Cochlear Implant
Ear Marks

Guido F. Smoorenburg

University Medical Centre Utrecht

guido@smoorenburg.com
http://audiologics.com

ISBN-10: 90-9021523-9
ISBN-13: 978-90-9021523-5

This monograph is dedicated to my teachers:

Prof. dr. M. A. Bouman
Sensory Physiology

Prof. dr. ir. R. Plomp
Audiology

Prof. J. E. Rose MD PhD
Neurophysiology

Prof. dr. E. H. Huizing
Oto-Rhino-Laryngology

and to

Deana
My beloved companion in life

Contents

Preamble

I first became acquainted with the cochlear implant during my visits to California in the early seventies. Fascinating research and development was underway by William House and Jack Urban in Los Angeles, Blair Simmons and Earl Schubert in Stanford, and Robin Michelson, Marc White, Michael Merzenich, and Robert Schindler in San Francisco. House and Urban developed a simple system with a single electrode and no internal electronics, thereby reducing the risk of system failure. They focussed on clinical development of the implant. In contrast, Simmons and Schubert, conducted animal research, investigating the possibility of selective electrical stimulation of cochlear nerve fibres via multiple electrodes placed in the cochlear modiolus. The San Francisco group studied electrical stimulation via arrays of intra-cochlear electrodes; their work indirectly evolved into the Clarion implant device. Thus, California represented a wide gamut of research and development on cochlear implants. Of course, California was not the only place where interest in cochlear implants was blooming: the Med-El device was being developed in Vienna (and later in Innsbruck), the Nucleus device in Melbourne and Sydney, the Symbion device in Utah, the MXM device in France and Banfai's device in Germany. The time was ripe for successful development of the cochlear implant.

However, no one could have foreseen that cochlear implantation would become such an important intervention in otology. Early on, hearing physiologists pointed out that electrical stimulation, even with an array of electrodes, would be very crude in relation to the impressively refined hearing organ. In addition, they warned that electrical stimulation of the auditory nerve might damage the neurones and that there was risk of infection when implanting electrodes in the cochlear turns or modiolus. However, animal research showed that these risks were small, and courageous work in clinics showed that electrical stimulation of the auditory nerve could produce useful auditory sensations. Even so, it was still unclear at that time how many deaf people would benefit from cochlear implantation and to what extent hearing could be restored. When we started our work in the early eighties, we felt we should first try electrical stimulation via an electrode placed on the promontory to test the candidacy of a deaf person for cochlear implantation. Now, this is done only in special cases (*e.g.*, suspected neuropathy, skull trauma). It is almost taken for granted that cochlear implantation can restore hearing to the extent that once-deaf people can understand speech again and that children implanted at an early age do well in regular schools. But given the relatively crude electrical stimulation the implant provides and the degeneration of the auditory pathway with deafness it is quite amazing that the outcomes are so generally good. While there have been great advances in cochlear implant technology, it is the plasticity of the brain that largely contributes to these positive outcomes. These two factors contribute to the success of the cochlear implant, with more than 100,000 implant users worldwide, to date.

The cochlear implant has not only benefited the deaf and profoundly hearing impaired but also hearing research. The field has grown significantly. This growth has been possible because of support from governmental health research organisations to relatively small clinical research budgets. But we should also acknowledge the support from implant manufacturers. Research has benefited greatly from the exemplary cooperation between the manufacturers and research groups. This support has contributed not only to questions related to cochlear implantation, but also to long-standing scientific questions such as the relative roles of spatial and temporal processing in hearing. The cochlear implant has provided the opportunity to independently control the spatial and temporal information sent to the brain. At the same time we are able to objectively measure evoked neuronal activity via compound action potentials and compare this activity to behavioural responses. Together, these opportunities contribute to our research regarding the relative roles of spatial and temporal processing in hearing. In addition, cochlear implantation has triggered research regarding fundamental questions about neuronal repair. Is it possible to stimulate growth and re-growth of cochlear neurones, and to control this growth so that it terminates once the neurones are stimulated adequately? On a larger scale, cochlear implantation has stimulated research regarding brain plasticity. To what extent does the brain adapt to new forms of stimulation, and in what time frame?

Despite these achievements, many challenges remain. Before the development of the cochlear implant, deaf persons were guided toward living with their deafness. Speech reading and sign language were important aspects of these guidance programs. Deaf people set up deaf communities. With cochlear implantation, the focus shifted toward aural rehabilitation. However, this also means that implant recipients depend entirely on their implants and, by extension, on clinical implant teams and manufacturers. This is a great responsibility, requiring adequate financial planning to support implant recipients not only over the short term, but over the whole of their lives. In turn, clinical implant teams must rely on continuity in production and service from implant manufacturers. As such, implant manufacturers have a great social responsibility. Even with the best planning and precautions, implant recipients are vulnerable to unexpected events. For example, what if there is a shortage of batteries (due to war)? What if the health care system collapses, and there is no clinical support if the implant or speech processor should fail? These scenarios, though far-fetched, should be kept in mind when deciding whether or not to train implant recipients in speech reading and sign language as a means of communication. After all, without a functioning implant, the recipients are deaf. Our research has shown that sign language does not interfere with aural rehabilitation; thus a “bilingual” guidance program may be the best option. Finally, while the old fears of infection and neuronal damage have not come to fruition, these remain very real risks, requiring caution and care when developing new electrode arrays and surgical techniques. Even advanced speech processor settings in which impulse rate or charge are increased should be evaluated with respect to any risks of neuronal damage.

Another serious challenge is the variability in cochlear implant recipient outcomes, as reflected in Chapters 4 and 5 of the present monograph. Even among subjects implanted in the past decade with the latest implant technology, there remains great variability in implant recipient performance. Aetiology of deafness contributes to these differences in outcomes, but only to a very small extent; age at onset of deafness and the duration of deafness (from onset to time of implantation) also contribute. However, these factors do not significantly contribute to the variability in outcomes (at least in statistical terms). Poor performance is, of course, a great disappointment for any implant recipient. It also can be very time- and resource-consuming for clinicians when trying to improve performance by optimising speech processing for each implant recipient. It seems to me that this problem deserves more attention. Typically, when developing new coding strategies good performers are usually selected for evaluation, as they can report very well on what they can and cannot hear. However, optimisation techniques based on good performers may not be appropriate for poor performers. For example, does increasing the stimulus rate improve the performance of poor performers? Or has neuronal responsiveness deteriorated to the point where low stimulation rates might provide better performance? Similarly, we may need to apply different charges per stimulus impulse or different stimulus waveforms to improve performance in those implant recipients not performing well with the present stimuli of choice. There are many possibilities, and it may be more fruitful if manufacturers and implant teams more fully consider poor performers in development and optimisation of speech, or more generally sound processing.

After two decades of steady progress, cochlear implant research and development seems to be slowing down. However, this perception has more to do with the difficult problems that remain to be solved, rather than reduced efforts on the part of clinicians, researchers and manufacturers. While the plasticity of the brain is quite large, plastic information processing stops if the information is not present in the periphery. It stops if information in auditory pathway channels is buried in noise. It stops if frequency components needed to identify sounds are smeared or distorted. Speech understanding in noisy environments and music perception remain the greatest challenges. It seems to me that these aspects of sound perception are mainly limited by the spectral resolution of current implant devices. The spread of electrical stimulation in the cochlea is large, even with the newest electrode arrays. This means that frequency resolution is very limited. The signal-to-noise ratio in a channel of the auditory pathway is bound to be very poor if the channel bandwidth is so wide that it captures a great deal of noise in addition to the desired signal. Classical psychoacoustic research has shown that pitch perception of complex tones is based primarily on the frequency information in the 3rd to 6th harmonic of the complex tone. Pitch perception of complex tones will be limited if these harmonics are not well separated because of poor frequency resolution. Thus, it seems to me that the next period of real progress in cochlear implants performance depends greatly on improving frequency resolution, *i.e.*, on reducing the spread of electrical excitation. Recent efforts to reduce the spread of excitation by adding compensating signals to adjacent electrodes (*e.g.* tri- or quadru-pole stimulation) have not been successful; it is likely to be very difficult to achieve adequate compensation at the desired sites of neuronal excitation. The most promising approach is to improve the electrode-neurone interface. Such an approach requires fundamental research on how to stimulate growth and re-growth of peripheral neurones as well as on how to control this growth so that it terminates once the neurones have made contact with the electrode, similar to the way neurones seek hair cells in the normally developing cochlea. If this approach is feasible, it will become useful to implant arrays with many more electrodes and frequency selectivity will be much improved.

Acknowledgement

I am very grateful to my former co-workers and members of our implant team of the University Medical Centre Utrecht who examined drafts of the chapters in this monograph and suggested many corrections and improvements: Eefje van den Broek, Bas van Dijk, Pieter Lamoré, Margreet Langereis, and Bert van Zanten. In addition, I am much indebted to John Galvin, John Graham and Bob Shannon who checked the drafts for proper usage of the English language. Any mistakes or misinterpretations that the reader may encounter, in spite of all this assistance, are, of course, solely my responsibility.

Printing and distribution of this monograph was supported by the Foundation of Scientific Research in Oto-Rhino-Laryngology Utrecht "ORLU", the Netherlands' Society of Audiology and Cochlear Europe.

Besse sur Issole, November 2006

Executive summary

Chapter 1 p. 15-34

The course of stimulation levels over time: optimising the time interval for speech processor readjustment

1.1 The most basic parameters in speech processor programming, *i.e.* the stimulation level at the threshold of hearing (T-level) and the highest level still comfortably loud (C-level), can change over months, even years, after implantation. This requires regular adjustments to these levels in the speech processor. But what is the optimal schedule for these adjustments? Frequent readjustment sessions may ensure that implant recipients are programmed with accurate stimulation levels, but they also require greater effort and expense on the part of the implant recipient and clinician. Statistical analysis of changes in T- and C-levels over time provides useful information toward optimising the readjustment schedule.

1.2 It is important, especially during the initial period of implant use, to keep the programmed stimulation levels in accordance with the perceptually measured T- and C-levels, which may shift over time. If, for example, the differences between the perceptual T- and C-levels and the programmed levels are to be kept within 6 current units, for 90% of implant recipients, adults should be reprogrammed once per week during the first 4 weeks of implant use, and children once per week for the first 6 weeks. After this initial adaptation period, reprogramming sessions can be less frequent. Three months after initial stimulation, adults should be reprogrammed after 6 weeks, and children after 4 weeks. Six months after initial stimulation, adults should be reprogrammed after 40 weeks and children after 10 weeks. Chapter 1 also presents the readjustment schedules for differences of 3, 9, 12, and 15 current units

1.3 If implant patients are never reprogrammed after 6 months of initial implant use, C-levels will shift by more than 6 current units for 13% of adult implant users and 41% of pediatric implant users.

1.4 During the initial adaptation period, T-levels are easier to measure than C-levels. Unfortunately, changes in T-levels are no good indicators of changes in C-levels. They cannot be used to estimate changes in C-levels.

Chapter 2 p. 35-54

T- and C-level profiles across the electrode array; fitting the speech processor by profile parameter adjustment

2.1 When programming the speech processor, T-levels are typically measured for individual electrodes independently. However, the correlation between these levels across the electrode array is high. The same is true for C-levels. For both levels the correlation coefficient rarely is smaller than 0.8. Adjusting and readjusting T- and C-levels for some 20 electrodes is quite demanding for both the implant recipient and clinician. In view of the high correlation it seems possible to reduce the number of measurements. We propose to adjust the "profile" of T- and C-levels across the electrode array, rather than adjusting the stimulation levels for individual electrodes. Profile adjustment has the great advantage of using speech signals (or other natural sounds with a broad frequency spectrum) for programming the speech processor.

2.2 Three parameters appear to govern the T- and C-profiles: 1) *shift* - a nearly parallel shift in stimulation levels for all electrodes with respect to the population average, 2) *tilt* - a change in the slope of the stimulation profile, and 3) *curvature* - a change from a peak-shaped to a valley-shaped profile. Proper adjustment of these three parameters yields profiles that are within 2-3 current units from those measured with individual electrodes.

2.3 The three profile parameters *shift*, *tilt*, and *curvature* may not relate to easily identifiable aspects of sounds, and therefore may not be suitable for interactive speech processor adjustment between implant recipient and clinician. However, more commonly identifiable parameters (*shift, bass, and treble*) can be used with only a 25% reduction in accuracy (*i.e.,* a 25% increase in the difference between the proper profile adjustment and the levels measured with individual electrodes).

2.4 In addition to the proposed profile adjustment method the three parameters can also be used to estimate T- and C-levels for individual electrodes across the full electrode array after measuring the T- and C-levels at two marginally located electrodes (*i.e.,* the basal and apical ends of the array) and at one electrode in the middle of the array. This approach could replace general interpolation schemes that are not based on T- and C-level statistics.

2.5 The values of the parameters related to the shape of the T- and C-profiles (*i.e., tilt* and *curvature*, or *bass* and *treble*), measured shortly after implantation and at a much later date, show a modest correlation (0.4 to 0.5). This implies that the shape of the profile may significantly change as implant recipients adapt to their device. Thus, readjustments cannot be limited to overall stimulation level (*i.e., shift*).

2.6 For all shape-related parameters (*i.e., tilt* and *curvature*, or *bass* and *treble*) the correlation between the T- and C-parameter values was 0.6 to 0.7. During the initial adaptation period, T-levels may be easier to measure than C-levels. The correlation suggests that the shape of the T-level profile somewhat predicts the shape of the C-level profile.

2.7 The relative roles of individual electrode level adjustment and profile adjustment should evolve in clinical practice. Profile adjustment, as suggested in this monograph, offers the possibility of interactive speech processor programming using natural sounds, rather than the tone bursts used in single electrode adjustments.

Chapter 3 p. 55-66

Thresholds of electrically evoked compound action potentials; relation to T- and C-levels

3.1 Instead of programming speech processors according to perceptually measured T- and C-levels, it is possible (in principle) to use thresholds of electrically evoked compound action potentials (ECAPs) for speech processor programming. However, the correlation between ECAP thresholds, measured mostly during surgery, and perceptually measured T- and C-levels is small. The correlation between the profile parameters (introduced in Chapter 2) of the ECAP thresholds and of the T- and C-levels was not much better; the highest correlation coefficient (0.48) was found between the *tilt* of ECAP thresholds and the *tilt* of C-level profiles.

3.2 The speech processor can be programmed by shifting the overall level of ECAP thresholds, without changing their shape, to auditory threshold (taking this as an alternative T-level profile) and to comfortable loudness level (taking this as an alternative C-level profile). Despite the small correlation between ECAP thresholds and perceptually measured T- and C-levels in terms of profile shape parameters (*tilt* and *curvature*, or *bass* and *treble*), simply shifting ECAP thresholds to obtain alternative T- and C-levels, using live speech, yielded speech perception scores equal to those found when programming the speech processor according to T- and C-levels, measured perceptually and independently for individual electrodes. This suggests that profile shape may not be critical to speech performance. However, other aspects of sound perception (*e.g.,* listening comfort, music perception, etc.) may require more accurate speech processor programming than speech perception. The concept of adjusting the shape of the profile (using, for example, the *bass* and *treble* parameters) may contribute substantially to a broader *sound* processing, rather than to *speech* processing.

3.3 Simplification of speech processor programming using only a few profile parameters provides the opportunity for implant recipients to adjust the processor themselves. This could be an important development, not only in terms of self-optimisation of the processor, but also with respect to clinical support, which can be limited in sparsely populated areas and developing countries with limited audiological infrastructure.

3.4 The ECAP threshold increases by about 5 current units if the stimulation rate is increased from 80 to 250 Hz. T-levels decrease by about 10 current units if the stimulation rate is increased from 250 Hz to the higher rates of 720-1200 Hz. These rate effects should be considered when comparing ECAP thresholds to T- and C-levels.

3.5 Decreasing the masker offset from 10 current units above probe level to the probe level itself did not affect the mean values of the ECAP thresholds. However, the variability in ECAP thresholds across the electrode array increased substantially. Thus, adequate masking is necessary in the ECAP measurement procedure when using ECAP thresholds for speech processor programming. ECAP measurement accuracy during surgery may be improved by priming the electrode impedance with electrical stimulation before starting ECAP measurements.

Chapter 4 p. 67-85

Development of speech perception and language acquisition in children over time; comparison of performance measures

4.1 Speech perception in children, 5 to 8 years of age at the time of implantation, continues to improve for a long period after implantation, even for four years after initial stimulation. After four years of implant use, mean phoneme scores for words presented in consonant-vowel-consonant (CVC) context were 55% correct when presented in auditory-only mode and 75% correct when presented in auditory-visual mode.

4.2 The range in performance among implant recipients is very large. While after four years mean phoneme recognition performance in auditory-only mode was 55% correct, 10% of 61 children scored less than 25% and another 10% of these children scored more than 80%. In auditory-visual mode, 10% of children scored below 60% and another 10% scored above 95%. After four years of implant use, there were only small differences in performance (< 10 percentage points) between different types of implant devices and speech processors marketed at different points in time. The large differences in performance may be explained in part by the age at onset of deafness and the duration of deafness from its onset to the time of implantation. In children, congenitally deaf or deafened before the age of 2, phoneme scores four years after implantation ranged between 50 and 95% correct in auditory-only mode and between 65 and100% in auditory-visual mode if implanted within two years of the onset of deafness. However, part of the large differences in outcomes remains unexplained, even after considering differences in type of implant device and differences in aetiology of deafness. The differences in performance are a great concern of clinicians.

