical output was set at its default value of 20, implying a 20% decrease in CL at –10 dB input level relative to the maximum level.

2.2.2 ECAP measurements

ECAP measurements were performed with the NRT™ (Neural Response Telemetry) software version 2.04 of the Cochlear Company. They were performed postoperatively in a single session of 1.5 - 2 h per subject. Subjects were seated in a chair without further restrictions. The ECAP growth functions were measured on all electrodes activated in the conventional fitting.

We used a modified version of the protocol described by Abbas et al. (1999). Recording electrodes were chosen two positions above the stimulation electrode, thus the second electrode, N+2, from the stimulation electrode, N, in the apical direction, except for electrodes 21 and 22, for which the recording electrodes were numbers 19 and 20, respectively. The stimulation mode was monopolar (MP1 mode, using the extracochlear ball reference electrode). In contrast to the protocol used by Abbas et al. (1999), which includes a fixed masker level, we used a masker level 10 current units above the probe level. However, when this offset of 10 current units yielded too loud a sensation, the offset was reduced to zero. This did not seriously affect the threshold estimates. Masker advance, which is the masker-probe interval, was fixed at 500 µs. As a rule, the sampling delay, i.e. the interval between stimulation and initiation of sampling, was set at

50 µs. If amplifier saturation occurred, the delay was increased stepwise to 70, 90, 120 or 140 µs until a satisfactory response was obtained. Initially the amplifier gain was set at 60 dB. However, when increasing the delay up to 140 µs did not remove amplifier saturation, the gain was decreased to 40 dB. With 60 dB amplifier gain the number of sweeps was 100, instead of 50 in the protocol of Abbas et al. (1999). With 40 dB gain it was increased to 200. In conformance with Abbas et al. (1999), the pulse duration was set at 25 µs per phase. Stimulation rate was 80 Hz.

2.2.3 ECAP threshold assessment

With each electrode we started with the test stimuli at approximately the T-level. We increased the probe level in steps of 5 current units until a neural response was seen on the baseline corrected low-resolution components screen of the NRT v2.04 software. The occurrence of the response was identified visually on the basis of a priori knowledge of its waveform (Lai, 1999). Subsequently, the amplitude growth function was measured for several increments of 5 current units. If excessive loudness growth did not allow for such an increase of stimulus level beyond the observed threshold level, we obtained at least one additional response after increasing the level by 2 current units in order to confirm the response. From repeated measurements we estimated measurement accuracy at 3 current units.

2.2.4 ECAP-based processor fitting

The ECAP threshold across all electrodes, the ECAP profile, was used to adjust the speech processor in SPEAK mode. The T-NEW levels across electrodes 3-22 were obtained by maintaining the profile, the relative current levels across the electrode array, while changing the overall current level. This was realized by changing the T- and C-levels equally while keeping their difference at 0. With this procedure each electrode delivers a binary stimulus; below a certain acoustic input level there will be no electrical stimulus, beyond this input level there will be an electrical stimulus with amplitude independent of the acoustic input level.

Using live voice the overall T-level was thus adjusted to the level at which the speech signal was just detected. After this measurement of the T-NEW levels we determined the C-NEW levels by shifting all C-levels upward by the same number of current units to the maximum comfortable loudness level for live voice. To this end the C-level was uncoupled from the T-level, the T-NEW level remaining untouched. Moreover, the profile was not changed. Thus, the T- and C-NEW curves differ by a constant number of current units across the electrode array. The volume and sensitivity settings were always at the default values of 9 and 8, respectively.

2.2.5 Speech perception measurements

Speech materials consisted of Dutch (linguistically meaningful) words, each consisting of the sequence consonant-vowel-consonant (CVC words). The test lists consisted of 12 words each. These lists were phonetically balanced. They were uttered by a female speaker (Bosman, 1989). The test was presented from compact disc at 65 dBA at the headset microphone of the implant's speech processor. The subjects were allowed to give any response. (The response set was open.) Eight lists (96 words) were used for each condition. A certain list was never presented twice to the same subject. A contra-lateral acoustic hearing aid, if present, was switched off.

We chose to use a speech test based on CVC words in order to avoid habituation effects as much as possible. Patient ethics prevented us from starting with non-standard fittings such as an ECAP-based fitting. Therefore, we recognized the problem that habituation to the conventional fitting might affect the results for an alternative fitting. We tried to reduce this effect by using the laboratory-like CVC test rather than using tests based on sentences or running speech, which are closer to everyday listening situations.

The CVC scores for both the initial (conventional) and the new ECAP-based fitting were determined during the first session in which also the ECAP-based T- and C-levels were determined. After this session we asked the subjects to use the ECAP-based fitting as much as possible during a period of about two weeks. Both fittings were available to the subjects through the processor's program switch. However, as mentioned before, the subjects were allowed to use the initial fitting whenever they felt the ECAP-based fitting was too unpleasant. Subjects recorded the number of hours they used the ECAP-based fitting and the quality of the sound. After this period we repeated the speech tests for both the conventional and the ECAP-based processor fittings.

2.3 Results

2.3.1 Success rate of the ECAP measurements

In 13 subjects (numbered 1-13 in Table 1), it was possible to determine ECAP thresholds at virtually all activated electrodes. Only in two subjects we did not get a clear neural response at one or two electrodes within acceptable loudness levels. In these cases we estimated the ECAP thresholds by interpolating threshold values from adjacent electrodes. In the remaining14 subjects (numbers 14-27 in Table 1) the first part of the measurements at electrodes 3, 5, 10, 15 and 20 showed that no clear ECAPs could be obtained across the full electrode array within acceptable loudness levels. These 14 subjects were excluded from the study. Our success rate of 13/27 (48%) may seem low but one should keep in mind that they refer to responses at all 20 electrodes. Higher success rates in the literature usually refer to ECAP responses restricted to only part of the electrode array, frequently to at least one electrode.

Analyzing the differences between the group with complete ECAP responses and the group with incomplete responses we found a statistically significant difference only for the C-levels (averaged across electrodes CL = 174 for the incomplete responses versus CL = 185 for the complete ones ($p = 0.02$ according to Student's t-test for independent samples). The T-levels, dynamic ranges (differences between the C- and T-levels) and impedances measured in all modes available did not differ significantly ($p > 0.05$). In the Cochlear field study, mentioned above, we noticed that the ECAP thresholds averaged across the subjects with incomplete responses (in so far they could be measured) did not differ from those with the complete responses. Thus, together these results suggest that when ECAPs are not found in a particular subject at an acceptable loudness level it is because the C-levels are relatively low rather than that the ECAP threshold would be relatively high.

subject	age (y)	deafness			duration of CI use (y;m)	electrodes / support rings	electrodes in MAP	mean		
		aetiology	duration (y;m)					DR (CL)	C-level (CL)	impedance (kΩ)
			ipsi	contra						
1	37	unknown	17	17	0;6	22 / 9	20	10	174	4.9
2	61	otosclerosis	36	36	2;6	22 / 6	18	19	182	6.5
3	44	hereditary	15	3	15/2;6 [a]	22 /10	20	58	192	3.8
4	45	Cogan	2	2	3	22 /10	20	34	179	4.5
5	47	hereditary	2	RH	1	22 / 8	20	25	173	2.5
6	43	meningitis	2	2	0;6	22 / 9	20	51	193	4.8
7	52	unknown	48	5	3	22 / 0	20	28	163	5.1
8	64	unknown	4	2	2	22 /10	16	25	185	7.3
9	72	hereditary	30	30	2	22 / 8	20	33	210	-
10	54	otosclerosis	10	RH	2	22 /10	20	23	209	2.7
11	46	meningitis	40	RH	2	22 / 3	20	5	184	6.9
12	47	unknown	12	12	0;6	22 /10	20	24	179	3.9
13	55	otosclerosis	9	9	0;6	22 /10	20	23	178	4.8
14	69	unknown	42	30	0;9	22 /10	19	38	175	-
15	70	unknown	50	21	0;6	22 /10	20	23	185	5.0
16	63	unknown	15	10	2	22 /10	20	11	176	3.0
17	14	meningitis	0;3	0;3	1	22 /10	20	36	148	8.0
18	31	unknown	4	RH	1	22 / 7	20	14	178	4.2
19	68	hereditary	17	RH	2;6	22 /10	20	12	173	4.7
20	41	unknown	3	3	2	22 / 9	20	17	170	5.6
21	47	unknown	18	21	2;6	22 /10	19	26	169	5.4
22	62	meningitis	3	3	1	22 /10	20	35	177	5.8
23	47	unknown	3	13	1;6	22 / 9	20	18	173	5.4
24	51	otosclerosis	12	RH	1	22 / 5	20	21	188	6.7
25	30	meningitis	25	25	2;6	22 /10	20	11	174	4.6
26	62	unknown	11	11	2	22 /10	20	17	169	5.1
27	52	otitis	11	4	0;3	22 /10	19	18	177	3.3

Table1: RH = Residual hearing; CI = cochlear implant; DR = dynamic range; CL = current level; - = not measured.
[a] Re-implantation ipsilateral.

2.3.2 Averaged profiles

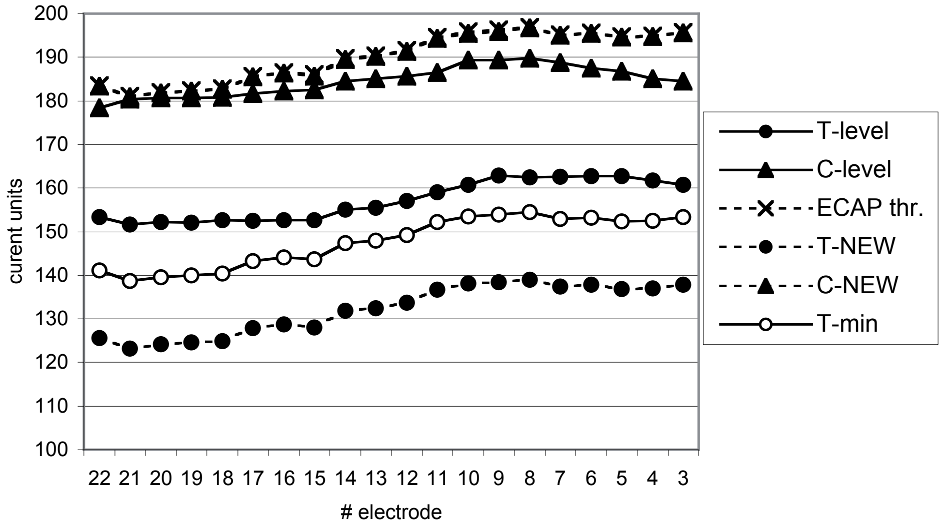


Figure 1: Conventional T- and C-levels, ECAP thresholds and T- and C-levels (T- and C-NEW) based upon the profile across electrodes of the ECAP thresholds adjusted with respect to their overall level using a subjective procedure including speech perception. T-min indicates a hypothetical result for T-NEW assuming that the speech detection threshold was based upon detection of the speech signal at one electrode whereas the speech signal at the other electrodes would still be too low to provide a detectable signal.

The conventional T- and C-levels, averaged across the 13 subjects, are presented in Fig. 1. The difference between these two levels ranges, on average, from 25 to 30 current units. (Above we mentioned that a difference of 34 units corresponds to 6 dB.) In addition, Fig. 1 shows the averaged ECAP thresholds. This average lies close to the averaged C-levels. However, at the basal electrodes, transmitting the high-frequency components of the acoustic stimulus (low electrode numbers), the individual ECAP thresholds tend to exceed the C-levels. This explains the limited success rate of measuring ECAPs; the ECAP threshold may be reached at an uncomfortably high loudness level. Also included in Fig. 1 are the averaged T- and C-levels based on the ECAP profiles, T-NEW and C-NEW. The T-NEW levels are found at considerably lower current levels than the conventional T-levels, the difference amounts from 25 to almost

30 current units in the averaged results. On average, the C-NEW levels are found very close to the ECAP thresholds.

In determining the T-NEW level, as described above, one might argue that a detection threshold will be found whenever only one electrode is stimulated (and the others not yet) when the overall level of the ECAP profile is increased starting at a subthreshold level. In view of this possibility we calculated, for each subject individually, the overall level of the ECAP profile at which one electrode is first reached, using the T-levels measured per electrode. The result, averaged across subjects, has been added to Fig. 1 (T-min). This result shows that the markedly low average T-NEW level does not originate with sound detection based on the above assumption. The T-min curve is found at about one-third of the total difference between the T-NEW level and the conventional T-level.

2.3.3 Principal components of the measurements across electrodes

Within the scope of this study it was appropriate to measure all electrodes. However, one may expect substantial correlations between the levels measured at adjacent electrodes. Therefore we analyzed the number of independent components describing the profiles using the statistical technique of principal components analysis. The analysis was conducted for the T- and C-levels and ECAP thresholds together and each one separately. The four analyzes yielded closely related results. The profiles can be described by two significant components only. These two components account for 97.7, 96.7, 95.8 and 96.4% of the variance for the combined results, the T-levels, the C-levels and the ECAP thresholds, respectively. For the ECAP thresholds the remaining variance amounts to 3 current units per electrode, which corresponds exactly to the estimated measurement error. This illustrates that the two components provide a full description of the data.

The first component, resulting from either the combined or the separate data, corresponds to the overall level of the profiles. (We did not apply any type of the rotations frequently used in principal components analysis.) This component accounts for 94.1, 90.0, 88.3 and 90.0% of the variance for the combined and separate results, respectively. Also the second component represents the same aspect of the data, irrespective of the data set. It corresponds closely to the slope, the tilt, of the profile. Since the principal components do not depend on the data set we will continue to use the solution for the combined data set. The interpretation of the two components is illustrated in Figs. 2 and 3. All individual data contributing to the averaged data in Fig. 1, except those for the theoretical T-min curve, have been included in Fig. 2, all individual T-, C- and ECAP profiles in Fig. 3. Fig.2 shows how closely the first component is related to overall level. One unit of the component score corresponds to about 20 current units. Although less close, Fig. 3 shows that the second component is related to the slope of the profiles. One unit of the score for this component corresponds to a slope of about 2 current units per 3 electrodes distance. (The units of the principle component scores are arbitrary in the sense that they are defined statistically.) In order to further illustrate the representation of the profiles in terms of the two principal components Fig. 4 shows five individual profiles and their corresponding component scores.

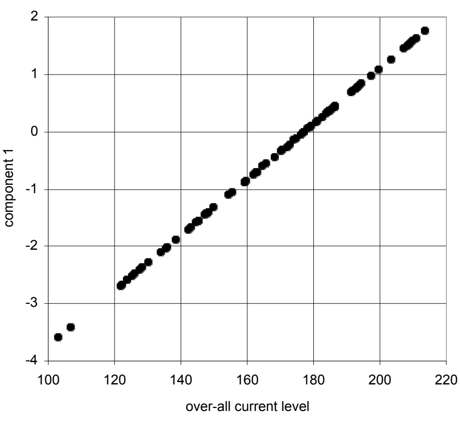


Figure 2: Relation between the scores for component 1 and the average level across all electrodes for all individual data contributing to Fig. 1 except those for T-min. An increase of the score for component 1 by one unit corresponds to about 20 current units.

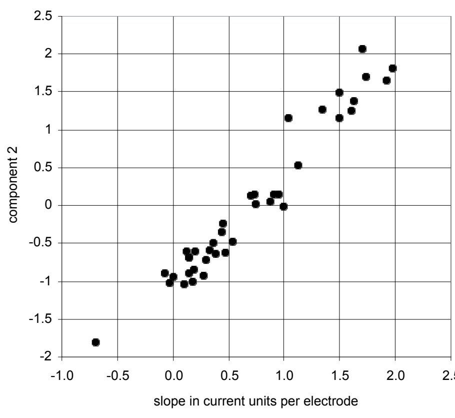
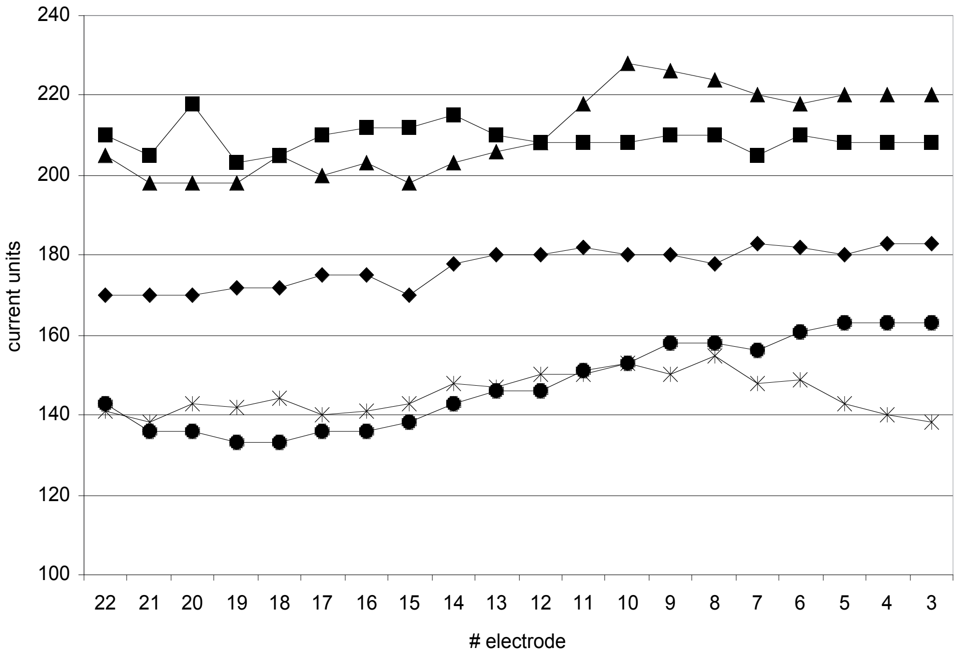


Figure 3: Relation between the scores for component 2 and the slope of the profiles across all electrodes for the individual T-, C- and ECAP profiles contributing to Fig. 1. An increase of the score for component 2 by one unit corresponds to a slope of about 2 current units over a 3-electrode distance.



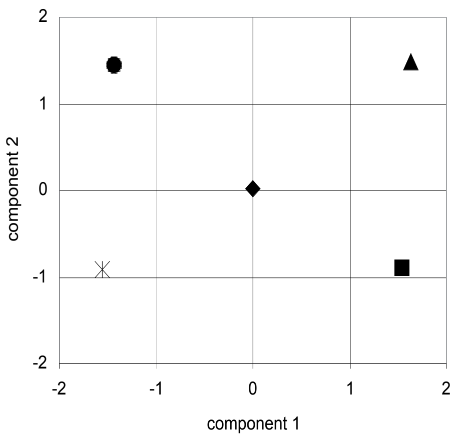


Figure 4: Five individual profiles in relation to their scores for components 1 and 2. Component 1 represents overall level, component 2 the tilt in the profiles. The symbols of the individual profiles correspond to those in the components plot.

2.3.4 Relation between ECAP thresholds and T- and C-levels

As mentioned in the section ' Introduction' the straight correlations per electrode between the ECAP thresholds and the T- and C-levels have been somewhat disappointing. Brown et al. (2000) and Hughes et al. (2000) concluded that the ECAP thresholds cannot be used to predict the T- or C-levels individually with a sufficient degree of accuracy. When studying these relations in terms of the principal components, we find the following correlation coefficients across the 13 subjects for ECAP threshold versus T-level and versus C-level, respectively: component 1 (overall level), r = 0.64 and 0.39; component 2 (profile tilt), r = 0.82 and 0.36. This result shows that the ECAP threshold cannot be used to predict the overall C-level or the tilt in the C-profile. Second, it shows a correlation between overall ECAP threshold and T-level of r = 0.64, which is within the range of 0.55 to 0.71 reported by Brown et al. (2000) and Hughes et al. (2000), both mentioned in the introduction. However, the result also shows that there is a clear relation (r = 0.82) between the tilt in the ECAP threshold and the tilt in the T-level.

Assuming that the T-profile, in particular the tilt in this profile, is important when fitting the speech processor, one could use the profile of the ECAP threshold if the overall level could be determined in another way. This has been the rational of introducing the T-NEW levels described above. Having no other clue from the ECAP measurements as to the C-levels, we decided to also use the ECAP threshold profile for C-NEW, limiting ourselves to a subjective adjustment of the overall level of this profile.

2.3.5 CVC scores for the conventional and ECAP-based speech processor fitting

The phoneme scores from the CVC tests administered with the conventional speech processor fitting, with the ECAP-based fitting immediately after the new fitting and with the ECAP-based fitting after about two weeks of requested practice are shown in Fig. 5. The mean scores across the thirteen subjects, presented on the right-hand side, show little difference between the three scores. This is a remarkable result considering the large differences between the conventional and ECAP-based fittings, in particular the T- versus T-NEW levels. The mean scores are 59, 49 and 53% for the conventional fitting and the first and second ECAP-based measurements, respectively. Although small, the differences between the scores for the conventional fitting and the first and second ECAP-based fittings were statistically significant at the $p = 0.01$ level. Also, the increase of the ECAP-based score after two weeks was significant ($p = 0.02$).

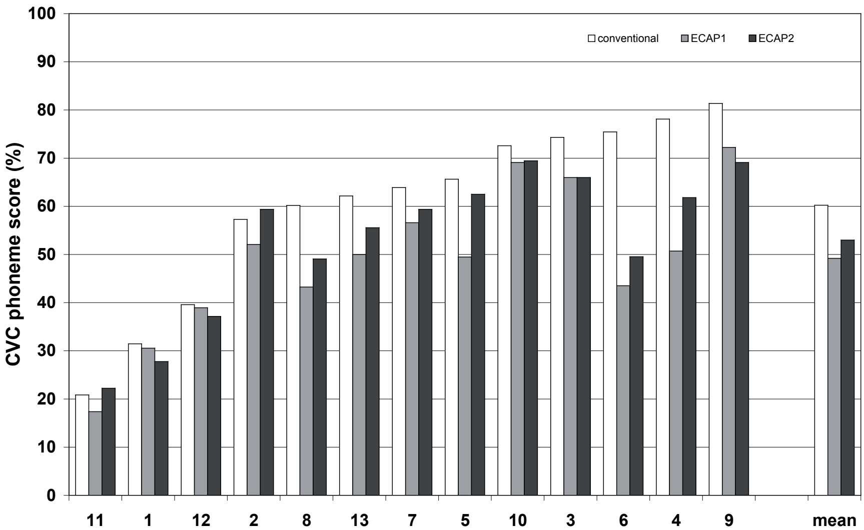


Figure 5: CVC phoneme scores for the 13 subjects individually and their mean value. The first bar indicates the score for the conventional processor fitting, the second one for the ECAP-based fitting immediately after adjustment (ECAP1), the third one for the ECAP-based fitting about two weeks later (ECAP2). Measurement error, based on two scores for the conventional fitting separated by two weeks, is 3%. Subjects are ordered according to the initial score.

The scores for the conventional fitting were determined in combination with both the first and the second ECAP-based measurement. These scores did not show a statistically significant change over the two-week period. From these two measurements we calculated a measurement error in the phoneme scores of only 3 percentage points. Fig. 5 shows the means of the first and second measurement.

Although the mean phoneme scores for the two fitting methods are fairly close to one another, Fig. 5 shows considerable individual differences. In the second session differences between the scores for the conventional and ECAP-based fittings may vary from 6% improvement to a 22% loss. Fig. 6 shows the amount of practice the subjects had with the ECAP-based fitting. Five subjects used the ECAP-based fitting for less than 60 hours. The four subjects (4, 6, 8, and 9) with phoneme scores for the conventional fitting 10% or more higher than the second ECAP-based measurement are among those five subjects. The number of days the ECAP-based fitting was tried in the two-week period (Fig. 6) shows that the five subjects with less than 60 hours of practice did try the new fitting on 5 - 13 days. Although they were perfectly willing to try the new fitting they had to limit its use from 5 to less than 0.5 h per day because sound perception was too unsatisfactory. In the next section these individual differences in performance will be compared to the differences in speech processor fitting.

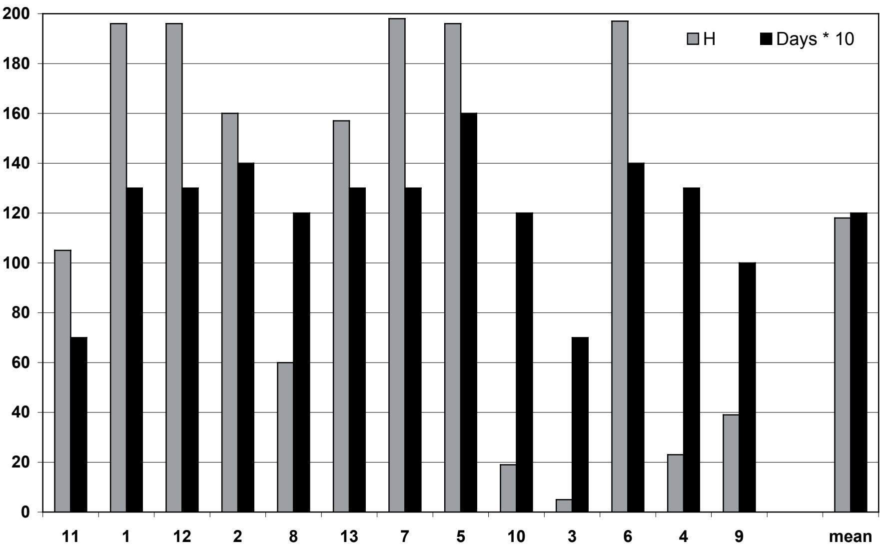


Figure 6: Use of the ECAP-based fitting in hours and number of days (times 10) the experimental fitting was tried by each subject in the two-week period between the two speech tests ECAP1 and ECAP2.

2.3.6 Speech perception performance in relation to the profile characteristics

Fig. 7 presents speech performance in relation to the differences between the scores for the conventional fitting and the first ECAP-based measurement on profile component 1, overall level. Fig 8 presents the same data for component 2, the tilt in the profiles. As for speech perception we distinguish between worse and similar performance. The criterion for worse performance is a decrease of more than 10% when changing from the conventional fitting to the first ECAP-based measurement. This concerns subjects 4, 5, 6, 8, and 13. The first ECAP-based CVC score was taken because this score, without habituation to the new processor fitting, best expresses the effect of changing processor fitting.

Fig. 7 shows that individual T-NEW scores can be almost two score units lower than the T-scores, which implies overall levels lower by 40 current units (Fig. 2). The overall C-NEW levels are up to 0.65 score units higher, thus 13 current units. In some subjects the dynamic range (C- minus T-level) may more than double in the ECAP-based fitting. With respect to speech perception Fig. 7 shows that this is not at all related to the difference between the T- and T-NEW or the C- and C-NEW levels, in spite of the large differences of up to 40 current units for the T-levels.

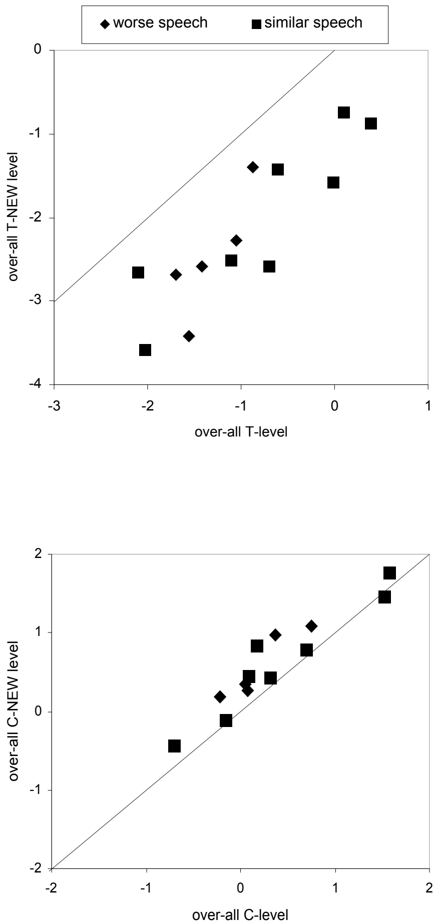
Fig. 8a shows that the T- and T-NEW tilts differ by up to 1.5 score units, which corresponds to changes in slope up to 1 current unit per one electrode distance (Fig. 3). Again, there is no clear relation between speech performance and the difference between the T- and T-NEW scores. The T-tilt in the ECAP-based processor fitting can be one score unit lower than in the conventional fitting (corresponding to a decrease in the slope of the T-profile of 2/3 current units per one electrode distance) without affecting the speech score. Fig. 8b shows that the C- and C-NEW tilts differ by up to 3 score units, corresponding to changes in slope of up to 2 current units per one electrode distance. Moreover, Fig. 8b shows a trend of speech perception performance with the difference in profile tilt. Speech performance deteriorates when the ECAP-based C-profile possesses a greater slope than the conventional C-profile. Thus, speech performance decreases as the ECAP-based fitting yields stronger high-frequency stimulation at the basal electrodes. The worse-speech data point most remote from the one-to-one line in Fig. 8a derives from a subject that is responsible for the data point most remote from the one-to-one line in Fig. 8b. Thus, not the difference in the tilt of the T-profile but the difference in the tilt of the C-profile may well have been responsible for worse speech performance. This suggests that differences in the slope of the T-profiles within 2/3 current unit per one electrode distance (one score unit for the second profile component) can be allowed without affecting speech perception performance.

2.4 Discussion

The major results of the present study are:

1. A substantial correlation ($r = 0.82$) between the ECAP thresholds and the T- and C-levels is found only between the slope of the ECAP thresholds and the slope of the T-levels across the electrode array. The coefficient of correlation between the overall ECAP and T-levels is $r = 0.64$. The C-levels are not related to the ECAP thresholds ($r < 0.4$).
2. The overall T-levels determined with wide-band running speech stimulation, T-NEW, were 25 to 30 current units below the conventional T-levels.
3. Using these lower T-levels in speech processor fitting does not affect the CVC scores.
4. An increase of the slope of the C-profile, such that basal electrodes receive more stimulation by the high-frequency components of the signal, results in a decrease of the CVC score.
5. The role of ECAP thresholds in speech processor fitting has yet to be determined. The high correlation between the slopes of the ECAP thresholds and of the T-levels across electrodes within the present data set does not allow a conclusion as to the importance of the slope in processor fitting. The negative effect on speech perception performance of ECAP-based C-levels with greater slopes across the electrode array than the conventional C-levels may be due to habituation to the conventional fitting.

Chapter 2



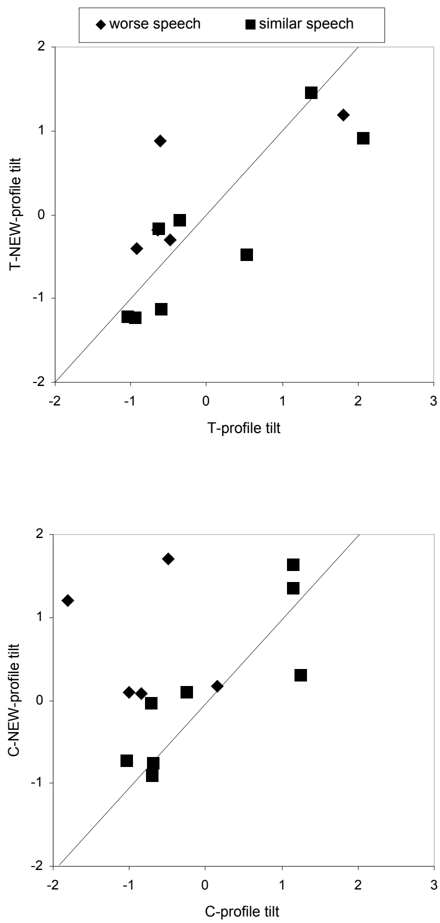


Figure 7: The scores for the first profile component, overall level, for the T- and C-profiles compared to the ECAP-based T-NEW and C-NEW profiles. Diamonds indicate the subjects with 10% worse speech in the first ECAP-based measurement. The subjects with ECAP-based speech performance within 10% of the scores for the conventional fitting are indicated by squares. The solid line represents the one-to-one relation.