4.3 Language acquisition in children after implantation does not show progression to normal. On the contrary, over the three-year period after implantation, the improvement in language skills and speech production of these children was 25 to 50% lower than the improvement observed with normal-hearing children; a point of concern considering the fact that the implanted children start at a substantial disadvantage. Thus, early cochlear implantation of deafened children is very important with respect to language acquisition.

4.4 One might expect that good speech reception would benefit language acquisition. However, language acquisition over the three-year period after implantation shows little correlation with speech reception six and twelve months after implantation.

4.5 Five tests were used to evaluate paediatric implant recipient performance: a Dutch CVC word reception test, the Erber word comprehension test, the Peabody Picture Vocabulary Test, the Reynell Verbal Comprehension test, and the Dutch Schlichting sentence production test. Taken together, the CVC word test and the Reynell Verbal Comprehension test were the most relevant indicators of performance; the Schlichting sentence production test could be viewed as a valuable addition to these tests.

Chapter 5 p. 87-105

Development of speech perception in adults over time; comparison of performance measures

5.1 Speech perception in about 150 adults, deafened at age 37 ± 20 years and with an average duration of deafness until the time of implantation of 15 years, improved over a period of about two years after cochlear implantation. After two years their performance stabilised at an average of 50 to 70% phonemes perceived correctly when presented with words consisting of a consonant-vowel-consonant (CVC) combination. The scores stabilised at an average of 60 to 90% syllables perceived correctly when presented with sentences consisting of 8 or 9 syllables. These values were found in a testing condition without speech reading. With speech reading the average syllable score reached 85 to 100%. The differences in average value depended mainly on the type of implant.

5.2 The results of the connected discourse test already stabilised one year after implantation. In auditory-only mode the implant recipients reached 45 to 70 words per minute perceived correctly while the score reached 70 to 85 words per minute when speech reading during the test was allowed. Normal-hearing people score over 80 words per minute, 100 words per minute may be adopted as the standard reference.

5.3 The scores differed seriously amongst the implant recipients. Two years after implantation 10% of the implant recipients scored below 15% syllables correct when sentences were presented in auditory-only mode, whereas another 10% of users scored above 95%. These were scores within one type of implant. Aetiology of deafness had a small effect on the scores. However, large differences remained unexplained. The marked possibility that an implant recipient will perform poorly is quite a note of concern.

5.4 Speech perception increased with technological development. One year after implantation we found in adults an average score of 45% for the Nucleus® CI24M, implanted since 1997, and an average score of 70% for the Nucleus® CI24R(CA) with the Contour Advance™ electrode, implanted since 2003. The COMBI 40®, implanted since 1995, reached an average score of about 60%, which was a relatively high score at that time.

5.5 The number of adult implant recipients who continued to use their acoustic hearing aid in the contralateral ear immediately after implantation decreased by 50 % over a three year post-implantation period. After three years this percentage did not change over another two years. The average phoneme score in auditory-only mode for the acoustic hearing aid with the implant switched off was 12%. The acoustic hearing aid contributed about 5% to the total score for the cochlear implant and the acoustic hearing aid together.

5.6 Considering the accuracy of measurement reached in a given duration of a measurement session, the phoneme-in-word score did substantially better than the syllable-in-sentence score and the connected discourse score. The CVC word test would be the primary choice if one wishes to reduce the number of tests. The connected discourse test could be viewed as a valuable addition.

Chapter 6 p. 107-113

Effect of cochlear implantation on speech production; vowel quality

6.1 Alongside speech perception, it also is important to evaluate the effect of cochlear implantation on the quality of speech production. Cochlear implantation generally improves speech production because implant users are better able to control voicing and articulation once they can hear their own voice. Clear speech production implies that implant users will be understood without too much effort on the part of the listener. With clear speech production, implant recipients will engage more easily in group conversations, providing further opportunities to improve speech perception, speech production, and language skills.

6.2 The quality of vowel production for vowels embedded in words in consonant-vowel-consonant (CVC) context can be assessed by measuring the frequencies at which the peaks in the frequency spectrum occur, *i.e.,* the frequencies of the formants F1 and F2. During the first two years after implantation, these formant frequencies shifted from relatively close positions to near-normal positions. Spectral contrasts among the vowels increased markedly, which reduces vowel confusions on the part of the listener.

6.3 In addition to improved spectral vowel contrast, there was also a substantial reduction in the variability of the formant frequencies when the same vowel was uttered repeatedly. This variability also reached near-normal values. The smaller variability further reduces vowel confusions on the part of the listener.

Postscript

This monograph presents a retrospective analysis of our clinical data about cochlear implantation. Such a retrospective analysis has advantages and disadvantages. One disadvantage is that these clinical results do not relate to well-balanced experimental designs. Lack of balanced designs may imply that seemingly independent variables can be mutually related. When, for example, speech performance improves with newer types of electrode arrays, the benefit may not necessarily be due to the new electrode design if other system properties, such as speech coding and stimulation rate, have also been changed. Another disadvantage in these clinical results is the problem of missing data. For example, when a speech recognition score is missing, it can either mean that the implant recipient could not complete the task (thus, the score was probably near zero), or that the examiner expected that the task would be too easy for the implant recipient (thus, the score was probably near 100%). Results of clinical trials may be more pronounced than those found in everyday practice because clinical trials generally aim to exclude many factors that may affect the results of a clinical trial in an uncontrolled way but that do affect performance in every day practice. The advantage of the present retrospective analysis is that the results do reflect factors playing a part in everyday practice. I tried to minimise the disadvantages of the present retrospective studies by carefully examining and discussing the possible pitfalls in all steps of the analyses presented.

Chapter 1

The course of stimulation levels over time; optimising the time interval for speech processor readjustment

Summary

Objective

T- and C-levels can change over time. This requires readjustment of the signal processor. Clinical management demands optimisation of the periods in between processor adjustment sessions. The intervals should be as large as possible in order to minimise the call on patient and clinician but without risking unacceptable maladjustments. This chapter presents guide lines for planning readjustment intervals derived from population statistics of the changes in T- and C-levels. All data were collected with the Nucleus® cochlear implant.

Conclusions

The results are presented for T- and C-levels that are not allowed to drift away by more than 3, 6, 9, 12, and 15 current units from the programmed values. The course of the T- and C-levels was calculated for electrodes 5, 13, and 20. C-levels showed the largest shifts and thus determined the readjustment intervals. When, for example, a clinic aims to limit the mismatch between the actual T- and C-levels and the programmed values to 6 current units for at least 90% of the implant recipients, one should perform the second adjustment within one week after the initial fitting at signal processor switch-on. Later on, the acceptable time interval between two successive readjustment sessions appears to be shorter for children than for adults. After a session at 3 months post switch-on an adult should be called back for processor readjustment within 6 weeks to keep the drift from the adjusted value within 6 current units but a child should be called back within 4 weeks. After half a year an adult should be called back within 40 weeks, a child within 10 weeks.

If implant patients are never reprogrammed after 6 months of initial implant use, C-levels will shift by more than 6 current units for 13% of adult implant users and 41% of pediatric implant users.

The correlation between the early shifts in T- and C-levels was low (R=0.5). Thus, the early shifts in T-level (usually easier to measure than the early shifts in C-level) cannot be used to predict the change in C-level. The shifts for electrodes 5, 13 and 20 in both the T- and C-levels were highly correlated (R = 0.77 to 0.87) while the magnitudes of the shifts were similar. This suggests that one may start a readjustment session by shifting the levels for all electrodes by equal amounts of current units.

1.1 Introduction

Adjustment of the signal processor of a cochlear implant to the sensitivity of the auditory system to electrical stimulation is typically based upon the threshold level at which an individual detects short stimulus bursts (T-level) and upon the highest level at which these bursts are still comfortably loud (C-level). This adjustment to T- and C-levels is repeated for each electrode of the electrode array. Clinical experience has shown that these basic levels may change over time so that readjustment of the processor becomes necessary. Good patient care may suggest that one frequently checks and readjusts the processor. However, this is quite demanding for the implant recipient. Moreover, frequently trying different adjustments may confuse the implant recipient who then has to cope with frequently changing sound templates. In addition, frequent checking imposes quite a demand on the capacity of cochlear implant teams, which nowadays is limited because of the rapidly growing number of implant recipients. Therefore, we have analysed the time course of T- and C-levels of our population of implant recipients over the years. The results are used to derive optimum readjustment intervals for a choice of shifts in level that may be considered just acceptable by the implant team: 3, 6, 9, 12, and 15 current units (CUs). These units are defined by the manufacturer of the Nucleus implant, Cochlear.

The time course of the T- and C-levels depends on several factors:

1) a physiological factor directly related to the electrophysiological properties of auditory nerve stimulation,
2) a habituation factor related to the sound percept created in the implant recipient by electrical stimulation of the auditory nerve, and
3) an interaction factor related to the communication between the implant recipient and the audiologist performing the processor adjustments.

The effect of the first two factors will be comparable from one clinic to the next. However, the influence of the third factor may depend on the attitude of a particular audiologist toward the implant recipient and the level adjustment procedure in general. More specifically, the audiologist may be inclined to push C-levels up quickly or he/she may give the implant recipient more time to slowly adapt to the new auditory sensations. Thus, in principle, the results of the analysis presented here may, to some extent, depend on the performance of the individual carrying out processor adjustments. This could mean that the results of the present analysis may be less relevant to another clinic. However, we believe that the influence of this factor in the present study is small because the data are taken from the records of four audiologists who worked over the years according to their own experience without a compulsory protocol.

1.2 Materials and methods

1.2.1 Fitting the data

In clinical practice processor adjustments are not conducted at exact time intervals across the population of implant recipients. Therefore, it is impossible to determine the statistics of the T- and C-levels at given time intervals after implant switch-on. This problem could be solved by interpolating the data. However, the magnitudes of the shifts in level change with time (first rapidly, then more slowly) and they depend on the individual. Thus, it is difficult to choose an interpolation scheme that is generally applicable. Moreover, implant

recipients were often sent home with two (or even three) different adjustments (so-called MAPs) between which they could choose. Retrospectively, it is impossible to determine to what extent the recipient used one or other adjustment. In those cases we have to include all levels programmed in the signal processor. This implies that the number of data points at a given sample time is variable. We solved the latter problem, and the interpolation problem, by accepting that all T- and C-levels change over time in accordance with an exponential growth curve. Only a very limited number of data did not comply with this exponential growth model. Some examples of deviant data are given in the results section. The exponential growth curve contained three parameters to be estimated: onset level, time constant and final, asymptotic level. The time constant, τ, is expressed as the time between processor switch-on and the time at which the level has reached a value of $1/e$ or 36.8% from the final level. The final level is expressed relative to the onset level, L_0, as the total level shift or the level range, R (R can be negative).
In formula:

$$L = L_0 + R\,(1-e^{-t/\tau})$$

This function was fitted to all data, including multiple data at a given sample time. The three parameters were estimated minimising the differences between the data and this function in the sense of least squares using a program written in Matlab®.

1.2.2 Limitations in fitting the data

The three-parameter fit became ambiguous whenever the time span of the data was restricted so that the data showed only a linear change over time (the initial part of the exponential growth curve given by the slope R/τ). In that case the fitting procedure did not converge but it produced a time constant increasing to infinity. This was prevented by interrupting the fitting procedure once the time constant had reached a value of 3000 days. Setting the time constant to this value did not invalidate the results to be reported below, because each fit was used only to represent the data within the time interval covered by the data. If the data showed only a linear change within the observation period then the fitting algorithm yielded the proper slope R/τ at τ = 3000 days. This slope was not used for extrapolating the data beyond the observation period. The exponential curves resulting from the fitting procedure were used to calculate the distributions of T- and C-levels across the implant recipients at weekly intervals after processor switch-on. As mentioned above, the results from the fitting procedure were restricted for each individual to the post-implantation period covered by the original data. This implies that the number of implant recipients included in the calculated distributions decreases with time after processor switch-on. An alternative approach could be to choose a certain analysis period of, for example, two years and to include only those implant recipients for whom data are available for at least this period. However, the shifts in level shortly after processor switch-on are the largest ones and therefore it is desirable to include the largest number of subjects possible in this initial period in order to achieve the highest accuracy. Accepting that we loose an increasing number of subjects with increasing time after processor switch-on we may introduce a bias. For example, the speech processor fitting procedure may have changed over time or the selection criteria for implantation may have changed. Implant recipients with data collected over the largest time intervals are those implanted earlier. Thus, the tails of the longitudinal data may then reflect the results of the early fitting and selection procedure while the initial part of the longitudinal data may reflect the results of these procedures as they were applied over the whole period studied, from 1997 to 2003. This possible bias will be examined carefully when the data are presented.

1.2.3 Materials

The results are presented for 146 subjects, 86 adults and 60 children. Children up to 16 years of age visited a separate hospital but were seen by the same implant team members. We included subjects only if their T- and C-levels were measured over a period of at least 6 months after implant switch-on. We could also have included subjects measured over a shorter period if they showed T- and C-levels that stabilised within this shorter period. However, this option was discarded because including these subjects and excluding the subjects not stabilising within this shorter period would introduce a bias in the data toward shorter time constants. Thus, the number of subjects in the population studied was constant, 146 in total, during the first 6 months after which it gradually decreased. Subjects with intermittent periods without T and C measurements were also excluded from the data set if these periods without data could result in an ambiguous fit.

All subjects received a Nucleus implant with the Sprint™ or Esprit™ signal processor. All recipients were stimulated in MP1+2 mode (one reference electrode beneath the temporal muscle and the implant housing used as a second reference electrode). Stimulus impulse width was always 25 µs. Stimulus rate was 250, 720, 900 or 1200 Hz. Exploratory analyses showed that there was virtually no difference between the results found for 720, 900, and 1200 Hz. Therefore, we limited ourselves to two rate categories; *low rate* (250 Hz) and *high rate* (720, 900, 1200 Hz). We excluded those subjects in whom the rate of stimulation had been changed during readjustments. The analysis was conducted for three electrodes: the numbers 5, 13, and 20. Since electrodes 1 and 2 are often not used the three electrodes included are considered to be a representative sample of the total array of electrodes from 3 to 22.

1.3 Results

1.3.1. Data fitting

Figure 1.1 (top left panel) shows a typical example of a proper fit. In this case, the C-level increases by about 70 CUs over time. Thus, the range ***R*** is about 70 CUs. In the analysis the first data point is always set at 0. All data points are expressed relative to this first measurement. The first C-level measured was in fact 150 CUs. It rose to about 220 CUs. The time constant of this measurement was 88 days. This example also shows multiple data points at certain sample times. The top right panel shows a near- linear increase (except for the first data point) over 1500 days. The T-level increased from 129 to 182 CUs. In this case the time constant could not be determined. It was set at 3000 days. The bottom-left panel shows the time course of a T-level that could not be fitted by the exponential growth curve because of “undershoot”. The T-level dropped from 174 to 138 CUs after which it increased and levelled off at 160 CUs. Finally, the bottom-right panel shows a case that could not be fitted because the T-level suddenly changed drastically. It changed from 155 to about 30 CUs. After a period of almost 1000 days it was reset to about 145 CUs. The three examples given above, which clearly could not be fitted, concern T-levels. Similar deviations from exponential growth were found for C-levels. These examples were presented to show the limitations of the postulate of exponentially growing T- and C-levels. However, the number of data tracks that could not be fitted to this function was very small. It amounted to 1.8% for T-levels and 2.6% for C-levels.

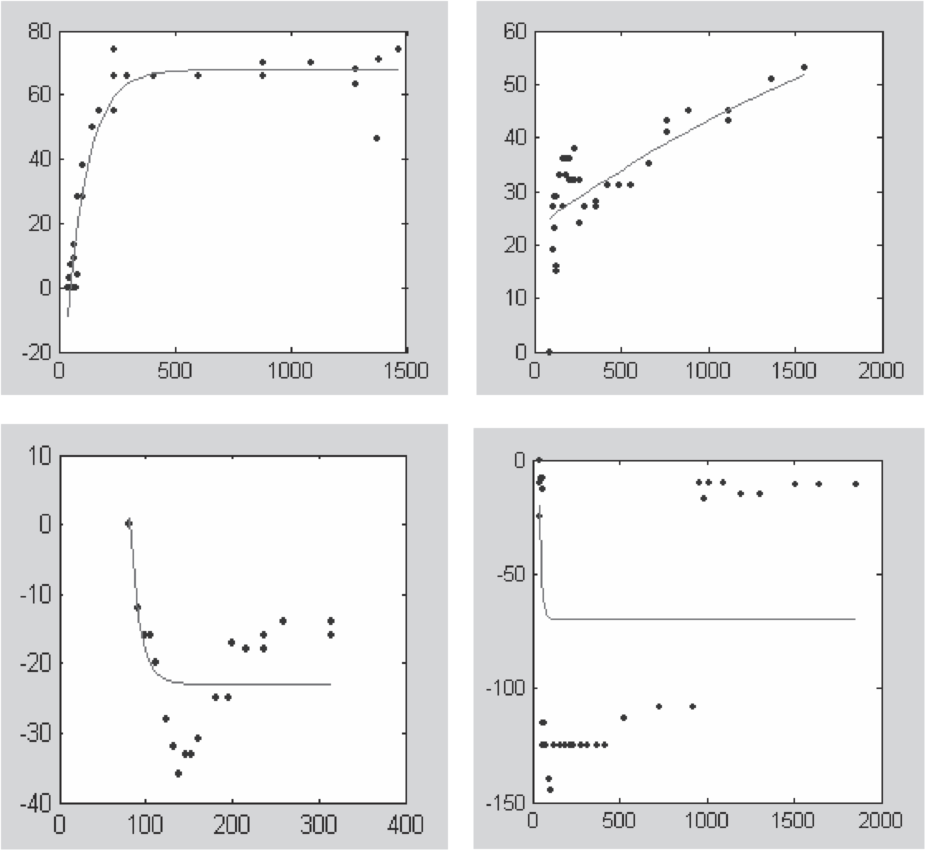

Figure 1.1. *Top-left panel:* example of a normal fit; *top-right panel:* example of a nearly linear increase within the observation period so that the time constant of the postulated exponential growth function could not be determined; *bottom-left panel:* example of data that could not be fitted by the exponential growth function because of “undershoot”; *bottom-right panel:* example of data that could not be fitted because of a drastic change in adjustment. *Abscissa:* time in days since processor switch-on, *Ordinate:* T-level in CU relative to the first level measured after processor switch-on.

1.3.2 Preliminary analysis at 6 months post switch-on

The parameters included in the analysis are:
Between subjects: type of subjects (adults or children) and rate of stimulation;
Within subjects: electrode number and type of level (T- and C-level).
In view of this number of parameters we first explored the data at 6 months post switch-on because we have the largest data set at this sample time.

Analysis of variance showed that the parameter *electrode number* itself and its interactions with the other three parameters was completely insignificant. Thus, there was no significant difference between the mean changes after 6 months in the T- and C-levels measured for the three electrodes, including in subsets of the data subdivided according to type of subject and rate of stimulation. Moreover, the standard deviations of the changes in the T- and C-levels for electrodes 5, 13 and 20 were very similar: 21, 20, and 23 CUs for T-levels and 22, 21, and 22 CUs for C-levels. Further, correlation coefficients for the changes in T- and C-levels between each pair of electrodes were very high, all within the range from 0.77 to 0.87. Finally, there was no difference between the dynamic ranges (C minus T-levels) found for the three electrodes: 36 ± 20, 37 ± 15, and 35 ± 16 CUs for electrodes 5, 13 and 20, respectively. Essentially the same result was found at a separate analysis of the data one year after switch-on. This result suggested that we may pool the data collected for the three electrodes.

The parameter *rate of stimulation* did not show a main effect in the analysis of variance but there was a highly significant ($p<0.01$) interaction with *type of subject*. The data collected for adults showed significantly lower mean changes in the T- and C-levels at the low stimulation rate of 250 Hz. However, this is not necessarily an effect of rate itself. In adults the data were collected at the low stimulation rate until about halfway the year 2000; thereafter the high stimulation rates were used. Thus, the rate effect could be a time effect related to build up of speech processor adjusting experience over the years. Children were measured some years later than adults, which might explain the insignificant effect that stimulation rate had on the results in children. Before, we mentioned that there was virtually no difference between the results found for 720, 900, and 1200 Hz. These rates were all used after mid 2000. Since there was no significant main effect of stimulation rate and since the arguments presented above suggest that the interaction effect may represent an effect of increased experience in speech processor adjusting we shall discard the parameter *rate of stimulation* but return to this aspect when analysing the time tracks.

This preliminary analysis shows that the data can be further analysed as a function of time, limiting ourselves to a segregation of the data according to the parameters *type of subject* (child or adult) and *type of level* (T or C-level).

1.3.3 Course of T- and C-levels over time

As mentioned before the data set comprises 86 adults and 60 children measured over a period of at least 6 months post switch-on. Beyond 6 months the number of subjects included in the data decreases. After 2 years the number of adults has decreased to 32, the number of children to 33. Since we are interested in the upper and lower deciles of the distributions of T- and C-levels, the analysis of the time course of the data has to be limited to this two year period. Pooling the data collected for the three electrodes we have 256 data points for adults and 180 data points for children up to 6 months. These numbers decrease gradually to 96 and 99 data points, respectively, at 2 years. The distributions of the changes in T-levels for adults and children calculated per week over the two year period are presented in figures 1.2 and 1.3, those of the C-levels in figures 1.4 and 1.5.

The middle curves in figures 1.2 through 1.5 show the median values of the shifts of the T- and C-levels with respect to the first measurement. All median shifts are positive. Thus, in the two-year period most levels move upward. The two curves above the median curve in each figure show the shifts in level exceeded by 25 % and 10 % of the data. 10 % of the measured T-levels change by 30 CUs or more in adults and by about 40 CUs or more in children. For the C-levels these shifts are even higher, almost 50 CUs and more than 60 CUs. The curve immediately below the median curve shows level shifts exceeded

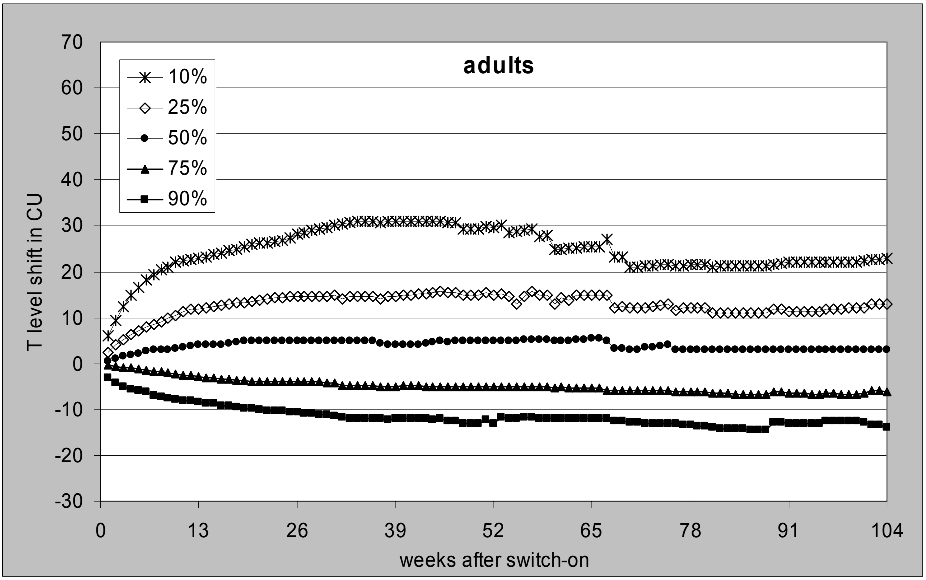

Figure 1.2. Distribution of T-levels in adults per week after switch-on of the signal processor. *Abscissa:* time in weeks up to 2 years (104 weeks), *Ordinate:* T-level in current units relative to the first measurement. *Parameter of the curves:* Percentage of data points above value indicated.

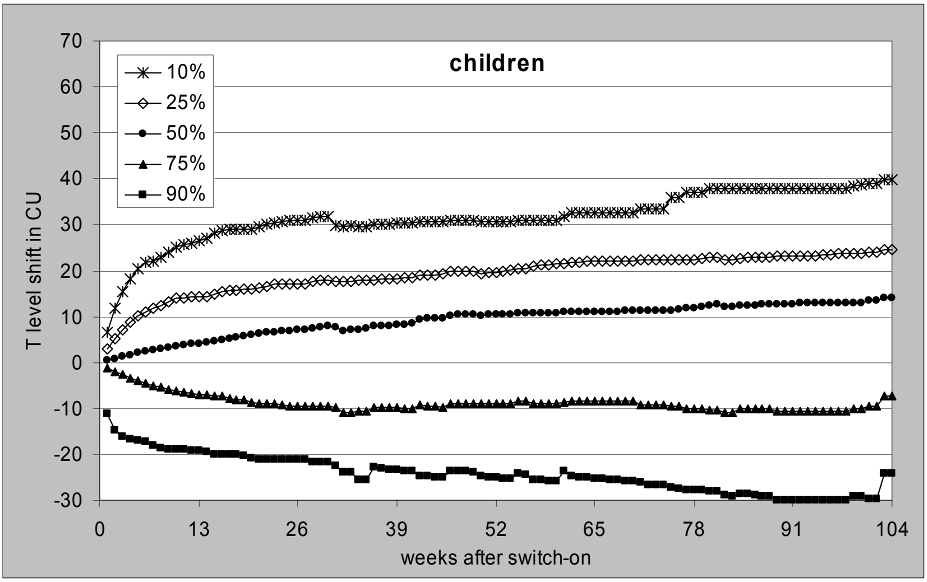

Figure 1.3. Distribution of T-levels in children per week after switch-on of the signal processor. *Abscissa:* time in weeks up to 2 years (104 weeks), *Ordinate:* T-level in current units relative to the first measurement. *Parameter of the curves:* Percentage of data points above value indicated.

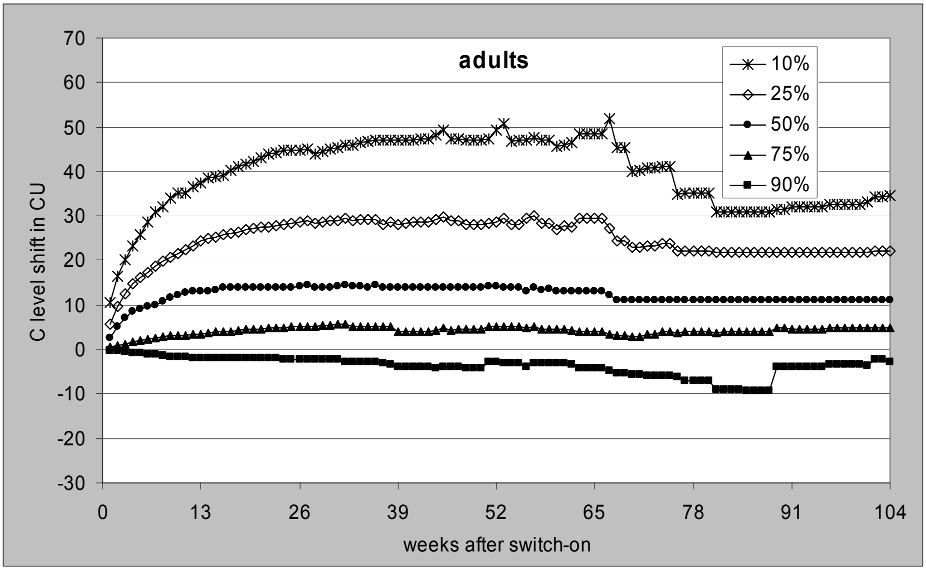

Figure 1.4. Distribution of C-levels in adults per week after switch-on of the signal processor. *Abscissa:* time in weeks up to 2 years (104 weeks), *Ordinate:* C-level in current units relative to the first measurement. *Parameter of the curves:* Percentage of data points above value indicated.

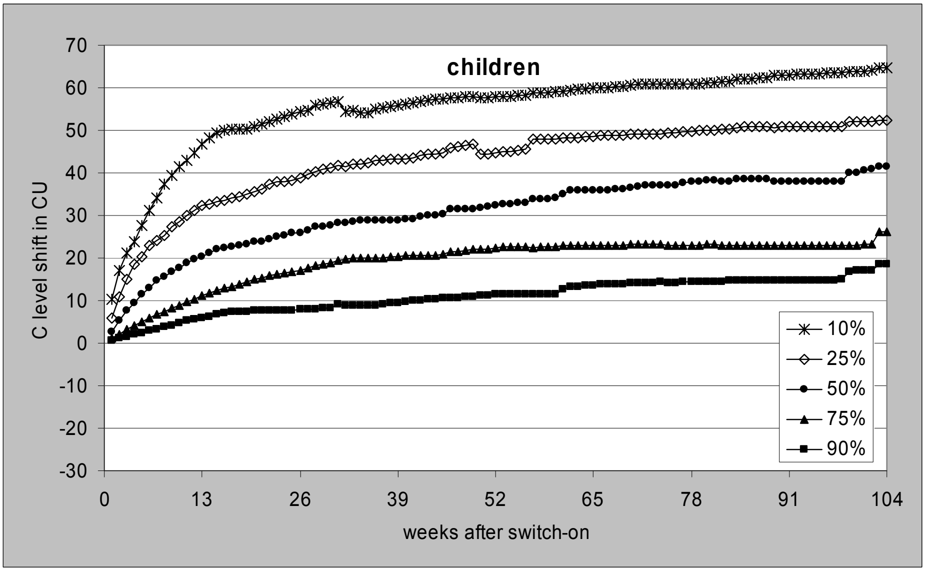

Figure 1.5. Distribution of C-levels in children per week after switch-on of the signal processor. *Abscissa:* time in weeks up to 2 years (104 weeks), *Ordinate:* C-level in current units relative to the first measurement. *Parameter of the curves:* Percentage of data points above value indicated.

by 75 % of the data points. These values are negative for T-levels in adults and children but positive for C-levels in both types of subjects. Thus, more than 75 % of the measured C-levels show an increase of level over time. For children (lowest curve of figure 1.5) this is true for even more than 90 % of the C-levels.

The T-levels show some stabilisation after about 40 weeks. However, after about 60 weeks the upper decile (10% curve) of the adults seems to decrease. This is most probably due to the sample of implant recipients included in the period after 60 weeks. As mentioned before, after 60 weeks there is a larger fraction of subjects implanted relatively early in our department. T- and C-levels in these subjects tended to be adjusted to lower values than the adjustments carried out later. This shows up particularly in the upper percentiles. Also the C-levels show some stabilisation after 40 weeks, except for the median values and higher percentiles found in children. Even after two years there still appears to be an upward trend of C-levels in more than 50 % of the children. This suggests that children should be followed for more than two years after speech processor switch-on.

The reader should note that the curves presented in figures 1.2-1.5 do not represent time tracks of individuals. Successive data points of one curve may stem from different subjects. The initial part of the upper decile, for example, may reflect level shifts in subjects with the shorter time constants while the latter part may reflect shifts in other subjects with larger time constants but also with larger ranges of shifts.

The exponential growth approach offers a good estimate of the magnitude by which T- and C-levels change initially (given by the initial slope $\boldsymbol{R/\tau}$ of the growth curve). If these slopes are highly correlated we may use the measurement of one slope to estimate the rate of change in the other level. More specifically, shortly after switch-on of the processor T-levels are usually easier to measure than C-levels. If the changes are highly correlated one could then use the change in T-levels to estimate the change in C-levels. However, the correlation coefficient appears to be limited to 0.5. Thus, a change in T-level is no good indicator of the change to be expected in C-level.

1.3.4 Determining the readjustment interval

The data of the previous section can be used to estimate the readjustment interval, given that one considers a certain shift in T- and C-level between two successive adjustment sessions acceptable. The calculations have been performed for mismatches limited to 3, 6, 9, 12, and 15 CUs, respectively, at the time of speech processor readjustment. Statistically, it is impossible to perform these calculations such that all shifts in level will stay within the specified limit. One has to accept that a small percentage of the shifts will exceed the criterion set. The smaller we set this percentage, the smaller the reliability of the result. The risk was set at 10%.

The time intervals for speech processor readjustments are presented in figures 1.6 and 1.7 for T-levels of adults and children, respectively, and in figures 1.8 and 1.9 for C-levels. The curves are calculated for mismatches of less than 3 to 15 CUs in either direction, positive or negative, and for the 10 % risk of an excess mismatch discussed before. The irregular course of the curves in these figures gives an impression of the accuracy of the result.

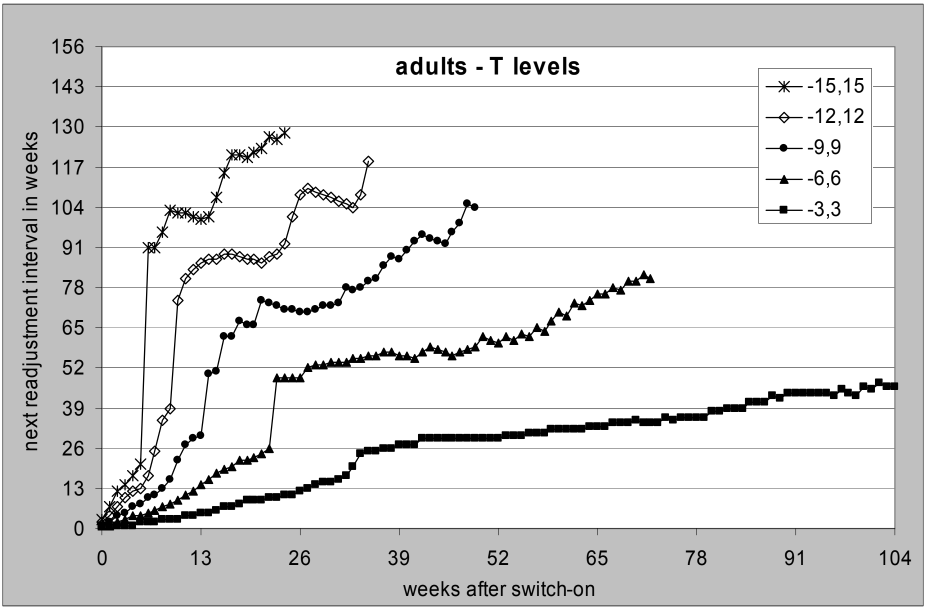

Figure 1.6. Period after which an adult should be called back for processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the T-level of the last adjustment and the actual T-level within the CU range given in the legend.