Figure 8: The scores for the second profile component, profile tilt, for the T- and C-profiles compared to the ECP-based T-NEW and C-NEW profiles. Diamonds indicates the subjects with 10% worse speech in the first ECAP-based measurement. The subjects with ECAP-based speech performance within 10% of the scores for the conventional fitting are indicated by squares. The solid line represents the one-to-one relation.

2.4.1 Relation between ECAP threshold and T- and C-levels

As mentioned before in the section 'Introduction', Brown et al. (2000) and Hughes et al. (2000) concluded that their correlations between the ECAP thresholds and the T- and C-levels were not strong enough to allow confident use of the ECAP thresholds in speech processor fitting. They showed that the correlation improves considerably when the ECAP thresholds are shifted per subject over a constant amount of current units where the size of this shift was based on a behavioral estimate of either the T- or C-level for one electrode (electrode No. 10). Their 'correction' corresponds to our method of measuring the T- and C-NEW levels. After this correction Brown et al. reported coefficients of 0.83 and 0.77 for the correlations between the predicted and measured T- and C-levels, respectively. Hughes et al. reported 0.85 and 0.89, respectively. However, these correlation coefficients are based on the variance across both subjects and electrodes within subjects. Therefore, it is impossible for the reader to evaluate the extent to which these two factors have contributed individually to the value of the coefficients.

In our approach, we analyzed the correlations across subjects based upon the two principal components providing, in a statistical sense, a full description of the data across the electrode array. The first component, representing overall level, showed coefficients of 0.64 and 0.39 for the correlations between the ECAP thresholds and the conventional T- and C-levels, respectively. These coefficients become 0.87 and 0.94, respectively, when the T- and C-levels are compared to the T-NEW and C-NEW levels. Thus, although differing in absolute value (Fig. 7, in particular the considerable downward shift of the overall T-NEW levels), there still is quite a good correspondence between the conventional and the ECAP-based T-NEW and C-NEW levels across subjects, including the fair amount of intersubject variability. The second component, representing the tilt in the profiles, showed coefficients of 0.82 and 0.36 for the correlations between the ECAP thresholds and the conventional T- and C-levels, respectively. In principle, these correlations do not change with the parallel shifts. The high correlation found for the

T-levels confirms the notion of Brown et al. (2000) and Hughes et al. (2000) that the ECAP threshold may be indicative of the contour of the T-level. The poor correlation found for the C-levels originates with two strong outliers. These outliers produced the worse speech results.

The ECAP thresholds were found on average near the C-levels, which implies that in a number of subjects they were higher than the C-levels. The largest difference was 29 current units above the C-level. This means that the subject concerned accepted a considerable higher loudness during the ECAP measurements than the C-level, representing the loudest comfortable level, in the initial fitting session. Part of this can be attributed to the lower impulse repetition rate used in the ECAP measurement: 80 Hz instead of 250 Hz in the standard fitting procedure. Using data from Shannon (1985, for 100-μs biphasic impulses), we estimate the effect of this change in impulse rate at 2 dB corresponding to about 10 current units. This implies that three of our subjects may have accepted clearly higher loudness levels during the ECAP measurements.

2.4.2 The low T-NEW levels

The T-NEW levels appeared to be lower than the conventional T-levels by 25 to 30 current units. The assumption that the first electrode responding to the speech signal, when the ECAP profile is raised from a sub-threshold level, would determine T-NEW accounted for a difference of about 10 current units. (This difference corresponds to the differences in the T- and ECAP profiles; they do not run perfectly in parallel, Fig. 8a.) With a correction of 10 current units a difference of 15 - 20 current units remains unexplained. This difference may be due to differences in impulse repetition rate. In SPEAK mode the repetition rate is, on average, 250 Hz per electrode; the value used for measurement of the T-level. The maximum rate across all electrodes is 2500 Hz. If there are few peaks in the spectrum the impulses will be distributed over fewer channels than the maximum of 10 and the impulse rate per electrode may increase proportionately. Moreover, if the peaks are broad such that a number of adjacent electrodes are stimulated nearly simultaneously, we may expect spatial integration in the ear. In both cases we may assume that when stimulating with speech the effective impulse rate at some cochlear location could increase to, say, 1200 Hz. Inspection of the data collected within the Cochlear concerted NRT field study (published after publication of this study: Cafarelli Dees et al., 2005) shows that with such an increase of the impulse rate behavioural thresholds may decrease by 19 current units. Hence, this may explain the low T-NEW levels. From this evaluation we may conclude that the T-levels determined for individual electrodes do not represent the threshold for wide-band stimuli.

2.4.3 Low T-levels do not affect the speech scores

A marked result of the present study is found in the absence of any noticeable effect of the low T-NEW levels on the CVC scores. Fig. 7a showed T-NEW levels up to 40 current units (2 score units) below the T-levels, whereas the effects on the CVC scores were smaller than 10% immediately after changing from the conventional T-levels to the ECAP-based T-NEW levels. Since the C-NEW levels are somewhat higher than the conventional C-levels (Fig. 7b), we should keep in mind that the levels at which the speech processor operates effectively in everyday practice might be quite similar with both fittings. The increase in the dynamic range of the ECAP-based fitting, in particular due to the lower T-NEW levels, mainly implies less compression of the speech signal when it is transformed into the electrical stimulus. Fu and Shannon (1998, 2000) and Zeng and Galvin (1999) showed that the compression factor has very little effect on phoneme recognition. This might well explain our results.

The present results imply that one may question the common focus on T-levels in fitting cochlear implant speech processors. For overall level our results clearly demonstrate that it is not self-evident to base processor fitting on the T-levels. But even the tilt in the T-level profile may be relatively unimportant. Since this tilt correlated highly with the tilt in the ECAP thresholds we had only limited differences between these two tilts (Fig. 8a). Whether or not larger differences will affect the speech score cannot be concluded from the results of this study. The insensitivity of the speech scores with respect to the overall T-levels suggest that adjustment of the T-level in processor fitting might be quite uncritical, at least for speech in quiet, the condition tested. For speech perception in noisy situations we expect that the lower T-NEW levels are not likely to reduce the speech scores. Signal-to-noise ratios of +5 to +10 dB, typical in implant use, imply that in the ECAP-based fitting the noise levels will fall in the region below the conventional T-levels. Thus, in the new fitting the noise will produce less electrical stimulation.

The subjective responses of the participants contributed little to the evaluation of the impact of the ECAP-based fitting. Some subjects appreciated the new fitting because it provided better discrimination between soft and loud sounds; others disliked the new fitting because soft sounds became too soft.

2.4.4 Effect of the slope of the C-profile

Fig. 8b clearly indicated that a marked increase of the slope of the C-profile, such that high-frequency stimulation at the basal electrodes is enhanced, might negatively affect the CVC scores. Here we should add that the enhanced slopes always implied an increase of the C-levels at the basal electrodes, not a decrease of these C-levels at the apical electrodes. Thus, the speech scores did not decrease because of a less efficient transfer of low-frequency information but because of more high-frequency stimulation at the basal electrodes. In addition, we should note that this increase was accepted by the subjects at the time C-NEW was measured without forcing the subjects to accept uncomfortably high C-levels. Nevertheless, the poorer speech performers were those who practiced the new fitting the least number of hours (Figs. 5 and 6). They complained that the sharp sounds were irritating when the new fitting was used for longer periods of time. The question arises as to what extent the quality judgement was based on habituation to the conventional fitting. A follow-up crossover experiment suggests itself. The order of the two fittings should be balanced in this experiment.

2.4.5 Final conclusions

The tilt in the ECAP threshold profiles is closely related to the tilt in the T-levels, whereas the overall levels are not. This suggests an ECAP-based fitting procedure using the profile across the electrode array from the ECAP measurements complemented by a subjective assessment of the overall level. Although this can be accomplished as we have presently shown, we have to face the question of how important this is when fitting the speech processor. The CVC scores appear to be quite insensitive to the T-levels.

With respect to the C-levels we found that steeper slopes in the new profiles may decrease the CVC scores. Rather than using the ECAP profile one may therefore consider using a fixed profile independently of the ECAP outcome. However, Fig. 8 b shows that the tilt in the C-profile may vary considerably from one subject to the next with similar covariance of the tilt in the ECAP-based C-profile. Thus, it is not advisable to neglect this tilt. The outcome of the present study does not provide a clue as to how too much tilt in the ECAP-based C-profiles can be avoided. Moreover, it is possible that the worse speech scores are caused by habituation to the conventional fitting. Therefore, further research based upon balanced experimental designs is needed. The results of the present study suggest that this can be done without risking maladjustments that may seriously delay hearing development with the cochlear implant.

In this study we used principal components analysis in order to identify the aspects of the T- and C-profiles important in speech perception. Two components accounted for all variance, measurement error excluded. In contrast to this statistical approach one may encounter in clinical practice individuals with T- or C-levels at certain electrodes clearly deviant from the neighbouring results; outliers that cannot be neglected. This should be taken into account when one considers to use an approach based on a few parameters describing the complete profile across the electrodes. However, we think this approach will be clinically very effective. One may start by estimating the profile parameters from a limited number of measurements and start a user-friendly speech-based processor fitting as suggested here. Later, one might, if necessary, include additional trimming of the T- and C-levels at individual electrodes.

Using the procedure proposed in this study, the ECAP-based CVC scores were remarkably close to scores based on the conventional fitting. Thus, whereas it is impossible to arrive at a conclusion with respect to the proper method of finding the optimal T- and C-levels, we may conclude that the proposed procedure provides a good starting point for ECAP-based speech processor fitting. Four subjects indicated that they wanted to keep the ECAP-based speech processor fitting. Three of them had no more than six months of cochlear implant experience. Six more subjects indicated that the ECAP-based fitting sufficed but that they preferred the conventional fitting. Given these results, promising in view of the probable habituation to the initial fitting, and the large advantage of a simplified fitting procedure, in particular for children, it seems worthwhile to further develop an alternative fitting procedure.

Acknowledgement

The authors are very grateful to drs. D. Cafarelli-Dees and A. F. van Olphen who gave valuable comments to a previous version of this manuscript.

References

Abbas, P.J., Brown, C.J., Shallop, J.K., Firszt, J.B., Hughes, M.L., Hong, S.H., Staller, S.J. (1999). Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear. 20(1), 45-59.

Bosman, A.J. (1989). Speech perception by the hearing impaired. Thesis. Utrecht University, the Netherlands, Chapter 3.

Brown, C.J., Abbas, PJ. and Gantz, B.J. (1990). Electrically evoked whole-nerve action potentials: Data from human cochlear implant users, J. Acoust. Soc. Amer. 88, 1385-1391.

Brown, C.J., and Abbas P.J. (1990). Electrically evoked whole-nerve action potentials: Parametric data from the cat, J. Acoust. Soc. Amer. 88, 2205-2210.

Brown, C.J., Abbas, P.J, Borland, J., Bertschey, M.R. (1996). Electrically evoked whole nerve action potentials in Ineraid cochlear implant users: Responses to different stimulating electrode configurations and comparison to psychophysical responses. J. Speech Hearing Res. 39, 453-467.

Brown, C.J., Abbas, P.J., Gantz, B.J. (1998). Preliminary experience with neural response telemetry in the Nucleus CI24M cochlear implant. Am. J. Otol. 19(3), 320-327.

Brown, C.J., Hughes, M.L., Luk, B., Abbas, P.J., Wolaver, A.A., Gervais, J.P. (2000). The relationship between EAP and EABR thresholds and levels used to program the Nucleus 24 speech processor: Data from adults. Ear Hear. 21(2), 151-163.

Cafarelli Dees, D., Dillier, N., Lai, W.K. et al. (2005). Normative findings of electrically evoked compound action potential (ECAP) measurements using the neural response telemetry (NRT™) of the Nucleus CI24M cochlear implant system. Audiol. Neurotol., 10(2), 105-16.

Fu, Q-J and Shannon, R.V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners, J. Acoust. So. Amer. 104, 2570-2577.

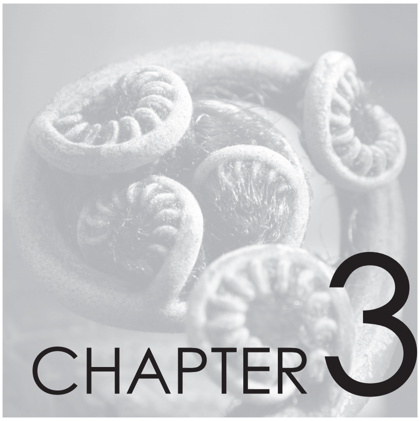
Fu, Q-J and Shannon, R.V. (2000). Effects of dynamic range and amplitude mapping on phoneme recognition in Nucleus-22 cochlear implant users, Ear Hear. 21(3), 227-235.

Hughes, M.L., Brown, C.J., Abbas, P.J., Wolavor, A.A., Gervais, J.P. (2000). Comparison of EAP thresholds with MAP levels in the Nucleus 24 cochlear implant: Data from children. Ear Hear. 21(2), 164-174.

Lai, W.K. (1999). An NRT Cookbook. Guidelines for making NRT measurements. ISBN 3-9521853-0-2.

Shannon, R.V. (1985). Threshold and loudness functions for pulsatile stimulation of cochlear implants. Hearing Res. 18, 135-143.

Zeng, F-G and Galvin III, J.J. (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners, Ear Hear. 20(1), 60-74.



Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T- and C-levels or on compound action potentials thresholds and live-voice speech in a prospective balanced crossover study

Christina Willeboer and Guido F. Smoorenburg
Ear and Hearing 2006, 27(6): 789-98.

Abstract

Objective: The objective of the present study is to improve the efficiency of the fitting procedure of cochlear implant processors by making use of measurements of the electrically evoked compound action potential (ECAP) and live-voice speech.

Design: In a balanced prospective crossover design we compare speech performance of eighteen adult subjects when following the

conventional fitting procedure to a procedure in which we use the profile of the ECAP threshold levels across the full electrode array measured intra-operatively. The overall level of the profile is shifted (by an equal amount of current units per electrode) until we find the threshold for live-voice speech (new T-levels) and the loudness comfort level (new C-levels). Each fitting procedure is tested for six weeks. The first fitting procedure is repeated at 12 weeks. Speech performance is measured in quiet and in noise every other week.

Results: The results show little difference between the scores (Dutch CVC words) for the conventional fitting procedure and the ECAP based fitting, although the T and C-levels may differ markedly.

Conclusion: The new fitting procedure is much faster and easier in the initial phase. Further improvement of performance may be obtained in a later stage of the fitting procedure by adjustment of the T and C-levels of individual electrodes.

3.1 Introduction

Fitting the speech processor of a cochlear implant (CI) is a time-consuming task. In the Nucleus CI24 system the conventional fitting requires a subjective estimation of the threshold level (T-level) and the comfortable loudness level (C-level) for each of the 22 intracochlear electrodes applying short sound bursts. In toddlers and infants it may be laborious to obtain these behavioural measurements, as the tone bursts may be meaningless to the child. Perception of these tone bursts is influenced by the child's cognitive maturation and the ongoing development of audition. Even with some adults, especially those who have been deaf for a long period of time, it may be difficult to get reliable responses within a restricted time.

Objective measures of the auditory system's response to electrical stimulation may be used to facilitate the speech processor fitting process. One of these measures is the electrically evoked compound action potential (ECAP) (Brown and Abbas, 1990), which can be measured with the neural response telemetry (NRT) system developed by Cochlear and the University of Zurich (Abbas et al., 1999, Dillier et al., 2002). This non-invasive method consists of sending an electrical signal to one of the intracochlear electrodes and recording the ECAP using one of the adjacent electrodes.

Several studies have focused on the prediction of behavioural T and C-levels from ECAP thresholds. The first large-scale studies indicated that in adults there is a significant, but moderate correlation between the ECAP thresholds and the T-levels (r = 0.55) and C-levels (r = 0.57) from a conventional fitting (Brown et al., 2000). In children, Hughes et al. (2000) found somewhat higher coefficients: r = 0.70 and 0.71 respectively. In both studies, ECAP thresholds showed relatively small variability across adjacent electrodes. These data were replicated in a number of studies. Thai-Van et al. (2001) found a good correlation between ECAP thresholds and behavioural levels for the apical electrodes (r = 0.70 – 0.90), whereas there was no significant correlation for the basal electrodes. Gordon et al. (2002) found only a weak correlation between ECAP thresholds and behavioural T- and C-levels.

The results of these studies imply that the relation between ECAP thresholds and behavioural responses is not strong enough to allow for an accurate prediction of behavioural T- and C-levels in individual cochlear implant users. More recent work focuses on how conventional and ECAP-based fittings affect speech perception. Three ECAP-based fitting methods have been incorporated in the Nucleus R126 fitting software (Cochlear). Seyle and Brown (2002) proposed the 'T/C offset fitting', in which ECAP thresholds are combined with the behavioural T- and C-level measurement at one electrode in the centre of the array or a set of distributed electrodes, to create an entire fitting. Speech perception scores showed no difference at a low speech level (55 dB SPL), but at 70 dB SPL the ECAP-based fitting was inferior to the conventional fitting. The second method is the 'Progressive pre-set fittings' by Almqvist (Reference Note 1). It makes use of four pre-set fittings based on the profile of the ECAP thresholds with increasing stimulus level that are sequentially presented to the patient. Results were obtained only for young children. Sound field aided thresholds were around 25-35 dB HL. Children fit with the progressive pre-set fittings method scored high on the Meaningful

Auditory Integration Scale (MAIS). The third method is the ‘Shift and tilt approach’, developed by our group (Smoorenburg et al., 2002). It was shown by principal component analysis that the profiles of ECAP thresholds and the conventional T- and C-levels across the full electrode array are governed by two factors. The first factor, overall level (termed shift), accounts for 90% of the variance. Inclusion of the second factor, roughly the slope (termed tilt), accounted for more than 95% of the variance.

Previously (Smoorenburg et al., 2002), we designed a fitting method, in which the profile of ECAP thresholds across the electrode array was used to determine new T- and C-levels using live-voice speech. The overall level of the profile was shifted (by equal amounts for each electrode) until we found the threshold (ECAP-based T-levels) and the loudness comfort level (ECAP-based C-levels) for live speech (see also the section below on the ECAP-based fitting procedure). The correlation between the overall level of ECAP thresholds and T- and C-levels across subjects was weak. However, the results for CVC word lists showed little change in speech perception when the T- and C-levels were switched from the conventional ones to the ECAP-based ones. On average, the speech scores decreased by about 7 percentage points. Most subjects indicated that the ECAP-based fitting sufficed. Those who appreciated the new fitting had been using their cochlear implant only for about six months prior to the study.

Since the majority of the subjects in this previous study had been using their conventional fitting for more than six months, whereas the acclimatisation period for the new ECAP-based fitting was only two weeks, lack of habituation to the ECAP-based fitting may have been a factor in the outcome of the speech perception scores. Taking this factor into consideration, the results of our initial study were sufficiently promising to proceed with a prospective study, particularly because the ECAP-based fitting is a fast and easy method. This study concerns a balanced crossover trial, in which new patients alternately start with either the conventional or the ECAP-based fitting and switch to the other fitting method after six weeks. Since in the previous study tilt was only a small factor in terms of explained variance, we decided to allow only shifts in the ECAP-based fitting. Outcome measures are speech perception scores as well as subjects' subjective reports.

3.2 Methods

3.2.1 Study outline

The present study concerns a prospective balanced crossover design, in which all adult patients who received a cochlear implant between February 2003 and March 2004 in the University Medical Centre Utrecht, with full insertion of the electrodes and ECAP measurements during surgery on at least 20 electrodes, were included. The study design is depicted in Table 1. Alternately subjects started with a conventional fitting or with an ECAP-based fitting. After six weeks of using the first fitting method, there was a crossover to the other fitting method, again for six weeks. Fittings were repeated every week. In order to alleviate the crossover to the different processor adjustment, patients received both fittings in their speech processor in the first week after the crossover. However, they were encouraged to use only the new fitting and they all complied. Speech perception was measured every other week, at 2, 4, and 6 weeks with the first fitting method and at 8, 10, and 12 weeks with the second fitting method. Primarily, the results at 6 and 12 weeks were compared. In view of the possibility of substantial habituation to electrical stimulation during the second period of six weeks we measured speech perception after twelve weeks not only using the second fitting method, but also repeating the first one (condition 12R). Prior to the speech test we repeated the fitting according to the first method.

week	CVC test			group 1	group 2
	65 dB quiet	65 dB noise	55 dB quiet	fitting	fitting
1				Con-ven-tio-nal	ECAP
2	x	x			
3					
4	x	x			
5					
6	x	x	x		
7				ECAP	Con-ven-tio-nal
8	x	x			
9					
10	x	x			
11					
12	x	x	x		
12R	x	x	x	conventional	ECAP

Table 1: Study design.

3.2.2 Subjects

Eighteen adult subjects entered the study. They received their Nucleus® 24 Contour™ cochlear implant in our centre and used it with an Esprit 3G speech processor. Subjects that started with the conventional fitting are named C1 to C9 and belong to 'group 1'. Subjects that started with the ECAP-based fitting are named E1 to E9 and belong to 'group 2'. Subject characteristics are listed in Table 2.

Group	Subject	Sex	Age (y)	Deafness					Contra lateral HA	Electrode anomalies	Electrode
				etiology	duration (y)		PTA 0.5, 1, 2 kHz (dB)				
					ipsi	contra	ipsi	contra			
1	C1	M	35	viral infection	2	14	120	>120	N		CS
	C2	F	42	hereditary	5	6	120	110	N		CS
	C3	F	73	unknown	4	4	>120	100	Y		CS
	C4	F	62	unknown	>25	>25	110	110	Y		CA
	C5	M	60	meningitis	1	5	100	>120	Y		CA
	C6	F	49	unknown	2	RH	105	90	Y		CA
	C7	F	27	hereditary	27	27	100	95	Y		CA
	C8	M	52	hereditary	15	15	>120	>120	N		CS
	C9	M	75	ototoxicity	2	2	90	85	Y		CA
2	E1	M	32	meningitis	7	RH	110	105	Y		CS
	E2	F	44	hereditary	1	4	>120	>120	N		CS
	E3	F	37	unknown	32	32	110	110	N		CS
	E4	F	56	hereditary	17	17	110	110	N	22 open 18-19 short	CS
	E5	F	74	unknown	1	3	100	90	N		CS
	E6	F	51	hereditary	10	10	105	100	Y		CA
	E7	F	41	unknown	RH	29	110	>120	N		CS
	E8	F	80	Meniere's disease	6	23	95	100	N		CA
	E9	F	30	Fechtner syndrome	5	5	120	110	Y		CA

Table 2: Subject characteristics. C1-9 started with the conventional fitting, E1-9 started with the ECAP-based fitting. RH indicates progressive hearing loss with residual hearing, CS indicates Nucleus 24R Contour, CA indicates Nucleus 24R Contour Advance.

3.2.3 ECAP measurements

Immediately after surgery, during the closing of the wound, impedance and ECAP measurements were performed on all electrodes using the NRT™ (Neural Response Telemetry) software version 3.0 from Cochlear.

ECAPs were measured following the standard procedure described by Lai (1999). They were measured for all electrodes using a recording electrode positioned two electrodes more apically than the stimulating electrode. For electrode 21 and 22, recording electrodes were 19 and 20, respectively. The stimulation mode was monopolar (MP1 mode, using the extracochlear ball reference electrode). When recording, the metal housing of the implant was used as the reference electrode (MP2 mode). A fixed masker offset level of 10 current units (CU) above the probe level was used. Pulse duration was set at 25 µs/phase and stimulation rate was 250 Hz. The number of sweeps was 100. Masker advance was fixed at 500 µs. The highest current level used was at least 15 CU above the visual ECAP threshold.

To select the optimal settings of the recording parameters gain and delay, they were varied in a series of recordings from electrode 17 while stimulating at electrode 15 at a clearly suprathreshold level. Settings were selected that yielded the largest and smoothest ECAP waveform without amplifier saturation, preferring a gain of 60 dB and a delay of 90-120 µs when the results for this setting differed little from other settings. The optimal gain and delay were unchanged during subsequent measurements.

To determine ECAP amplitudes, markers were manually set at the negative N1 and positive P1 peaks in the response pane with both low and high resolution baseline corrected components. An amplitude growth function was determined per electrode. At low stimulation levels there is a deviation from linear growth, meaning that the amplitude of the ECAP decreases only slightly with decreasing stimulation levels (Cafarelli Dees et al., 2005). Therefore, we defined the stimulus level at which an ECAP amplitude of 40 µV is reached as the ECAP threshold, instead of the intercept of the linear part of the amplitude growth function with the

X-axis. The amplitude of 40 μV implies a response just above the noise level. The shape of the ECAP thresholds across the whole electrode array, irrespective of overall level, is called the *profile* of the ECAP thresholds.

3.2.4 General aspects of the fitting procedure

Fitting the speech processor was done using the Cochlear Windows Diagnostic and Programming System (WinDPS R126 version 2.0) software with the Clinical Programming System (CPS). All subjects were given the ACE strategy with a per-channel stimulation rate of 900 Hz. Pulse duration was 25 μs/phase. Stimulation mode was monopolar using both extracochlear reference electrodes (MP1+2). Stimulation amplitude is expressed in current level (CL), a quantity defined by Cochlear. The CL ranges from 1 to 255 CU, which corresponds to electrical currents from 10 μA to 1.75 mA. The relation between CU and electrical current is approximately logarithmic with 34 CU corresponding to a factor of 2 in electrical current or 6 dB.

3.2.5 ECAP-based fitting procedure

In the first ECAP-based fitting, the profile of the ECAP thresholds for electrodes 3 – 22 was manually entered into the T- and C-levels fields of the 'Psychophysics' tab sheet of the WinDPS R126 software. During the whole procedure, the shape of this profile across the electrode array was maintained; only vertical shifts were applied. The C-profile was initially set at the ECAP profile and the T-profile was set 1 CU lower, resulting in a dynamic range of 1 CU. Subsequently, the T and C-profiles were shifted down by equal amounts of current units, so that the lowest T-level was 30 CU. Next, in live mode, the T and C-profiles were shifted upwards by equal amounts of current units until the subject reported that live voice speech was just audible. The ECAP-based T-profile was set at this level. Next, the C-profile was shifted upward until a comfortable loudness level was reached, again using live voice speech as the input signal for the speech processor. The ECAP-based C-profile was set at this level. In subsequent fittings, the threshold level for live voice speech was determined again using a dynamic range of 1 CU, after which the C-profile was shifted upward until a comfortable loudness level was reached.

3.2.6 Conventional fitting procedure

In the conventional fitting, subjective T- and C-levels were measured for all electrodes 3 – 22 individually stimulating with 500 ms bursts of pulses repeated at 900 pulses/s. The T-level was defined as the lowest stimulus level per electrode that elicits a very soft, but consistent hearing sensation. The C-level was defined as the maximum stimulus level per electrode that still produces a comfortable loudness sensation. In addition the C-levels were balanced across electrodes for equal loudness.

3.2.7 Principal components analysis

As in the previous study (Smoorenburg et al., 2002), we analysed the number of independent components describing the fitting data using the statistical method of principal components analysis (PCA). The previous study showed that with this technique the whole fitting data set of 20 T-levels and 20 C-levels per fitting can be reduced to two factors describing the shape of the T- and C-levels across the electrode array. Using the PCA for fitting data reduction allows for a comparison of the speech results and the principal characteristics of the T- and C-levels across the electrode array described by component 1 and component 2. In the previous study component 1 accounted for 90% of the variance and related closely to the overall level of the profile, that is the average CL over the electrodes for a specific profile. Since the PCA units are rather abstract we relate them to current units. One unit of component 1 corresponded to approximately 20 CU. Component 2 related closely to the slope of the profile. One unit of the score for this component corresponded to a slope of about 0.67 CU per electrode distance.

3.2.8 Speech perception measurements

Speech material consisted of Dutch (linguistically meaningful) words, each consisting of the sequence consonant-vowel-consonant (CVC words), uttered by a female voice. The material was played from a compact disc. Subjects were allowed to give any answer and were strongly encouraged to respond to each word presented. These words were presented in three conditions: at 65 dBA in quiet, at 65 dBA in 55 dBA continuous noise with a speech-shaped spectral energy distribution, and at 55 dBA in quiet. Speech and noise were presented via the same loudspeaker. Presentation levels were calibrated at the microphone position of the speech processor. The first and second condition were presented every other week, whereas the third one was only presented at six and twelve weeks. For each condition 8 lists of 12 words were used. As there were 45 word lists available, lists were repeated after six weeks. A contralateral acoustic hearing aid, if present, was switched off.

3.2.9 Questionnaire

At the beginning of each visit, the audiologist interviewed the subject about the previous week's experiences following a standard questionnaire. The questionnaire covered questions about the loudness and quality of sounds with different spectral energy distributions, *e.g.* running water, traffic, crockery and cutlery. Additionally, there were questions about the duration of daily use and the frequency of the patient's volume adjustments.

3.3 Results

3.3.1 T- and C-levels

The T- and C-levels of the conventional and the ECAP-based fittings, and the ECAP thresholds, averaged across all eighteen subjects, are displayed in Figure 1. The T- and C-levels represent the results of each fitting method at 12 weeks (second fitting or repeated first fitting). Standard deviations for conventional T- and C-levels range from 20 to 27 CU, respectively.

The mean ECAP thresholds and the conventional and ECAP-based C-levels roughly coincide, except for the most apical and basal electrodes. At the apical and basal end of the electrode array, averaged conventional C-levels are approximately 10 CU below average ECAP-based C-levels. Across the whole electrode array the averaged ECAP-based T-levels are approximately 30 CU below the averaged conventional T-levels.

Figures 2 and 3 display individual T- and C-levels for all active electrodes in the conventional and the ECAP-based fitting at 12 weeks. All but one ECAP-based T-levels are below the conventional ones, with a maximum of 64 CU difference. The maximum differences between conventional and ECAP-based C-levels are -43 CU and 28 CU.

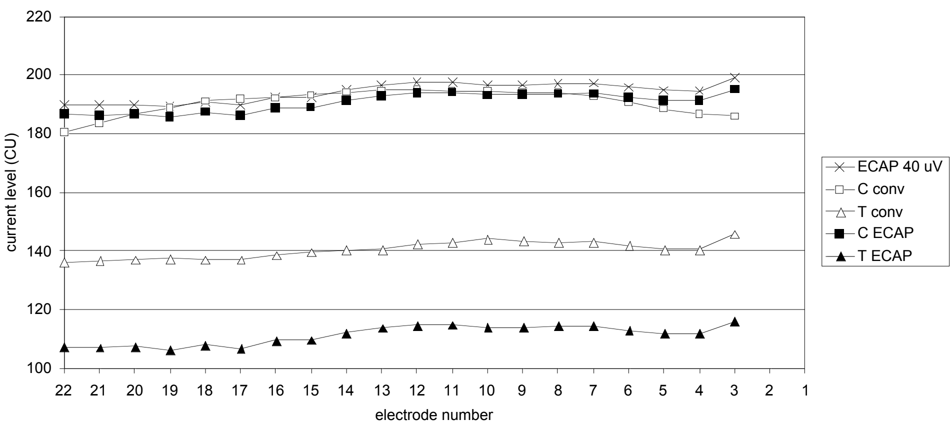


Figure 1: Conventional T- and C-levels, ECAP-based T- and C-levels and ECAP thresholds (amplitude of 40 μV) averaged over eighteen subjects.