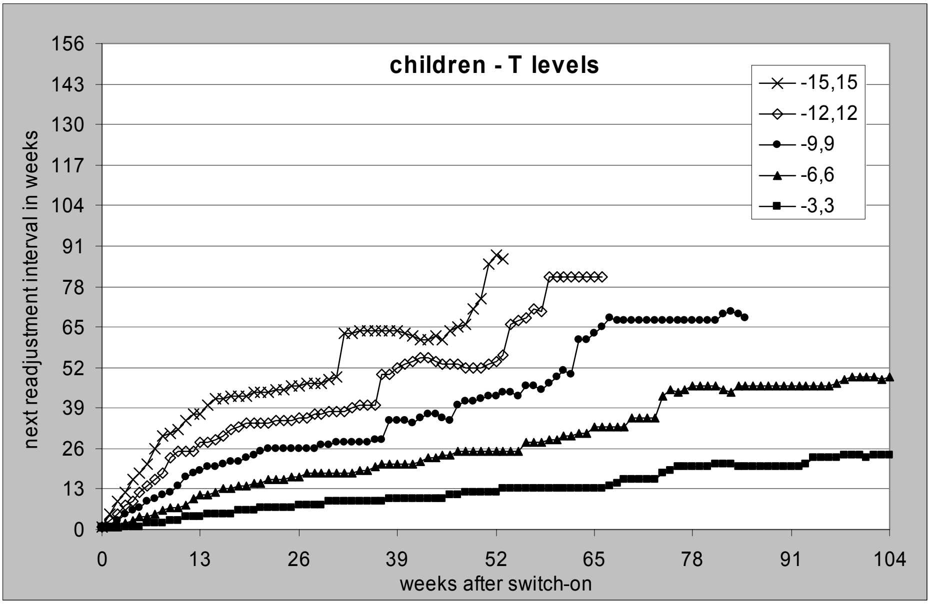

Figure 1.7. Period after which an child should be called back for processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the T-level of the last adjustment and the actual T-level within the CU range given in the legend.

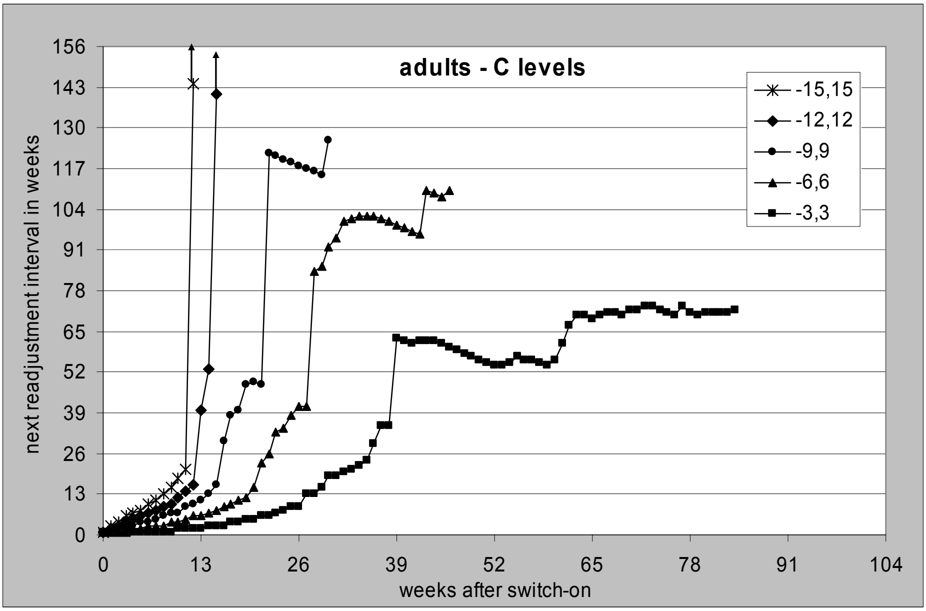

Figure 1.8. Period after which an adult should be called back for processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the C-level of the last adjustment and the actual C-level within the CU range given in the legend.

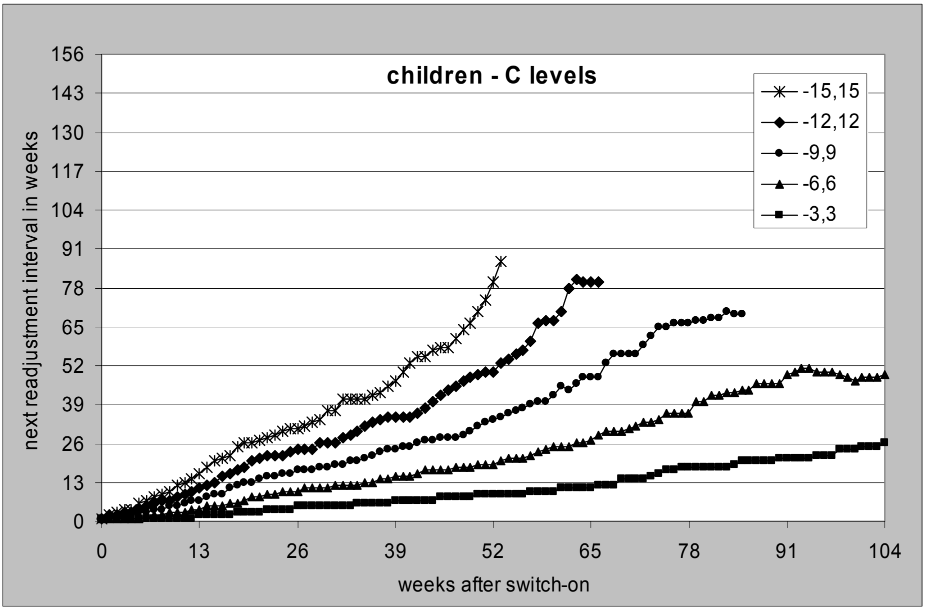

Figure 1.9. Period after which an child should be called back for processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the C-level of the last adjustment and the actual C-level within the CU range given in the legend.

The required readjustment intervals for children appear to be substantially shorter than those for adults. In addition, we find that virtually all readjustment intervals found for C-levels in children are somewhat smaller than those found for T-levels. In adults the same is found initially. However, once the readjustment intervals exceed 13 weeks we find a large increase in the length of the intervals, even more for C-levels than for T-levels. The long intervals are related to the smaller shifts in T- and C-levels found in the first group of adult implant recipients; those with the longest observation periods. This was already apparent in figures 1.2 and 1.4.

The data are presented separately for T- and C-levels although, of course, these levels are measured in one session. The shorter of the two intervals calculated for T- and C-levels will determine the readjustment interval. These intervals are presented in figure 1.10 for adults and in figure 1.11 for children, together with Tables 1.1 and 1.2, which give the results up to week 27 in more detail. The calculations were performed over a period of three years (156 weeks). The empty cells at weeks 25-27 in Table 1.1 for mismatches of 15 CUs indicate that the time of the last adjustment session (25-27 weeks) plus the readjustment interval exceeds 156 weeks.

The results for T- and C-levels together show that in order to limit the mismatches to, for example, ± 6 CUs one should plan for four weekly readjustment sessions with adults and for six weekly sessions with children after the first adjustment of the speech processor at switch-on (week 0). At 13 weeks (3 months) the readjustment interval becomes six weeks for adults and four weeks for children in order to keep 90% of the mismatches smaller than 6 CUs. Half a year after switch-on these intervals become 41 and 10 weeks, respectively.

As a final step the readjustment intervals are combined to provide a complete schedule of readjustment sessions. These schedules are presented in Tables 1.3 and 1.4, again for mismatches smaller than 3 to 15 CUs in 90% of the cases. The tables show that in the first year, keeping the shift within ± 6 CUs, one should anticipate 9 sessions (including the session at switch-on, week0) with adults and 13 sessions with children. The periods between successive sessions increase gradually.

1.3.5 Consequences of final readjustments at 3, 6, or 12 months

Nowadays the adjustment procedure develops smoothly in most implant recipients and T- and C-levels stabilise well after some time. This raises the question of whether or not it is safe to limit the interval in which recipients are called back for readjustments, after which readjustments would not be needed. In response to this question we have analysed the consequences of a follow-up period limited to 3, 6, or 12 months. Therefore, we calculated the long-term asymptotic shifts of the T- and C-levels to be expected with respect to the final readjustment. Thus, in contrast to the previous calculations, the data are extrapolated in this analysis, using the exponential growth function beyond the interval spanned by the data. However, these calculations were limited to only those T- and C-level tracks with calculated time constants smaller than the time intervals spanned by the data. This allows for an accurate estimate of the three parameters involved and, consequently, a proper estimate of the long-term asymptotic T- and C-levels to be expected.

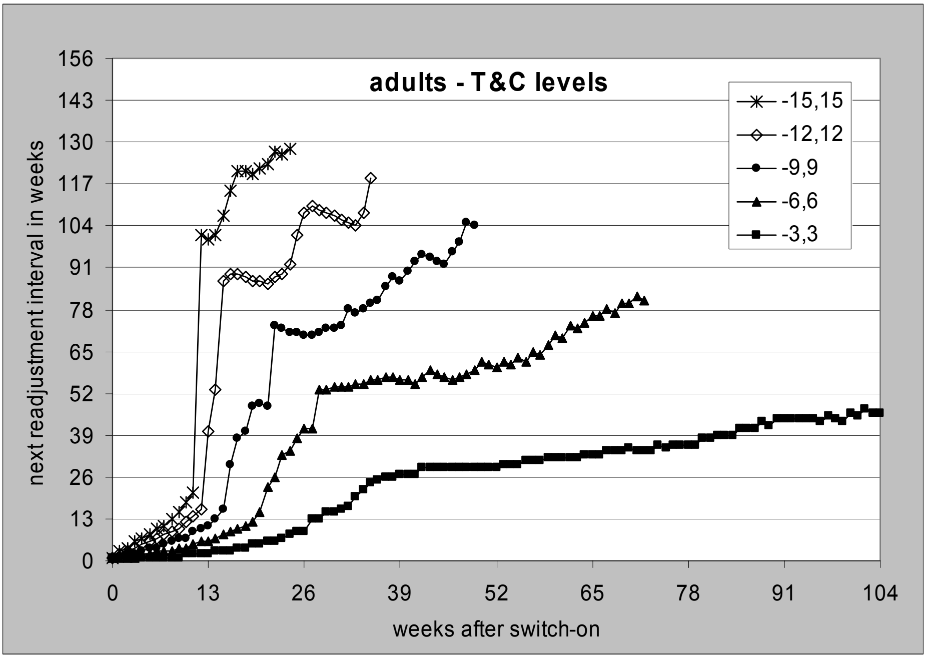

Figure 1.10. Period after which an adult should be called back for speech processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the T- and C-levels of the last adjustment and the actual T- and C-levels within the CU range given in the legend.

week	0	1	2	3	4	5	6	7	8	9	10	11	12	13
-15,15	1	3	4	6	7	8	10	11	13	15	18	21	101	100
-12,12	1	2	3	4	5	6	7	8	9	10	12	14	16	40
-9,9	1	1	2	2	3	4	4	5	6	7	7	9	10	11
-6,6	1	1	1	1	2	2	3	3	3	4	4	5	6	6
-3,3	1	1	1	1	1	1	1	1	1	1	2	2	2	2
week	14	15	16	17	18	19	20	21	22	23	24	25	26	27
-15,15	101	107	115	121	121	120	122	123	127	126	128			
-12,12	53	87	89	89	88	87	87	86	88	89	92	101	108	110
-9,9	13	16	30	38	40	48	49	48	73	72	71	71	70	70
-6,6	7	8	9	10	11	12	15	23	26	33	34	38	41	41
-3,3	3	3	3	4	4	5	5	6	6	7	8	9	9	13

Table 1.1. Detail of the data presented in figure 1.10; the period in weeks after which an adult should be called back for processor readjustment with respect to the time of the last adjustment. Week 0 denotes the first adjustment session at switch-on of the speech processor. Week 1 denotes last adjustment session 1 week after switch-on, etc. The data were calculated for a period of three years (156 weeks). Absence of data at sessions 25-27 weeks post switch-on for shifts within 15 CUs implies that the time of the last adjustment session (25-27 weeks) plus the readjustment interval exceeds 156 weeks.

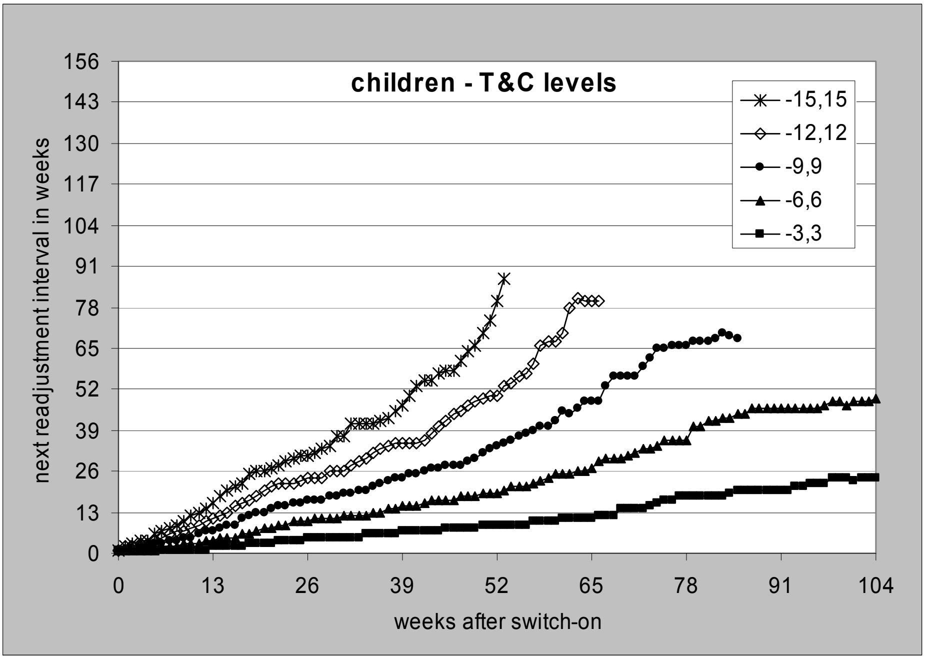

Figure 1.11. Period after which an child should be called back for speech processor readjustment with respect to the time of the last adjustment post switch-on in order to keep the mismatch between the T- and C-levels of the last adjustment and the actual T- and C-levels within the CU range given in the legend.

week	0	1	2	3	4	5	6	7	8	9	10	11	12	13
-15,15	1	2	3	4	4	6	7	8	9	10	12	13	14	16
-12,12	1	2	2	3	3	4	5	6	7	7	8	9	10	11
-9,9	1	1	1	2	2	3	3	4	4	5	5	6	7	7
-6,6	1	1	1	1	1	1	2	2	2	3	3	3	4	4
-3,3	1	1	1	1	1	1	1	1	1	1	1	1	1	2
week	14	15	16	17	18	19	20	21	22	23	24	25	26	27
-15,15	18	20	21	22	25	26	26	27	28	29	30	31	31	32
-12,12	12	13	15	16	17	18	20	21	22	22	22	23	24	24
-9,9	8	9	9	11	12	13	13	14	15	15	16	16	17	17
-6,6	5	5	5	6	6	7	8	8	9	9	10	10	10	11
-3,3	2	2	2	2	3	3	3	3	4	4	4	4	5	5

Table 1.2. Detail of the data presented in figure 1.11; the period after which a child should be called back for processor readjustment with respect to the time of the last adjustment. Week 0 denotes the first adjustment session at switch-on of the speech processor. Week 1 denotes last adjustment session 1 week after switch-on, etc. The data were calculated for a period of three years (156 weeks).

session	0	1	2	3	4	5	6	7	8	9	10	11	12	13
-15,15	0	1	4	11	32	>156								
-12,12	0	1	3	7	15	102	>156							
-9,9	0	1	2	4	7	12	22	95	>156					
-6,6	0	1	2	3	4	6	9	13	19	31	85	>156		
-3,3	0	1	2	3	4	5	6	7	8	9	10	12	14	17
session	14	15	16	17	18	19	20	21	22	23	24	25	26	27
-15,15														
-12,12														
-9,9														
-6,6														
-3,3	21	27	40	67	101	146	>156							

Table 1.3. Readjustment intervals cumulated for adults. The table presents the number of weeks after switch-on at which one should plan the successive readjustment sessions in order to keep the mismatch between the T- and C-levels of the last adjustment and the actual T- and C-levels within the CU range given in the left column. Week 0 denotes the first fitting session at signal processor switch-on. Whatever mismatch is considered acceptable, the data show that the first readjustment session should take place one week after switch-on. The number of refitting sessions within one year (52 weeks) after switch-on required for mismatches to be limited to 3, 6, 9, 12, and 15 CUs are 16, 9, 6, 4, and 4, respectively.

session	0	1	2	3	4	5	6	7	8	9	10	11	12	13
-15,15	0	1	3	7	15	35	76	>156						
-12,12	0	1	3	6	11	20	40	75	>156					
-9,9	0	1	2	3	5	8	12	19	32	51	84	153	>156	
-6,6	0	1	2	3	4	5	6	8	10	13	17	23	32	44
-3,3	0	1	2	3	4	5	6	7	8	9	10	11	12	13
session	14	15	16	17	18	19	20	21	22	23	24	25	26	27
-15,15														
-12,12														
-9,9														
-6,6	61	86	130	>156										
-3,3	15	17	19	22	26	31	36	42	49	57	67	79	97	119

Table 1.4. Readjustment intervals cumulated for children. The table presents the number of weeks after switch-on at which one should plan the successive readjustment sessions in order to keep the mismatch between the T- and C-levels of the last adjustment and the actual T- and C-levels within the CU range given in the left column. Week 0 denotes the first fitting session at signal processor switch-on. Whatever mismatch is considered acceptable, the data show that the first readjustment session should take place one week after switch-on. The number of refitting sessions within one year (52 weeks) after switch-on required for mismatches to be limited to 3, 6, 9, 12, and 15 CUs are 22, 13, 9, 6, and 5, respectively.

As a consequence of this additional constraint, the number of tracks decreased from the original number of 258 for adults (86 subjects, 3 electrodes) and 180 for the 60 children to 225 T-level and 239 C-level tracks for adults and 157 T-level and 171 C-level tracks for children. Thus, time constants larger than the time interval spanned by the data are found somewhat more frequently for T-levels than for C-levels. However, the number of remaining tracks suffices to present reliably the risk of shifts of more than 3 to 15 CUs when one decides to have a final readjustment session at 3, 6, or 12 months (13, 26, or 52 weeks) after speech processor switch-on.

The results are presented in Tables 1.5 and 1.6 for adults and children, respectively. In line with the previous results we find the highest risk of long term mismatches for C-levels rather than for T-levels. Thus, taking the data for C-levels and focussing again on mismatches up to 6 CUs we note that when the last readjustment is performed at 3, 6, and 12 months, the risk of level shifts in excess of 6 CUs is 29, 13 and 5% (6% for T-levels in the latter case) in adults and 68, 41 and 18% in children, respectively. The risk found for children strongly suggests that one should continue to follow these implant recipients for more than one year after speech processor switch-on, even for the smallest mismatch considered to be acceptable.

weeks	13	26	52	13	26	52
shift	T-levels			C-levels		
-15,15	9.1	3.7	1.6	12.4	5.2	0.8
-12,12	12.6	4.8	2.2	16.1	6.8	1.1
-9,9	16.2	7.1	3.3	20.5	9.6	3.9
-6,6	23.9	12.2	6.0	28.6	13.1	4.9
-3,3	36.1	22.2	8.4	40.1	21.5	8.2

Table 1.5. Risk in percent for adults of a long term shift in level larger than indicated in the left column when one performs the final readjustment of the speech processor 13, 26, or 52 weeks after processor switch-on.

weeks	13	26	52	13	26	52
shift	T-levels			C-levels		
-15,15	20.8	10.9	6.1	40.6	18.8	10.7
-12,12	28.2	13.9	8.9	48.8	22.3	12.9
-9,9	32.6	21.9	10.8	61.1	28.6	15.0
-6,6	40.2	28.6	15.2	68.2	41.1	18.3
-3,3	48.0	35.5	21.7	75.8	56.7	24.7

Table 1.6. Risk in percent for children of a long term shift in level larger than indicated in the left column when one performs the final readjustment of the speech processor 13, 26, or 52 weeks after processor switch-on.

1.3.6 *A special case*

One implant recipient did not show an exponential growth of the T- and C-levels with time, nor deviations from exponential growth as shown in figure 1.1. In fact, her T- and C-levels did not stabilise over the years. The T- and C-levels changed roughly in parallel by more than her dynamic range. Hence, readjustment of the signal processor was frequently required. She finally learned to perform the required readjustments herself and received the apparatus needed to do so. After several years she received a cochlear implant in the other ear. T- and C-levels did not stabilise in that ear either. There appeared to be a clear correlation between the fluctuations in the two ears. Thus, a (central) neurological disorder may have played a part in this case. The course of the T- and C-levels for electrode 13 in the first ear implanted is presented in figure 1.12.

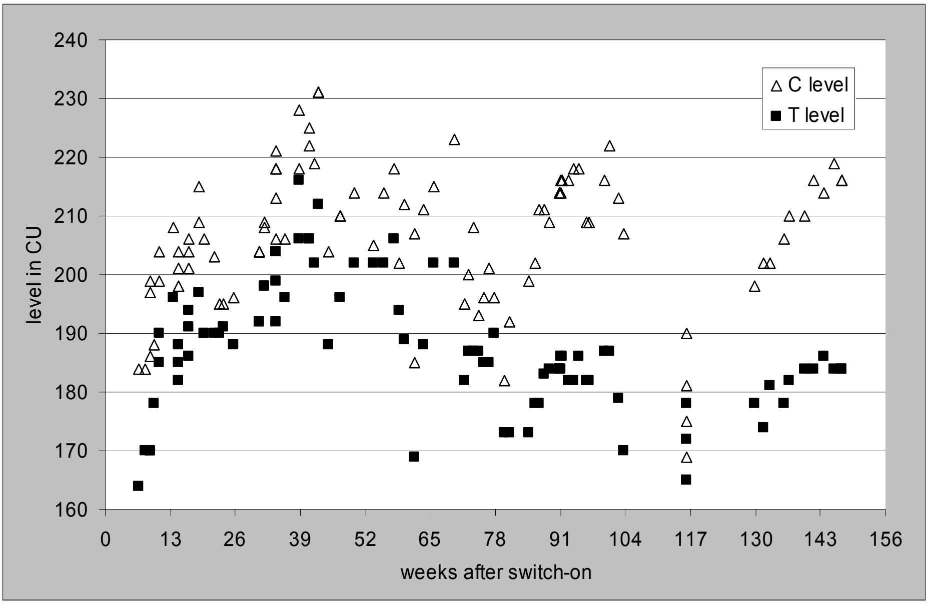

Figure 1.12. A special case where T- and C-levels changed considerably in an irregular fashion. C-levels could occasionally drop below T-levels measured at another time.

1.4 Discussion

The present analysis was initiated because we aimed to optimise the time intervals for speech processor readjustments. The approach was essentially statistical in nature, considering a group of implant recipients. The history of individuals was not taken into account. For example, when calculating the interval at which an implant recipient should be called back for speech processor readjustment we looked at the distribution of the shifts in level without attempting to predict the shift for a particular individual. If T- and C-levels in a particular individual seem to have stabilised at a point in time where the statistics still show changes in levels (for example the C-levels in children half a year after switch-on, figure 1.5) a clinician may choose a longer time interval until the next visit to the clinic than the statistics tell us. However, estimating the degree of stabilisation is a subjective matter, depending on insight and experience of the clinician performing processor adjustments. This certainly is a factor in clinical practice but it is very difficult to include a criterion in the present analysis that accurately represents the clinician's judgement. Therefore, this analysis had to be restricted to a statistical approach. It implies that in clinical practice readjustment intervals may be longer than the ones calculated here. In that respect one may conclude that the present results may overestimate the frequency at which implant recipients should be called back. However, in view of the difficulty in estimating the development of T- and C-levels in a particular implant recipient in clinical practice we feel that the present data are very helpful in planning the capacity of an implant team. In addition, we should take into account that the implant recipients for which the stimulation rate had to be changed at some point in time after processor switch-on needed extra care of the clinician, which somewhat increases the capacity to be planned.

The calculations were performed for maximum acceptable shifts from 3 to 15 CUs in both directions. Of course, it is up to an implant team to set its own standard with respect to the acceptable shift. The range from 3 to 15 CUs was chosen with the intention of covering all standards that may be acceptable to implant teams. In our presentation of the results, we focussed on a maximum allowable shift of ± 6 CUs. Previous analyses of our data have shown that measurement accuracy amounts to about 3 CUs for T-levels and to 2-3 CUs for C-levels. A maximum allowable shift equal to measurement error may be considered too small. In a few cases a 2-3 CUs increment in C-levels may result in overstimulation. However, the general trend over time is an increase of C-levels so that the programmed levels will tend to be lower, rather than higher, than the actual levels. Therefore, we did not take the ±3 CUs criterion to illustrate our results but the next one of ±6 CUs. The criterion of ±9 CUs is obviously too large with respect to C-levels in many subjects. The present results showed that in almost all cases C-levels determined the length of the intervals at which the speech processor has to be readjusted. Moreover, clinical experience has taught us that T-levels are not very critical in processor programming. Although we have presented our view about the maximum acceptable mismatch it is, once again, of course the responsibility of the implant team to set this standard.

The results showed that data from the first series of adults, measured before mid-2000, were somewhat deviant from the latter ones because especially the changes in C-level tended to be smaller than those measured later. Evidently, there was initially a certain reluctance to push up C-levels. However, this had only a limited effect on the present results. The readjustment intervals in adults may have been somewhat overestimated. We mentioned earlier (Sec. 1.3.2) that the smaller changes in C-levels of the first series of adults could be related to an effect of stimulation rate, a low rate being associated with the first implant recipients who were measured over the longest periods. In principle, it is impossible to distinguish between these two effects. However, the absence of a main effect of rate and the absence of this effect in children measured later, at low and high rates, suggest strongly that the effect of stimulation rate as revealed by analysis of variance was essentially an effect of a gradual change in the way C-level measurements were performed.

Finally, the results have shown that the correlation between the initial shifts in T- and C-levels is low. C-levels may be difficult to assess in the initial stage of speech processor programming. In particular, it may be difficult to decide to what extent they can be increased in a responsible way. The low correlation between the T- and C-levels suggests, however, that the changes in C-level cannot be estimated from changes in T-level. However, this examination was helpful with respect to optimising the adjustment procedure as there appeared to be a high correlation between the shifts in T- and C-levels across electrodes while the magnitudes of the shifts were similar. This suggests that one may start a readjustment session by shifting the levels in all electrodes by equal amounts of current units, an approach accommodated by present-day speech processor programming software.

Chapter 2

T- and C-level profiles across the electrode array; fitting the speech processor by profile parameter adjustment

Summary

Objective

Adjusting and readjusting T- and C-levels for some 20 electrodes is quite demanding for implant recipients and it takes a lot of the implant team's time. Moreover, it is based on the perception of unnatural sounds because one has to use bursts of electrical impulses presented to individual electrodes independently. Natural sounds require the representation of the spectral distribution of these sounds across an array of electrodes. Therefore, it is interesting to investigate whether or not it is possible to find certain parameters governing the profiles of the T- and C-levels measured across the electrode array. These parameters can then be implemented as controls in speech processor programming software, enabling processor adjustment while using natural sounds. If the number of parameters required is substantially lower than the number of electrodes, this approach could also improve the efficiency of the fitting procedure. Many T- and C-profiles collected over the years enable an analysis meeting this objective.

Conclusions

Three parameters appeared to govern T- and C-profiles. These parameters are *shift*, - a nearly parallel shift in current units of the stimulation levels for all electrodes with respect to the population average, *tilt*, - a change in the slope of the stimulation profile, and *curvature*, - a change from a peak- to a valley-shaped profile. Proper adjustment of these three parameters yields profiles within 2 to 3 current units from the T- and C-levels measured originally with individual electrodes. However, the two parameters determining the shape of the profile, *tilt* and *curvature*, are not related to easily identifiable aspects of the sounds implant recipients perceive. Therefore, they may not be suitable for interactive speech processor adjustment between implant recipient and clinician.

An extended analysis yielded two new profile-shape related parameters, *bass* and *treble*, that do correspond to identifiable aspects of sounds. Together with the *shift* parameter these two parameters can be used in an adjustment procedure applying natural sounds.

Proper adjustment of these three parameters yielded profiles with differences between the profile levels and the T- and C-levels measured originally that were only 25% greater than those found for the *tilt* and *curvature* parameters.

In addition to the proposed profile adjustment method, the three parameters can also be used to estimate T- and C-levels for individual electrodes across the full electrode array after measuring the T- and C-levels at two marginally located electrodes (*i.e.* the basal and apical ends of the array) and at one electrode in the middle of the array. This approach could replace general interpolation schemes that are not based on T- and C-level statistics.

Parameter extraction facilitates the comparison of T- and C-profiles measured at several points in time after implantation. The values of the parameters related to the shape of the T- and C-profiles (*i.e. tilt* and *curvature*, or *bass* and *treble*), measured shortly after implantation and at a much later date, showed a modest correlation (0.4 to 0.5). This implies that the shape of the profile may significantly change as implant recipients adapt to their device. Thus, readjustments cannot be limited to overall stimulation level (*i.e., shift*).

Parameter extraction also facilitates the comparison of T- to C-profiles. For all shape-related parameters (*i.e., tilt* and *curvature*, or *bass* and *treble*) the correlation between the T- and C-parameter values was 0.6 to 0.7. During the initial adaptation period, T-levels may be easier to measure than C-levels. The correlation suggests that the shape of the T-level profile somewhat predicts the shape of the C-level profile.

2.1 Introduction

Threshold and loudness comfort levels (T- and C-levels) are commonly adjusted for individual electrodes. With up to 22 electrodes per array and the need to frequently readjust these levels, particularly shortly after implantation (see Chapter 1), these readjustments take much time. Moreover, within a fitting session the time consuming level adjustments for individual electrodes impede a procedure in which different adjustments are tried out interactively with the implant recipient. Therefore, the question arises of whether or not it is feasible to adjust certain aspects of the level profile across the electrode array, adjusting stimulus levels for different electrodes simultaneously rather than individually. This requires some a-priori knowledge of the profiles. Experience in fitting speech processors, accumulated over years, shows that there is certain conformity in level profiles. This can be concluded statistically but there also are physiological reasons to expect certain profile properties. For example, the extent to which neurones are degenerated will show a longitudinal correlation across the full electrode array from base to apex. In addition, there is considerable longitudinal spread of excitation, particularly with present day monopolar stimulation, which contributes to a correlation between the levels adjusted for adjacent electrodes. Thus, stimulus levels are adjusted for individual electrodes although these levels are not independent of one another.

An important reason to adjust stimulation levels individually is the chance to hit an outlier, an exceptionally low or high level considering the levels adjusted for adjacent electrodes, or to encounter an electrode that is not delivering a proper sound. However, in view of the previous physiological considerations and our experience chances are great that such an outlier is due to technical problems in the implant itself. After initial checks such as impedance measurements, integrity tests (checking the actual stimulus waveform via recording electrodes placed on the scalp) and recordings of the electrically evoked compound action potential do not show abnormalities, we may expect that the risk of outliers is small.

The concept of adjusting profile parameters rather than individual electrode levels is not solely based on time efficiency considerations. Spread of excitation, mentioned above, implies that adjacent electrodes excite to a large extent the same neurones. Thus, if the stimulus levels for individual electrodes are adjusted to the threshold of audibility then (near-) simultaneous stimulation of adjacent electrodes will evoke sounds above threshold. With simultaneous stimulation the threshold may be found at some 10 Current Units (CUs, as defined by Cochlear, 1 CU ≈ 0.18 dB) below the thresholds for the electrodes stimulated individually. Also, if individual electrode levels are set at loudness comfort level then spread of excitation and loudness summation may cause loudness levels to rise above comfort level when the electrodes are stimulated simultaneously.

Another aspect of simultaneous stimulation is more important than the above summation effects. Stimulating an individual electrode implies that natural sounds cannot be used but that abstract sounds like tone “beeps” have to be presented when programming the speech processor. Natural sounds have a broad frequency spectrum, which can only be presented to the implant recipient by (near-) simultaneous stimulation of a number of electrodes. Profile adjustment, however, presents the possibility to program the speech processor using natural sounds like speech stimuli rather than tone beeps. This facilitates the fitting procedure in children substantially. It renders the procedure more users friendly. Moreover, it contributes to arriving quickly at the proper adjustment which is important for two reasons: (1) one should avoid that implant recipients adapt to a maladjusted processor in the initial stage of speech processor fitting, which later might be very confusing, and (2) it is very important for language development in children that they are adequately stimulated as soon as possible after deafness commenced.

The present chapter presents a statistical analysis of T- and C-levels of 215 implant recipients, children and adults, collected in almost 4000 sessions over a period of up to 6 years after implantation. The results are presented in terms of the most important parameters describing the T- and C-level profiles across the electrode array. The results suggest how level *profiles* can be adjusted, rather than levels for individual electrodes. The profile parameters can be included in processor fitting software to enable parameter-based processor adjustment. The present results may support clinicians in adjusting the speech processors efficiently. Moreover, a profile-based speech processor fitting procedure is easier to perform than the procedure in which levels for individual electrodes have to be adjusted, which presents the possibility of implant recipients readjusting their processor themselves.