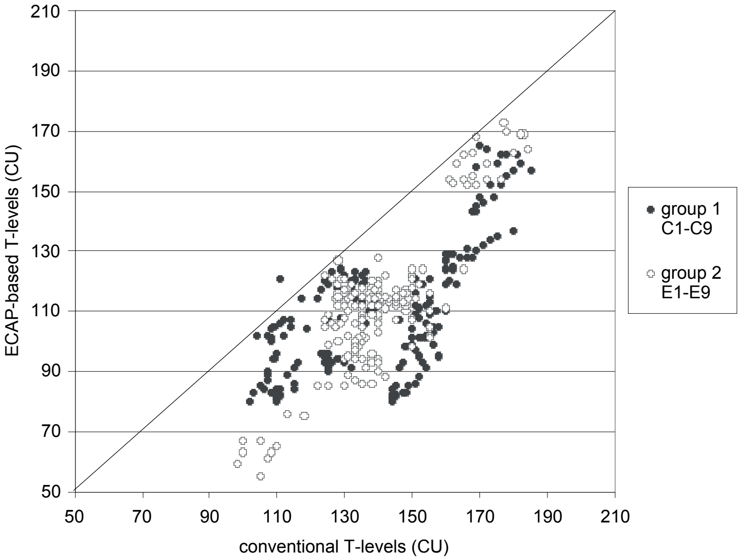


Figure 2: Individual T-levels of all active electrodes in the conventional and ECAP-based fitting at 12 weeks (second fitting or the repetition of the first fitting).

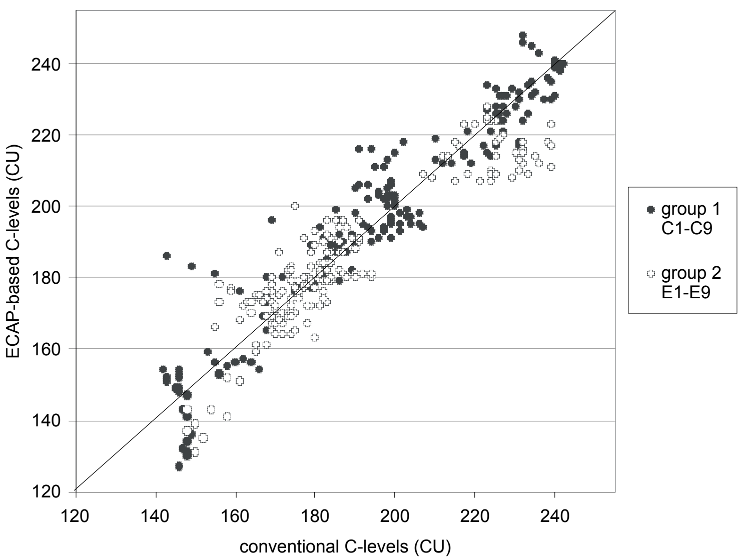


Figure 3: Individual C-levels of all active electrodes in the conventional and ECAP-based fitting at 12 weeks (second fitting or the repetition of the first fitting).

3.3.2 Principal components analysis of the T- and C-levels and ECAP thresholds

Analyzing the ECAP thresholds and the conventionally measured T- and C-levels collected at 12 weeks (either with the second fitting or the repeated first fitting) using PCA it shows that two components describe the fitting data to a high degree. These two components explain 92, 95 and 98% of the total variance of the ECAP thresholds and the T- and C-levels, respectively. As in the previous study (Smoorenburg et al., 2002), the first component closely relates to the overall level of the data. It accounts for 75, 82, and 94% of the variance, respectively. The second component again relates approximately to the slope of the profiles across the array, determined by a linear regression fit. These results are very similar to those of the previous study.

In order to compare all T- and C-levels we repeated the principal components analysis on the complete set of T-levels collected at 12 weeks (the ECAP-based ones and the conventional ones, either with the second fitting or the repeated first fitting) and on an identical complete set of C-levels. The result was similar: 97% of the variance in the T-levels and 98% in the C-levels was explained by two components of which 92% and 94%, respectively, by the first component. The first component shows a perfect correlation with the average level across the electrode array. One unit of this component corresponds to a shift in overall level of 5.4 CU for the T-levels and 5.6 CU for the C-levels. The second component shows a strong correlation with the slope of the T and C-profiles (although these profiles are not linear). One unit of component 2 corresponds to a change in slope of 0.87 and 0.89 CU per electrode distance in the T and C-profiles, respectively. In the section below 'Speech perception in relation to the profile characteristics' we will use the components to analyse the effect of the differences in the T and C-profiles of both fitting methods on the speech perception scores.

3.3.3 Speech test performance

Figure 4a shows the speech perception scores at the 65 dBA in quiet condition of all eighteen subjects measured at 2, 4, and 6 weeks with the first fitting method, at 8, 10 and 12 weeks with the second fitting method and at 12 weeks with the repetition of the first fitting method. Figures 4b and 4c show the results for the 65 dBA in noise and 55 dBA condition, respectively. Of the eighteen subjects, fifteen were able to perform speech perception tests at 65 dBA in quiet and at 65 dBA in 55 dBA speech noise at all six test moments. Two subjects, C4 and C5, were not able to perform the CVC test in any of the speech conditions at 2 weeks. One subject, E3, was not able to perform the CVC test during any of the 12 weeks of the study. Due to an early onset and a very long period of deafness without residual hearing (i.e. without speech discrimination and without using hearing aids), her speech perception in the first three months after cochlear implantation was too poor to perform the tests. This subject was excluded from statistical analysis.

Figures 4a-c show a clear learning effect, which is to be expected in the first weeks after the initial fitting. Statistical analysis of the crossover design (Armitage and Berry, 1994) shows that the mean difference over all subjects between the 6-week and 12-week results is highly significant in all three speech conditions (p = 0.0008, 0.0002 and 0.02 for the speech conditions 65 dBA in quiet, 65 dBA in noise and 55 dBA in quiet, respectively). In the crossover design this is the ‘order’ effect. The average learning effect amounts from 7.0 percentage points in the 65 dBA in quiet condition to 8.5 percentage points in the 65 dBA in noise condition.

In Figure 4a-c the average phoneme scores of group 2 (ECAP-based fitting first) seem to be higher than those of the group 1 (conventional fitting first), both with the first and the second fitting method. However, this difference appears to be statistically insignificant at the 5% level for all of the three speech conditions.

Analysis of the difference between average phoneme scores with the ECAP-based fitting and the conventional fitting shows that in the 65 dBA speech condition, both in quiet and in noise, there is no significant difference between the speech results with both fitting methods on all test moments. In terms of the crossover design, there is no ‘treatment’ effect in these conditions. Also in the 55 dBA in quiet condition, there is no significant difference between the phoneme scores with the first fitting method at 6 weeks and the second fitting method at 12 weeks. However, Figure 4c shows that in group 2, when changing from the conventional measurement at 12 weeks to the repetition of the ECAP measurement at 6 weeks, the ECAP-based fitting phoneme scores are substantially lower than the conventional fitting scores of the same group. In the analysis of the average phoneme scores at 12 weeks, the difference between the ECAP-based and the conventional fitting method reaches the 5% significance level. The conventional fitting yielded the highest scores, whereas the repetition of the ECAP fitting yielded a substantially lower score than the conventional fitting at 12 weeks.

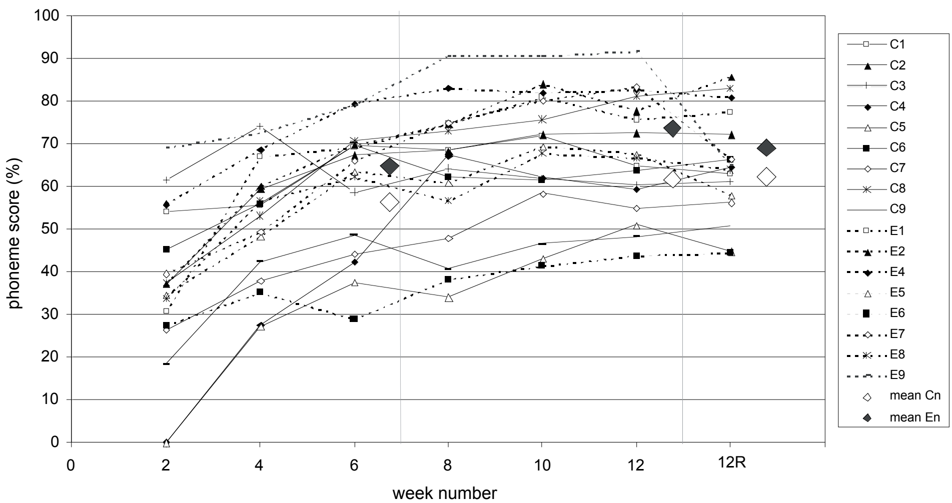


Figure 4a

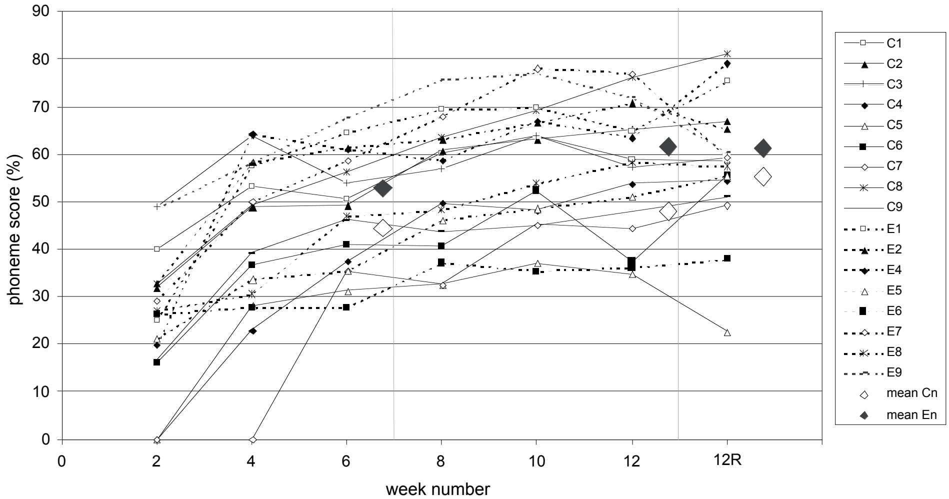


Figure 4b

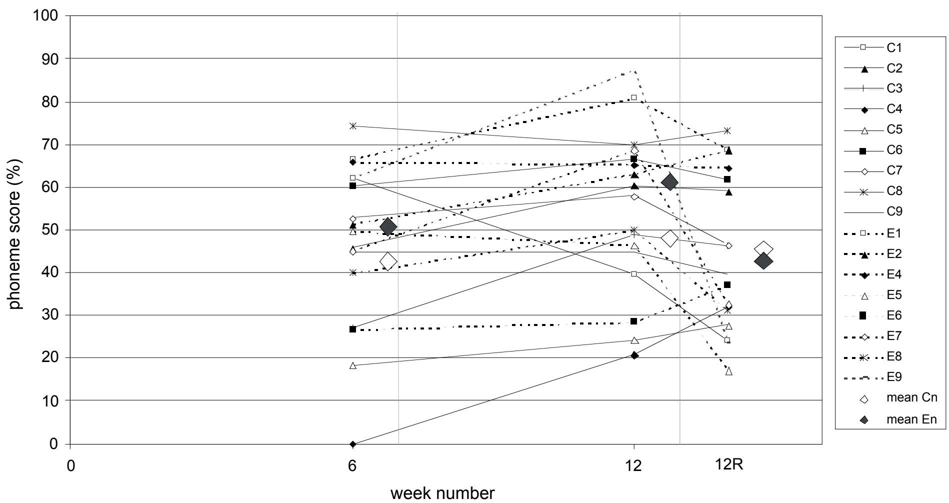


Figure 4c

Figure 4: Individual CVC phoneme scores measured at 2, 4, and 6 weeks with the first fitting method, at 8, 10, and 12 weeks with the second fitting method and at 12 weeks with the repetition of the first fitting method (12R). Solid lines indicate subjects in group 1 (conventional fitting first, subjects C1-C9 in Table 2), dashed lines indicate subjects in group 2 (ECAP-based fitting first, subjects E1-E9 in Table 2). Diamonds indicate the average phoneme score of each group at various moments.
Panel a shows the results of the 65 dBA in quiet condition, panel b shows the results of the 65 dBA in noise condition, and panel c shows the results of the 55 dBA condition.

3.3.4 Speech perception in relation to the profile characteristics

From Figure 1 it is clear that the overall level, i.e. component 1, of the average ECAP-based T-levels is much lower than the overall level of the average conventional T-levels. The differences in individual phoneme scores found for the two fitting methods, after correction for the learning effect, could be related to differences in overall level, i.e. component 1, in particular those found for speech at 55 dBA in quiet. To correct for the learning effect when comparing the result at 6 and 12 weeks, we added the average difference over the group between both measurements for each speech condition separately to the individual results at 6 weeks. However, comparing the corrected phoneme scores in the 55 dBA in quiet speech condition at 6 and 12 weeks, we found no significant correlation with component 1 of the T and C-profiles. Thus, low speech scores at 55 dBA presentation level are not related to low overall T-levels. Considering component 2 in the same way, there were no significant correlations at the 5% level. Thus, differences in tilt did not affect the speech scores at 6 and 12 weeks in a systematic way. Also, there was no significant correlation with component 1 or component 2 and the phoneme scores collected at 12 weeks with the second fitting and the repetition of the first fitting.

3.3.5 Subjective appreciation

With respect to sound quality most subjects appreciated both fitting methods. The crossover to the second fitting method generally gave no problems. If after the crossover from the ECAP-based fitting to the conventional fitting there were any habituation problems mentioned, they related to soft sounds being a bit louder in the conventional fitting. Two subjects reported that speech understanding in noise was somewhat more difficult immediately after the crossover, however, after two weeks of habituation, this was no longer apparent in the CVC test. After the crossover from the conventional fitting to the ECAP-based fitting, four subjects reported that sound quality, independently of speech perception, decreased. One subject reported a booming sound quality with the ECAP-based fitting, whereas three subjects indicated that the ECAP-based fitting had a sharper sound quality. These complaints disappeared after one or two weeks of using the ECAP-based fitting. In addition, there were no differences between both fitting methods neither in the duration of daily use, nor in the number of volume adjustments they made during the day.

After the entire study period of twelve weeks, the subjects could choose which fitting they preferred for continued use. As speech perception was generally almost equal in the two fittings, both in the CVC test and in the subject's experience of real-life situations, most subjects made a choice based on sound quality. Fourteen of the eighteen subjects continued using the second fitting. In eight subjects this was the conventional fitting, in six the ECAP-based fitting. Two subjects continued using both fittings. Two subjects, one from each group, switched to the first fitting again after the end of the study period, mainly for sound quality reasons. Thus, there was no clear preference for either the conventional or the ECAP-based fitting.

3.4 Discussion

3.4.1 T- and C-profiles

In accordance with our previous study (Smoorenburg et al., 2002), the most striking difference between the conventional and the ECAP-based fitting is the low overall level of the ECAP-based T-levels. Wide-band running speech, which stimulates multiple channels nearly simultaneously resulting in an effectively higher repetition rate due to integration across electrodes, leads to lower thresholds than an impulse train of 900 pulses/s that is presented to one single channel in the conventional fitting procedure. When determining ECAP-based T-levels using live-voice speech, all subjects showed a large range of CU over which speech was detectable but very soft. After the profile of the T-levels was set at the lower edge of that range, C-levels could be increased over a large range of CU with only little, but nevertheless discernable, loudness growth.

The ECAP thresholds, conventional C-levels, and ECAP-based C-levels in this study roughly coincide, whereas in other studies (e.g. Cafarelli Dees et al., 2005, Dillier, Reference Note 2), ECAP thresholds are found at about 70% of the dynamic range in a stable fitting. The difference may be due to a number of factors. First, the short study period of only twelve weeks after the first fitting implies that the fitting may not be stable yet. The C-levels may continue to rise. Second, in most studies ECAP thresholds are determined through visual inspection or as the intercept of the linear part of the amplitude growth function with the baseline, whereas the criterion applied in the present study is the stimulus level at which an ECAP amplitude of 40 µV is reached. Our definition results in higher ECAP thresholds. Third, in this study ECAP thresholds were measured at a stimulation rate of 250 Hz, whereas the most previous studies used 80 Hz. Higher rates result in smaller amplitudes and therefore higher thresholds (Charasse et al., 2004). However, the normalized effect was only up to 4% of the CU of the ECAP thresholds.

We showed that the electrodes at the array boundaries have higher C-levels for the ECAP based fitting than for the conventional fitting. Subgroup analysis shows that this finding is restricted to group 1 (conventional fitting first), and is strong enough to show up as an effect in the whole group. At electrode 22 the average ECAP-based C-level in group 1 is 8 CU above the average conventional C-level in the same group. At electrode 3 the average ECAP-based C-level in this group is 15 CU above the average conventional C-level. In group 2, none of these differences are found. These results suggest that due to the shape of the profile of the ECAP thresholds subjects get used to stronger low and high frequency stimulation in the ECAP-based fitting as a result of which they have a higher tolerance for the lowest and highest frequencies in the subsequent conventional fitting.

3.4.2 Longitudinal aspects

During the study period, the growth of the speech perception scores followed on average the same pattern for both fitting methods. Most subjects achieve a more or less stable high score after six to ten weeks, independently of the fitting method they started with. At the crossover moment after 6 weeks there was little change in the speech scores and the same was found for speech at 65 dBA in quiet and in noise when switching acutely from one fitting method to the other at 12 weeks. At the end of the study period, most subjects chose to continue using the fitting they were last familiar with, as speech performance was in general equal with both fitting methods.

3.4.3 Speech perception

Comparing the corrected phoneme scores we could not show a difference between groups nor at 6 weeks nor at 12 weeks using the last fitting and using the first fitting method (12R) for the 65 dBA speech level, both in quiet and in noise. This result was found in spite of large differences in overall current levels reaching up to 60 CU for T-levels and 25 CU for C-levels and differences in slope of 2 and 1.4 CU per electrode distance for the T- and C-levels, respectively. The results for the 55 dBA in quiet condition were different. There was no difference between groups in the speech perception scores at 6 weeks using the first fitting method and at 12 weeks using the second fitting method, but the switch at 12 weeks to the repetition of the first fitting resulted in significantly lower scores for the ECAP-based fitting. The previous analysis showed that this was specifically due to switching back from the conventional fitting to the ECAP-based fitting (Figure 4c). We feel that this is due to the lower T-levels found in the ECAP-based fitting procedure although we could not reach a secure level of statistical significance for this relation. Other studies, like Skinner et al. (1999) found that raising the T-levels of a conventional fitting results in a better speech perception score at low presentation levels. However, we saw the decrease in speech perception using the ECAP based fitting only in the group that switched from the conventional to the ECAP based fitting at 12 weeks and not in the group that had the ECAP based fitting at 12 weeks and switched to the conventional fitting.

3.4.4 Conclusion

In this group of eighteen consecutively implanted patients we showed that the ECAP-based fitting method yields speech perception scores that are equal to those obtained with a conventional fitting. Subjective sound quality was acceptable to good in both the conventional and the ECAP-based fitting. To our experience the ECAP-based method is fast and easy. Therefore, it is appropriate to start with the ECAP-based fitting to quickly obtain an adequate fitting. Further improvement of performance may be obtained in a later stage of the fitting procedure. Improvement may result from increasing the T-levels and by adjustment of the T- and C-levels of individual electrodes.

Acknowledgement

The authors gratefully acknowledge critical comments and suggestions given by Dr. G. A. van Zanten and Dr. B. van Dijk and contributions from Drs. N. Dillier and P. van de Heyning to the design of the experiment.

References

Abbas, P.J., Brown, C.J., Shallop, J.K. et al. (1999). Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear., *20,* 45-59.

Armitage, P. and Berry, G. (1994). Statistical Methods in Medical Research *(*3rd edition). Blackwell Scientific Publications, pp. 245-49.

Brown, C.J. & Abbas, P.J. (1990). Electrically evoked whole-nerve action potentials: Parametric data from the cat. J. Acoust. Soc. Am., 88*,* 2205-2210.

Brown, C.J., Hughes, M.L., Luk, B. et al. (2000). The relationship between EAP and EABR thresholds and levels used to program the Nucleus 24 speech processor: Data from adults. Ear Hear., 21*,* 151-163.

Cafarelli Dees, D., Dillier, N., Lai, W.K. et al. (2005). Normative findings of electrically evoked compound action potential (ECAP) measurements using the neural response telemetry (NRT™) of the Nucleus CI24M cochlear implant system. Audiol. Neurotol., 10(2), 105-16.

Charasse, B., Chanal, J.M., Berger-Vachon, C. et al. (2002). Influence of stimulus frequency on NRT recordings. Int. J. Audiol., *43(4),* 236-44.

Dillier, N., Lai, W.K., Almqvist, B. et al. (2002). Measurement of the electrically evoked compound action potential via a neural response telemetry system. Ann.Otol. Rhinol.Laryngol*., 111(5 Pt 1),* 407-14.

Gordon, K.A., Ebinger, K.A., Gilden, J.E. et al. (2002). Neural response telemetry in 12- to 24-month-old children. Ann.Otol.Rhinol.Laryngol Suppl., 189*,* 42-8.

Hughes, M.L., Brown, C.J., Abbas, P.J. et al. (2000). Comparison of EAP thresholds with MAP levels in the Nucleus 24 cochlear implant: Data from children. Ear Hear., 21, 164-174.

Lai, W.K. (1999). An NRT Cookbook. Cochlear AG.

Seyle, K., Brown, C.J. (2002). Speech perception using maps based on neural response telemetry measures. Ear Hear., 23(1 Suppl), 72S-79S.

Skinner, M.W., Holden, L.K., Holden, T.A. (1999). Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. J. Speech Lang. Hear. Res., 42, 814-28.

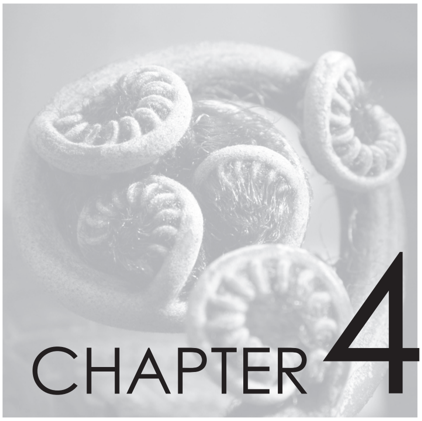
Smoorenburg, G.F., Willeboer, C., van Dijk, J.E. (2002). Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds. Audiol. Neurotol., 7(6), 335-47.

Thai-Van, H., Chanal, J.M., Coudert, C. et al. (2001). Relationship between NRT measurements and behavioral levels in children with the Nucleus 24 cochlear implant may change over time: preliminary report. Int. J. Ped. Otorhinolaryngol., 58(2), 153-62.

Reference notes

Almqvist, B. (2004). The Almqvist/Lund procedure for adjusting a speeck processor for a congenitally deaf child using objective measurements. Cochlear Nucleus Report July/Agust, 3-4.

Dillier, N. (2000). Objective measures in cochlear implantation: methods and applications. Plenary session, 5th European Symposium on Paediatric Cochlear Implantation, Antwerp, Belgium, June 2000.



Parametric adjustments to the ECAP-based fitting by cochlear implant recipients during everyday life

Christina Willeboer[1], Guido F. Smoorenburg[1], Bas van Dijk[2], René D. Roos[2], Gijsbert A. van Zanten[1]

Re-submitted to Ear and Hearing after revisions

1: University Medical Center Utrecht, The Netherlands
2: Cochlear Technology Centre Europe, Belgium

This study was supported by Cochlear Europe.

Abstract

Objective: Previous research has shown that a fitting procedure based on thresholds of the electrically evoked compound action potential (ECAP) and live speech is a fast and easy alternative to the conventional fitting procedure, which is based on subjective thresholds (T-levels) and loudness comfort levels (C-levels), determined for each electrode (Willeboer and Smoorenburg, 2006). T-levels, C-levels, and ECAP thresholds across the electrode array, the "profiles", can be largely described by two parameters; shift (corresponding to overall level) and tilt (corresponding to the slope of the profile). The objective of this study is improving the ECAP-based fitting procedure by giving the cochlear implant recipients themselves the possibility to adjust shift and tilt of the ECAP-based C-profile during everyday use.

Methods: Eighteen cochlear implant (CI) recipients received a research speech processor provided with a data-logging system. During a period of three weeks, they used the ECAP-based fitting in which the audiologist applied only a shift of the ECAP thresholds (ECAP-shift fitting). This fitting could not be changed by the CI recipient. During a second period of three weeks, the recipients themselves were asked to adjust shift and tilt of the C-profile programmed by the audiologist, optimizing their perception. The result was a self-fitting. Speech perception was tested presenting CVC words in quiet at 65 and 55 dB SPL and in noise at 65 dB and a signal to noise ratio (S/N) of +10 dB. Sentences were also presented, both in quiet and in noise at 60 dB SPL (S/N of +5 dB). A questionnaire was used to assess subjective appreciation.

Results: The correlation between the tilt of the conventional C-profile and the ECAP-profile was fairly good (r = 0.70). After self-adjustment of the tilt it increased to r = 0.90. Basal stimulation levels were higher in self-fitting than in ECAP-shift and conventional fitting. CVC phoneme scores measured in quiet at 65 dB SPL were on average 4 percentage points higher after self-adjustment of the processor, compared to both conventional fitting and ECAP-shift fitting performed by the audiologist. The other speech test conditions did not show significant differences between the three fittings. Ten participants appreciated the possibility of self-adjustments, whereas eight were insecure about how to adjust the processor. Subjects indicated that self-adjustment of the speech processor can be accomplished in a limited period of time of up to three weeks.

Conclusion: Adjusting shift and tilt of the ECAP-based C-profile by CI users themselves leads to a small but significant improvement of CVC phoneme scores at 65 dB SPL in quiet relative to both conventional fitting and ECAP-shift fitting performed by an audiologist. No improvement after self-adjustment is found for the other speech conditions. However,

in some individuals self-adjustment of shift and tilt of the ECAP-based C-profiles yielded a large improvement in the speech perception scores. Subjects try to find one fitting of their speech processor that suffices in all listening situations, instead of varying the adjustment from one situation to the next.

4.1 Introduction

Although performed by the audiologist, conventional fitting of the speech processor of a cochlear implant (CI), relies to a large extent on an implant recipient's subjective response. For each of the electrodes, i.e. 22 in the Nucleus CI system (Cochlear, Lane Cove, Australia), the recipient has to indicate the threshold level of sound perception (T-level) and the most comfortable loudness level (C-level) while stimulated with pulse trains. Obtaining these 44 behavioral measurements is a time-consuming task. It requires cooperation and considerable effort of the CI recipient. Especially in adults that have been deaf for many years, it can be laborious to obtain these behavioral measurements.

Using objective measures of the auditory system's response to electrical stimulation may reduce the dependency on the recipient's subjective feedback during the fitting procedure. One of the objective measures that have been investigated in this respect is the electrically evoked compound action potential (ECAP) (Brown and Abbas, 1990). ECAP measurements can easily be performed during surgery or post-operatively. ECAP thresholds measured across the electrode array can be used to start the speech processor fitting procedure.

One of the methods developed for ECAP-based fitting is the 'shift and tilt' procedure (Smoorenburg et al., 2002; Willeboer and Smoorenburg, 2006; Smoorenburg, 2007). Principal components analysis (PCA) showed that ECAP thresholds as well as the conventional T- and C-levels across the full electrode array, the 'profiles', can be largely described by only two parameters. The first parameter, overall level (termed shift), accounted for 90% of the intersubject variance in the profiles. Including the second parameter, corresponding to roughly the slope of the profile or tilt, increased the explained variance to at least 95%. An increase in shift means an increase of the levels at all electrodes by nearly the same amount. If the increase is larger than the just noticeable difference, it results in an increase in loudness. An increase in tilt implies a steeper slope of the profile, by definition higher stimulation levels at the basal electrodes and lower levels at the apical ones. An electrode located centrally in the array is the pivot remaining at constant level. If large enough, an increase in tilt subjectively results in a sharper sound.

Previously, we compared conventional to ECAP-based fittings in a prospective randomized and balanced cross-over study (Willeboer and Smoorenburg, 2006). In this study only a shift of the ECAP thresholds was applied to obtain the threshold of detection (ECAP-based T-profile) and loudness comfort level (ECAP-based C-profile) using live speech. The tilt of the profiles remained unchanged. This ECAP-based fitting procedure was performed by the audiologist and resulted in an ECAP-shift fitting. There was no significant difference between the average speech perception scores found for the conventional and the ECAP-shift fitting. Subjective sound quality was acceptable to good in both the conventional and the ECAP-shift fittings. Some recipients suggested that the ECAP-shift fitting could probably be optimized by allowing for other changes in the profile than shift only. However, in adjusting shift there are two clearly defined markers: the threshold of hearing and loudness comfort level, which both can be indicated by the recipient. In adjusting tilt, it is not trivial what one should ask an implant recipient in order to find the optimal setting and obvious guidelines are lacking.

The ECAP-shift fitting procedure was faster and hence less demanding to the recipient as it required only two subjective responses using live speech instead of 44 subjective responses to abstract pulse trains. It implied two advantages. First, the most important sound in the acoustic environment of the CI recipients is speech. Natural sounds, like speech, have a broad frequency spectrum that can only be presented to the CI recipient by (near-) simultaneous stimulation of a number of electrodes. Moreover, speech is a dynamic stimulus. Thresholds and loudness comfort levels measured per electrode for sound bursts might be different from those measured for dynamic stimuli, as loudness growth of dynamic stimuli is different from that of steady-state stimuli (Zeng and

Shannon, 1995; Zeng and Galvin, 1999). Second, the ECAP-shift fitting method is much faster than the conventional fitting method. The speech processor is in live mode from the start, which may provide enhanced communication with the recipient during the first fitting and a quicker start in CI usage.

Another drawback of current fitting procedures is that speech processor fitting is performed in a clinical environment, i.e. in a quiet room with only a limited number of different voices and sounds. This does not resemble everyday conditions in which the CI recipient experiences the effects of various acoustic environments, background noises and different voices on sound perception. Currently, the recipient is only able to make changes to the fitting by adjusting volume and microphone sensitivity settings within a restricted range, or by choosing one of the other fittings programmed in the speech processor. However, these programs are always prefabricated by the clinician in the clinical environment and are therefore only the clinician's suggestion of an optimal setting during everyday life. It would be of great interest to offer the recipients themselves the opportunity to make adjustments to their fitting during everyday use.

The present study evaluates the effects of recipients' adjustments to the ECAP-shift fitting during daily life on speech perception. Volunteers received a research processor enabling them to adjust shift and tilt of the ECAP-shift fitting themselves during everyday use and to experience directly the perceptual effects of the adjustments.

4.2 Materials and methods

4.2.1 Subjects

Subjects were eighteen adult, postlingually deafened CI recipients with between 6 and 36 months of CI experience, who volunteered to participate. They all had received a Nucleus CI 24R implant, with either the Contour or Contour Advance electrode array. One subject had a Straight array. There were no other inclusion criteria than duration of CI use and implant type given above. Before participating in this experiment all subjects had been using the Esprit 3G speech processor with a conventional fitting as opposed to an ECAP-based fitting. All subjects gave written informed consent according to the Declaration of Helsinki. The study was approved by the Medical Ethics Committee of the University Medical Center Utrecht. Subject characteristics are listed in Table 1.