2.2 Methods and materials

2.2.1 Methods

The present analysis is primarily based upon the statistical technique of principal components analysis (PCA). This technique analyses the matrix containing the correlations between the levels for all pairs of electrodes and yields a number of components, substantially smaller than the number of electrodes, providing a good description of the data. These principal components are the important parameters of the T- and C-profiles. Effects of stimulation rate and type of implant device on the profiles are examined by analysis of variance (ANOVA) of these parameters. In addition, these parameters are used to analyse changes in the profiles over time. The Statistica software package was used for statistical analyses (Statsoft Inc., release 7.1).

2.2.2 Materials

The study included 79 children and 136 adults (together 215 subjects). All implants were Nucleus® devices of the types CI24M (N=78), CI24R(CS) (N=124), CI24R(CA) (N=3) and CI24(ST) (N=10). SPEAK and ACE strategies were used, which implied stimulation rates of 250 (N=90) and 720-1200 (N=125) impulses per second. Impulse duration was 25 μs/phase except in five recipients with durations from 37 to 63 μs/phase. The impulse duration parameter was not included for further analysis because five deviant impulse durations is too small a number to assess a possible impulse duration effect. All stimulation was programmed in monopolar mode, mostly using both the separate external electrode and the implant housing as reference electrodes (MP1+2 mode), in a few cases using only the reference electrode (MP1, N=10). The speech processors were of the types Sprint™ (body worn) and Esprit™ (behind the ear).

Implant recipients were included only if they had complete level profiles from electrode number 3 to 22. The basally located electrodes 1 and 2 were not included in the analysis because they were mostly not used. PCA was mainly based upon the most recent T- and C-profile of each implant recipient and not on all profiles measured post implantation.

Including all profiles would result in an unequal representation of implant recipients because the number of profiles per implant recipient varies considerably across recipients. The most recent profile was chosen aiming at profiles that had stabilised as much as possible. The most recent profiles had to be measured at least 6 months post implantation. Comparing the PCA results for all 4000 profiles to those found for the most recent profiles showed that there was no essential difference between the principal components found for these two sets of profiles. When comparing profiles measured shortly after implantation to the most recent ones we avoided comparison of profiles measured at different stimulation rates. If the rate had been changed at some point in time after implantation we included only the initial part of a series of fittings conducted at the same rate (if longer than six months) or the second part conducted at the other rate (if the initial part was short and the latter part sufficiently long).

2.3 Results

2.3.1 Basic analysis

Principal components analysis of T- and C-levels separately showed that there was one dominating component. This component represented 92.8% of the total variance in T-levels and 93.7% in C-levels. For both, T- and C-profiles, this major component was closely related to overall level. The coefficient of the correlation between this component and overall profile level (or average level) was R=1.000. It suggests that the major difference among individual profiles is simply a shift of all stimulation levels across the electrode array by nearly the same amount of current units. In spite of this high percentage of explained variance there are important differences in the shape of the profiles, in addition to the differences in overall levels. The second component explained 4.3% additional variance in the T-levels and 3.7% in the C-levels. This component was highly correlated with the slope of the profiles, R=0.997 for T-and R=0.999 for C-levels. Previously, we have called these two components *shift* and *tilt*. In this analysis we included a third component. This component explained 1.4% of the variance in the T-levels and 1.5% in the C-levels. The third component represents the *curvature* in the profiles: convex versus concave or peak- versus valley-shaped. Together, the three components explained 98.4% of the variance in the T-levels and 98.9% in the C-levels. These high percentages demonstrate that all profiles can be described accurately by only three components.

The intersubject differences in the three parameter values are characterised by the lower decile, the median and the upper decile of their distributions. The results are presented in figures 2.1-2.4 for T- and C-levels and the combinations *shift*-and-*tilt* and *shift*-and-*curvature*. Each figure shows three sets of three curves. The middle curve of each set (circles) shows the result with no *tilt* and *curvature*. Hence, the middle curve of the lower set of three curves presents the values exceed by 90% of the 215 implant recipients (the lower decile) without taking into account *tilt* and *curvature*, the middle one the median values and the upper one the values exceeded by 10% of the recipients. The two remaining curves of each set of three (triangles and squares) indicate the 10% borders for *tilt* (figures 2.1 and 2.3) and *curvature* (figures 2.2 and 2.4). Thus, each profile in these figures is given by the *shift* parameter (with overall T-levels varying between about 118 and 171 CUs and C-levels varying between about 157 and 213 CUs for 80% of the population) modified by a certain *tilt* and *curvature* presented in figures 2.1 through 2.4 in terms of the 80% range.

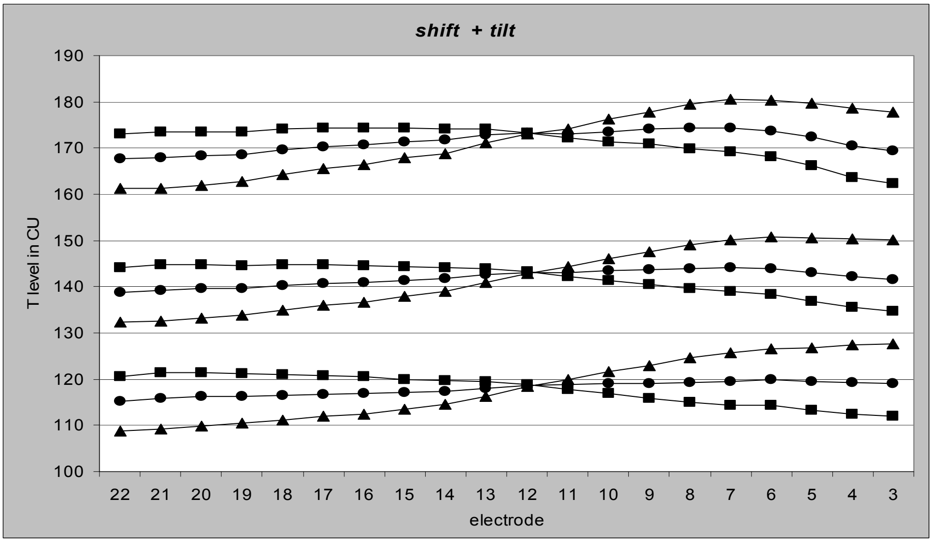

Figure 2.1. T-level profiles across the electrode array from the most apical electrode 22 (stimulated by the lowest frequencies) to the most basal one activated, number 3 (stimulated by the highest frequencies). The middle curve of each set (circles) presents the profile corresponding to only the first principal component. The middle curve of the lowest set shows the profile exceeded by 90% of 215 subjects (the lower decile), the one of the middle set the median profile (exceeded by 50%) and the one of the upper set the profile exceeded by 10% (the upper decile). The two remaining curves of each set (squares and triangles) show the range of *tilts* between the upper and lower deciles.

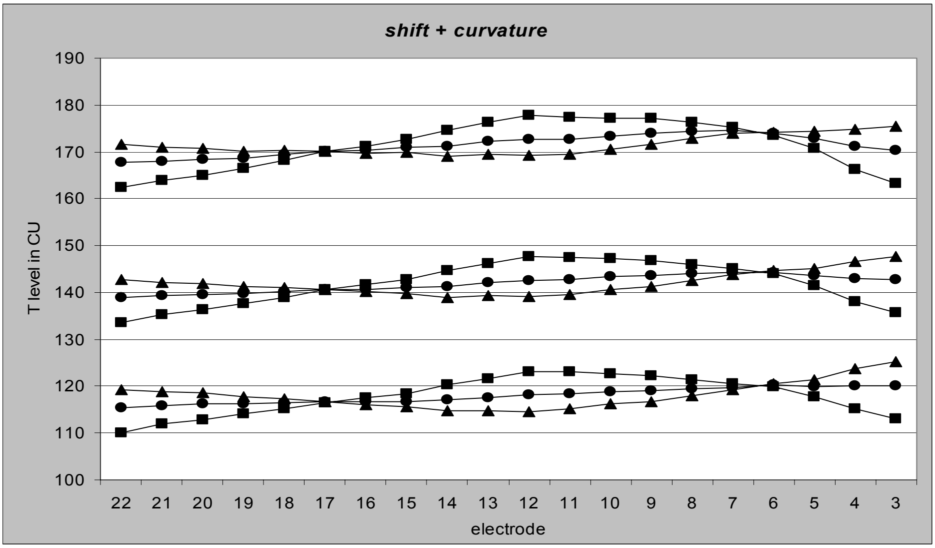

Figure 2.2. As figure 2.1 for the *shift* and *curvature* components.

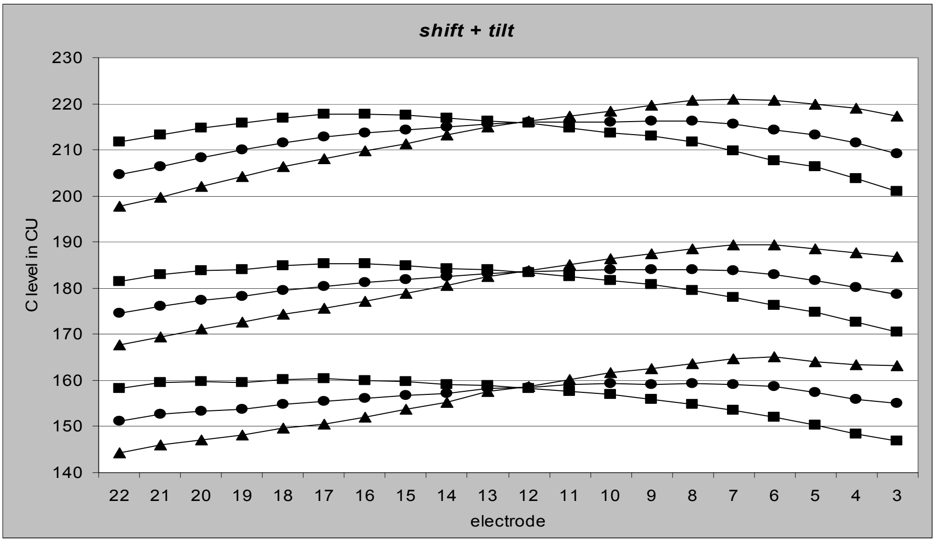

Figure 2.3. C-level profiles across the electrode array from the most apical electrode 22 (stimulated by the lowest frequencies) to the most basal one activated, number 3 (stimulated by the highest frequencies). The middle curve of each set (circles) presents the profile corresponding to only the first principal component. The middle curve of the lowest set shows the profile exceeded by 90% of 215 subjects, the one of the middle set the median profile (exceeded by 50%) and the one of the upper set the profile exceeded by 10%. The two remaining curves of each set (squares and triangles) show the range of *tilts* between the 10% borders.

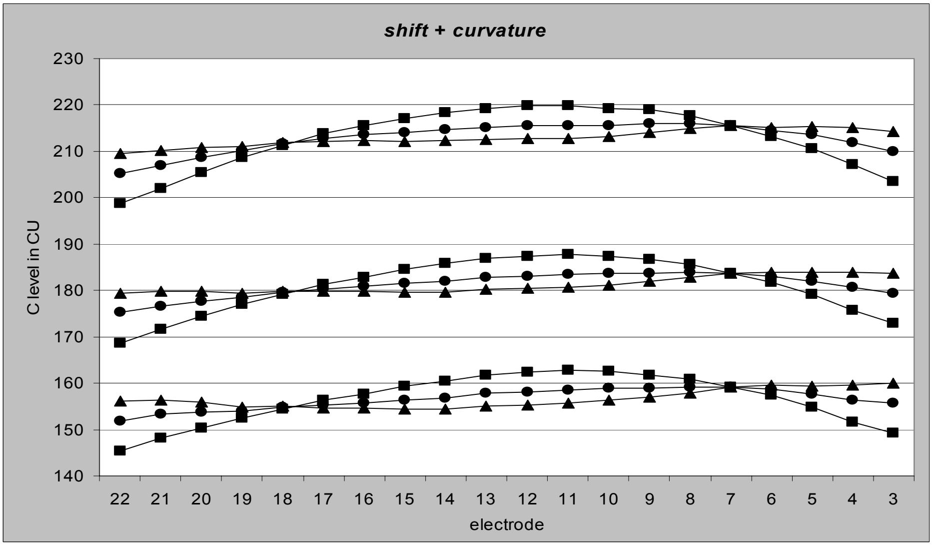

Figure 2.4. As figure 2.3 for the *shift* and *curvature* components.

2.3.2 Mismatch between original measurements and the principal components approach

The mismatch that one has to face if level profiles are restrained to a limited number of principal components rather than measured for individual electrodes can be calculated by comparing the profiles measured originally to the best fitting profiles derived from the principal components. Figures 2.1-2.4 showed quite smooth profiles whereas measured profiles may be more irregular. The calculations were performed in terms of the standard deviations of the differences between the levels of the measured profiles and those of the best fitting principal component profiles. These differences are denoted the *error*. The errors were calculated for the one-component approximation *shift*, the two-component approximation *shift* + *tilt* and the three component approximation *shift* + *tilt* + *curvature*. Figure 2.5 shows the error per electrode for T-levels, figure 2.6 for C-levels. These figures show that after adding the third component the error becomes 2-3 CUs for T-levels and 2 CUs for C-levels, except for electrodes 22 and 3 at the borders of the array. An error of 2 CUs is about the accuracy that one may expect when adjusting C-levels for individual electrodes (except perhaps in few cases with very small dynamic ranges). Adjusting T-levels is somewhat less accurate. In addition to examining the errors per electrode, for all implant recipients together, it is also interesting to examine the error distribution across implant recipients. Figure 2.7 shows this error distribution, for all electrodes together. This figure shows a clear shift of the distribution toward more subjects with smaller errors when the number of principal components contributing to the profile is increased from two to three. Still, the result for three components shows that the number of subjects with an error across the whole electrode array above 3 CUs is 46 for T-levels and 24 for C-levels in the group of 215 subjects.

2.3.3 Comparison of the principal components mismatch between the profiles measured shortly after implantation and the most recent ones

Inspecting the profiles as they developed after implantation it seemed that they became smoother as fitting of the speech processor progressed. Therefore, it is interesting to examine the mismatch error as a function of time after implantation. Taking the first complete profile of each implant recipient when there are no electrodes switched off or some individual levels set at very low values, we found an overall error across electrodes and subjects of 3.6 CUs for T-levels and 3.3 CUs for C-levels. For the most recent profiles these errors were only 2.8 and 2.3 CUs. These results were found when both the set of initial profiles and the set of most recent profiles were analysed using the principal components of the final set. Therefore, this result may be due to the choice of principal components; the principal components derived from the set of most recent profiles may be less suited to accommodate the initial profiles. However, when using the principal components found for the merged sets of initial and final profiles, the analysis yielded essentially the same errors of 3.6 CUs and 3.3 CUs for the T- and C-levels of the initial profiles and 2.8 CUs and 2.3 CUs for the T- and C-levels of the most recent profiles. This result suggests that the initial profiles were more irregular than the most recent ones. Thus, one might obtain better profiles to start with (profiles more like the final ones) when one starts with the principal components approach. The initial profiles were collected between 42 and 140 days after implantation, the most recent ones between 180 and 1500 days after implantation (lower and upper deciles).

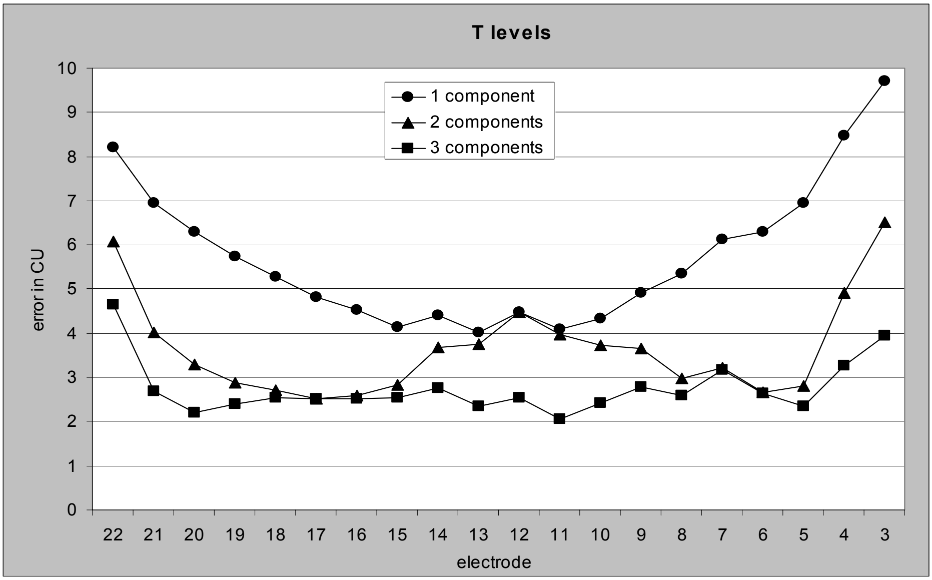

Figure 2.5. Error in current units between the T-levels measured originally and the profiles best fitting these T-levels calculated with 1, 2 and 3 principal components (*shift*, *shift*+*tilt* and *shift*+*tilt*+*curvature*). The error is expressed as the standard deviation of the differences per electrode between the levels of the measured profiles and those of the best fitting principal component profiles, calculated across all implant recipients.

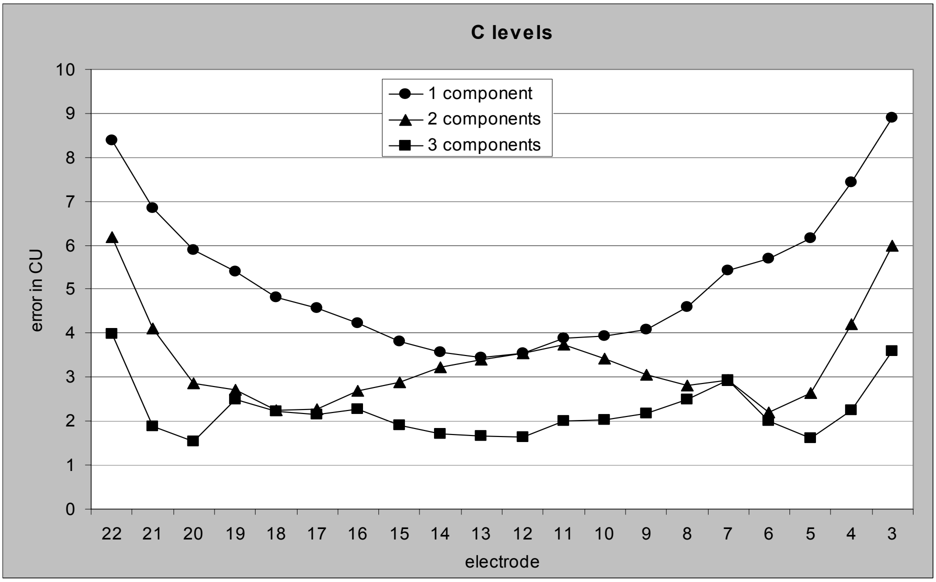

Figure 2.6. As figure 2.5 for C-levels.

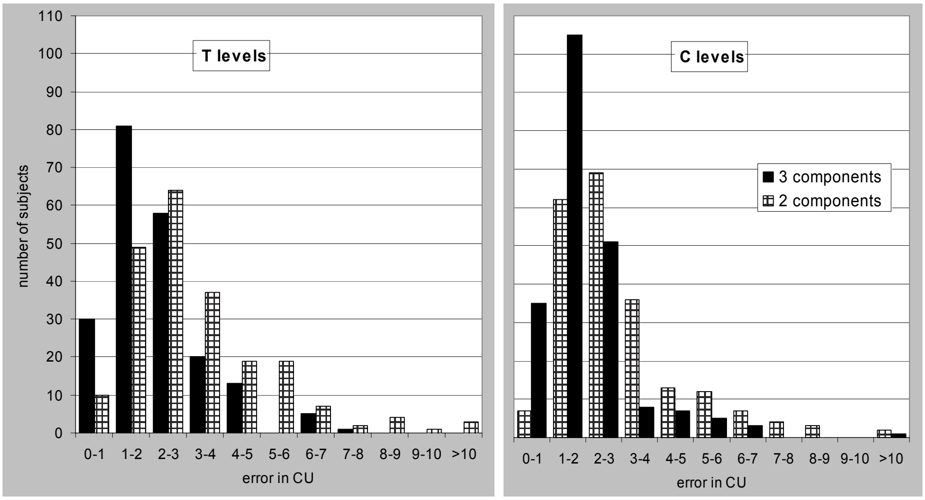

Figure 2.7. Distribution across 215 implant recipients of the error in current units, calculated in terms of the standard deviation across all electrodes of the differences between the levels measured originally and those of the best fitting profiles. Results for best fitting profiles calculated with two and three principal components (*shift*+*tilt* and *shift*+*tilt*+*curvature*).

2.3.4 Effects of stimulation rate, type of implant device and type of reference electrodes

Both modes of reference electrodes, MP1 and MP1+2, were used only with the CI24M device and only at an impulse rate of 250 Hz. An ANOVA for this restricted set of data showed no significant difference between the T-levels measured in these two modes (p>0.5) and the C-levels in the two modes (p>0.9). Therefore, this parameter was excluded from further analysis.

The effects of *stimulation rate* and *type of implant device* on T- and C-levels were subjected to a two-factorial ANOVA with the T- and C-levels across electrodes as two sets of repeated measures. In addition, ANOVA was performed for the three principal components *shift*, *tilt* and *curvature*. *Stimulation rate* was specified in two categories: the low rate of 250 Hz and high rates of 720-1200 Hz. The CI24R(CA) device was excluded from this analysis since there were data for only three implant recipients. The repeated measures ANOVA of the T-levels across electrodes showed significant effects of *stimulation rate* (p<0.005), *type of implant device* (p=0.005), and *electrode number* (p<0.001). All interaction terms were insignificant (p>0.1). ANOVA for the three principal components of the T-levels yielded results completely in line with the previous analysis: the first component, *shift*, showed significant effects of *stimulation rate* and *type of implant device*, at exactly the same significance levels as the main effects in the previous analysis, and no interaction between these two effectors. The other principal components, *tilt* and *curvature,* did not show significant effects of *stimulation rate* and *type of implant device*. Thus, a clear result emerged: T-levels depended on electrode number (on average T-profiles are not flat), their profiles shifted parallel wise when stimulation rate was changed, and the overall level differed with the type of implant device. In contrast to these significant effects, there was no significant effect of *stimulation rate* and *type of implant device* on the shape of the profiles. Repeated measures ANOVA of the C-levels yielded no significant main effects of *stimulation rate* (p>0.6), *type of implant device* (p>0.4), and no significant interaction between these two effectors (p>0.6). However, there was a significant effect of *electrode number* (like T-profiles, C-profiles are not flat), and a significant interaction between *electrode number* and *stimulation rate* (p=0.05). The latter effect reappeared in ANOVA of the principal components of the C-levels. The third component, *curvature*, was affected significantly (p=0.05) by *stimulation rate*. This was the only significant effect found for the three principal components of C-profiles, in line with the results of the repeated measures ANOVA of the C-levels for individual electrodes.

In order to obtain unbiased estimates of the effects of *stimulation rate* and *type of implant device* ANOVA was repeated excluding the interaction term *stimulation rate * type of implant device* from the statistical model and calculating the best fitting means in terms of least squared differences. The significance levels closely matched those of the previous analysis. The effect of *stimulation rate* on T- and C-levels is shown in figure 2.8 and the effect of *type of implant device* on these levels in figure 2.9.

Figure 2.8 shows how the average profile of T-levels measured at high stimulation rates follows a course parallel to the profile measured at low rate, in line with the absence of an interaction term in ANOVA. Also in line with the previous ANOVA results, figure 2.8 shows that C-levels measured at high rate had somewhat more *curvature* than those measured at low rate. Note that there was almost no effect of *stimulation rate* on C-levels whereas there was an effect of -9.6 CUs on T-levels when stimulation rate was increased from 250 to 720-1200 Hz. A decrease in threshold with increasing stimulation rate is to be expected. However, a decrease in the C-levels could also be expected. The present result suggests that the increase in stimulation rate hardly affects the number of action potentials elicited by the electrical stimulus at C-level. The major effect of increasing the stimulation rate could be an increase in the stochasticity of the neuronal discharges rather than an increase in discharge rate.

Figure 2.9 shows that the mean profile of T-levels of the CI24R(CS) device follows a course parallel to the mean profile of the CI24M. T-levels of the CI24R(CS) device were 10.2 CUs lower than those of the CI24M. T-levels of the CI24R(ST) device were on average 6.5 CUs lower than those of the CI24M. Figure 2.9 suggests that the profiles of the CI24R(ST) implant were more nearly flat than the other ones. However, this interaction between *type of implant device* and *electrode number* was not statistically significant, perhaps because the number of CI24R(ST) implants was only 10. The results for C-levels are qualitatively similar to those for T-levels but the effects are smaller and, as mentioned above, statistically insignificant.

2.*3.5 Practical application of the principal components approach*

The previous sections showed how well T- and C-profiles can be described by the three parameters *shift*, *tilt* and *curvature*. These parameters can be implemented in fitting software as buttons controlling these aspects of the T- and C-profiles. However, while *shift* can be adjusted easily in threshold or comfortable loudness level measurements, it is not at all trivial what one should ask an implant recipient in order to be able to adjust *tilt* and *curvature*. When performing a threshold measurement one could search for the *tilt* that yields the highest *shift* value. In principle, this would yield a correct adjustment but adjusting two parameters interactively is not an attractive option. Moreover, this does not yet include the third parameter, *curvature*. A practical adjustment procedure would re-

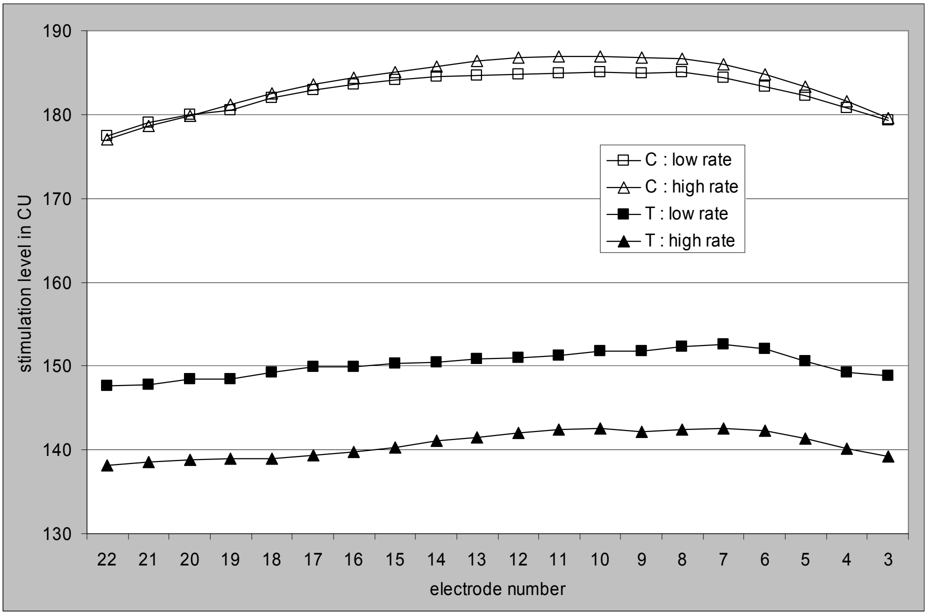

Figure 2.8. Effect of *stimulation rate* on T- and C-levels. Low rate implies 250, high rate 720-1200 impulses per second. The mean values are presented.

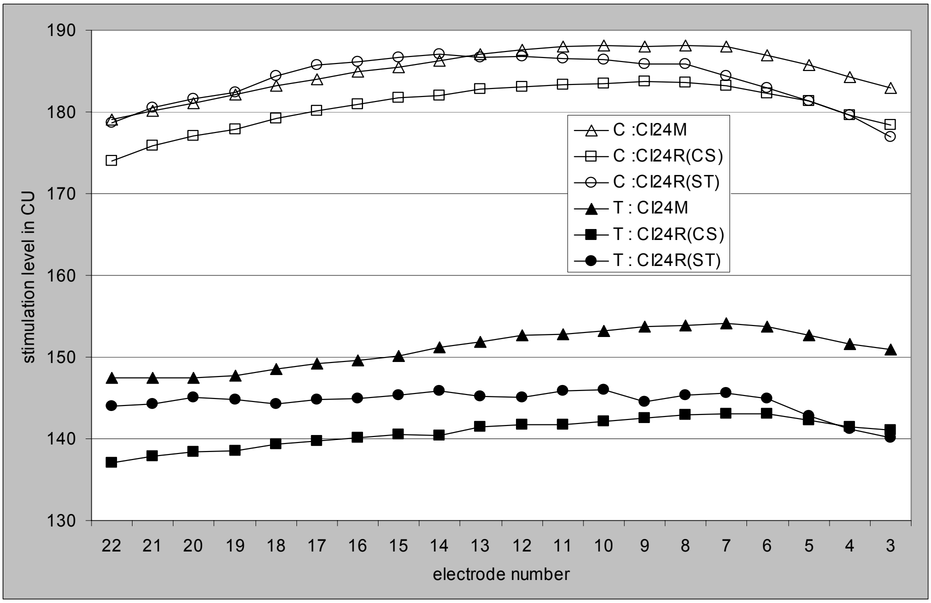

Figure 2.9. Mean T- and C-levels for three types of implant devices.

electrode	22	21	20	19	18	17	16	15	14	13
T average	140.1	140.6	141.0	141.2	141.8	142.3	142.6	143.0	143.4	144.3
T shift	0.955	0.949	0.948	0.953	0.966	0.974	0.980	0.988	0.988	0.997
T tilt	-0.762	-0.786	-0.747	-0.690	-0.629	-0.570	-0.515	-0.419	-0.337	-0.200
T curvature	-0.759	-0.575	-0.472	-0.302	-0.179	-0.006	0.125	0.246	0.467	0.566
C average	175.8	177.4	178.7	179.6	180.9	181.8	182.6	183.2	183.8	184.6
C shift	0.922	0.927	0.949	0.971	0.980	0.992	0.997	0.996	1.000	0.991
C tilt	-0.858	-0.830	-0.782	-0.708	-0.647	-0.602	-0.496	-0.379	-0.232	-0.088
C curvature	-0.990	-0.760	-0.502	-0.226	-0.051	0.153	0.296	0.449	0.570	0.619
electrode	12	11	10	9	8	7	6	5	4	3
T average	144.9	145.0	145.4	145.7	146.0	146.2	146.1	145.2	144.3	143.7
T shift	0.993	0.988	0.991	1.000	0.999	0.999	0.977	0.963	0.932	0.916
T tilt	-0.020	0.135	0.309	0.453	0.616	0.723	0.793	0.882	0.961	1.000
T curvature	0.712	0.660	0.547	0.462	0.282	0.106	-0.059	-0.299	-0.708	-1.000
C average	185.0	185.2	185.4	185.5	185.5	185.1	184.2	183.0	181.5	180.0
C shift	0.995	0.984	0.979	0.987	0.984	0.973	0.961	0.965	0.959	0.936
C tilt	0.028	0.158	0.293	0.408	0.553	0.692	0.798	0.844	0.931	1.000
C curvature	0.653	0.657	0.575	0.449	0.272	0.001	-0.195	-0.437	-0.741	-1.000

Table 2.1. Average T- and C-levels and the *shift*, *tilt* and *curvature* weighting factors. The largest weighting factor of each parameter is set to –1.000 or +1.000. Per electrode the levels of each profile are found by adding the values of the *shift*, *tilt* and *curvature* parameters, V_{shift}, V_{tilt}, and $V_{curvature}$, multiplied by the respective weighting factors, to the levels of the average profile.

quire a few parameters that are related to easily identifiable aspects of sounds, and also parameters that are mutually independent in the adjustment procedure. The *shift*, *tilt*, and *curvature* parameters do not seem to meet this requirement. However, the results can be used in another way. Since the number of independent variables is effectively only three one can reduce the number of T- and C-level measurements for individual electrodes to only three and use the principal components to estimate the T- and C- levels for the other electrodes across the array.

Intuitively the best result will be obtained when one measures T- and C-levels for three electrodes that are well separated, because the levels measured for these electrodes will be least correlated. The best set appears to consist of electrodes 22, 12, and 3. Since level measurements for marginally located electrodes are sometimes problematic the results are also given for electrodes 21, 12, and 4. The weighting factors for the three principal components are presented in Table 2.1. These factors are normalised by setting the coefficient with the largest magnitude to -1.000 or +1.000, depending on the sign of the coefficient. They have to be multiplied by the values of the *shift*, *tilt*, and *curvature* parameters, V_{shift}, V_{tilt}, and $V_{curvature}$, in order to find the contributions of these parameters to the T- and C-levels. The calculation of the parameter values from the levels, L_n, measured for the *n*-th electrode is given by:

For T-levels measured for electrodes 22, 12, and 3:

$$V_{shift} = +0.258*L_{22} + 0.567*L_{12} + 0.208*L_{3}$$
$$V_{tilt} = -0.613*L_{22} + 0.097*L_{12} + 0.535*L_{3}$$
$$V_{curvature} = -0.377*L_{22} + 0.616*L_{12} - 0.275*L_{3}$$

For C-levels measured for electrodes 22, 12, and 3:

$V_{shift} = +0.230*L_{22} + 0.623*L_{12} + 0.180*L_{3}$
$V_{tilt} = -0.542*L_{22} - 0.001*L_{12} + 0.535*L_{3}$
$V_{curvature} = -0.327*L_{22} + 0.582*L_{12} - 0.297*L_{3}$

For T-levels measured for electrodes 21, 12, and 4:

$V_{shift} = +0.295*L_{21} + 0.488*L_{12} + 0.252*L_{4}$
$V_{tilt} = -0.602*L_{21} + 0.060*L_{12} + 0.550*L_{4}$
$V_{curvature} = -0.428*L_{21} + 0.724*L_{12} - 0.336*L_{4}$

For C-levels measured for electrodes 21, 12, and 4:

$V_{shift} = +0.262*L_{21} + 0.551*L_{12} + 0.217*L_{4}$
$V_{tilt} = -0.569*L_{21} - 0.018*L_{12} + 0.568*L_{4}$
$V_{curvature} = -0.375*L_{21} + 0.692*L_{12} - 0.355*L_{4}$

2.3.6 From tilt and curvature to bass and treble control

Although the previous section showed that the full profile can be estimated from T- or C-levels measured for three well chosen electrodes, it remains attractive to seek profile parameters that are related to easily identifiable aspects of sounds, so that one can adjust these profile parameters interactively with the implant recipient. A solution is found in combining the *tilt* and *curvature* parameters into two different parameters. If *tilt* is increased in the sense of more high-frequency and less low-frequency stimulation and *curvature* is increased in the sense of a more pronounced peak (less low- and high frequency stimulation) then the two changes together result in much lower low-frequency stimulation levels and no change in high-frequency stimulation levels. In other words: the result amounts to a simple *bass* control. Vice versa, a *treble* control is obtained if *tilt* is decreased and *curvature* increased. This procedure can be optimised within principal components analysis by so-called rotation of the *tilt* and *curvature* components. The biquartimax procedure is the most suitable one among the rotation procedures commonly available in statistical packages. The best solution is found if the biquartimax procedure is applied to all 20 principal components (not just the three most important ones) and if oblique, non-orthogonal, factors are allowed. A draw back of this approach is that the amount of variance explained by the three components may decrease. Weighing this draw back against the improved relation between the three new components and easily identifiable aspects of sounds, the latter option appeared to be the best one. The variance explained decreased from 98.4 to 97.4% for T-levels and from 98.9 to 98.3% for C-levels. This corresponded to an increase in the total mismatch error from 2.8 to 3.5 CUs for T-levels and from 2.3 to 2.9 CUs for C-levels, an increase of about 25%.