Subject	Sex	Age (y)	Deafness					Contra lateral hearing aid	Electrode array	CI use (mo)
			etiology	duration (y)		PTA 0.5, 1, 2 kHz (dB)				
				ipsi	contra	ipsi	contra			
1	M	36	hereditary	4	3	95	95	Y	ST	6
2	F	58	unknown	RH	15	95	>120	N	CA	8
3	M	45	unknown	RH	RH	100	100	Y	CA	7
4	M	71	unknown	10	10	>120	>120	Y	CA	9
5	M	55	otosclerosis	2	35	>120	>120	N	CA	9
6	F	69	hereditary	12	4	>120	105	N	CA	7
7	F	57	unknown	8	8	100	>120	N	CA	8
8	F	57	meningitis	50	50	110	110	Y	CA	10
9	F	52	barotrauma	4	4	>120	>120	N	CA	18
10	F	80	otitis	1	30	>120	>120	N	CA	21
11	M	78	ototoxicity	2	2	90	85	Y	CA	26
12	F	45	hereditary	5	6	120	110	N	CS	36
13	F	50	hereditary	45	45	100	110	N	CA	26
14	F	52	unknown	2	RH	105	90	Y	CA	30
15	F	30	hereditary	27	27	100	95	Y	CA	30
16	M	38	viral infection	2	14	120	>120	N	CS	36
17	M	63	meningitis	1	5	100	>120	N	CA	25
18	F	65	unknown	30	30	110	110	Y	CA	27

Table 1: Subject characteristics. RH indicates residual hearing. ST indicates Nucleus 24R Straight, CS indicates Nucleus 24R Contour, CA indicates Nucleus 24R Contour Advance electrode array.

4.2.2 ECAP measurements

Determination of the profile of ECAP thresholds was performed using the NRT software 3.0 and 3.1 from Cochlear in sixteen subjects immediately after electrode insertion, during the closing of the wound, and in subjects 9 and 10 at the beginning of the first visit of this study. This difference in timeline is not important for the study, as the shape of the profile of ECAP thresholds measured intraoperatively is stable over time (Lai et al., 2004; van Wermeskerken et al., 2006).

ECAPs on at least all evenly numbered electrodes were measured following the standard procedure described by Lai (1999) and using the same protocol as in our previous study (Willeboer and Smoorenburg, 2006). The stimulation and recording mode was monopolar (MP1 and MP2 respectively), recording from an electrode located two positions more apically than the stimulating electrode. Stimulation amplitude is expressed in current level (CL), a quantity defined by Cochlear. The CL ranges from 1 to 255 current units (CU), which projects on the logarithm of the amount of electrical current, ranging from 10 µA to 1.75 mA. A masker offset level of 10 CU above the probe level was used. Pulse duration was set at 25 µs/phase and stimulation rate was 250 Hz. The number of sweeps was 100. Masker advance was fixed at 500 µs. The highest CL used was at least 15 CU above the visual ECAP threshold. Gain and delay settings were optimized running the 'optimizing gain and delay' series of the software, stimulating at electrode 15 and recording from electrode 17. The optimal gain and delay found for these electrodes were kept unchanged during subsequent measurements.

To determine ECAP amplitudes, markers were manually set at the negative N1 and positive P1 peaks in the response pane with both low and high resolution baseline corrected components. An amplitude growth function was determined per electrode. The ECAP threshold was defined as the stimulus level at which the ECAP amplitude reached 40 µV. For the Nucleus CI 24R family cochlear implants, with noise floors in the order of tens of microvolts, we consider this method to be more accurate than determining the intersection of the linearly extrapolated amplitude growth function and the 0 µV base line (Cafarelli Dees et al., 2005). In five subjects ECAPs were measured only at the evenly numbered electrodes. The ECAP thresholds were interpolated to obtain values for all electrodes.

4.2.3 Speech processor fitting procedure

For the duration of the study, the participants received a specially programmed Cochlear L34 body worn speech processor designed for research purposes and provided with a built-in data logging system. It was programmed using the Nucleus Programming Environment Software 1.4.0. Subjects were fitted using the ACE strategy with a per-channel stimulation rate of 900 Hz. Pulse duration was 25 µs/phase. Stimulation mode was monopolar using both extracochlear reference electrodes (MP1+2). Before the study subjects had been using the Esprit 3G speech processor, in which a maximum of 20 of the 22 electrodes can be activated. In all subjects electrode 1 and 2, the most basal electrodes, had been disabled in the Esprit 3G. Therefore, electrode 1 and 2 were switched off in both the conventional and the ECAP-based fittings during this study.

Conventional fitting using the Custom Sound 1.2 software was performed by the audiologist after implantation on a regular basis. All subjects had a stable fitting for at least three months before the start of the present study. The conventional fitting had been created by measuring subjective T- and C-levels for all electrodes 3 – 22 individually, stimulating with 500 ms bursts of pulses at a rate of 900 pulses/s. The T-level was defined as the lowest stimulus level per electrode that elicits a very soft, but consistent hearing sensation. The C-level was defined as the maximum stimulus level per electrode that still produces a comfortable loudness sensation. In addition the C-levels were balanced across electrodes for equal loudness. The T- and C-profiles resulting from this procedure were copied in the L34 using the Nucleus Programming Environment software without further adjustments.

The ECAP-shift fitting procedure was also performed by the audiologist and was identical to the procedure used in our previous study (Willeboer and Smoorenburg, 2006). The C-profile was initially set at the

ECAP-profile and the T-profile was set 1 CU lower, resulting in a dynamic range of 1 CU. Subsequently, both profiles were shifted down by equal amounts of current units at each electrode while keeping the difference between the T- and C-profiles at 1 CU, until a T-level on one or several electrodes reached 30 CU. Then the processor was switched on in live mode. Next, in live mode, the T- and C-profiles were shifted upwards by equal amounts of CU until the subject reported that live speech was just audible. The ECAP-shift T-profile was set at this level. Finally, the C-profile was shifted upward until a comfortable loudness level of live speech was reached. The ECAP-shift C-profile was set at this level. Thus, in this ECAP-shift fitting procedure only shift was varied by the audiologist, whereas the tilt, and in fact the whole shape of the ECAP-profile, was copied in the ECAP-shift T- and C-profiles.

4.2.4 Experimental design

Before the current study we performed a pilot study in four subjects to optimize the experimental design. First of all we studied the effect of switching from the Esprit 3G to the L34 speech processor while keeping the conventional fitting. This change did not affect the speech scores or perceived sound quality, justifying the use of the L34. Secondly, the feasibility of self-adjustment of both shift and tilt of the ECAP-shift T- and C-profiles was studied. As in the pilot study adjustments to the T-profile appeared to be not noticeable, for the present study self-adjustments were limited to shift and tilt of the ECAP-shift C-levels only. The results of this pilot study are further described in the Discussion section.

The present study was designed as a block trial that consisted of two blocks of three weeks each. In the first block, subjects used the ECAP-shift fitting created by the audiologist. Manual volume adjustments were disabled, but manual microphone sensitivity adjustments were permitted so that the subjects could cope with all acoustic environments. In the second block, subjects themselves could adjust shift and tilt of the C-profile of the ECAP-shift fitting provided by the audiologist. The result of these actions will be referred to as the self-fitting. Manual volume and microphone sensitivity adjustments were disabled, so that subjects were forced to use shift and tilt to adjust their processor.

After the first block of three weeks speech perception was measured using the conventional fitting and the ECAP-shift fitting. After the second block speech perception was measured using the ECAP-shift fitting provided by the audiologist and the final self-fitting by the CI recipient. Also, at the end of the study, subjects filled in a questionnaire concerning their subjective experiences during the study. The participants were strongly encouraged to use only the ECAP-based fittings. However, they could return to the conventional fitting in the L34 processor or to their own Esprit 3G processor, if so desired.

4.2.5 Shift and tilt parameters

As described above, the outcome of the PCA showed that the ECAP-profile can be largely described by two parameters; shift and tilt (Smoorenburg et al., 2002; Willeboer and Smoorenburg, 2006). These studies incorporated ECAP-profiles of 13 and 18 subjects, respectively. An extension of the PCA on clinical data of 100 ECAP-profiles and 200 sets of conventional T- and C-profiles yielded essentially the same results (Smoorenburg, 2007). This large data analysis provided weighting factors per electrode for shift and tilt used in the present study. Weighting factors were normalized by setting the coefficient with the largest magnitude to +1.00, and maintaining the sign of the coefficient. Because the weighting factors for shift and obviously also for tilt were not equal across the electrode array, the changes in CU varied over the electrodes when adjusting the shift and tilt values. The shift and tilt weighting factors per electrode are presented in Table 2. Increasing the value of the shift parameter by 10 implied that the C-level of electrode 22 increased by 9 CU, of electrode 12 by 10 CU, and of electrode 3 by 9 CU. Increasing the value of the tilt parameter by 10 implied that the C-level of electrode 22 decreased by 9 CU, the C-level of electrode 12 remained unchanged, and the C-level of electrode 3 increased by 10 CU. The effective changes per electrode were rounded to whole current units for each adjustment action. However, for the next adjustment action the exact changes were used as the starting point. For example, if the first adjustment resulted in a change of 4.2 CU, the C-level was increased by 4 CU. If the second adjustment implied a change of 4.4 CU, the C-level was increased by 5 CU, as the result of both actions together is 8.6 CU and hence 9 CU. Changing the tilt parameter results nearly in a rotation of the C-profile with electrode 12 as the pivot. An increase in tilt "rotates" the profile counterclockwise. Thus, an increase in tilt implies more high-frequency stimulation. Therefore the weighting factors are increasingly more negative for higher numbered electrodes and increasingly more positive for the lower numbered electrodes.

Electrode	delta C-level (CU)	
	shift	tilt
22	0.92	- 0.86
21	0.93	-0.83
20	0.95	-0.78
19	0.97	-0.71
18	0.98	-0.65
17	0.99	-0.60
16	1.00	-0.50
15	1.00	-0.38
14	1.00	-0.23
13	0.99	-0.09
12	1.00	-0.03
11	0.98	0.16
10	0.98	0.29
9	0.99	0.41
8	0.98	0.55
7	0.97	0.69
6	0.96	0.80
5	0.97	0.84
4	0.96	0.93
3	0.94	1.00

Table 2: Shift and tilt weighting factors per electrode: change in C-level per electrode in CU when increasing the shift or tilt parameter value with 1 unit.

Subjects had two independent buttons at their disposal; one for adjusting shift and one for adjusting tilt, which could be used from the start of the second block. In the ECAP-shift fitting provided by the audiologist, shift and tilt were set at the default values of 50 for each individual participant. The parameter values ranged from 1 to 99, of which one step corresponded with the change in CU presented in Table 2. This range covered more than plus-minus two standard deviations of the distribution of the shift and tilt values across CI recipients, according to the previous large data analysis (Smoorenburg, 2007). In order to make the parameters comprehensible to the subjects, shift and tilt were termed the Dutch equivalent of ‘loudness’ and ‘pitch’, respectively.

To protect the subjects from excessive stimulation during self-adjustments of shift or tilt, loudest acceptable presentation levels (LAPLs) per electrode were programmed in the L34 processor as the upper limit of stimulation. If the CL at one or more electrodes reached the LAPL, further increase when adjusting shift and tilt was denied. At the start of the second block, LAPLs were determined for all electrodes in live mode using three paradigms for increasing stimulation level. First, the C-profile was shifted upward in live mode until the subject reported that live speech was at the loudest acceptable level. Second, the C-profile was tilted positively, i.e. more basal stimulation, and third negatively, i.e. more apical stimulation, until the loudest acceptable level of live speech was reached. For each electrode the highest of these three levels was programmed in the L34 processor as the LAPL.

4.2.6 Speech perception measurements

Speech perception was measured using both sentences and words presented in quiet and in noise. All speech materials were DVD-recorded, Dutch female-spoken, and played in auditory-only mode. Subjects were to repeat the items perceived and were strongly encouraged to respond to each item presented. Speech and noise were presented via the same loudspeaker. Presentation levels were calibrated at the microphone position of the speech processor. A contra lateral acoustic hearing aid, if present, was switched off.

The sentence test (Versfeld et al., 2000) consisted of in total 39 lists of 13 mutually unrelated sentences each. Per list the final 10 sentences were used to determine the percentage of syllables repeated correctly. The sentences were presented in two conditions: at 60 dB SPL in quiet and at 60 dB SPL in 55 dB SPL continuous noise with a speech-shaped spectral energy distribution. Per condition two lists were presented, of which the syllable scores were averaged.

The word test (Bosman and Smoorenburg, 1992) consisted of 45 lists, each a set of 12 mutually unrelated, linguistically meaningful words. The words consisted of the sequence consonant-vowel-consonant (CVC). The test is scored in terms of the percentage of phonemes responded correctly, excluding the response to the first word and averaging the results for three lists. The stimuli were presented in three conditions; at 65 and 55 dB SPL in quiet and at 65 dB SPL in 55 dB SPL continuous noise with a speech-shaped spectral energy distribution.

4.2.7 Questionnaire

At the end of the second block, subjects filled in a questionnaire that covered items such as the time the L34 speech processor was used, the ease of making adjustments and of hearing differences when making adjustments, the perceived benefit of adjustments made during daily use, the time frame of the study and whether or not the subjects would like to continue using their self-fitting in their Esprit 3G after the end of the study.

4.3 Results

4.3.1 Usage of L34 speech processor

All subjects commented on the size of the L34 speech processor, which hindered daily activities. In spite of this drawback eleven of the eighteen subjects used the ECAP-based fitting in the L34 speech processor continuously during both blocks of the study. Five subjects used the L34 processor mainly in the evening and in weekends, because the body worn L34 processor was too inconvenient during daytime. Two subjects

(4 and 7) tried to use the ECAP-shift fitting in the L34 processor during the first block, but preferred their conventional fitting. Also in the second block they preferred the conventional fitting. However, they tried self-adjustment of the fitting alternated with the conventional fitting during this block.

All subjects indicated that handling the L34 speech processor was easy. They said to have understood the procedure for adjusting the shift and tilt parameters, which was supported by the records of the adjustments in the log-file (see below 'Number of self-adjustments'). Subjects indicated that it was easy to hear differences between different settings of shift and tilt. A change in shift was subjectively perceived as a change in loudness. An increase in tilt was perceived as an increase in sharpness of the sound, whereas a decrease was perceived as "deepening" of the sound.

4.3.2 Number of self-adjustments

Figure 1 shows a representative series of shift and tilt settings over time for subject 5. The open symbols indicate the settings tried within one adjustment action, the filled symbols represent the result at the end of that action. Fig. 1 shows that most adjustments were made in the first week, and that the adjustments converged to a final setting. The final setting in this example (an increase in shift of 10 and in tilt of 12 units) resulted in a decrease of the C-level of electrode 22 of 1 CU, an increase of the C-level of electrode 12 of 10 CU, and an increase of the C-level of electrode 3 of 21 CU.

For each subject we counted the total number of shift and tilt settings tried over the three weeks period. The average over subjects was 103 (minimum 12, maximum 289) for shift and 137 (minimum 11, maximum 452) for tilt. The standard deviations were 89 and 125, respectively. Hence, there were large interindividual differences. In the first week subjects tried different settings about once a day. On average, 9 shift settings and 13 tilt settings were tried within one adjustment action. In the second week adjustment actions were seen once in two days, each action consisting of, on average, 4 different settings for shift and tilt. In the third week adjustment actions decreased to once in four days, each consisting of 2 different settings for shift and tilt. Almost always shift and tilt were adjusted concurrently.

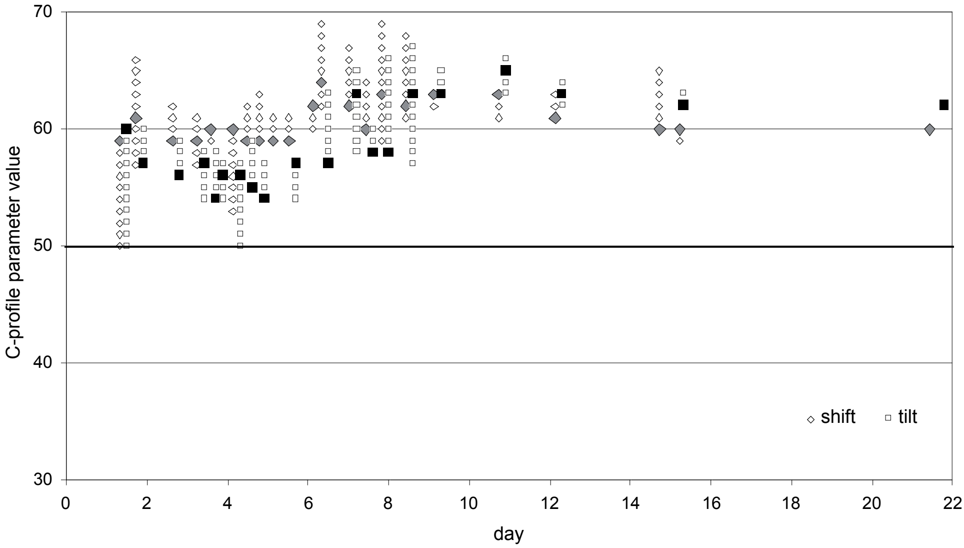


Figure 1: Shift and tilt adjustments for subject 5 during the second block of the study. Open symbols indicate the settings tried within one adjustment action, and the filled symbols are the result at the end of each action.

4.3.3 T- and C-levels

In accordance with our earlier findings (Smoorenburg et al., 2002; Willeboer and Smoorenburg, 2006) the present results show that the T-levels yielded in the ECAP-shift fitting procedure were on average 30 CU lower than those found in the conventional fitting ($p < 0.01$ in paired samples T-test on the averages per electrode over subjects). The dynamic range was on average 50 CU in the conventional fitting and 82 CU in the ECAP-based fittings. Figure 2 shows the averaged C-profiles of the conventional fitting and both ECAP-based fittings. The first transition, from the conventional to the ECAP-shift fitting adjusted by the audiologist, yielded C-levels that were on average 4 CU higher at the apex, while these levels were 3 to 4 CU lower for electrodes 6-11, and 7 CU higher at the base. Standard deviations of these differences range from 5 CU for the middle electrodes to 11 CU at the apex and base. During self-fitting the ECAP-shift C-levels for the apical electrodes remained on average unchanged. For the basal electrodes self-fitting resulted in an increased stimulation level of on average 4 CU. Thus, although the average basal ECAP-shift C-levels were already higher than the basal conventional C-levels, the inclusion of tilt in the self-fitting yielded even higher high-frequency (basal) stimulation levels. The self-fitting C-level of the most basal electrode 3 was, on average, almost 11 CU (with a standard deviation 11 CU) higher than that of the conventional fitting, which corresponds to 25% of the average dynamic range of electrode 3 in the conventional fitting.

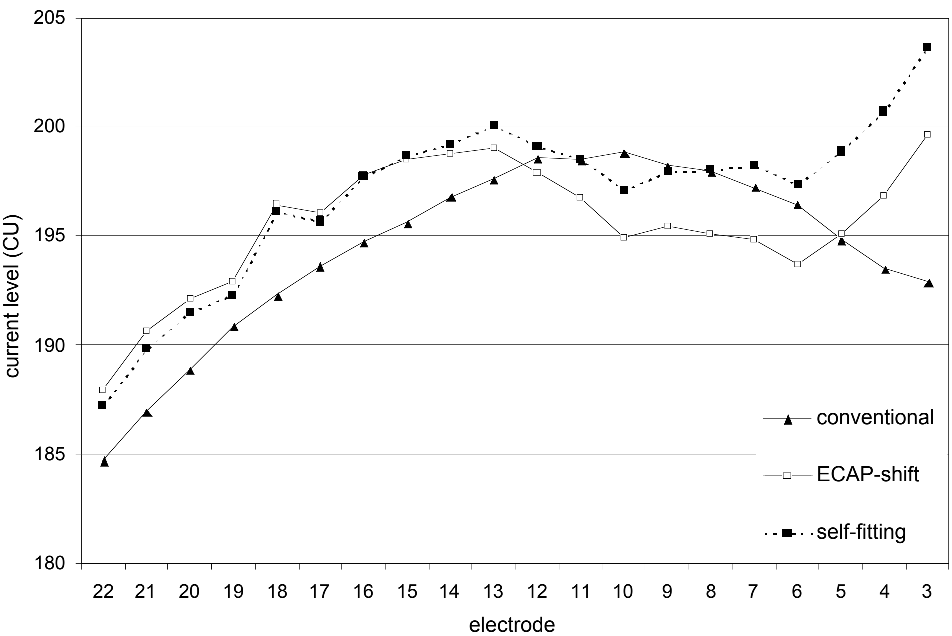


Figure 2: Averaged C-profiles of the conventional fitting, ECAP-shift fitting, and self-fitting. Electrode 22 is the most apical, and electrode 3 is the most basal contact.

In Figure 3 the averaged differences across the electrode array between the C-profile of the self-fitting, including adjustments of shift and tilt, and that of the ECAP-shift fitting by the audiologist are shown, as well as individual results. There were large differences between the individual results. In our experience, a change in C-level larger than ± 2 CU is a clinically relevant difference. Applying this definition, we can categorize the subjects according to the changes they made to their C-profile. Five subjects made no clinically relevant adjustment to any of the electrodes. Eight subjects increased their C-levels of the most basal electrodes, either with unchanged (three subjects) or decreased (five subjects) apical C-levels. Three subjects chose lower stimulation levels for the most basal electrodes, of up to -9 to -14 CU at electrode 3, either in combination with unchanged (two subjects) or increased (one subject) apical stimulation. One subject decreased only the apical C-levels up to -5 CU at electrode 22, while the last subject increased the C-levels of all electrodes by 4 CU.

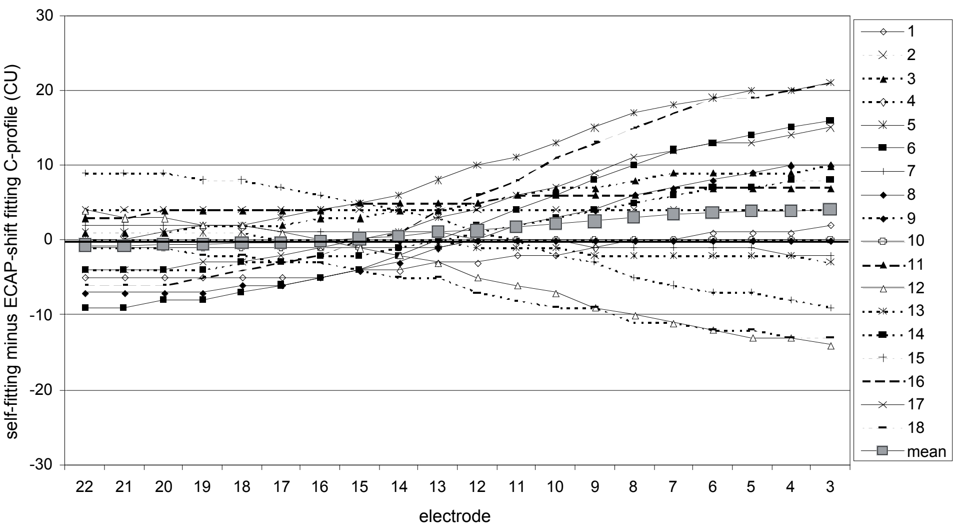


Figure 3: Individual and averaged differences between C-profiles of the self-fitting and of the ECAP-shift fitting provided by the audiologist.

Since it is difficult to compare eighteen individual C-profiles from three different fittings on an electrode-by-electrode basis, further comparisons of the C-profiles are based on the two profile parameters shift and tilt. The shift and tilt coefficients were calculated for each profile. Figure 4a shows horizontally the shift coefficients (which corresponds to the average C-level for all electrodes in CU) for the conventional fitting and vertically those for both the ECAP-shift fitting performed by the audiologist and self-fitting. Shift was roughly similar in the three fittings; on average the shifts were 1.0 and 2.4 CU higher in the ECAP-shift and self-fitting, respectively, than in the conventional fitting. The coefficients of the correlation across subjects between the shifts of the conventional fitting on the one hand and the ECAP-shift and self-fitting on the other hand were significant and high: for both $r = 0.96$ and $p < 0.001$. Fig. 4b shows that the correlation between the tilt parameter in the ECAP-shift fitting and the conventional fitting was fairly good ($r = 0.70$, $p = 0.002$). The tilt in the ECAP-shift fitting was on average 0.2 CU/el (per electrode) smaller than in the conventional fitting, primarily due to the lower C-levels found for electrodes 6-11 (see Fig. 2). In self-fitting subjects increased the tilt by 0.3 CU/el. This resulted in an average tilt 0.1 CU/el higher than in the conventional fitting. With self-fitting the correlation across subjects between the tilts in self-fitting and conventional fitting increased markedly to $r = 0.90$ ($p < 0.001$).

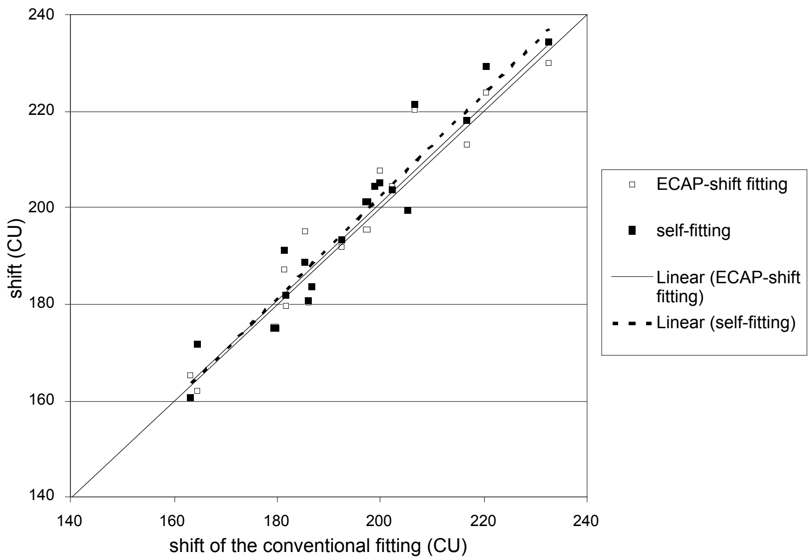


Figure 4a

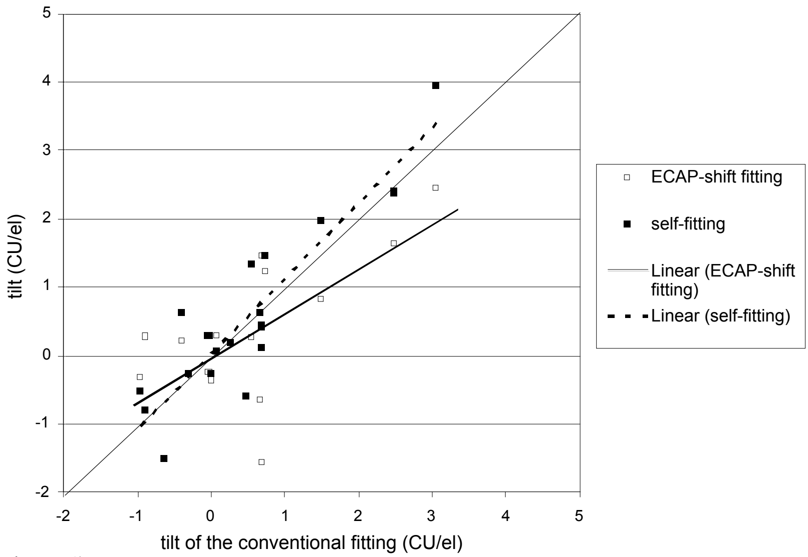


Figure 4b

Figure 4: Correlation across subjects between the conventional fitting and both the ECAP-shift fitting and self-fitting with regard to the shift (a) and tilt (b) parameter. The shift parameter corresponds to the average C-level over the electrodes, and the tilt corresponds to the slope of the profile. The thick solid and dashed lines represent linear least square fits through the ECAP shift-fitting and self-fitting respectively. The thin solid line represents a one-to-one linear relation.

4.3.4 Speech perception scores

Figures 5a and b show the individual speech perception scores from the CVC and sentence test respectively, using the ECAP-shift fitting provided by the audiologist and the shift-and-tilt adjusted self-fitting, measured after the second three-weeks block. For most subjects, CVC phoneme scores at 65 dB SPL in quiet were roughly similar for both ECAP-based fittings. However, three subjects (7, 8, and 15) improved their phoneme scores at 65 dB SPL in quiet by fifteen percentage points or more adjusting the speech processor themselves. The CVC phoneme scores at 65 dB SPL in noise showed more variation across subjects than in quiet, but on average the results for both fittings were comparable. In all but two subjects CVC phoneme scores at 55 dB SPL in quiet were equal to or higher in self-fitting than in the ECAP-shift fitting. Individual results from the sentence test (Fig. 5b) showed that speech perception scores from 11 subjects collected in quiet and in noise were comparable for both ECAP-based fittings. In the other 7 subjects, higher as well as lower scores were found, evenly distributed.

Figures 5c and d show comparisons of the CVC phoneme and sentence scores found for the self-fitting by the recipient on the one hand and for the conventional fitting performed by the audiologist on the other hand. At the presentation level of 65 dB SPL in quiet the CVC scores appeared to be comparable for most subjects (Fig. 5c). However, five subjects (2, 6, 8, 17, and 18) showed an increase of 10 percentage points or more using the self-fitting rather than the conventional fitting. On average the phoneme scores collected at 65 dB SPL in noise and at 55 dB SPL in quiet were comparable using both fitting methods, although they showed more variation than the scores in quiet. Sentence syllable scores collected at 60 dB SPL in quiet showed a ceiling effect in the majority of subjects (Fig. 5d). However, three subjects (4, 7, and 15) showed a clear decrease of 15 percentage points or more when using the self-fitting rather than the conventional fitting. The syllable scores collected at 60 dB SPL in noise were comparable, but showed more variation over subjects than in quiet.

Statistical analysis of the speech perception data was performed using a Repeated Measures Analysis of Variance (ANOVA) for each of the five speech perception tests, with the conventional fitting, the ECAP-shift fitting provided by the audiologist and the shift-and-tilt adjusted self-fitting as the repeated measures. Speech perception using the ECAP-shift fitting was tested after the first as well as the second block. There were no statistical significant differences between the speech results at both moments using the ECAP-shift fitting ($p > 0.05$ for all five paired samples T-tests). Only the data last collected were used in the subsequent ANOVA comparing the conventional, ECAP-shift and self-fitting, allowing for the longest period of habituation to the ECAP-based fittings. The CVC phoneme scores at 65 dB SPL in quiet were significantly higher for self-fitting than for both the conventional fitting and the ECAP-shift fitting (means were 80.7%, 76.2% and 77.2% respectively, $p < 0.05$). With the other speech perception scores ANOVA showed no significant difference between the three fittings.

Linear regression analysis showed that increases in tilt were positively correlated with improvements in the CVC phoneme score at 65 dB SPL in noise ($r = 0.54$, $p = 0.02$) and the improvement in sentence syllable scores at 60 dB SPL in quiet ($r = 0.71$, $p = 0.001$). There were no other significant correlations between changes in shift or tilt and changes in speech perception scores.

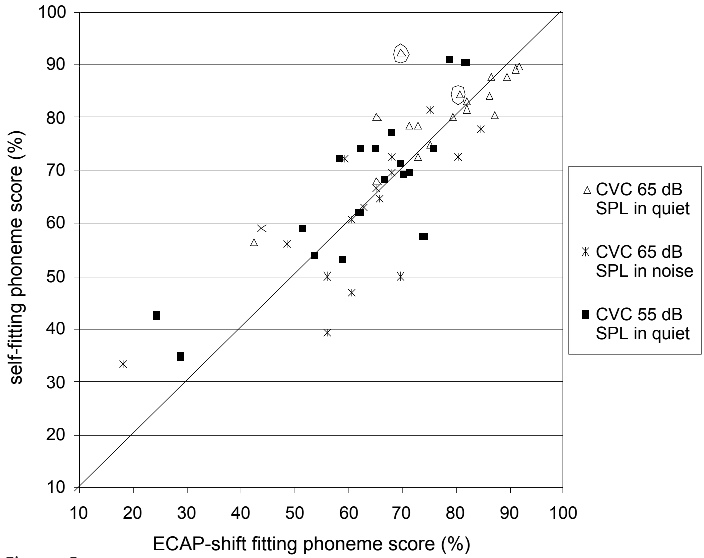


Figure 5a

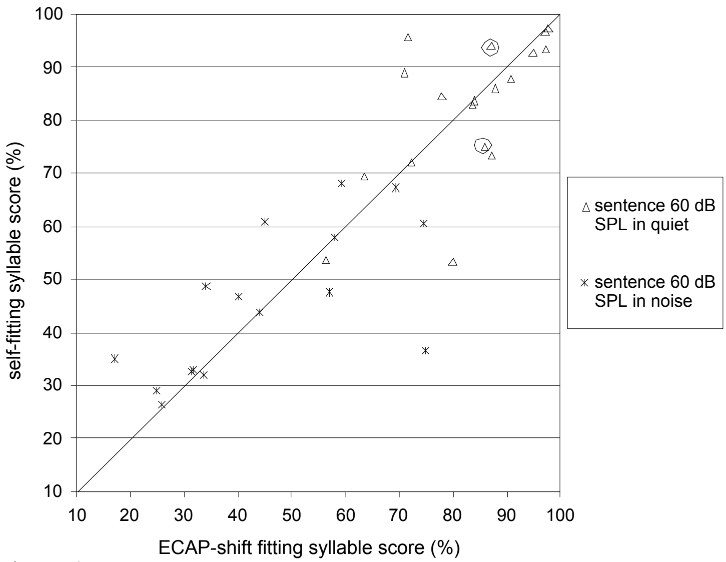


Figure 5b

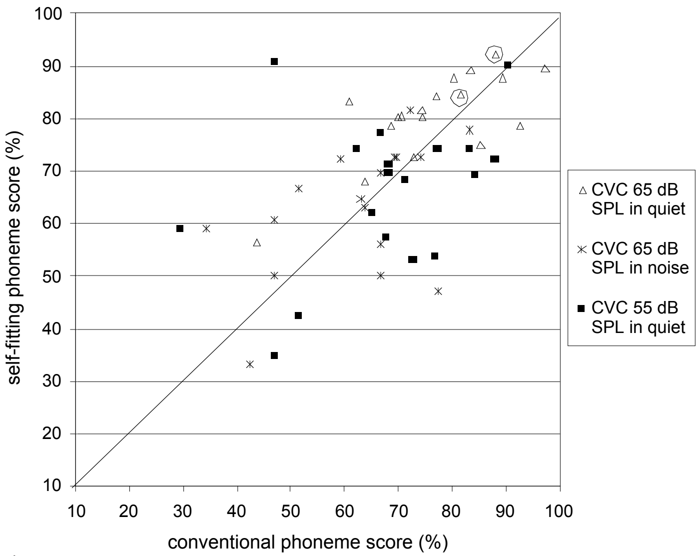


Figure 5c

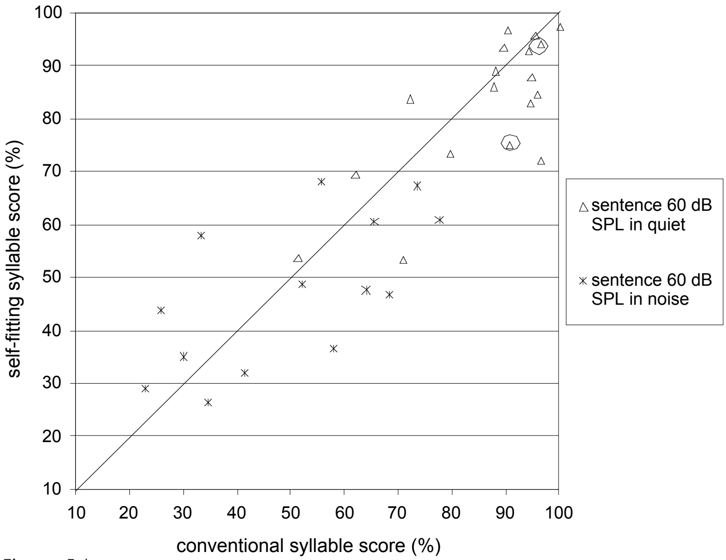


Figure 5d

Figure 5: Speech perception scores using the ECAP-shift fitting and self-fitting in the CVC test (a) and sentence test (b), and using the conventional fitting and self-fitting in the CVC test (c) and sentence test (d). The CVC test was performed at presentation levels of 65 and 55 dB SPL in quiet and 65 dB SPL in noise (S/R +10 dB). The sentence test was performed at a presentation level of 60 dB SPL in quiet and in noise (S/R +5 dB). Subjects 4 and 7, who used the ECAP-based fitting less than the other subjects, are encircled in each panel.

4.3.5 Subjective appreciation

All subjects indicated that they tried to optimize speech understanding, rather than sound comfort, by adjusting shift and tilt. In our view, their statement is supported by their adjustments resulting in an increased basal stimulation, because more high-frequency stimulation is often experienced as less comfortable. They aimed for one fitting that should suffice in all listening situations encountered, rather than trying specific adjustments of the speech processor per listening situation. The results of the questionnaire are listed in Table 3. Eight subjects indicated that they felt insecure about which adjustments were best for speech understanding; they did not know how to compare the different fittings. Therefore, these subjects suggested that it would be best to have an audiologist providing the fitting in the clinical environment. Yet, the other ten subjects preferred self-fitting. Nine of them preferred a couple of weeks to optimize the fitting themselves, after which they wanted to keep that fitting unchanged. One subject preferred self-fitting continuously during CI use. At the end of the study six subjects chose to have a copy of their shift-and-tilt adjusted ECAP-based self-fitting in their Esprit 3G. The other subjects mentioned that if they had a copy of their self-fitting, they would barely use it because they thought sounds were more natural with their conventional fitting. However, they reported spontaneously that both the ECAP-shift and the self-fitting would have sufficed in the initial stage of rehabilitation.

If you had the fittings of both blocks of the study at your disposal, how often would you use the ECAP-shift fitting?	Never **9**	Some-times **8**	Often **1**
If you had the fittings of both blocks of the study at your disposal, how often would you use the self-fitting?	Never **3**	Some-times **10**	Often **5**
Would you prefer to obtain a copy of the self-fitting in your Esprit 3G processor after the end of the study?	No **12**	Yes **6**	
How do you value the possibility of adjusting the fitting yourself?	Not valuable **2**	Impartial **10**	Very valuable **6**
Why do you think it is valuable or not valuable to be able to adjust the fitting yourself (choose one of the options per item)?			
Sound quality of the self-fitting is better than of the conventional fitting.	**9**		
Sound quality of the self-fitting is not better than of the conventional fitting.	**9**		
It is valuable to be able to adjust the fitting in various environmental situations.	**10**		
It is disturbing to adjust the fitting all the time.	**8**		
Self-fitting enables me to determine myself what I hear and what I don't hear.	**10**		
Self-fitting is disturbing, because I am not sure which fitting is best.	**8**		
It is better to let the audiologist decide which fitting is best for me.	**8**		
It is better to let me decide which fitting is best, because I am the one who can hear it.	**10**		
It is better to adjust the fitting at home or at work, instead of (only) in the clinic.	**10**		
It is better to adjust the fitting in the clinic, in a quiet room with familiar voices.	**8**		
It is best to adjust the fitting myself during a fixed period of time, after which the fitting should be kept unchanged.	**9**		
It is best to always be able to adjust the fitting myself.	**1**		
It is best to have only the clinician make the fitting.	**8**		
If you would be able to use the L34 processor for another year, how often would you adjust the self-fitting?	Never **5**	Some-times **7**	Often **6**
Learning to make adjustments was	Difficult	Moderate	Easy **18**
After getting used to making adjustments it was	Difficult	Moderate	Easy **18**

Table 3: Results of the questionnaire. The number of respondents per item is indicated in bold numbers after each item.

4.4 Discussion

4.4.1 Pilot study

Four subjects participated in a pilot study, which was conducted before the main study. The duration of the pilot was eight weeks. It started with a two week period in which subjects used the conventional fitting transferred to the L34 speech processor, followed by a three week period of using ECAP-shift fitting provided by the audiologist, and a final three week period in which the participants themselves could adjust shift and tilt of the ECAP-shift T- and C-profiles. Speech perception was tested with the conventional fitting at the start of the study, after two weeks, and after five weeks. Speech perception using the ECAP-shift fitting was tested after five weeks and after eight weeks, and using the self-fitting only after eight weeks.

Results of this pilot showed that there was no difference between the L34, programmed with the conventional fitting, and the subjects' own Esprit 3G, neither in speech understanding scores, nor in subjective sound quality. Therefore, we concluded that we did not need to validate the L34 processor, comparing the two processors with the same fitting, in the main experiment. Further, the speech perception scores using the conventional fitting were equal at all three test moments, indicating that there was no effect from using the ECAP-shift fitting for three weeks before measuring speech perception with the conventional fitting. With respect to self-adjusting shift and tilt of the ECAP T- and C-profiles, subjects reported that their adjustments of shift and tilt of the ECAP C-profile resulted in clearly audible changes to sound quality. However, adjustments to the ECAP T-profile were barely noticeable. They became confused making these adjustments and switched from one extreme to the other without noticing. We concluded that self-adjustment in the main experiment had to be limited to adjusting shift and tilt of the C-profile.

4.4.2 Usage of L34 speech processor

A prerequisite for success in an experimental design that relies heavily on the cooperation of subjects is that subjects understand their task and that they are able to carry it out well. A study, in which CI recipients adjust their C-levels themselves, has not been described in literature before. It is therefore important to investigate whether or not the above requirements are met. The results of the questionnaire regarding the controls of the L34 processor indicate that subjects did not encounter any problems adjusting the shift and tilt parameters themselves. The numbers of adjustments, even in the subject who made the least adjustments, support the view that subjects were able to operate the L34 processor with respect to adjusting shift and tilt. Subjects indicated that they did hear changes in sound when adjusting shift and tilt. Therefore, we consider the study design valid for investigating the effect of self-adjusting shift and tilt of ECAP-shift C-profiles.

Most subjects, however, commented on the size of the L34 body worn processor. This hindered in their daily activities, which in five subjects led to the decision of wearing the L34 processor mainly during the evening and weekends. Comparing the number of adjustments made by these subjects to the results of the other participants, it shows that two of them made the average number of adjustments. One subject made more adjustments and two subjects made fewer adjustments. With respect to speech perception four subjects followed the average pattern with all fittings. The speech scores for only one subject were higher for the conventional fitting than for self-fitting. We may therefore conclude that the limited usage of the L34 processor in these five subjects did not affect their results in another way than found in the other participants.

4.4.3 Number of adjustments

The number and range of adjustments was largest in the first week of the three-week period of self-fitting. Subjects first tried a large range of settings for shift and tilt, in order to obtain an impression of the effects on sound quality. These settings tried were clustered in, on average, only one adjustment action per day in the first week. In the second and third week, the number and range of adjustments gradually decreased; shift and tilt converged to a certain value. This objective result is in agreement with the opinion of most subjects that it was best to optimize the fitting during a certain amount of time and then keep it unchanged. The majority of the subjects mentioned that constantly trying to optimize sound quality in each listening situation encountered is too distressing and troublesome, especially because they had to perform the adjustment action on the body worn L34 processor. An alternative design in which adjustments can be made using a remote control and on a behind-the-ear speech processor might suffer less from this troublesome factor.

4.4.4 C-profile characteristics

Although the average C-levels of electrodes 3 and 4 were higher in the ECAP-shift fitting than in the conventional fitting, the tilt was 0.2 CU/el smaller. This was due to lower C-levels for electrode 6-11 in the ECAP-shift fitting. By making the adjustments themselves, subjects increased the C-levels of electrodes 6-11 to the levels in the conventional fitting, resulting in an increased tilt of 0.1 CU/el higher than in the conventional fitting. The difference between the self-adjusted C-level and the conventionally determined C-level increased on average, to almost 11 CU for electrode 3 and decreased to 2 CU for electrode 22 (Fig. 2). Although the ECAP-shift fitting provided by the audiologist had already higher stimulation levels for the most basal electrodes than the conventional fitting, most subjects made sound quality even sharper in their own adjustments. This finding is in accordance with the subjects' indication that they tried to optimize speech perception. The amount of the increase of the C-level at electrode 3 corresponds to 25% of the average electrical dynamic range of electrode 3 in the conventional fitting, which is a remarkable finding. Hearing aid patients generally prefer a comfortable fitting with less high-frequency gain than provided by the dispenser, even when sound clarity is better with increased high-frequency gain (Munro and Lutman, 2005).

Approximation of the conventional C-levels for electrodes 6-11 as well as for electrodes 3-5 would only have been possible if subjects had increased apical C-levels largely, or by the introduction of a third parameter next to shift and tilt. This parameter would enable the recipient to decrease and increase the C-levels of the middle electrodes independently from apical and basal electrodes. Mathematically spoken, this parameter would describe the curvature of the C-profile.

The starting point for self-fitting was the ECAP-shift fitting made by the audiologist, as the profile of ECAP thresholds provide basic information on the auditory system's functioning. Moreover, previous research has shown that the ECAP-shift fitting is fast and easy, and yields good speech perception scores. However, self-fitting might also be performed by making adjustments of shift and tilt of, for example, a flat C-profile. In that case, the best approximation using a least-square method of the self-fitting C-profile obtained in this study would on average have resulted in a C-level of 4 CU higher at electrode 22 and 2 CU lower at electrode 3. These are relatively small differences. This suggests that using shift and tilt to adjust a flat (or population average) profile into a desired self-fitting might gain as much possibilities as starting with an ECAP-profile and would be well worth trying.

4.4.5 Relation between C-level adjustment and speech perception

Despite large differences in T- and C-levels between the ECAP-shift and conventional fittings, the previous prospective trial has shown that there is no difference between the speech perception scores found for both fitting methods (Willeboer and Smoorenburg, 2006). Therefore, one might expect that any improvement by self-fitting would rather be reflected by increased subjective appreciation during everyday use than by increased speech perception scores. However, several subjects did prove to be able to significantly increase their CVC phoneme score at 65 dB SPL in quiet by adjusting shift and tilt of the ECAP-shift C-profiles themselves. The increase was on average only 4 percentage points, and therefore of minor clinical importance. In the other speech test conditions, there were no statistically significant differences between the scores found for the three fittings. Some individuals however did show a large improvement in speech perception on one or more of the other speech test conditions after self-adjustment of the speech processor. Together the results suggest that it is useful to try self-adjustment of shift and tilt of ECAP-shift C-profiles, because there are individuals who are able to improve their scores considerably in this way.

Self-adjustments of the speech processor characterized by an increase in tilt of the C-profile were positively correlated with an increase in the speech perception scores of sentences in quiet and words in noise. Probably, increased basal (high-frequency) stimulation leads to better perception of especially consonants. The effect was not observed using words in quiet. However, perception of words in quiet was already very good using the ECAP-shift fitting provided by the audiologist. There may have been too little variance in these scores to obtain a significant correlation between the changes in tilt and the speech scores after self-adjustment. Adding noise to the words, the variance in speech perception scores increased and the scores decreased. Using sentences in noise this correlation between tilt and speech scores was not found. A possible explanation is that word boundaries are lost due to the noise, which cancels the positive effect of better phoneme perception with higher tilt.

4.4.6 Subjective appreciation of self-adjustments

The evaluation of subjective appreciation of self-fitting showed less satisfaction than would be expected from the speech perception scores. The disadvantage of the body worn processor to make adjustments has been discussed above. However, also satisfaction of the effect of self-fitting was moderate. Only a third of the subjects wanted to keep their self-adjusted ECAP-based fitting after the end of the study, whereas the others indicated that, although differences were small, subjectively their conventional fitting outperformed their self-fitting. Apparently, although the speech tests were elaborate, using words and sentences, in quiet and in noise, they were unable to cover all elements of perceived sound quality during everyday use. All subjects were experienced CI users, who were used to their conventional fitting and therefore may have lacked the flexibility to get accustomed to their self-fitting within the time frame of the study. Five subjects reported spontaneously that both ECAP-based fittings would have sufficed in the initial stage of rehabilitation, i.e. before they got accustomed to their current fitting. Conducting the same study in fresh CI recipients may eliminate this habituation effect. However, another habituation factor may then be introduced, namely the effect of existing habituation to the sound quality of an acoustic hearing aid they might have used. In most patients this quality is less sharp than that perceived with their CI. This may affect self-fitting of the CI speech processor.

Whereas ten subjects preferred to optimize their speech processor fitting themselves, the other eight subjects indicated that they preferred to have an audiologist provide the fitting. Analyzing the speech perception scores, it showed that the first group consisted of slightly poorer performers when using their conventional fitting. Their CVC phoneme scores at 65 dB SPL in quiet were on average 71%, compared to 83% in the second group ($p = 0.05$ in a T-test for independent samples). In the first group, switching to the ECAP-shift fitting provided by the audiologist yielded average phoneme scores of 74%, which they were able to significantly increase to 79% by self-fitting ($p < 0.01$ in a Repeated Measures ANOVA). In the second group CVC phoneme scores at 65 dB SPL in quiet were high from the start with the conventional fitting (83%). They did not change significantly when using the ECAP-shift and the self-fitting. There were no significant differences between these two groups of subjects in the other speech perception tests. Also, there were no significant differences between the two groups for shift and tilt of the three fittings, nor were there differences in the changes in shift and tilt when adjusting the fitting. Further, the pre-study duration of CI use was not significantly different between both groups. Thus, the relatively poorer performers had a preference for self-fitting instead of fittings provided by an audiologist. They were able to increase their CVC phoneme score at 65 dB SPL in quiet with on average 8 percentage points. Apparently, these relatively poorer performers take advantage from fine-tuning their C-profiles themselves.

Acknowledgements

This study was supported by Cochlear Europe. The authors Bas van Dijk and René Roos are employees of Cochlear Ltd.

References

Abbas PJ, Brown CJ, Shallop JK et al. (1999) Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear. 20:45-59.

Bosman AJ, Smoorenburg, G.F. (1992). Word lists for speech audiometry. Compact disc released under the auspices of the Dutch Society of Audiology.

Brown CJ, Abbas PJ. (1990) Electrically evoked whole-nerve action potentials: Parametric data from the cat. J. Acoust. Soc. Am. 88:2205-2210.

Cafarelli Dees D, Dillier N, Lai WK et al. (2005) Normative findings of electrically evoked compound action potential (ECAP) measurements using the neural response telemetry (NRT™) of the Nucleus CI24M cochlear implant system. Audiol. Neuro-otol. 10(2):105-16.

Dillier N, Lai WK, Almqvist B et al. (2002) Measurement of the electrically evoked compound action potential via a neural response telemetry system. Ann. Otol. Rhinol. Laryngol. 111(5 Pt 1):407-14.

Lai WK. (1999) An NRT Cookbook. Cochlear AG.

Lai WK, Aksit M, Akdas F et al. (2004). Longitudinal behaviour of neural response telemetry (NRT) data and clinical implications. Int. J. Audiol. 43(5): 525-63.

Munro KJ, Lutman ME. (2005) Sound quality judgements of new hearing instrument users over a 24-week post-fitting period. Int J Audiol. 44(2):92-101.

Smoorenburg GF, Willeboer C, van Dijk JE. (2002) Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds. Audiol. Neuro-otol. 7(6):335-47.

Smoorenburg GF. (2007) Cochlear Implant Ear Marks, University Medical Center Utrecht, ISBN-13: 978-90-9021523-5, 1-113 (available at: http://www.audiologics.com).

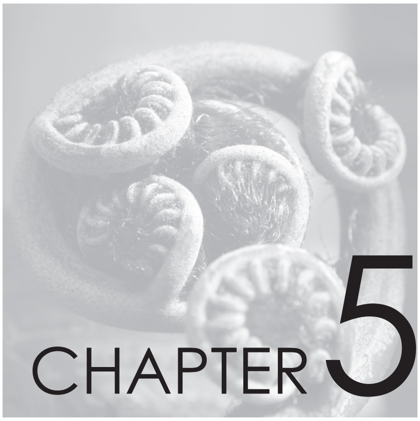
Versfeld NJ, Daalder L, Festen JM et al. (2000) Method for the selection of sentence materials for efficient measurement of the speech reception threshold. J. Acoust. Soc. Am. 107(3): 1671-84.

Wermeskerken, van GK, Olphen, van AF, and Zanten van, GA. (2006). A comparison of intra- versus post-operatively acquired electrically evoked compound action potentials. Int. J. Audiol. 45(10): 589-94.

Willeboer C, Smoorenburg GF (2006) Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study. Ear Hear. 27(6):789-98.

Zeng FG, Shannon RV (1995) Loudness of simple and complex stimuli in electric hearing. Ann. Otol. Rhinol. Laryngol. 104(Suppl. 166):235-8.

Zeng FG, Galvin JJ. (1999) Amplitude mapping and phoneme recognition in cochlear implant listeners. Ear Hear. 20(1): 60-74.



Speech perception in cochlear implant recipients using various parts of the electrical dynamic range and amplitude mapping functions

Christina Willeboer, Maaike Bierman, Guido F. Smoorenburg, and Gijsbert A. van Zanten

Abstract

Objective: This study aims at answering the question to what extent speech perception in cochlear implant (CI) recipients is affected by manipulating its electrical representation in two ways; (1) presenting speech only in the lower, middle or upper one-third of the subject's electrical dynamic range (EDR), or (2) presenting speech while mapping the upper 10 dB of the acoustic dynamic range into either the upper 10, 20, or 30% of the EDR.

Methods: Twenty-five experienced postlingually deafened adult

Nucleus 24M or 24R CI recipients participated. Their phoneme score using their conventional take-home fitting was at least 50% at a presentation level of 65 dB SPL. Five experimental fittings, maps, were created and compared to a standard conventional fitting. In map L only the lower third of the EDR was used, in map M the middle third, and in map H the upper third. In map Q10 the upper 10 dB of the acoustic input range was mapped onto the upper 10% of the EDR, whereas in map Q30 this was mapped onto the upper 30%. Speech perception was measured in quiet at presentation levels from normal to very low.

Results: Using map L still yielded two-third of the original phoneme score at 60 dB SPL. Maps M and H yielded an increase in phoneme scores at low presentation levels and a slight decrease in maximum phoneme scores. Map Q10 showed increased phoneme scores at low presentation levels and no effect on maximum scores. Map Q30 did not have any effect on speech perception in quiet.

Conclusions: Subjects are able to extract recognizable speech information from stimulation levels in the lower third of the EDR. However, the upper part of the EDR proves to be most important for speech perception. Raising T-levels enhances low-level speech perception. Mapping the upper 10 dB of the acoustic input in the upper 10% of the EDR enhances speech perception at low stimulation levels. Overall, the perceived loudness appears a major determinant for speech perception in quiet.

5.1 Introduction

A cochlear implant (CI) can partially restore hearing by electrical stimulation of the auditory nerve in patients with severe hearing loss or deafness. The implant per se is driven by an external part of the CI-system, the speech processor. It needs to be adjusted to the recipient's sensitivity for electrical stimulation. In the Nucleus CI 24 system (Cochlear, Lane Cove Australia) the conventional fitting requires an estimation of the threshold level (T-level) and the comfortable loudness level (C-level) for each of the 22 intracochlear electrodes, applying short bursts of electrical pulses and using psychophysical methods determining the threshold of perception and loudness discomfort level. However, some recipients, specifically those with a long duration of deafness or with a disturbing tinnitus, may show inconsistent reactions to low stimulation levels. It may be difficult to exactly determine their T-levels. Also, the determination of C-levels may be inconsistent as loudness comfort is a rather subjective judgment.

The fitting procedure can be facilitated by measuring objectively the response of the auditory system to electrical stimulation, such as the electrically evoked compound action potential (ECAP). In the ‘Shift and tilt fitting method’ (Smoorenburg et al. 2001), the profile of ECAP-thresholds across the electrode array is used as a starting point for the fitting. The overall level of the profile is shifted using a parallel shift until the threshold for live speech (ECAP-based T-levels) and the loudness comfort level for live speech (ECAP-based C-levels) are found. When comparing this ECAP-based fitting procedure to the conventional procedure, it appeared that T-levels were significantly lower in the ECAP-based fitting than in the conventional fitting, whereas C-levels were almost the same (Willeboer and Smoorenburg 2006). The electrical dynamic range (EDR), i.e. the range between the T- and C-level, was hence about one-third larger in the ECAP-based fitting than in the conventional fitting. Despite this large difference in T-levels, there was no difference in the speech perception scores for both speech presented in quiet and in noise at 65 and 55 dB SPL. Still, these markedly lower T-levels in the ECAP-based fittings raise questions as to the contribution of the lower parts of the EDR to speech perception.

In the fitting procedure T- and C-levels, EDR, and amplitude mapping all are factors that potentially affect speech perception. Several investigators have focused on one or more of these factors. Skinner et al. (1999) found in eight Nucleus 22 CI recipients that T-levels above the threshold of audibility were associated with an improvement in recognition of soft speech. This positive effect was found despite the fact that raising the T-levels narrows the EDR. Zeng and Galvin (1999) investigated in four Nucleus 22 recipients the effect of reducing the EDR by (1) raising the T-levels to 75% of the original EDR, and (2) raising the T-levels to 75 % of the EDR and reducing the C-levels to 76% of the EDR per electrode, thus creating a binary representation. Speech was presented via direct audio input and was set at a comfortable loudness level by the recipient. Their results showed that vowel recognition was marginally but significantly reduced by EDR reduction, whereas consonant recognition was not reduced. They also investigated the effect of varying the acoustic to electric amplitude mapping in four steps from the default mapping, in which the upper 10 dB of the acoustical input range is projected onto the upper 20% of the EDR, to mapping this upper 10 dB range onto the upper 50% of the EDR. There were no differences in either vowel or consonant recognition at comfortably loud presentation levels in quiet and in noise between these conditions. Fu and Shannon (1998, 2000) found that consonant and vowel recognition were equally but only mildly affected when applying a strong compression or an expansion. Loizou et al. (2000b) found that in cochlear implant recipients the division of the EDR in only eight discrete CU steps was sufficient to reach asymptotic performance in consonant recognition. In normal-hearing subjects, who listened to a simulation of cochlear implant signal processing, compression affected vowel recognition to a larger degree than consonant recognition (Loizou et al., 2000a).

The above studies show that a slight raise in T-levels is beneficial to soft-speech perception, but lowering T-levels is not associated with a decrease in speech perception at presentation levels of 55 dB SPL and up. Variation in amplitude mapping has only mild effects on speech perception. The above studies were predominantly performed in Nucleus 22 CI recipients using the SPEAK strategy and with speech presented at comfortable loudness levels. Nowadays, the ACE strategy offers more freedom in amplitude mapping, since it is possible to map the upper 10 dB of the acoustical input in less than the upper 20% of the EDR. This flexibility in mapping may have a positive effect on speech perception scores at low presentation levels. However, trying each combination of settings in T- and C-level and amplitude mapping in each individual recipient is very time-consuming.

This study aims at answering the question to what extent speech perception is affected by manipulating its electrical representation in two ways; (1) by presenting speech only in the lower, middle or upper one third of the subject's EDR or (2) by mapping the upper 10 dB of the acoustic speech range into the upper 10, 20, or 30% of the EDR. The present research question is of a fundamental nature. It aims at more insight into the respective contributions of parts of the EDR. These fittings are not intended for everyday use. Therefore, the experiments are acute and speech perception is measured at the lowest level of recognition, namely at phoneme level. Presentation levels varied from normal to very low.

5.2 Methods

5.2.1 Subjects

Twenty-five postlingually deafened adult cochlear implant recipients with between one and five years of implant experience participated. Inclusion criteria were a Nucleus CI 24M or CI 24R implant and a phoneme score using their conventional fitting of at least 50% for monosyllabic words at a presentation level of 65 dB SPL in quiet. Subjects had been using the body worn Sprint or behind-the-ear Esprit 3G speech processor and the SPEAK or ACE speech processing strategy. Subject characteristics are listed in Table 1. A contralateral acoustic hearing aid, if present, was switched off during the measurements.

5.2.2 Speech processor fitting characteristics

Speech processing strategies such as SPEAK and ACE (Skinner et al., 1994; Seligman and McDermott, 1995; Skinner et al., 2002) use a spectral maxima selection process to stimulate between 6 and 10 electrodes per analyzed time frame. The acoustic signal from the microphone of the speech processor is band-pass filtered into a number of channels corresponding to the number of electrodes. For each analyzed time frame of the audio signal, M electrodes with the largest amplitudes are selected. The number of maxima M is fixed per patient and usually set between 6 and 10. Finally, the selected electrodes are stimulated using a compression function to project acoustic amplitudes onto the recipient's EDR for these selected electrodes. The shape of the acoustical-to-electrical mapping function, named loudness growth function (LGF) by the Cochlear Company, is determined by the Q-parameter value and the base level (Fig. 1). The Q value indicates the upper percentage of the EDR into which the upper 10 dB of the acoustic input signal is mapped. The base level parameter controls the acoustic input dynamic range and thereby the lowest input level that results in a stimulus at T-level. Stimulation amplitude is expressed in current level (CL), a quantity defined by Cochlear. The CL ranges from 1 to 255 current units (CU), which projects on the logarithm of the electrical current, ranging from 10 µA to 1.75 mA. An increase by 34 CU corresponds to an increase in current by a factor of 2.

Subject	Phoneme score [%]	Sex	Age [y]	Implant type	Implant use [y]	Strategy	Rate [Hz]	Active electrodes	Max	EDR [CU]
1	89	M	44	R (CS)	3	ACE	900	20	8	50
2	80	M	57	R (CS)	3	ACE	900	20	8	72
3	80	M	60	R (CS)	4	ACE	900	19	8	48
4	56	F	44	M	5	ACE	900	20	8	17
5	57	M	68	M	5	ACE	720	21	8	19
6	75	F	83	R (CA)	2	ACE	900	20	8	40
7	87	M	63	R (CS)	4	ACE	900	20	8	30
8	70	M	63	R (CS)	4	ACE	900	20	8	48
9	84	F	64	M	5	ACE	720	20	6	45
10	81	M	55	R (CA)	2	ACE	1200	20	8	39
11	72	F	73	R (CS)	5	ACE	900	20	8	31
12	76	F	20	M	5	ACE	1200	20	8	81
13	74	M	43	R (CS)	4	ACE	900	19	8	31
14	73	M	45	R (CS)	2	ACE	900	20	8	73
15	76	F	58	R (CA)	2	ACE	1200	20	9	56
16	68	M	44	R (CS)	4	ACE	900	20	8	23
17	78	M	36	R (ST)	1	ACE	1200	20	8	46
18	75	F	70	R (CA)	2	ACE	1200	20	8	52
19	72	F	57	R (CA)	2	ACE	900	20	8	47
20	56	M	76	M	5	SPEAK	900	20	8	25
21	74	M	70	M	5	SPEAK	900	20	8	40
22	54	F	42	R (ST)	4	ACE	900	20	8	40
23	63	F	57	R(CA)	1	ACE	1200	20	8	41
24	56	M	75	M	5	SPEAK	900	20	8	31
25	78	F	54	M	5	SPEAK	900	19	8	45

Table 1: Subject and fitting characteristics. The phoneme score is determined at 65 dB SPL in quiet using the subject's own routine fitting. Implant type is Nucleus 24M or R, with electrode types Straight (ST), Contour (CS), or Contour Advance (CA). Max is the number of maxima. EDR is the electrical dynamic range in current units. The pulse width of subject 8 was 100 _s, of the other subjects it was 25 _s.

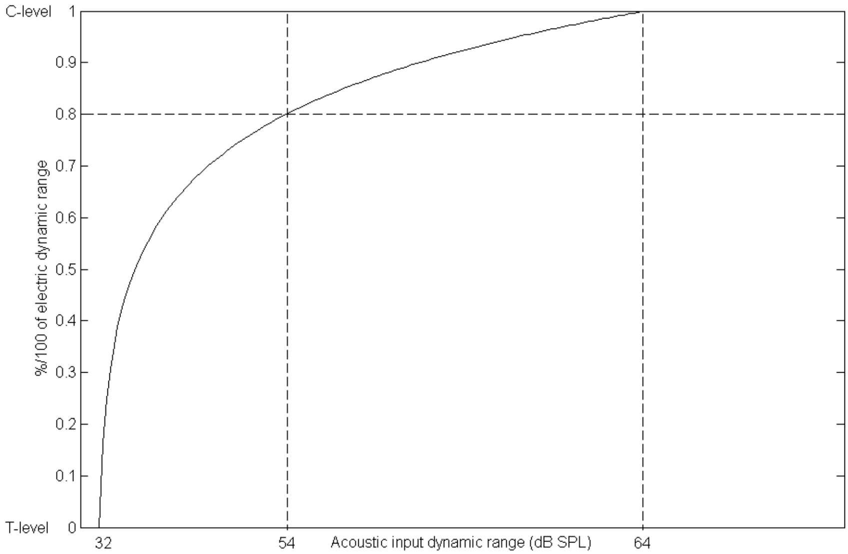


Figure 1: Acoustic-to-electric amplitude mapping according to the standard loudness growth function (LGF). The Q-parameter value determines the percentage of the EDR from C-levels downwards onto which the upper 10 dB of the acoustic speech range is mapped. In this example the Q value is at its default setting of 20. The base level determines the acoustic input dynamic range and thereby the lowest input level that results in a stimulus at T-level.

For all subjects speech processor fittings were created in the Custom Sound 1.2 software using the Sprint speech processor and ACE strategy, resulting in a *map*. For all subjects the same Sprint processor and microphone were used for the experiments. If subjects had been using the Esprit 3G processor, their maps were converted to Sprint using the conversion function in the Custom Sound software. Channel stimulation rate, pulse width and number of maxima were copied from the subject's own map and varied among subjects. Map characteristics are listed in Table 1. If subjects had been using the SPEAK strategy, a new map was created using the ACE strategy and a channel stimulation rate of 900 Hz. Stimulation mode was monopolar using both extracochlear reference electrodes (MP1+2). Fitting characteristics are listed in Table 1. All fittings, either pre-existing, converted, or newly created, were checked conventionally. That is, T- and C-levels were adjusted for electrodes 3, 7, 12, 17, and 22 individually stimulating with 500 ms bursts of impulses at a rate according to the channel stimulation rate. The T-level was defined as the lowest stimulus level per electrode that elicits a very soft, but consistent hearing sensation. The C-level was defined as the maximum stimulus level per electrode that still produces a comfortable loudness sensation. T- and C-levels were interpolated for the electrodes in between the measured electrodes. Volume and microphone sensitivity were set at default levels of 9 respectively 12. The Q value and base level were set at default value of 20 respectively 4. The combination of a base level of 4 and a microphone sensitivity of 12 resulted in theory in a lowest acoustic input level to be processed by the speech processor of 32 dB SPL. However, after the experiments the microphone characteristics were determined and the lowest input level appeared to be 29 dB SPL, which is within specification limits. The acoustic dynamic range is about 32 dB, which means that presentation levels above about 61 dB SPL were all projected at C-level. The fitting created following the above protocol was the standard map, map S, during the experiments.

Before starting the experiments, speech perception with the standard map was measured using CVC words, presented at 65 dB SPL in quiet (see “Speech audiometry” below). To verify that the conversion to another speech processor and/or speech processing strategy had not affected maximum speech perception scores, the phoneme score obtained was compared to the subject’s phoneme score found in the latest regular clinical evaluation.

Five experimental maps were created based on the standard map S (Fig. 2). In map L, M, and H, the full EDR was divided into three sections: low, middle, and high. They were evaluated separately. In map L, T-levels were unchanged, and C-levels were set at 33% of the full EDR for each electrode. In map M, they were set at 33% respectively 67%, and in map H at 67% respectively 100% of the full EDR for each electrode. The full acoustic dynamic range of 32 dB was projected on these parts of the EDR. In the final two maps, the Q value was varied, while keeping the standard map’s T- and C-levels. The Q value was set at 10 in map Q10 and at 30 in map Q30, while it was 20 in map S.

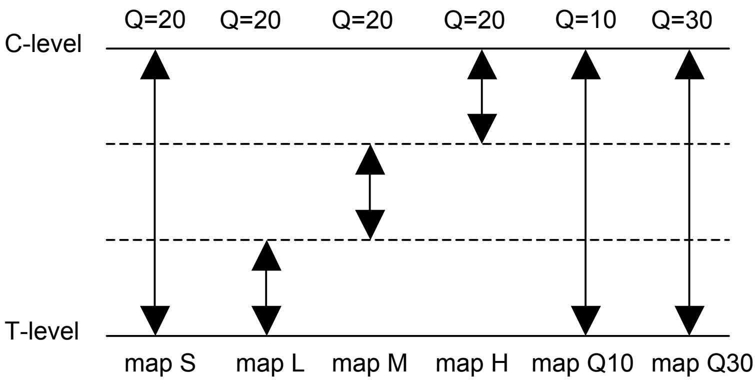


Figure 2: Standard map S and the five experimental maps L, M, H, Q10, and Q30. In map S, L, M, and H a Q value of 20 was used. In map L, the lower one-third of the EDR was used, in map M the middle, and in map H the upper third. In map Q10 and Q30 the full EDR was used, and the Q value was decreased to 10 in map Q10 and increased to 30 in map Q30.

5.2.3 Speech audiometry

Performance-intensity (P-I) functions (speech audiograms) were measured for each of the six maps using words presented in quiet. The words consisted of the sequence initial consonant - vowel - final consonant (CVC words). The word test consisted of 45 lists, each list containing a set of 12 unrelated linguistically meaningful words. They were Dutch female-spoken and played in the free field from a CD player, connected to an Interacoustics AC-40 audiometer. Subjects were to repeat the items perceived and were strongly encouraged to respond to each word presented. Excluding the first word, 11 words of a list were used to determine the percentage of phonemes responded correctly. Presentation levels were varied in 10 dB steps from 70 to 40 dB SPL. Per presentation level, the phoneme scores of two lists were averaged, which results in a standard deviation of the obtained phoneme score of less than 7 percentage points for phoneme scores roughly between 30 and 70% (Bosman, 1989). When the phoneme score dropped below 20%, measurements were terminated rather than trying lower presentation levels. If subjects did achieve a phoneme score higher than 20% at 40 dB SPL, words were also presented at a level of 35 dB SPL. The total number of word lists required for some subjects was more than 45, which implied that a limited number of word lists had to be used twice. The order of the conditions and word lists was varied at random over subjects.

5.2.4 Statistical analysis

For each of the experimental maps, individual phoneme scores per presentation level were compared to those obtained with the standard map. Because of the non-normal distribution of the scores, especially at high presentation levels due to ceiling effects, and at low presentation levels due to floor effects, the non-parametric Wilcoxon test for two related samples was used (SPSS 12.0, SPSS Inc., Chicago). In addition, perception scores were determined for the initial consonant, the vowel, and the final consonant separately, for each of the maps for normal and low presentation levels (60 respectively 40 dB SPL). For each of the three types of phonemes the score at either 60 or 40 dB SPL using the experimental map was compared to the score using the standard map, also using the Wilcoxon test for two related samples.

5.3 Results

5.3.1 Averaged speech audiometry curves

Using the standard map all subjects obtained a phoneme score within the measurement error of the score found at the yearly evaluation. Hence, the conversion to another speech processor and/or strategy had no effect on speech perception in quiet. Using map S all 25 subjects were able to obtain phoneme scores above 20% at presentation levels of 50, 60, and 70 dB SPL. At presentation levels of 40 and 35 dB SPL 15 respectively 7 subjects were able to achieve phoneme scores higher than 20%. Table 2 lists the number of subjects that achieved a phoneme score above 20%, for each of the maps and presentation levels.

Map	Presentation level (dBSPL)		
	35	40	50, 60, 70
S	7	15	25
L	1	15	22
M	8	21	25
H	13	24	25
Q10	8	24	25
Q30	6	18	25

Table 2: Number of subjects that achieved a phoneme score above 20%, which is significantly above 0, for each of the maps and presentation levels.

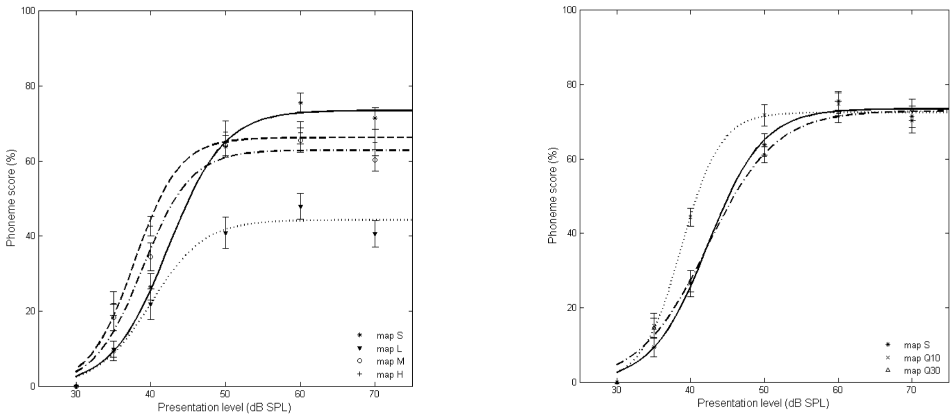


Figure 3: Speech perception scores averaged over subjects and the standard error of the mean. (a) Using the full EDR of the standard map S, and the maps using only one third of the EDR: L, M, and H. (b) Using the full EDR and a Q value of 20 in map S, of 10 in map Q10, and of 30 in map Q30. Matching of the curves and the data points was performed based on curve fitting with a logistic function.

Fig. 3a shows the speech perception scores averaged over subjects and the standard error of the mean using the full EDR of the standard map S, and the maps using only one third of the EDR: L, M, and H. Matching of the curves and the data points was performed based on curve fitting with a logistic function. There were large interindividual differences in the speech audiograms for each map. In map S this resulted in an interquartile range of about 20 percentage points at each of the presentation levels. In M and H the interquartile ranges were between 20 and 25 percentage points at each of the presentation levels. In map

L the variance in performance was higher, as three subjects were not able to perform the test. Interquartile ranges were 20 to 30 percentage points over the presentation levels. The shape of the average speech audiograms differs between the maps. Maps M and H show a left-shift of the presentation level at which 50% of the maximum score is achieved, compared to map S. The slopes of the curves in this point are about equal to the slope of map S. Map L shows a smaller slope of the curve in the 50% point than map S.

Fig. 3b shows the speech perception scores averaged over subjects and the standard error of the mean of map S, Q10, and Q30. Interquartile ranges were between 15 and 25 percentage points. Using map Q10 subjects achieve 50% of their maximum score at a lower presentation level than using map S. Using map Q30 the 50% point is about equal to map S. Map Q10 displays a steeper slope and map Q30 a weaker slope than map S.

5.3.2 Overall phoneme scores per level

Map	Presentation level (dB SPL)				
	35	40	50	60	70
L-S	-5 *	-5 *	-23 *	-28 *	-31 *
M-S	3	8 *	0	-10 *	-11 *
H-S	12 *	16 *	4	-8 *	-6 *
Q10-S	4	18 *	8 *	-1	-1
Q30-S	-2	1	-3	-3	2

Table 3: Averaged difference in phoneme score in percentage points between the experimental maps and the standard map, for each of the presentation levels. Significant differences in the non-parametric Wicoxon test for two related samples ($p < 0.02$) are denoted with an asterisk.

Table 3 displays the averaged percentage points difference for each of the presentation levels between the phoneme score of the experimental maps and the standard map. Significant differences in the Wilcoxon test are indicated with asterisks. One obvious result is that using map L, the phoneme scores at all presentation levels were significantly lower than using map S. Increasing the Q value to 30 did not affect phoneme scores at any of the presentation levels.

Maximum phoneme scores, i.e. those at 60 and 70 dB SPL, significantly decreased when using map L, M, and H compared to map S. The decrease in maximum phoneme score was largest, about 30 percentage points, when using map L. Decreasing the Q value to 10 did not affect maximum phoneme scores.

At a presentation level of 50 dB SPL the phoneme scores using map Q10 were higher than those using map S, whereas using map M and H, there were no differences compared to map S.

At low presentation level, i.e. 40 dB SPL, phoneme scores using the middle and upper third of the EDR were on average up to 16 percentage points higher than using the standard map. Changing the Q value to 10 resulted in an increase in phoneme scores at a low presentation level of on average up to 18 percentage points.

5.3.3 Consonant and vowel scores per level

Perception scores for the initial consonant, vowel, and final consonant using the standard map showed that reducing the presentation level from 60 to 40 dB SPL affected both initial and final consonant recognition, to a larger extent than vowel recognition. The relative decrease in vowel score was significantly smaller than de relative decrease in both the initial and final consonant score (means 56, 71, and 73% respectively, $p < 0.001$ in the Wilcoxon tests for two related samples for both comparisons between the vowel and consonants).

At presentation levels of 60 and 70 dB SPL, phoneme scores using map L, M, and H were significantly lower than using map S. Fig. 4a shows the initial consonant, vowel, and final consonant score for map S, L, M, and H at 60 dB SPL. Significant differences in the Wilcoxon test for two related samples between the scores using the experimental map compared to the standard map are indicated with asterisks. Using the lower, middle, or upper third of the EDR the vowel scores and both consonant scores were decreased compared to vowel and consonant recognition using map S.

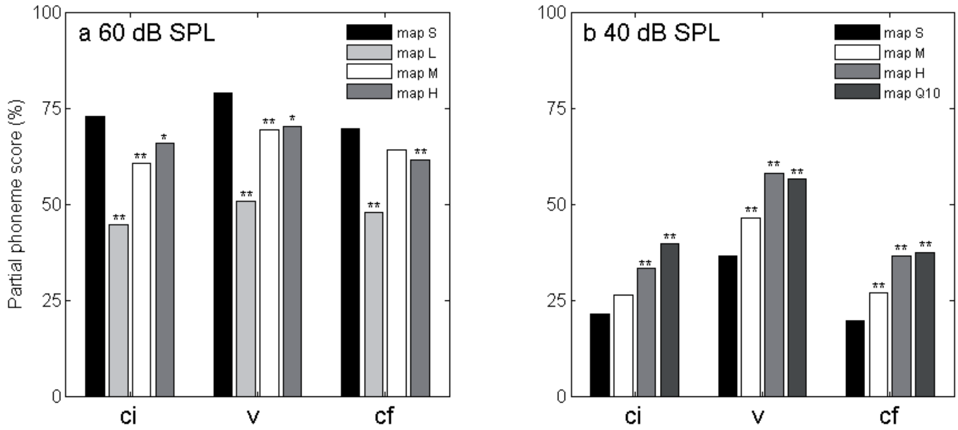


Figure 4: Initial consonant, vowel, and final consonant scores: (a) for map S, L, M, and H at 60 dB SPL, and (b) for map S, M, H, and Q10 at 40 dB SPL. Asterisks indicate a significant difference ($p < 0.02$) in the Wilcoxon test for two related samples, with the standard map S as the reference.

Fig. 4b shows the perception scores for initial consonant, vowel, and final consonant at 40 dB SPL obtained with map S, M, H, and Q10. At 40 dB SPL the initial consonant score using map M was equal to the score using map S, whereas the vowel and final consonant score was significantly higher using map M. Vowel and both consonant scores using the experimental maps H and Q10 were equally increased compared to those using the standard map.

5.4 Discussion

5.4.1 Speech audiometry using the standard map

Using the standard map a reduction of the presentation level from 60 to 40 dB SPL resulted in a decreased phoneme score primarily due to lower initial and final consonant scores. As vowels have higher intensities than consonants, their audibility is less affected compared to consonants. This finding is in accordance with results in cochlear implant recipients using the SPEAK strategy when the presentation level decreases from 70 to 50 dB SPL (Skinner et al., 1997). Bosman and Smoorenburg (1995) found the same effect in normal hearing subjects. However, they found in hearing impaired subjects that consonant and vowel recognition are equally affected by reducing presentation levels.

Using map S, but also using map L, M, and Q10, the phoneme scores at a presentation level of 60 and 70 dB SPL differed significantly ($p < 0.05$, in each Wilcoxon test for two related samples). This 'roll-over' effect was also reported by Skinner et al. (1997) and Donaldson and Allen (2003). It is attributed to the compression of the highest speech peaks by the processor's automatic gain control.

In research regarding speech perception with a CI, often presentation levels of 50 or 55 dB SPL are used to determine the perception of soft speech. However, Fig. 3a shows that at these presentation levels speech perception scores are still close to the maximum score. When using the default sensitivity setting of 12, it seems better to use a presentation level of 45 dB SPL to examine the perception of low-level speech.

5.4.2 Quantifying speech audiometry curves

It would be elegant to quantify the speech audiometry curves by curve fitting. This would allow for an estimation of the shift of the presentation level at which 50% of the maximum score is achieved and the slope of the curve in this point, and thus for a quantitative comparison of the shape of the curves between maps. In Fig. 3 we used a logistic function. This function was suitable to fit speech audiometry curves using the standard map S and obtaining the two measures described above. However, especially using map L, this function did not appear to give accurate results for shift and slope, as the shape of the fitted curve had a more logarithmic rather than a logistic shape in eleven subjects. Since lack of proper fitting resulted in missing data for at least eleven subjects, we did not use the curve fitting in quantifying the shape of the speech audiometry curves, but described the comparison of phoneme scores between the maps for each presentation level.

5.4.3 Speech perception using one-third of the EDR

In the current study measuring speech perception using a part of the EDR is different from measuring the contribution of that part to the whole in a fitting using the full EDR. In our study the full acoustic input range was presented to only a part of the EDR, whereas in a stimulation map using the full EDR that part of the EDR would only receive information from a specific range of the acoustic information. Therefore only relative contributions of the lower, middle, and upper third of the EDR can be measured. The fundamental design of our study allows us to measure and compare the processing capacities of the three parts of the EDR.

An important finding is that when using only the lower third of the EDR and speech is presented at 60 dB SPL, subjects are able to obtain on average two-third of their speech perception score using the full EDR. Apparently, stimuli at low levels contain a substantial amount of recognizable speech information. Using the middle or upper third of the EDR, maximum speech perception scores are lower than using the full EDR. One factor playing a part in this effect is the largely reduced EDR, whereas another factor is a change in perceived loudness of speech when using map M or H compared to map S. The compression of speech in the reduced EDR is for all maps the same and is a negative factor in speech perception. The perceived loudness of normal-level speech, however, is decreased in map M and increased in map H compared to map S. In map M, the smaller EDR and lower perceived loudness both affect speech perception negatively. In map H, the smaller EDR has a negative effect, but the higher perceived loudness might have a positive effect on speech perception scores.

The answer to the question which factor weighs more heavily on speech perception scores, either a reduced EDR or an increase in loudness, comes from analyzing speech perception using map M and H at a presentation level of 40 dB SPL. The reduction in EDR is still a negative factor for speech perception. The perceived loudness of low-level speech is higher in both maps M and H, because on average stimuli are presented at higher levels. This is a positive factor in speech perception. Using only the middle or the upper one-third of the EDR, speech perception at 40 dB

SPL is improved compared to using the full EDR. This implies that a higher stimulation level, and hence perceived loudness, is the most important factor in the effects on speech perception.

The effect of reduced audibility due to lowering the presentation level from 60 to 40 dB SPL is smaller using map M and H than when using the standard map, especially when using the upper third of the EDR. Vowels and consonants gain equally from this effect.

The studies described in the 'Introduction' found at most a small effect on vowel and consonant recognition when reducing the EDR (Zeng and Galvin, 1999, Loizou, 2000a). They did not fix the presentation level, but made the subjects adjust the presentation level to a comfortable perceived loudness. As perceived loudness proved to be the most important factor in speech perception scores, the effects found in the above studies were rather small.

5.4.4 Adjusting the Q value

Decreasing the Q value to 10 leads to an increased audibility of stimuli when presentation levels are low. The phoneme scores at low presentation levels are higher than with the standard map. The increased audibility favors both vowels and consonants. Increasing the Q value to 30 does not affect speech perception scores at any of the presentation levels. This is in accordance with the results of Zeng and Galvin (1999), who did not find an effect of increasing the Q value above 20. However, speech was presented only at a comfortably loud presentation level and not at lower levels.

5.4.5 Implications for the speech processor fitting procedure

Using only the lower third of the EDR subjects are able to obtain on average two-third of their speech perception score using the full EDR when speech is presented at 60 dB SPL. Subjects are able to extract a substantial amount of recognizable speech information from these low electrical stimulation levels. However, the upper part of the EDR proves to be most important for speech understanding. Raising T-levels in general improves speech perception in quiet when presentation levels are low. However, it reduces the EDR and eliminates the use of the lower part of the EDR, which does convey information, as shown above. Decreasing the Q value to 10 also improves speech perception in quiet at low presentation levels and does not have a negative effect on the maximum phoneme scores.

The ECAP-based fitting procedure comes with T-levels that are on average 30 CU lower than in the conventional fitting and thus with an increased EDR. Due to the lower T-levels, the perceived loudness at very low presentation levels is smaller compared to the conventional fitting. This may imply lower phoneme scores for low-level speech. Therefore, it might be beneficial to compensate that with a Q value of less than 20, e.g. 10. Another suggestion would be to set the T-levels well above the threshold of perception to improve low-level speech perception. The optimal setting of the T-levels should then be determined while presenting speech at a low-level. However, the effects of both procedures (decreasing the Q value and raising T-levels) may be different in noisy conditions from the quiet condition. This needs to be investigated.

References

Bosman, A.J. (1989). Speech perception by the hearing impaired. Thesis. Utrecht University, the Netherlands, Chapter 3.

Bosman, A.J. and Smoorenburg, G.F. (1995). Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. Audiol. 34: 260-84.

Donaldson, G.S. and Allen, S.L. (2003). Effects of presentation level on phoneme and sentence recognition in quiet by cochlear implant listeners. Ear Hear. 24(5): 392-405.

Fu, Q.J. and Shannon, R.V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. J. Acoust. Soc. Am. 104(5): 2570-7.

Fu, Q.J. and Shannon, R.V. (2000). Effects of dynamic range and amplitude mapping on phoneme recognition in Nucleus-22 cochlear implant users. Ear Hear. 21(3): 227-35.

Loizou, P.C., Dorman, M. and Fitzke, J. (2000a). The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants. Ear Hear. Feb 21(2): 25-31.

Loizou, P.C., Dorman, M., Poroy, O. et al. (2000b). Speech recognition by normal-hearing and cochlear implant listeners as a function of intensity resolution. J. Acoust. Soc. Am. 108(5): 2377-87.

Seligman, P.M and McDermott, H.J. (1995). Architecture of the Spectra 22 speech processor. Ann. Otol. Rhinol. Laryngol., 104(Suppl. 166), 172–175.

Skinner, M.W., Clark, G.M., Whitford, L.A. et al. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. Am. J. Otol. 15(Suppl. 2), 15–27.

Skinner, M.W., Holden, L.K., Holden, T.A. et al. (1997). Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants. J. Acoust. Soc. Am. 101: 3766-82.

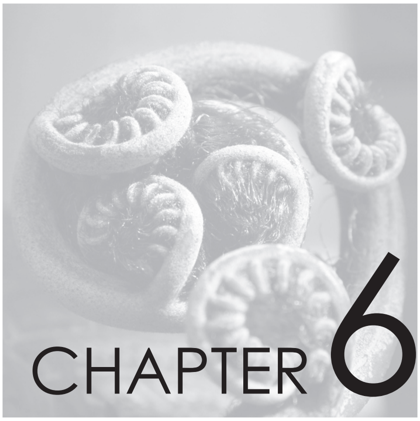
Skinner, M. W., Holden, L.K., Holden T.A. et al. (1999). Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. J. Speech Lang. Hear. Res. Aug 42: 814-28.

Skinner, M.W., Arndt, P.L., Staller, S.J. (2002). Nucleus 24 advanced encoder conversion study: performance versus preference. Ear Hear. Feb 23(1 Suppl): 2S-17S.

Smoorenburg, G.F., Willeboer, C., van Dijk, J.E. (2002). Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds. Audiol. Neuro-otol. 7(6):335-47.

Willeboer, C., Smoorenburg, G.F. (2006). Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study. Ear Hear. 27(6):789-98.

Zeng, F.G. and Galvin, J.J. III (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners. Ear Hear. Feb 20(1): 60-74.



Effect on speech perception of eliminating low-level stimuli from the electrical dynamic range in cochlear implant recipients

Christina Willeboer, Maaike Bierman, Guido F. Smoorenburg, and Gijsbert A. van Zanten

Abstract

Objective: This study's objective is to investigate the importance for speech perception of electrical stimulation in the lower part of the electrical dynamic range (EDR) in cochlear implant recipients by eliminating these low-level speech elements from the EDR.

Methods: Twenty-four experienced postlingually deafened adult Nucleus 24M or 24R CI recipients participated. Their phoneme scores using their conventional take-home fitting were at least 50% at a presentation level of 65 dB SPL. A simulation model of the speech processor was created using the Nucleus Matlab Toolbox 3.02 and Nucleus Interface Communicator software (both Cochlear). This allowed for manipulation of the signal processing strategy beyond the possibilities of changing clinical parameters in the fitting software. The loudness growth function (LGF), which determines the projection of the acoustic information onto the EDR, was manipulated by the introduction of a cut-off level below which all stimuli were presented at T-level. Speech perception was measured in quiet at a presentation level of 60 dB SPL with cut-off levels ranging from 30 to 55 dB SPL corresponding to 19 to 80% of the EDR.

Results: Cutting-off low-level stimuli up to 35 dB SPL or 43% of the EDR did not affect speech perception scores. Higher cut-off levels did significantly decrease phoneme scores. However, the decrease was small for cut-off levels up to 45 dB SPL. For example, a cut-off level of 40 dB SPL corresponding to 58% of the EDR resulted in a phoneme score of 90% of the original phoneme score obtained with the standard EDR at 60 dB SPL presentation level. Consonants were affected more than vowels. Surprisingly, subjects with a relatively low phoneme score obtained with the standard EDR improved after cutting-off at 30 dB SPL or 19% of the EDR.

Conclusions: For speech perception at 60 dB SPL low-level stimuli are less important than might have been assumed generally. This explains why decreasing T-levels below the conventional values does not change speech perception at normal presentation levels. Cutting-off low-level stimuli removes a larger part of the consonants' electrical representation than of the vowels', as their intensity is lower. Poor performers gain from cutting-off up to 19% of the EDR, probably because the gaps between the electrical representation of speech segments become wider and the temporal structure of the stimulation pattern thus becomes more distinct. Whether this effect holds for lower presentation levels than the 60 dB SPL used in our study needs to be investigated. Eliminating low-level stimuli may especially be beneficial in noisy conditions, as it does not affect normal level speech perception, but may reduce disturbing noise.

6.1 Introduction

Fittings of a cochlear implant (CI) require the determination of the threshold level of detection (T-level) and of the comfortable loudness level (C-level) using electrical pulse trains for each of the intracochlear electrodes. The process of adjusting the T- and C-levels per electrode is interactively done by the audiologist and recipient. The recipient has to listen to sounds evoked by stimulation with electrical pulse trains presented to a single electrode. Below, this fitting procedure is referred to as conventional fitting. The acoustic information from the microphone of the speech processor is projected onto the range between T- and C-levels, the electrical dynamic range (EDR). The microphone sensitivity setting determines the lowest input level that is processed and that results in stimulation at T-level. In the Nucleus CI 24M or CI 24R implant system (Cochlear, Lane Cove Australia) the lowest acoustic input level at a default sensitivity setting is around 32 +/- 5 dB SPL across microphones. As the input dynamic range is about 32 dB SPL, acoustic stimulation levels above around 64 dB SPL result in stimulation at C-level, i.e. there is infinite compression of sound above 64 dB SPL. The amplitude mapping function, named loudness growth function (LGF) by Cochlear, determines the projection of the acoustic information onto the EDR. Fig. 1 shows the standard LGF. The upper 10 dB of the acoustic input is projected onto the upper 20% of the EDR. Detailed information about amplitude mapping in the Nucleus system is given in Zeng and Galvin (1999).

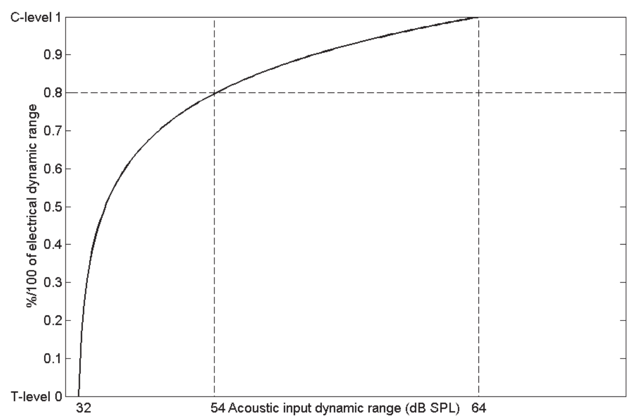


Figure 1: Acoustic-to-electric amplitude mapping according to the standard loudness growth function (LGF).

Previously, we investigated a fitting method not based on electrical pulse trains presented per single electrode, but on live speech presented across the whole electrode array (Smoorenburg et al, 2002; Willeboer and Smoorenburg, 2006). Wide-band running speech stimulates a number of frequency channels nearly simultaneously and results in an effectively higher pulse presentation rate due to summation of stimulation across electrodes. This leads to lower thresholds than when a pulse train is presented to one single channel in the conventional fitting procedure. For recipients of the Nucleus 24 implant system T-levels appeared much lower using our live speech fitting procedure, whereas C-levels were about equal to the conventional ones. Hence, the EDR was about one-third larger than using the conventional fitting method. Speech perception scores with this downward enlarged EDR were nevertheless equal to the scores for the conventional fitting, even at presentation levels as low as 55 dB SPL.

This result led to question how much low-level speech elements do contribute to speech perception. Our previous study (Willeboer et al., not published yet) showed that, when using only the lower one-third of the EDR in the conventional fitting and speech presented at normal presentation levels, CI recipients are still able to obtain a phoneme score of, on average, two-third of their phoneme score with the full EDR. Evidently, stimuli in the lower third of the EDR contain relevant speech information, at least when stimulating with the full acoustic input range of between 32 and 64 dB SPL. However, when mapping the acoustic input range of 32 to 64 dB SPL in the full EDR and not only the lower third, the LGF determines that only speech elements between roughly 32 and 36 dB SPL are projected in the lower third of the EDR. In this study the contribution to speech perception of elements in the lower part of the EDR is evaluated by eliminating these stimuli from the lower part of the EDR.

Eliminating stimuli from the lower part of the EDR in electric hearing might be compared with center-clipping in acoustic hearing (Licklider, 1946). Center-clipping eliminates the center parts of the acoustic wave and passes only the peaks. The amount of center-clipping is indicated by the number of dB by which the peak amplitude of the wave is reduced. A difference between center-clipping in acoustic hearing and eliminating stimuli from the lower part of the EDR is that the first is performed on the full range of frequencies in the acoustic signal at once, i.e. on the fine-structure as well as the envelope of the sound waveform. In CI only the envelope of band-pass filtered sound is coded for presentation per electrode. Therefore eliminating stimuli from the lower part of the EDR will be related to the sound envelope, not the fine-structure. Licklider (1946) showed in normal hearing subjects that acoustic center-clipping is much more detrimental to speech perception of words than peak-clipping. He concluded that the effect is caused by the fact that center-clipping strips out the weak consonant sounds and leaves only the vowels. With regard to speech intelligibility, the vowels are less important than the consonants. Drullman (1995) investigated a form of center-clipping in which not the peaks were reduced, but in which the troughs of the signal were filled at a certain level. This may be considered as creating an artificial noise-floor. Presenting sentences in quiet to normal hearing subjects, center-clipping led to only a modest degradation in phoneme recognition until the clipping exceeded 50% of the entire amplitude range.

In the present study the effect on speech perception of eliminating stimuli from the lower part of the EDR is investigated in CI recipients. In a simulation model of the speech processor, created with the Nucleus Matlab Toolbox (NMT) and the Nucleus Interface Communicator (NIC), the LGF was manipulated to eliminate stimuli from the lower part of the EDR, while the projection of the acoustic dynamic range onto the EDR for higher-level stimuli remained unchanged. Speech perception was measured in quiet at the lowest level of recognition, namely at phoneme level.

6.2 Methods

6.2.1 Subjects

Twenty-four postlingually deafened adult cochlear implant recipients with between one and five years of implant experience participated. Additional inclusion criteria were a Nucleus CI24M or CI24R implant and a phoneme score of at least 50% for monosyllabic words at a presentation level of 65 dB SPL in quiet. Subjects had been using the SPEAK or ACE speech processing strategy and the Sprint or Esprit 3G speech processor. Subject characteristics are listed in Table 1. For this study all subjects used the ACE stimulation strategy and the same Sprint speech processor or the simulated one. A contralateral hearing aid, if present, was switched off during the experiments.

Subject	Phoneme score [%]	Sex	Age [y]	Implant type	Implant use [y]	Strategy	Rate [Hz]	Active electrodes	Max	EDR [CU]
1	89	M	44	R (CS)	3	ACE	900	20	8	50
2	80	M	57	R (CS)	3	ACE	900	20	8	72
3	80	M	60	R (CS)	4	ACE	900	19	8	48
4	56	F	44	M	5	ACE	900	20	8	17
5	57	M	68	M	5	ACE	720	21	8	19
6	75	F	83	R (CA)	2	ACE	900	20	8	40
7	87	M	63	R (CS)	4	ACE	900	20	8	30
8	70	M	63	R (CS)	4	ACE	900	20	8	48
9	84	F	64	M	5	ACE	720	20	6	45
10	81	M	55	R (CA)	2	ACE	1200	20	8	39
11	72	F	73	R (CS)	5	ACE	900	20	8	31
12	76	F	20	M	5	ACE	1200	20	8	81
13	74	M	43	R (CS)	4	ACE	900	19	8	31
14	73	M	45	R (CS)	2	ACE	900	20	8	73
15	76	F	58	R (CA)	2	ACE	1200	20	9	56
16	68	M	44	R (CS)	4	ACE	900	20	8	23
17	75	F	70	R (CA)	2	ACE	1200	20	8	52
18	72	F	57	R (CA)	2	ACE	900	20	8	47
19	56	M	76	M	5	SPEAK	900	20	8	25
20	74	M	70	M	5	SPEAK	900	20	8	40
21	54	F	42	R (ST)	4	ACE	900	20	8	40
22	63	F	57	R(CA)	1	ACE	1200	20	8	41
23	56	M	75	M	5	SPEAK	900	20	8	31
24	78	F	54	M	5	SPEAK	900	19	8	45

Table 1: Subject and fitting characteristics. The phoneme score is determined at 65 dB SPL in quiet using the subject's own take-home fitting. Implant type is Nucleus 24M or R, with electrode types Straight (ST), Contour (CS), or Contour Advance (CA). Max is the number of maxima. EDR is the electrical dynamic range in current units. The pulse width of subject 8 was 100 _s, of the other subjects it was 25 _s.

6.2.2 ACE speech processing strategy in the Sprint speech processor

The ACE speech processing strategy uses a spectral maxima selection process to stimulate between 6 and 10 electrodes per analyzed time frame (Skinner et al., 2002). It consists of four components: the front-end, the filter bank, the sampling-and-selection component, and the amplitude mapping component.

In the front-end the acoustic signal from the microphone of the speech processor is amplified by a 30 dB gain preamplifier. The preamplifier has an equalizer that can be enabled as needed. Peaks in the acoustic signal are limited to values that eventually translate into stimulation at C-level by a process that the Cochlear Company calls automatic gain control (AGC), but in fact is peak clipping. The resulting range of acoustic input that produces stimulation between the T- and C-levels is about 32 dB. The sensitivity control sets the acoustic level above which the AGC will become active.

In the filter bank, the signal is band-pass filtered into a number of channels corresponding to the number of electrodes. In the sampling-and-selection component of the speech processor, M electrodes with the largest envelope-amplitudes are selected for each analyzed time frame of the audio signal. The number of maxima M is fixed per patient and is usually set between 6 and 10.

Finally, in the amplitude mapping component of the speech processor the selected electrodes are stimulated, using the LGF to project acoustic amplitudes onto the recipient's EDR. The Q-parameter value indicates the upper percentage of the EDR for each electrode onto which the upper 10 dB of the input signal is mapped. The base level controls the acoustic input dynamic range and thereby the lowest input level that results in a stimulus at T-level. Stimulation strength is expressed in current level (CL), a quantity defined by Cochlear. The CL ranges from 1 to 255 current units (CU), which projects on the logarithm of the electrical current, ranging from 10 µA to 1.75 mA. An increase by 34 CU corresponds to an increase in current by a factor of 2.

6.2.3 Speech processor fittings

Speech processor fittings were performed using the Custom Sound 1.2 software, using the Sprint speech processor and ACE strategy. For all subjects the same Sprint processor and microphone were used for the experiments. If subjects had been using the Esprit 3G processor, their fittings were converted to the Sprint using the conversion function of the Custom Sound software. Channel stimulation rate, pulse width and number of maxima were copied from the subject's own fitting and differed among subjects. If subjects had been using the SPEAK strategy, a new fitting was created using the ACE strategy at a channel stimulation rate of 900 Hz. Stimulation mode was monopolar using both extracochlear reference electrodes (MP1+2). Volume and microphone sensitivity were set at default levels of 9 respectively 12. The Q value and base level were set at default value of 20 respectively 4. Fitting characteristics are listed in Table 1. All fittings, either pre-existing, converted, or newly created, were checked conventionally. That is T- and C-levels were adjusted for electrodes 3, 7, 12, 17, and 22 individually stimulating with 500 ms bursts of impulses at a rate according to the channel stimulation rate. The T-level was defined as the lowest stimulus level per electrode that elicits a very soft, but consistent hearing sensation. The C-level was defined as the maximum stimulus level per electrode that still produces a comfortable loudness sensation. T- and C-levels were interpolated for the electrodes in between the measured electrodes.

Before starting the experiments, speech perception with this fitting was measured using CVC words in quiet, presented in the free field at 60 dB SPL (see 'Speech perception measurements' below). In order to verify that the conversion to another speech processor and/or speech processing strategy had not affected maximum speech perception scores, the phoneme score obtained was compared to the subject's phoneme score found in the latest regular clinical evaluation.

6.2.4 Sprint simulation model

The Nucleus Matlab Toolbox 3.02 (NMT) (Cochlear) is a research toolbox using the Matlab environment (The Mathworks, Natick, Massachusetts). The toolbox contains a simulation model of the Sprint speech processor and the ACE speech processing strategy. The toolbox allows for manipulation of the signal processing strategy beyond the possibilities of changing clinical parameters in the fitting software. Patient specific fitting parameters like T- and C-levels, channel stimulation rate, and LGF were imported into the model. The Nucleus Interface Communicator software 2.0 (NIC) (Cochlear) was used as a link between the NMT and the L34 research speech processor. The L34 speech processor directly feeds to the Nucleus 24 implant like the commercial speech processors do.

To a large extent the components of the model in the NMT were similar to the components of the ACE strategy in the Sprint processor described above. A difference was found in the front end. The model did not use a microphone to process an acoustic signal, but digital audio files in WAV-format instead. The WAV-files were re-sampled to 16 kHz and filtered according to the frequency response of the Sprint microphone, i.e. amplification between 1 and 6 kHz. An input level parameter was assigned to the WAV-file to determine the presentation level. In the NMT the sensitivity control was available, but the AGC was not available as a component. Another difference was found in the amplitude mapping component. In this component, the LGF could be manipulated in order to delete low-stimulation current components (see ‘Manipulation of the amplitude mapping component’ below).

6.2.5 Calibration of the Sprint simulation model

The output of the Sprint simulation model was calibrated according to the output of the Sprint speech processor used for the fitting in Custom Sound and for the speech perception measurements in the free field condition. Four fittings with EDRs within the range of those of the study group were used to calibrate the output of the simulation model to that of the Sprint processor.

In a sound insulated box the microphone of the Sprint speech processor was placed in a pure tone sound field of 1 kHz at levels from 25 to 90 dB SPL. The stimulation pattern created by the Sprint and sent to the coil as radiofrequency (RF) signals was captured and decoded by a Clinical Programming System (CPS, Cochlear), connected to a PC with the software program RF Statistics 1.03 (Cochlear). As the LGF is equal for all electrodes, only the output from electrode 16 was used for comparison with the simulation model.

In the simulation model WAV-files of a pure tone of 1 kHz were presented at simulated presentation levels of between 25 and 90 dB SPL by varying the input parameter of the model. The coil of the L34 processor was connected to the CPS and the RF signals for electrode 16 were captured in the same way as using the Sprint.

The captured current levels of the Sprint speech processor and the simulation model were compared. The input presentation levels of the simulation model were adjusted in order to obtain results equal to those using the Sprint processor. Using the above procedure, also the characteristics of the Sprint microphone used could be determined. With this result, the cut-off levels for eliminating low-level stimuli in dB SPL can be related to a cut-off percentage of the EDR (see 'Manipulation of the amplitude mapping component' below).

6.2.6 Manipulation of the amplitude mapping component

In order to eliminate stimuli from the lower part of the EDR, the LGF was manipulated by varying the lower boundary for the envelope amplitude. As an example Fig. 2 shows the manipulated LGF with a cut-off level of 45 dB SPL, corresponding to 69% of the EDR. The result of the manipulation was that stimuli with amplitudes below the cut-off level were all presented at 0% of the EDR, i.e. at T-level. As an example Fig. 3 shows the electrodogram (this is the stimulation pattern in time for each electrode) of the word /d œ n/ presented at 60 dB SPL using the simulation model with standard LGF and the LGF-manipulated strategy at cut-off levels of 40 and 50 dB SPL, corresponding to 58 respectively 80% of the EDR. It shows that with increasing cut-off level, more stimuli disappear and the gaps between pulse trains become wider.

A limitation of manipulating the LGF is that it does not fully eliminate low level stimuli. It results in presenting all stimuli below the cut-off level at T-level. Stimulation at T-level results in an audible, though very weak, sensation. Using manipulation of the LGF, however, this method comes as close as possible to eliminating low-level stimuli.

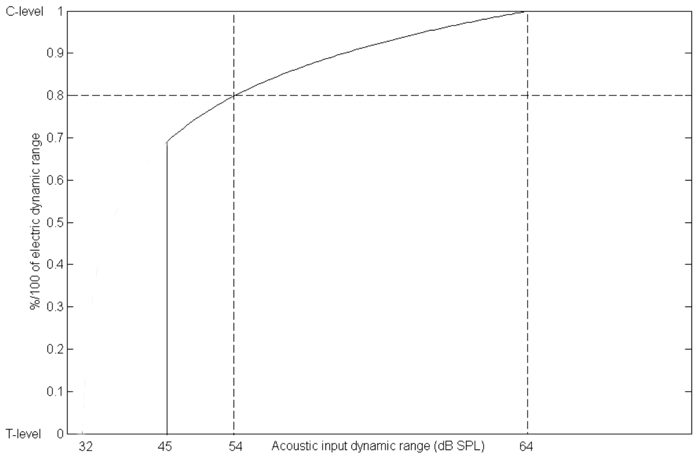


Figure 2: Acoustic-to-electric amplitude mapping according to the manipulated LGF. In this example a cut-off level of 45 dB SPL, corresponding to 69% of the EDR, is shown.

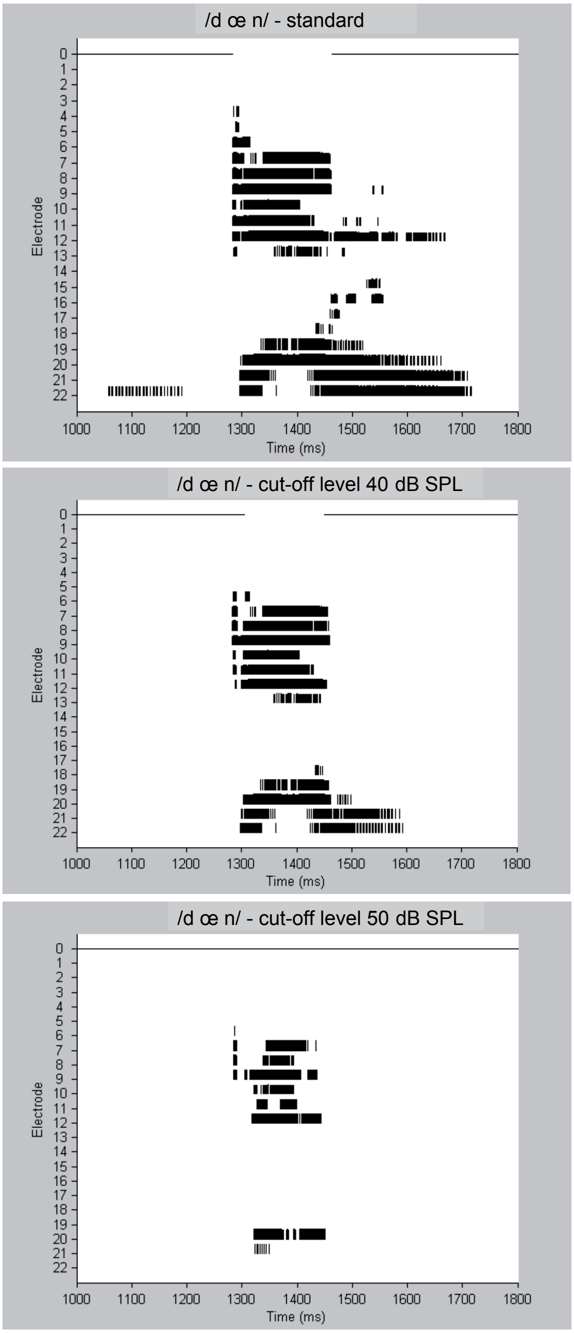


Figure 3: Electrodogram (this is the stimulation pattern in time for each electrode) of the word /d œ n/ presented at 60 dB SPL using the simulation model with standard LGF and the LGF-manipulated strategy at cut-off levels of 40 and 50 dB SPL, corresponding to 58 respectively 80% of the EDR.

6.2.7 Speech perception measurements

Speech perception was measured using words consisting of the sequence initial consonant - vowel - final consonant (CVC words). The word test consisted of 45 lists, each list containing a set of 12 unrelated linguistically meaningful Dutch female-spoken words. Subjects were to repeat the items perceived and were strongly encouraged to respond to each word presented. Excluding the first word, 11 words of a list were used to determine the percentage of phonemes responded correctly. In each condition, the phoneme scores of two lists were averaged. Presentation of two lists per condition results in a standard deviation of the obtained phoneme score of less than 7 percentage points for scores roughly between 30 and 70% (Bosman, 1989). When the phoneme score dropped below 20%, measurements were terminated rather than trying lower presentation levels or higher cut-off levels. Subsequently, the phoneme score was set at 0% for lower presentation levels or higher cut-off levels. Next to the overall phoneme scores per condition, also perception scores for the initial consonant, vowel and final consonant separately were determined for each condition.

After creating the fitting in the Custom Sound software, the CVC test was performed using the Sprint speech processor in the free field condition. Words were played from a CD player connected to an Interacoustics AC-40 audiometer at a presentation level of 60 dB SPL.

In the simulation model the words were presented as WAV-files, drawn from the CD, to the front end of the model in the NMT. The output of the NMT was sent to the L34 speech processor connected to the laptop. First, a performance-intensity (P-I) function was measured using the simulation model with the standard LGF. Presentation levels were 70, 60, 50, 40, and 35 dB SPL. Second, the LGF-manipulated strategy was used. Words were presented at 60 dB SPL and cut-off levels were varied from 30 to 55 dB SPL in 5 dB steps.

6.2.8 Data analysis

In the P-I function using the simulation model, the individual vowel and consonant scores were compared for each presentation level. Because of the non-normal distribution of the scores, especially at high presentation levels due to ceiling effects, and at low presentation levels due to floor effects, the non-parametric Wilcoxon test for two related samples was used (SPSS 12.0, SPSS Inc., Chicago).

The phoneme scores using the various cut-off levels were normalized with respect to the phoneme score obtained with the standard LGF simulation model at a presentation level of 60 dB SPL. An identical normalization was performed on the perception scores for the initial consonant, vowel, and final consonant separately. This allowed for graphically displaying the average of the phoneme scores over the whole group. As statistical analysis was a within-subject analysis, the scores without normalization were used. In Wilcoxon tests for two related samples the phoneme scores obtained with each cut-off level were compared to the scores using the standard LGF simulation model at 60 dB SPL. The same analyses were performed for the initial consonant, vowel, and final consonant separately.

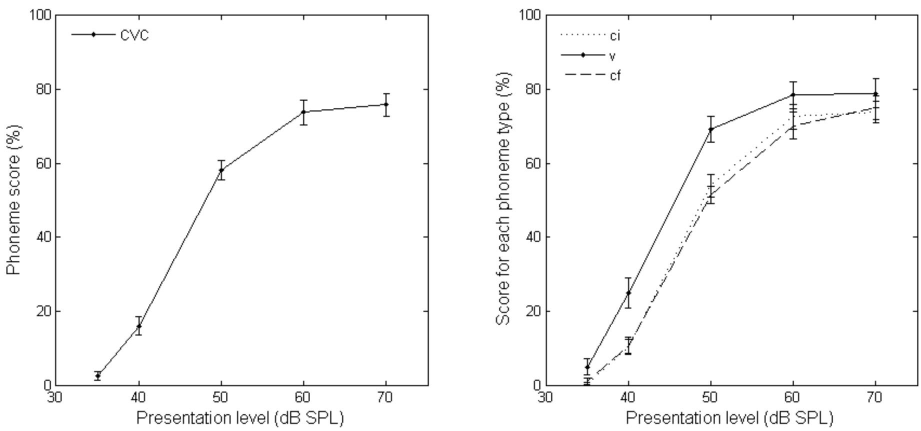


Figure 4: Averaged P-I function and the standard error of the mean (SEM) using the simulation model with the standard LGF. (a) Phoneme score averaged over all three phonemes and (b) perception scores for the initial consonant, the vowel, and the final consonant separately.

6.3 Results

6.3.1 Calibration of the model

The Sprint microphone used in the first section of the experiments, i.e. during fitting and the speech perception measurement in the free field, appeared to have a lowest input level of 29 dB SPL when using a default microphone sensitivity of 12. With this result, the cut-off levels in dB SPL were converted to the corresponding percentage of the EDR. They are listed in Table 2.

Cut-off levels	
dB SPL	% of the EDR
30	19
35	43
40	58
45	69
50	80
55	90

Table 2: Conversion of cut-off levels in dB SPL to percentage of the EDR, based on the calibration of the Sprint processor microphone used.

6.3.2 Speech perception using the standard LGF simulation model

Fig. 4a shows the averaged P-I function and the standard error of the mean (SEM) using the simulation model with the standard LGF. In eight subjects the lowest presentation level yielding a score above 20% was 50 dB SPL, in fifteen subjects it was 40 dB SPL, and in one subject it was 35 dB SPL. Fig. 4b shows the averaged P-I function and SEM for the initial consonant, the vowel, and the final consonant using the simulation model with the standard LGF. At presentation levels of 60, 50, and 40 dB SPL the vowel scores were significantly higher than both the consonant scores ($p < 0.02$, Wilcoxon test for two related samples). Hence, with decreasing presentation level the consonant scores decreased significantly more than the vowel score.

6.3.3 Speech perception using the LGF-manipulated model

All subjects were able to perform the speech perception measurements using cut-off levels from 30 to 45 dB SPL, corresponding to 19 to 69% of the EDR. In six subjects, a cut-off level of 45 dB SPL was the highest cut-off level yielding a phoneme score above 20%. In seventeen subjects it was 50 dB SPL (corresponding to 80% of the EDR) and in one subjects the highest cut-off level yielding a phoneme score above 20% was 55 dB SPL (corresponding to 90% of the EDR). The phoneme score obtained at 60 dB SPL using the standard LGF simulation model was significantly, though modestly, correlated with the highest cut-off level yielding a phoneme score above 20% ($r = 0.58$, $p < 0.01$).

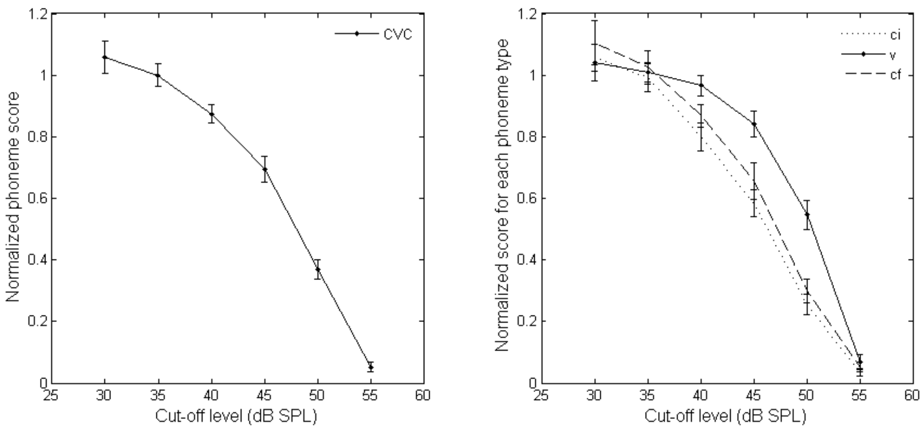


Figure 5: Averaged normalized phoneme scores and the standard error of the mean (SEM) at the various cut-off levels using the LGF-manipulated model. (a) Normalized phoneme score averaged over all three phonemes and (b) normalized perception scores for the initial consonant, the vowel, and the final consonant separately. Normalization is performed with respect to the score using the standard LFG at a presentation level of 60 dB SPL.

Fig. 5a shows the averaged normalized phoneme scores with their SEM using the LGF-manipulated model for each of the cut-off levels. Cut-off levels of 30 and 35 dB SPL did not significantly affect phoneme scores. Higher cut-off levels did significantly decrease phoneme scores ($p < 0.001$, Wilcoxon tests for two related samples). However, the decrease was small for cut-off levels up to 45 dB SPL. For example, a cut-off level of 40 dB SPL, corresponding to 58% of the EDR, resulted in a phoneme score of 87% of the original phoneme score obtained with the standard EDR at 60 dB SPL presentation level.

Fig. 5b shows the averaged normalized perception scores for the initial consonant, vowel, and final consonant separately using the standard LGF for each of the cut-off levels. The initial and final consonant scores did not significantly differ for any of the cut-off levels. At each cut-off level the vowel score was significantly higher than the consonant score ($p < 0.05$). Hence, both initial and final consonant scores showed a larger decrease than the vowel scores when cutting-off low-level electrical stimuli. For example, at a cut-off level of 50 dB SPL, corresponding to 80% of the EDR, both the initial and final consonant score were reduced to less than a third of the score using the standard LGF. On the contrary, the vowel score was, on average, still half the vowel score using the standard LGF at 60 dB SPL.

In Fig. 5a the phoneme score using a cut-off level of 30 dB SPL, corresponding to 19% of the EDR, seems slightly higher than using the standard LGF simulation model, although this difference was not significant. However, when analyzing the phoneme score obtained at a cut-off level of 30 dB SPL as opposed to the score using the standard LGF, it appeared that poor performers scored higher using a cut-off level of 30 dB SPL than using the standard LGF. Fig. 6 shows the relation between the individual phoneme score obtained with the standard LGF model at 60 dB SPL and the phoneme score obtained with a cut-off level of 30 dB SPL. There is a significant correlation ($r = 0.65$, $p = 0.001$). There are three subjects that are 'outliers' in this figure, namely the two poorest performers (phoneme score around 40%) and the best performer (phoneme score 95%) using the standard LGF. One of them showed an increase in phoneme score after cutting-off 30 dB SPL, whereas the other two showed a decrease.

Omitting these three subjects from determining the correlation, either separately or together, does not alter the significance of the correlation. Also, the slope of the regression line is smaller than 0.75 in all cases. This means that relatively poorer performers gained from cutting-off low-level speech information.

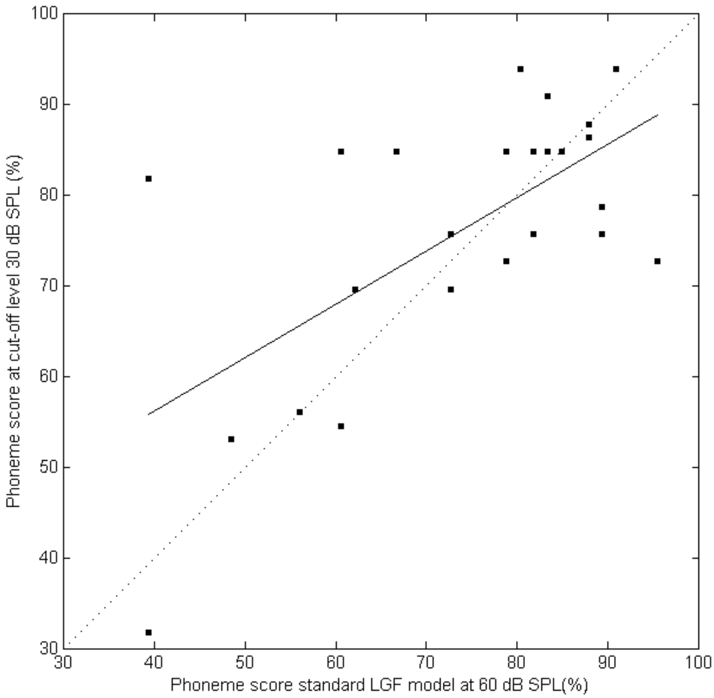


Figure 6: Relation between the individual phoneme score obtained with the standard LGF model at 60 dB SPL and with a cut-off level of 30 dB SPL. The solid line represents a linear least square fit. The dashed line represents a one-to-one linear relation.

6.4 Discussion

6.4.1 Speech perception using the standard LGF simulation model

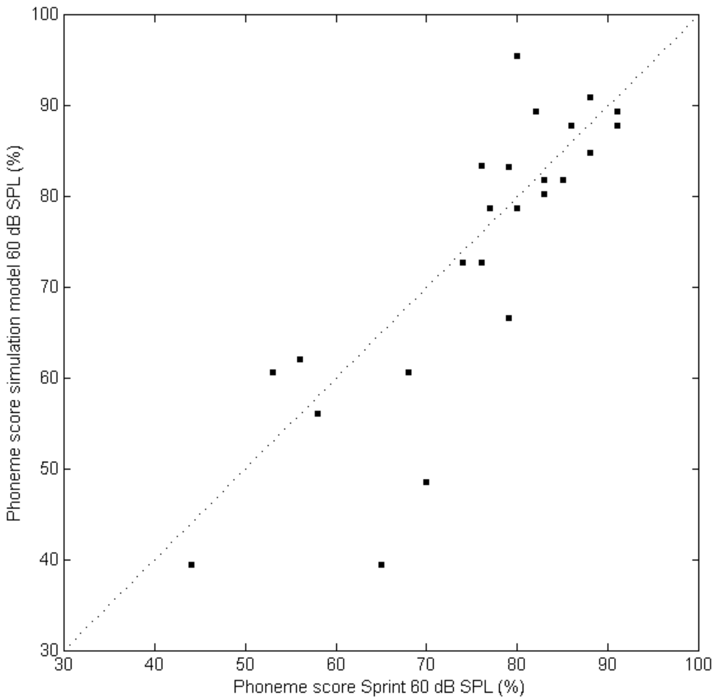


Figure 7: Phoneme scores at 60 dB SPL using WAV-files in the simulation model and those obtained in the free field with the Sprint processor.

Fig. 7 shows the speech perception scores at 60 dB SPL using WAV-files in the simulation model against those obtained in the free field with the Sprint processor. In twenty of the 24 subjects the difference between these scores was within seven percentage points, which is within one standard deviation when applying two word lists. The Wilcoxon test for two related samples did not show a significant difference (p = 0.54). Linear regression analysis showed that the slope and the intercept of the regression line were not significantly different from 1 and 0 respectively (both within one standard deviation). Thus, the simulation model sufficiently approximated the Sprint processor at a presentation level of 60 dB SPL. For the questions to be answered, only comparisons within the simulation model were relevant; an exact resemblance of results of the Sprint processor and the simulation model was not necessary.

The phoneme scores using the standard LGF simulation model did not significantly differ between the presentation levels of 70 and 60 dB SPL ($p = 0.48$, Wilcoxon test for two related samples). This difference was significant using the real Sprint speech processor in our previous study (Willeboer et al., not published yet). Other studies using the Sprint speech processor (Skinner et al., 1997; Donaldson and Allen, 2003) showed this 'roll-over' effect too, i.e. significantly lower scores at 70 dB SPL than at 60 dB SPL. The effect in the Sprint processor is attributed to compression of the highest speech peaks by the processor's AGC. In our simulation model the AGC was not available as a component, and the roll-over effect was not seen. In quiet conditions, the AGC thus reduces high-level speech perception rather than that it enhances perception. The electrical stimulation is limited by the C-levels, which prevents from overstimulation. An extra limiter at the front-end, i.e. acoustically, proves, in quiet conditions, to be detrimental for high-level speech perception. This is probably due to the introduction of distortion in acoustic peak-clipping.

6.4.2 Speech perception using the LGF-manipulated model

The vowel score was less affected by cutting-off low-level stimuli than both consonant scores. For example, when using a cut-off level of 50 dB, corresponding to 80% of the EDR, the vowel score was, on average, only halved compared to using the standard LGF, whereas consonant scores were reduced to less than one-third. The intensity of vowels is higher than of consonants, and therefore the effect of diminishing low level stimuli is less detrimental for vowels. The intensity of consonants is lower. Cutting-off low-level stimuli affects consonant perception more, as a larger part of the consonant's electrical representation is removed. This finding is in accordance with the results of acoustic center-clipping described in the 'Introduction'. Acoustic center-clipping is performed also on the sound fine-structure, whereas eliminating low-level stimuli from the EDR is performed only in M frequency bands after filtering and thus on the envelope. However, the implications for speech perception are essentially the same. The results show that the electrical representation of vowels using the ACE stimulation strategy is much clearer and more robust than that of consonants. Unfortunately, consonants are most important for speech understanding. This is still a challenge for future CI speech processing strategies.

Poor performers gain from cutting-off stimuli in the lowest 19% of the EDR. This might be contributed to clearer speech segment borders. The temporal as well as the spectral structure of the stimulation pattern is more distinct as the gaps in stimulation, within and between electrodes, get more distinct. This effect is also shown in Fig. 3; the gaps between the pulse trains become larger and more pronounced.

Finally, we found a significant correlation between the phoneme score obtained at 60 dB SPL using the standard LGF simulation model and the highest cut-off level used. Good performers can cope with higher cut-off levels and still achieve a phoneme score above 20%.

6.4.3 Implications for the speech processor fitting procedure

The results of the experiments show that stimuli in the lower part of the EDR are not essential for speech perception of words presented at 60 dB SPL in quiet. Eliminating speech information in the M frequency bands below the level of 35 dB SPL, corresponding to 43% of the EDR, did not significantly affect the phoneme score at 60 dB SPL in quiet. These results explain why the lower T-levels found in the live-speech fitting procedure (Smoorenburg et al, 2002; Willeboer and Smoorenburg, 2006) did not result in decreased speech perception. The EDR was downward enlarged, but either using or not using this part of the EDR does not alter speech perception scores at normal presentation levels.

For relatively poor performers a low cut-off level of 30 dB SPL corresponding to 19% of the EDR did increase speech perception scores at a presentation level of 60 dB SPL compared to using the standard LGF simulation model. The effect of a low cut-off level at lower presentation levels than the 60 dB SPL used in this study might be different and needs to be investigated. The effect might be explained by the enhanced temporal and spectral structure at low cut-off levels (see Fig. 3). Additional research is needed to further investigate the relation between enhanced temporal and spectral structure and speech perception in poor performers.

Eliminating low-level stimuli may especially be beneficial in noisy conditions, as it does not affect normal level speech perception, but may reduce disturbing noise. It would be interesting to investigate speech perception in noise using a programming option in which low-level stimuli are eliminated and only the upper part of the EDR is used.

References

Bosman, A.J. (1989). Speech perception by the hearing impaired. Thesis. Utrecht University, the Netherlands, Chapter 3.

Donaldson G.S. and Allen, S.L. (2003). Effects of presentation level on phoneme and sentence recognition in quiet by cochlear implant listeners. Ear Hear. 24(5): 392-405.

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. J. Acoust. Soc. Am. 97(1): 585-92.

Licklider, J.C.R. (1946). Effects of amplitude distortion upon intelligibility of speech. J. Acoust. Soc. Am. 18(2): 429-34.

Skinner, M.W., Holden, L.K., Holden, T.A. et al. (1997). Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants. J. Acoust. Soc. Am. 101: 3766-82.

Skinner, M. W., Holden, L.K., Holden T.A. et al. (1999). Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. J. Speech Lang. Hear. Res. Aug 42: 814-28.

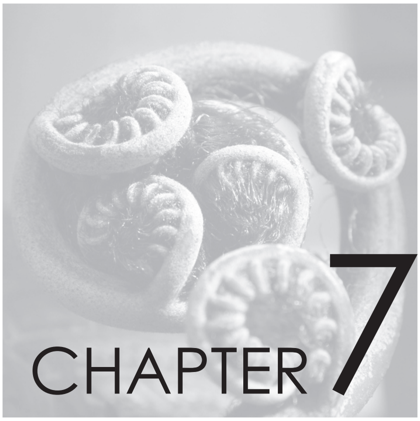
Skinner, M.W., Arndt, P.L., Staller, S.J. (2002). Nucleus 24 advanced encoder conversion study: performance versus preference. Ear Hear. Feb 23(1 Suppl): 2S-17S.

Smoorenburg, G.F., Willeboer, C., van Dijk, J.E. (2002). Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds. Audiol. Neuro-otol. 7(6):335-47.

Willeboer, C. and Smoorenburg, G.F. (2006). Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study. Ear Hear. 27(6):789-98.

Willeboer, C., Bierman, M., Smoorenburg, G.F, van Zanten, G.A. (not published yet). Speech perception using various parts of the electrical dynamic range and amplitude mapping functions in cochlear implant recipients

Zeng, F.G. and Galvin, J.J. III (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners. Ear Hear. Feb 20(1): 60-74



General discussion

7.1 Evaluation of goals

The main objective of this thesis was to develop a fitting method that is less demanding for the cochlear implant (CI) recipient and less time-consuming than the conventional fitting procedure. In the conventional fitting procedure behavioral responses are measured for each electrode in the array, the array consisting of 22 electrodes in the Nucleus CI system (Cochlear, Lane Cove, Australia) and at least 12 in the other commercial systems currently significant on the market. The recipient has to indicate the stimulus level at which a sound is just perceived (threshold or T-level) and the most comfortable loudness level (C-level) while stimulated with pulse trains. With this method electrodes giving a percept deviant from that of the other electrodes will be noticed. Also, this method reveals outliers; electrodes with remarkably low or high levels considering the levels found for adjacent electrodes. In the formerly used CI22 system, specifically with a bipolar instead of monopolar stimulation mode, there was quite a chance of finding outlier electrodes. An alternative fitting procedure would (initially) have to meet the same safety requirements as the conventional fitting with regard to finding outliers and electrodes with deviant sounds. The fitting method developed in the present study is based on the profile of the thresholds of the electrically evoked compound action potential (ECAP) across the electrode array. The ECAPs contain information on the functioning of each of the electrodes and the auditory system, and therefore will reveal a deviating electrode.

The profiles of ECAP thresholds, as well as T- and C-levels, proved to be largely governed by only two factors: *shift* and *tilt*. Shift represents a change in level, equal at all electrodes, and tilt represents a change in the slope of the profile. The 44 subjective responses to pulse trains in the conventional fitting have in the present ECAP-based fitting been replaced by only two subjective responses to live speech: the level at which speech becomes just audible and the level at which it is comfortably loud. Chapters 2 and 3 have shown that there are three clear advantages of an ECAP-based fitting performed by the clinician and based upon only a shift of the ECAP threshold to the threshold of audibility and the comfortable loudness level of live speech. First, the ECAP-based fitting method requires less effort of the recipient as less subjective responses are needed. In the population of CI recipients some recipients have co-morbidities, for example kidney diseases or brain damage due to meningitis or encephalitis, which may imply a weak condition. Moreover, now that the CI systems have proven their success, also aged deaf people are being implanted with a CI (Chatelin et al., 2004, Djalilian et al., 2002). Therefore, a fitting method that is less strenuous for the recipient will be beneficial to a large segment of the CI recipient population. Second, it is faster than the conventional fitting method. This means that within the restricted fitting time, determined by the clinic, there is more time for patient counseling and giving advice on the rehabilitation process. Specifically in the early start of rehabilitation this is very valuable. Third, the ECAP-based fitting method is easy for the clinician as well, as he does not have to interpret the recipient's verbal explanation of his auditory percept as much as in the conventional fitting. There are only two clearly defined markers: the threshold and the comfortable loudness level of live speech. The simplification of the fitting process implies that the ECAP-based fitting method is well suitable for countries with less experience in CI fitting.

However, the shift-adjusted ECAP-based fitting described in Chapters 2 and 3 had the disadvantage of being inflexible. Even if the recipient would give clear feedback on the fitting, it was not possible to adjust individual C-levels within the protocol of the shift-adjusted ECAP-based fitting. In Chapter 2 we already suggested that the shift-adjusted ECAP-based fitting made by the clinician could be used during the initial stage of rehabilitation to provide a fast and easy start of CI use. In a later stage, additional adjustments of the T- and C-levels at individual electrodes might, if necessary, be included.

An improvement of the shift-adjusted ECAP-based fitting performed by the clinician is the introduction of the second factor tilt in the fitting procedure. However, in adjusting shift there are two clearly defined markers, but in adding tilt it may be difficult to ask the recipient the appropriate questions to be able to find the optimal setting. Therefore in Chapter 4 we offered CI recipients the opportunity to optimize their shift-adjusted ECAP-based fitting *themselves* by adjusting shift and tilt during everyday life. This fine-tuning by the recipients themselves bypassed the need for interpretation by the clinician of the recipient's description of his auditory percept. The results of Chapter 4 showed that the recipient did not adjust the fitting to make a 'comfortable' sound, but actually tried to optimize speech perception by increasing basal stimulation with the tilt parameter and hence making the sound sharper. On average, speech perception scores using words at 65 dB SPL in quiet improved slightly but significantly by making adjustments. In Chapter 2 increased basal stimulation seemed to deteriorate speech perception. However, regarding the results of Chapter 4, this finding has solely been due to habituation problems, as the duration of the study in Chapter 2 was only two weeks. The results of self-fitting proves that the current assumption that the clinician knows what is best for the recipient is a bit pretentious, at least regarding experienced CI users. It is well possible that new CI users need more guidance towards accepting the new sound, sharp sound of a CI. Our group of eighteen experienced users proved to be able to determine themselves the adjustments for optimal speech perception. However, eight of the eighteen subjects were insecure about their adjustments. They had on average a better speech perception score from the start using the conventional fitting than the other ten subjects. Relatively poorer performers thus seem to profit most from self-adjusting their fitting.

The good results on speech perception measurements in Chapter 4 were not always associated with a positive subjective appreciation of the fitting. The question arises what aspects are important to CI recipients when judging a fitting, e.g. speech perception in quiet or in noise, music perception, clarity of environmental sounds, or comfort aspects of sound in general. Our speech perception test battery is elaborate, but clearly does not cover all aspects of sound in real life situations. Further research is needed to evaluate this aspect and to develop test materials (other speech perception tests, music perception tests, or questionnaires) that more extensively measure the CI user's benefit during daily life.

7.2 Other applications of self-adjustments in the shift-and-tilt fitting procedure

Self-adjustments by recipients may prove their benefit in other applications than used in our study. Self-adjustments allow the recipients themselves to fine-tune the fitting in case of slight changes, and may thereby increase the time between subsequent fittings. This may be very beneficial in countries in which the distance from the recipient's home to the clinic is large. Also, when recipients themselves are able to fine-tune their fitting, it may improve the results of CI in countries with less experience in CI fittings.

Performing the experiments described in Chapter 4 in new CI recipients would eliminate the effect of being accustomed to the conventional fitting. It would be interesting to study how inexperienced CI recipients change their fitting over time. Of course, they would need much more counseling than our experienced study group, so it would be necessary to counsel them every week and to frequently assess their progress in making adjustments and in speech understanding. Another difference to our study is that determining the upper level of stimulation, as prevention from overstimulation, would be difficult in new CI users. During habituation to stimulation with the CI the tolerance for higher stimulation levels will increase. Therefore the clinician would have to estimate the upper limit of stimulation. However, the inexperienced CI recipient might be more flexible in trying different settings, which may result in outcomes differing from those of our study.

Another group of CI recipients that would be interesting to consider for self-fitting is the group using an electrical-acoustic-stimulation (EAS) system. In EAS acoustic residual hearing is on the ipsilateral side combined with electrical stimulation by a short electrode. In this way, low frequencies are presented acoustically, whereas high frequencies are presented electrically through the CI. The electrical stimulation needs to be adjusted to the residual acoustic hearing. Therefore, these CI recipients themselves might be best able to optimize the CI fitting, as they can judge the sound produced by electrical stimulation in direct relation to the acoustic percept.

7.3 Further improvements of the shift-and-tilt fitting procedure

A further improvement to the shift-and-tilt fitting procedure may come from adding a third factor to obtain more precise fine-tuning. Using only shift and tilt, it is not possible to change stimulation levels at electrodes in the middle of the array independently of those at apical and basal electrodes. In the principal components analysis of the T- and C-profiles the third factor, after shift and tilt, explains an additional 1.5% of variance. It is termed 'curvature' (Smoorenburg, 2007). This factor represents a convex versus concave shape of the profile. Adding this third factor 'curvature' would enable the recipient to adjust stimulation levels at various electrodes more independently from each other.

The adjustments to the fitting made by the CI recipients themselves during everyday life may also be improved by making the design more practical. In our study, subjects used a body worn speech processor with which they could make adjustments to the fitting. On average they made adjustments at the most once a day, instead of trying to optimize speech perception in each listening situation encountered. A design in which the adjustments can be performed using a remote control and, preferably, a behind-the-ear speech processor might eliminate the practical issues and give less biased results regarding the recipients' attitude to changing the fitting during everyday use.

A last suggestion for future research is starting the adjustments of shift and tilt not from the ECAP profile, but from a flat profile. In that case the information about the auditory system's functioning will be removed. However, the current CI systems offer fast impedance measurements. Non-functioning electrodes will be noticed during impedance checks and can be switched off in the fitting. Moreover, behavioral T- and C-levels in the current CI systems are highly correlated between adjacent electrodes due to considerable spread of excitation using monopolar stimulation. In the last few years experience has been gained with conventional fittings based on T- and C-levels measured at 5 electrodes and interpolation of these levels at all other electrodes without actually measuring those (Plant et al., 2005). Hence, setting T- and C-levels without psychophysical or electrophysiological information for all individual electrodes is not regarded precarious anymore. Starting from a flat profile further shortens the time required for a fitting, as the ECAP measurements need not be performed. Using three factors shift, tilt, and curvature, the recipient has many opportunities to create a fitting of his choice.

7.4 Additional recommendations regarding the speech processor fitting procedure

The results of the ECAP-based fitting experiments show that the conventional fitting is not to the golden standard, as we used to think in the last decades. Behavioral T- and C-levels determined on the individual electrodes, if so, at least do not exclusively lead to the optimal fitting. The ECAP-based fitting procedure yields much lower T-levels and higher basal C-levels than the conventional fitting procedure, but the speech perception scores at presentation levels of 55 dB SPL and up are equal. It shows us that we need to be more flexible in our vision on speech processor fitting.

The results of all experiments in this thesis show that the most important action in the fitting procedure is to make sounds well audible.

Audibility is the first requirement for speech understanding. Plant et al. (2005) investigated in six CI users speech perception using a 'flat map'. In this fitting the profiles of T- and C-levels were completely flat. A behavioral T-level was determined at one electrode, after which all other T-levels were set at the same level. C-levels were set at a level gaining comfortable loudness of live speech with an equal EDR for each electrode, thus also with a flat profile. Results of this fitting with no fine-tuning at all, solely based on audibility of broad band sound, showed that CI users were on average still able to reach about 75% of their speech perception score using the conventional fitting.

The results of Chapter 5 also show that audibility is crucial; the upper part of the electrical dynamic range (EDR) is most important in speech understanding. Using this part of the EDR yielded the highest speech perception scores at normal and low presentation levels, compared to using the middle or the lower part of the EDR. The results show that raising T-levels does increase speech perception scores at low presentation levels of speech. This result is in accordance with the results of Skinner et al. (1999). However, it should be noted that this as well as Skinner's result holds only for the Nucleus device, as its acoustic input range, limited to 32 dB, might be of crucial influence. Other devices, like the Medel device, have an input dynamic range of 55 dB. The estimated loss of information by reducing T-levels to 0 μA is only 9 dB of the 55 dB input range. With the Medel device Spahr and Dorman (2005) showed that there is no difference between speech perception using fittings with T-levels set at 0 μA as opposed to behavioral T-levels. However, the lowest presentation level in that study was 54 dB SPL, which might not be low enough to demonstrate a decrease of speech perception at lower levels, e.g. 45 dB SPL (see Chapter 5, Fig. 3).

Cutting-off low level information, up to 35 dB SPL, from speech stimuli did not affect speech perception scores for speech presented at 60 dB SPL (Chapter 6). This finding explains why the downward expansion of the EDR due to lower T-levels in the ECAP-based fitting did not affect speech perception at presentation levels of 55 dB SPL and up. However, regarding the results of Chapter 5, a decrease of speech perception at lower presentation levels (45 dB SPL and down) might be expected for the ECAP-based fitting. For speech understanding in quiet (results in noise may be different), it is therefore better to set the T-levels not at the lowest level of live speech audibility, but at a higher level. Low-level speech perception combined with intra-subject loudness comparison, for instance with Acalos (Brand and Hohmann, 2002), might be used for determining the optimal setting for the T-levels. The exact position of the T-levels, however, seems of minor importance. This finding in Chapter 5 is supported by the results of the pilot study in Chapter 4. In this pilot, subjects were able to adjust shift and tilt of the C-profile as well as the T-profile. Subjects proved not to be able to make useful adjustments to the T-profile. Actually, they did not hear any difference when adjusting the T-profile. They became confused making their adjustments and switched from one extreme to the other without noticing.

Increase of low-level speech perception can also be achieved by increasing the steepness of the loudness growth function (LGF), i.e. decreasing the Q-parameter value, as defined by Cochlear, to values lower than 20, e.g. to 10. However, the effect in noisy conditions may again be different from the quiet condition, and needs to be investigated. Eliminating low-level stimuli may specifically be beneficial in noisy conditions, as it does not affect normal level speech perception, but may reduce disturbing noise. It would be interesting to investigate speech perception in noise using a programming option in which low-level acoustic information is eliminated and only the upper part of the EDR is used.

Chapter 7

References

Brand, T. and Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. J. Acoust. Soc. Am. 112(4): 1597-1604.

Chatelin, V., Kim, E.J., Driscoll, C. et al. (2004). Cochlear implant outcomes in the elderly. Otol. Neurotol. 25(3): 298-301.

Djalilian, H.R., King, T.A., Smith, S.L. et al. (2002). Cochlear implantation in the elderly: results and quality-of-life assessment. Ann. Otol. Rhinol. Laryngol. 111: 890-5.

Plant, K., Law, M.A., Whitford, L. et al. (2005). Evaluation of streamlined programming procedures for the Nucleus cochlear implant with the Contour electrode array. Ear Hear. 26(6): 651-68.

Skinner, M. W., Holden, L.K., Holden T.A. et al. (1999). Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. J. Speech Lang. Hear. Res. Aug 42: 814-28.

Spahr, A.J. and Dorman, M.F. (2005). Effects of minimum stimulation settings for the Medel Tempo+ speech processor on speech understanding. Ear Hear. 26(4 Suppl): 2S-6S

Nederlandse samenvatting (Summary in Dutch)

Een cochleair implantaat (CI) kan door middel van directe elektrische stimulatie van de gehoorzenuw een gedeelte van de hoorfunctie herstellen in mensen die doof of zeer ernstig slechthorend zijn. Met wereldwijd meer dan 150.000 CI gebruikers is het CI na de pacemaker het meest toegepaste systeem voor functionele elektrische stimulatie in de mens. Het CI bestaat uit een inwendig en uitwendig gedeelte. Het implantaat wordt inwendig onder de huid achter het oor geplaatst, waarbij de aangehechte elektroden in de cochlea geplaatst worden. Uitwendig bevindt zich de spraakprocessor, met de microfoon die het geluid opvangt en de processor die het geluid analyseert en codeert voor verzending naar het implantaat. Communicatie tussen het uit- en inwendige gedeelte vindt plaats door radiofrequente overdracht door de huid tussen de uit- en inwendige spoel.

Het CI moet voor elke gebruiker afzonderlijk worden ingesteld, het zogenaamde afregelen, zodat binnenkomend geluid hoorbaar is, maar niet onaangenaam luid. Hiertoe wordt volgens de conventionele afregelmethode voor elk van de in het Nucleus systeem (Cochlear, Lane Cove, Australië) 22 intracochleaire elektroden de drempel (T-level) en aangenaam luidheidsniveau (C-level) bepaald met behulp van pulstreinen. Het bepalen van deze 44 subjectieve responsen van de gebruiker op de pulstreinen is tijdrovend en belastend voor de gebruiker. Het doel van de studies beschreven in dit proefschrift is het vereenvoudigen van de afregelprocedure, door gebruik te maken van objectieve metingen van de reactie van het auditieve systeem op elektrische stimulatie. De objectieve meting die gebruikt werd als basis voor de afregeling is de drempel van de *electrically evoked compound action potential* (ECAP) over de elektrode array: het profiel van ECAP-drempels.

Hoofdstuk 2 beschrijft de ontwikkeling van de ECAP-gebaseerde afregelprocedure. De ECAP-gebaseerde afregeling werd gemaakt door het profiel van ECAP-drempels in live mode parallel te verschuiven tot de drempel van geluidswaarneming van spraak (nieuwe T-levels) en het aangename luidheidsniveau van spraak (nieuwe C-levels).

Dit impliceert slechts twee subjectieve responsen in plaats van 44. De studie is uitgevoerd bij 13 CI gebruikers met tussen een half en drie jaar ervaring met hun CI. De deelnemers kregen gedurende twee weken de ECAP-gebaseerde afregeling mee naar huis, waarna het verstaan van consonant-vowel-consonant (CVC) woorden werd gemeten. De resultaten lieten zien, dat de nieuwe T-levels veel lager waren dan de conventionele T-levels. De nieuwe C-levels waren voor de meest basale elektroden hoger dan conventionele C-levels, terwijl deze voor de overige elektroden gelijk waren. Ondanks de grote verschillen in T- en C-levels tussen de ECAP-gebaseerde en de conventionele afregeling, was het spraakverstaan met de ECAP-gebaseerde afregeling gemiddeld slechts 7 procent punten lager dan met de conventionele afregeling. Aangezien de deelnemers gewend waren aan hun conventionele afregeling en slechts twee weken de ECAP-gebaseerde afregeling gebruikten, was dit resultaat bemoedigend. Een ander resultaat van de studie was dat in een principale componenten analyse de sets van ECAP-drempels en T- en C-levels over de elektrode array, de zogenaamde profielen, beschreven bleken te kunnen worden met twee factoren. De eerste factor, shift, verklaarde 92 % van de totale variantie en was sterk gerelateerd aan het gemiddelde niveau van het profiel. De tweede factor, tilt, verklaarde nog eens 3 % van de variantie en bleek gerelateerd aan de helling van het profiel. De in deze studie gebruikte methode voor een ECAP-gebaseerde afregeling is daarom de 'Shift-and-tilt' methode genoemd en is als zodanig geïmplementeerd in de Cochlear afregelsoftware.

Gezien het bemoedigende resultaat van de ECAP-gebaseerde afregeling in ervaren CI gebruikers, werd besloten tot het uitvoeren van een studie naar de ECAP-gebaseerde afregeling in nieuwe CI gebruikers. Achttien CI gebruikers namen deel aan deze cross-over studie. Negen deelnemers startten bij de eerste afregeling met de conventionele afregeling, terwijl de andere negen startten met de ECAP-gebaseerde afregeling. Na zes weken vond de cross-over plaats. Het spraakverstaan werd elke week gemeten met de in die week gebruikte afregeling. Na twaalf weken werd het spraakverstaan met beide afregelmethoden gemeten. De resultaten lieten dezelfde verschillen in T- en C-levels zien als de studie beschreven in hoofdstuk 2. In de spraakverstaanscores waren er echter geen verschillen, noch op normaal, noch op laag aanbiedingsniveau (65 respectievelijk 55 dB SPL).

In hoofdstuk 2 en 3 werd alleen een shift van de profielen gebruikt, maar nog geen tilt. Het is lastig om de juiste vragen aan de CI gebruiker te stellen om tot een optimale instelling van de tilt te komen. In hoofdstuk 4 wordt een studie beschreven, waarin CI gebruikers *zelf* de shift en tilt van het C-profiel van hun ECAP-gebaseerde afregeling kunnen aanpassen in de dagelijkse situatie. Achttien CI gebruikers namen deel aan de studie. Zij bleken gemiddeld genomen hun basale C-levels te verhogen. Hoewel het geluid van de ECAP-gebaseerde afregeling vergeleken met de conventionele afregeling al scherper is, werd dit door de gebruiker zelf nog scherper werd gemaakt. Het spraakverstaan van woorden op normaal aanbiedingsniveau in stilte bleek significant toegenomen. Voor de woorden in ruis en zinnen in stilte en ruis bleken er geen verschillen. De deelnemers probeerden een instelling te maken die in alle luistersituaties voldeed, in plaats van telkens in elke situatie opnieuw aanpassingen te doen. Tien mensen waardeerden de mogelijkheid tot het zelf aanpassen van de afregeling, terwijl acht mensen zich onzeker voelden over hun aanpassingen.

In hoofdstuk 5 en 6 werden twee meer fundamentele studies beschreven. De bijdrage van stimuli in verschillende delen van het dynamisch bereik tussen T- en C-level aan het spraakverstaan werd bepaald. Veruit het belangrijkste voor het spraakverstaan is het bovenste gedeelte van het dynamisch bereik. Het gebruik van alleen het bovenste 1/3 deel van het dynamisch bereik verhoogt de verstaanscore van zachte spraak. Het weghalen van stimuli in het onderste gedeelte van het dynamisch bereik, tot wel 43 % van het dynamisch bereik, bleek geen negatief effect te hebben op het verstaan van spraak van normaal niveau (60 dB SPL). Dit verklaart waarom de lage T-levels in de ECAP-gebaseerde afregeling ten opzichte van conventionele T-levels geen invloed heeft op het verstaan van spraak vanaf een aanbiedingsniveau van 55 dB SPL. Voor het verstaan van zeer zachte spraak in stilte is het beter de T-levels bóven de drempel van geluidswaarneming te plaatsen.

Hoofdstuk 7 beschrijft tenslotte een algemene discussie en samenvatting van de onderzoeken.

Dankwoord

Het magische moment is aangebroken; het proefschrift is af. Dit was niet gelukt zonder de steun van vele mensen. Enkele hiervan wil ik graag op deze plaats wil bedanken.

Prof. dr. G.F. Smoorenburg, beste Guido, vanaf het begin ben jij enthousiast en betrokken geweest. Shift en tilt (met bijbehorende arm-zwaai-bewegingen) was dan ook een beetje jouw kindje. Jij had het in een heel vroeg stadium al steeds over "hoofdstuk 1 van je boekje", ook op momenten dat ik nog dacht "welk boekje?". Mede door jouw inzet is het er nu dan ook van gekomen, waarvoor mijn grote dank.

Dr. G.A. van Zanten, beste Bert, op het moment dat jij co-promotor werd (en gelukkig slechts twee deuren verder zat), kwam er meer vaart in het onderzoek. Je hebt er, met de toch wel krappe personele bezetting, steeds voor gezorgd dat mijn onderzoekstijd niet volledig werd opgeslokt door klinisch werk. Jouw inbreng was altijd helder en je schrijfstijl is zonder opsmuk, wat de leesbaarheid van de stukken zeer bevordert. Je bent een heel eerlijke wetenschapper. Dank voor al je inspanning!

Prof. dr. G.J. Hordijk, prof. dr. F.W.J. Albers, helaas veel te vroeg overleden, en dr. A.F. van Olphen dank ik voor hun vertrouwen in mij en de kans die ze mij gegeven hebben dit proefschrift te schrijven.

Maaike, dank voor al het werk dat je verricht hebt. Je bent enthousiast en weet van aanpakken. Ik vond het erg leuk met je samen te werken. Veel plezier en succes met je opleiding tot klinisch fysicus audioloog!

Alle collega's en oud-collega's van het CI team dank ik voor de fijne samenwerking. Adinda, Anne, Bas v D, Bas F, Bert, Birgit, CC, Dieuwertje, Erwin B, Erwin D, Eefje, Hans, Kati, Lonneke v G, Lonneke K, Margreet, Patrick, Pauline, Rinus, Ruth, Selma, Tania, Vera en Yvonne, het is heel fijn te werken met zoveel mensen met ieder hun eigen expertise. Samen werken we op het leukste gebied van de gezondheidszorg!

Ook alle andere collega's van het Audiologisch Centrum (al mag dat niet meer zo heten, het dekt toch de lading), Kind & communicatie en de research dank ik voor de samenwerking. In het bijzonder Anne K, Anne te L, CC, Liesbet en Merlijn: het is heerlijk om met elkaar te lachen, klagen, eten, drinken en te dansen. Heel erg bedankt voor jullie luisterend oor en jullie steun. Jullie weten precies hoe het hier toe gaat. Helaas werken we niet meer bij elkaar en wordt de afstand wat groter. Maar zodra CC in het land is, moeten we ons eetrondje afmaken!

Alle proefpersonen die meegewerkt hebben aan het onderzoek wil ik bedanken voor hun inzet en tijd. Zeker de laatste twee onderzoeken waren voor de proefpersonen tijdrovend en leverden voor hen zelf niet direct iets op. Toch vonden zij het de moeite waard om mee te doen, "voor alle CI gebruikers na ons", zoals vele zeiden. Hartelijk dank!

Mijn paranimfen Margreet en Natalie, naast oud-collega's zijn jullie vooral vriendinnen geworden. Een half woord is vaak genoeg om even te spuien over het werk en de afdeling, en dat helpt enorm! Natalie, bedankt ook voor je inwijding in de wereld van Matlab. Dat was altijd iets heel magisch, maar dank zij jou kan ik het zelf nu ook (een beetje). Ik ben blij dat jullie beiden naast mij staan op de grote dag, dat geeft een vertrouwd gevoel.

Mijn vrienden, in het bijzonder Anneke, Mascha en Tjetske, heel erg bedankt voor de etentjes, borrels, koffies, lunches, het samen shoppen en alle andere leuke dingen die we samen doen, maar vooral voor de fijne gesprekken. Het leven is zo veel meer dan promoveren en dat is goed om bij jullie telkens weer op positieve wijze te ervaren!

Lieve familie, thuis heb ik geleerd door te zetten en dat is me de afgelopen jaren goed van pas gekomen. Thuis is ook de plek waar het altijd druk is, gezellig, waar het krioelt van de kinderen. Dank voor jullie interesse, maar ook voor de afleiding en voor de liefdevolle omgeving!

Lieve Tjeerd, jouw steun, optimisme en wanneer het moest strengheid zijn doorslaggevend geweest in de afronding van dit boekje. Je hebt zoveel voor me gedaan! Dankjewel dat je er altijd voor me bent. Samen beginnen we nu een nieuw avontuur!

Curriculum Vitae

Christina (Krista) Willeboer werd op 29 december 1977 te Arnhem geboren. In 1995 haalde zij op het Christiaan Huygens College te Eindhoven haar eindexamen VWO. Hierna studeerde zij Medische Biologie aan de Universiteit Utrecht. De afstudeerstage vond in 2000 plaats onder leiding van Prof. Dr. G.F. Smoorenburg op de afdeling Experimentele Audiologie, afdeling Keel-, Neus- en Oorheelkunde van het Universitair Medisch Centrum Utrecht (UMCU). Het resultaat van deze stage is te lezen als het eerste hoofdstuk van dit proefschrift. Aansluitend aan de afstudeerstage trad de auteur in dienst van het UMCU als klinisch medewerker binnen het cochleair implantatie team. In 2002 begon zij aan haar opleiding tot klinisch fysicus audioloog, welke zij in 2006 afrondde. Sindsdien is zij als klinisch fysicus audioloog in dienst van het UMCU. In de klinische werkzaamheden ligt een groot accent op cochleaire implantatie. Tevens is zij waarnemend coördinator van de Zorggroep Cochleaire Implantatie.