Figure 2.10 shows the results in terms of weighting factors for the first set of parameters *shift*, *tilt* and *curvature*, figure 2.11 for the second set *shift*, *bass* and *treble*. All curves in figures 2.10 and 2.11 are normalised with respect to the magnitude of the largest weighting factor. This factor is arbitrarily set at -1.000 or +1.000, depending on the sign of the original weighting. Figures 2.10 and 2.11 show that the *shift* parameter does not represent a completely parallel shift. The levels for the electrodes at the margins of the electrode

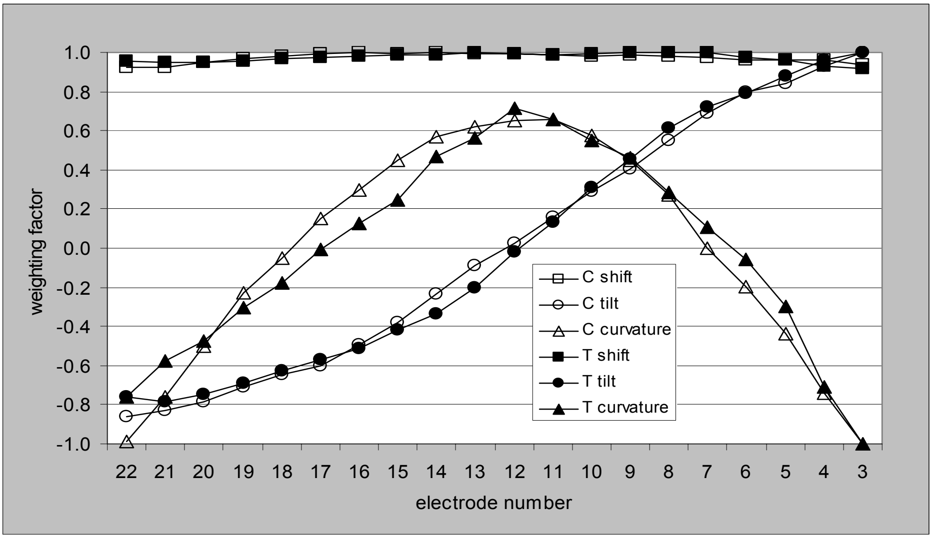

Figure 2.10. S*hift*, *tilt* and *curvature* weighting factors for T- and C-levels. The largest weighting factor of each parameter is set to -1.0 or +1.0. Per electrode the levels of each profile are found by adding the values of the *shift*, *tilt* and *curvature* parameters, V_{shift}, V_{tilt}, and $V_{curvature}$, multiplied by these weighting factors to the levels of the average profile.

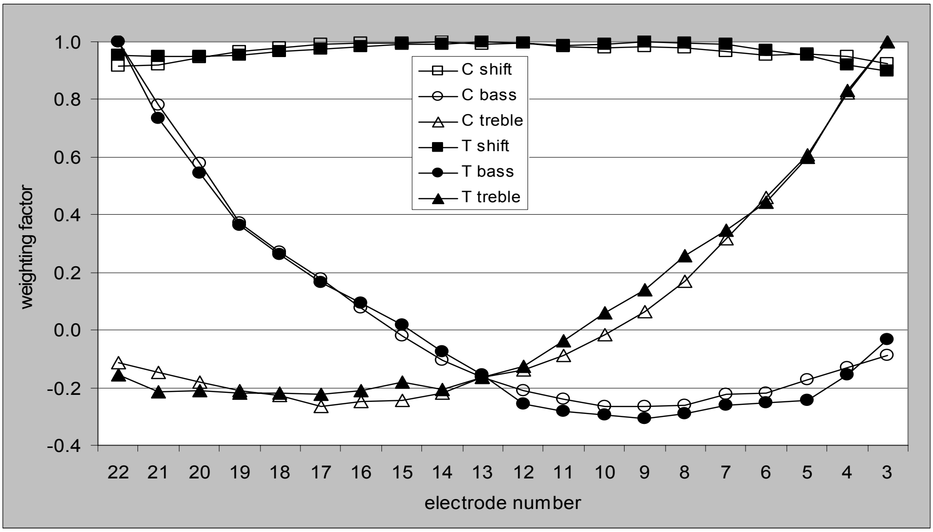

Figure 2.11. S*hift*, *bass* and *treble* weighting factors for T- and C-levels. The largest weighting factor of each parameter is set to -1.0 or +1.0. Per electrode the levels of each profile are found by adding the values of the *shift*, *bass* and *treble* parameters, $V_{shift,}$, V_{bass}, and V_{treble}, multiplied by these weighting factors to the levels of the average profile.

array change less than those for the middle of the array. This could be noticed also in figures 2.1 through 2.4. The *tilt* and *curvature* coefficients in figure 2.10 show a nearly symmetric behaviour with respect to the centre of the array, specifically electrode 12. In addition, note in figures 2.10 and 2.11 that the T- and C-coefficients are nearly the same.

Regrettably, the *bass* and *treble* weighting factors in figure 2.11 are not restricted to the *bass* and *treble* sides of the electrode array. The weightings are not zero but negative at the opposite side. This is related to the *shift* component. Positive weighting on one side and zero weighting on the other side would imply a change in overall level when *bass* or *treble* is changed, but the overall level is incorporated in the *shift* component. The weightings of figure 2.11 show, for example, that if the *bass* value is increased then low-frequency stimulation at the apical electrodes (high electrode numbers) will increase but high-frequency stimulation at the basal electrodes will decrease a little. This may be perceived as a disadvantage of the present approach but, in practice, the level changes at the opposite side are small. For the *bass* parameter we find that its contribution to the T-level of electrode 22 stays within ±10.3 CUs for 80% of the implant recipients. The largest co-variant contribution at the *treble* side is only 3.1 CUs for electrode 9. For C-levels these values are 11.0 CUs for electrode 22 and 2.9 CUs for electrode 9, respectively. For the *treble* parameter and T-levels these values are 13.1 CUs for electrode 3 and 2.9 CUs for electrode 17; for C-levels 11.9 CUs for electrode 3 and 3.1 CUs for electrode 17. Thus, this side effect is limited to about 3 CUs for the majority of the implant recipients.

We may conclude that the three parameters *shift*, *bass* and *treble* can be used for profile adjustment. They are related to easily identifiable aspects of sounds so that adjustment can be performed effectively in interaction with the implant recipient. Moreover, this simplified approach (rather than adjusting many electrodes individually) opens the way to implant recipients fitting the processor themselves. The weighting factors for the three parameters *shift*, *bass* and *treble* are presented in Table 2.2.

The *shift*, *bass*, and *treble* parameters can be used for the three-electrode approach described in section 2.3.5. Calculation of the values of these parameters, once T- or C-levels are measured for electrodes 22, 12, and 3 or electrodes 21, 12, and 4, is presented below:

For T-levels measured for electrodes 22, 12, and 3:

$$\begin{aligned} V_{\text{shift}} &= +0.187*L_{22} + 0.717*L_{12} + 0.119*L_{3} \\ V_{\text{bass}} &= +0.799*L_{22} - 0.789*L_{12} + 0.026*L_{3} \\ V_{\text{treble}} &= -0.143*L_{22} - 0.671*L_{12} + 0.894*L_{3} \end{aligned}$$

For C-levels measured for electrodes 22, 12, and 3:

$$\begin{aligned} V_{\text{shift}} &= +0.165*L_{22} + 0.742*L_{12} + 0.121*L_{3} \\ V_{\text{bass}} &= +0.840*L_{22} - 0.764*L_{12} - 0.008*L_{3} \\ V_{\text{treble}} &= -0.078*L_{22} - 0.753*L_{12} + 0.888*L_{3} \end{aligned}$$

For T-levels measured for electrodes 21, 12, and 4:

$$\begin{aligned} V_{\text{shift}} &= +0.250*L_{21} + 0.620*L_{12} + 0.158*L_{4} \\ V_{\text{bass}} &= +1.015*L_{21} - 1.063*L_{12} + 0.104*L_{4} \\ V_{\text{treble}} &= -0.086*L_{21} - 0.887*L_{12} + 1.050*L_{4} \end{aligned}$$

For C-levels measured for electrodes 21, 12, and 4:

V_{shift} = $+ 0.204*L_{21} + 0.674*L_{12} + 0.149*L_{4}$
V_{bass} = $+ 1.024*L_{21} - 0.968*L_{12} + 0.022*L_{4}$
V_{treble} = $- 0.076*L_{21} - 0.930*L_{12} + 1.048*L_{4}$

Electrode	22	21	20	19	18	17	16	15	14	13
T average	140.1	140.6	141.0	141.2	141.8	142.3	142.6	143.0	143.4	144.3
T shift	0.953	0.949	0.949	0.954	0.968	0.976	0.983	0.991	0.992	1.000
T bass	1.000	0.733	0.543	0.363	0.263	0.166	0.092	0.018	-0.076	-0.154
T treble	-0.156	-0.216	-0.209	-0.217	-0.217	-0.221	-0.210	-0.182	-0.207	-0.163
C average	175.8	177.4	178.7	179.6	180.9	181.8	182.6	183.2	183.8	184.6
C shift	0.914	0.920	0.944	0.967	0.977	0.991	0.996	0.996	1.000	0.991
C bass	1.000	0.782	0.577	0.371	0.272	0.178	0.075	-0.021	-0.104	-0.165
C treble	-0.115	-0.147	-0.179	-0.211	-0.225	-0.265	-0.249	-0.243	-0.219	-0.165
Electrode	12	11	10	9	8	7	6	5	4	3
T average	144.9	145.0	145.4	145.7	146.0	146.2	146.1	145.2	144.3	143.7
T shift	0.996	0.989	0.991	0.998	0.995	0.993	0.970	0.954	0.919	0.900
T bass	-0.256	-0.282	-0.296	-0.305	-0.292	-0.263	-0.254	-0.244	-0.156	-0.032
T treble	-0.125	-0.038	0.058	0.140	0.259	0.346	0.444	0.598	0.830	1.000
C average	185.0	185.2	185.4	185.5	185.5	185.1	184.2	183.0	181.5	180.0
C shift	0.995	0.984	0.978	0.985	0.980	0.968	0.954	0.957	0.949	0.924
C bass	-0.209	-0.242	-0.264	-0.265	-0.261	-0.222	-0.217	-0.174	-0.128	-0.089
C treble	-0.137	-0.086	-0.015	0.064	0.169	0.315	0.461	0.608	0.822	1.000

Table 2.2. Average T- and C-levels and the *shift*, *bass* and *treble* weighting factors. The largest weighting factor of each parameter is set to -1.000 or + 1.000. Per electrode the levels of each profile are found by adding the values of the *shift*, *bass* and *treble* parameters, $V_{shift,}$, V_{bass}, and V_{treble}, multiplied by these weighting factors, to the levels of the average profile.

2.3.7 Relation between the profile parameters determined shortly after implantation and the most recent ones

Chapter 1 showed that overall levels of T- and C-profiles may change considerably over quite a period after implantation. The question arises whether or not the shape of the profile remains the same. If there is little change in profile shape then one may start a speech processor readjustment by simply changing overall level or, in terms of profile parameters, if the profile-shape related parameters shortly and late after implantation are highly correlated then readjustments can be started by simply shifting the profile followed by some trimming of the shape related parameters. The correlation coefficients were calculated between the first complete profile of each implant recipient, when there are no electrodes switched off or some individual levels set at very low values (*cf.*, Sec. 2.3.3), and the most recent profile for the *shift*, *tilt*, *curvature*, *bass*, and *treble* parameters. Table 2.3 presents the results. All correlations were modest. The correlations were more evenly divided between *bass* and *treble* than between *tilt* and *curvature*. Yet, none of these correlation coefficients assumed a value that would indicate that one does not need to readjust *tilt*, *curvature*, *bass*, and *treble* after one has acquired the first full profile; later readjustments of the shape of the profile may be necessary. The scatter diagrams of the shape-related parameters (not presented) showed an ordinary elliptic silhouette. Scatter diagrams for the *shift* parameter are presented in figure 2.12. Overall T-levels increased or decreased about evenly between the initial and last profile except for some ten implant recipients in whom the initial adjustment was markedly lower than the last one. Overall C-levels showed a systematic trend toward higher levels. This suggests that, particularly shortly after implantation, one may start a new adjustment by shifting the profile upward, succeeded by adjustment of the two profile-shape related parameters. The present results correspond well to the findings reported in Chapter 1.

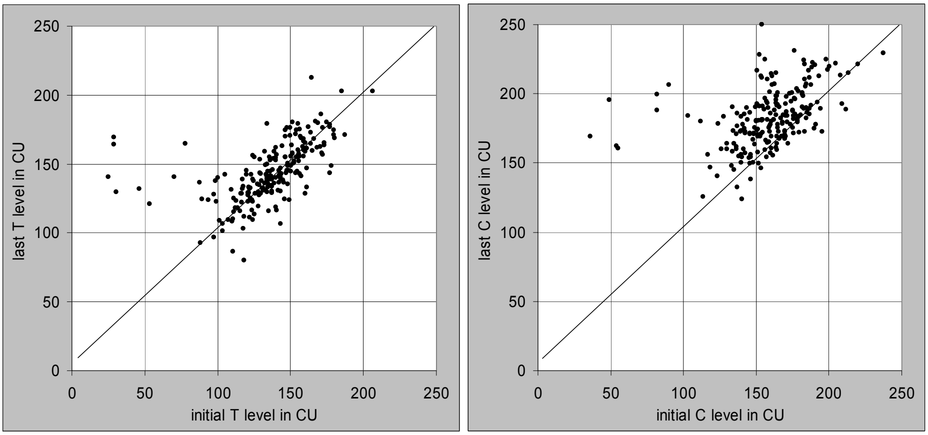

Figure 2.12. Scatter diagram of the overall levels measured shortly after implantation (abscissa) versus those of the most recent measurement. Note the systematic increase of overall C-level.

	shift	tilt	curvature	bass	treble
T-levels	0.55	0.56	0.42	0.49	0.50
C-levels	0.45	0.52	0.37	0.45	0.42

Table 2.3. Correlation between the values of the *shift*, *tilt*, *curvature*, *bass*, and *treble* parameters of the initial full profiles and the most recent profiles of 215 implant recipients.

2.3.8 Relation between T- and C- profiles

Once the principal parameters of T- and C-profiles are known it is interesting to examine whether or not there is a correlation between the values found for the T- and for the C-parameters. If, for example, the *bass* and *treble* values found for T-profiles are correlated with those found for C-profiles then one could decide to start adjusting the C-profile by simply shifting upward an already measured T-profile. Analysis of the correlation between the parameters of T- and C-profiles requires that these parameters are derived from the complete set of T- and C-profiles whereas, up to this point, the T- and C-profiles have been treated separately. Calculation of the principal components for the whole set of T- and C-profiles yielded essentially the same parameters *shift*, *tilt*, *curvature* and, after bi-quartimax rotation, *shift*, *bass*, and *treble*. The coefficients of the correlation between the T-and C-values of these parameters calculated across 215 implant recipients are presented in Table 2.4. The resulting coefficients, from 0.6 to 0.7, suggest that once the two parameters representing the shape of the profiles, *tilt* and *curvature* or *bass* and *treble*, are known for a T-profile of an individual they provide some indication of the shape to be expected for the C-profile of that individual and *vice versa*.

shift	tilt	curvature	Bass	treble
0.69	0.69	0.62	0.63	0.67

Table 2.4. Correlation between the values of the *shift*, *tilt*, *curvature*, *bass*, and *treble* parameters of the T- and C-profiles for 215 implant recipients.

2.4 Discussion

Measuring the threshold and comfortable loudness levels for 20 electrodes is a time consuming task. It takes much time of an implant team, particularly when the speech processor has to be readjusted frequently (see Chapter 1). In clinical practice the fitting procedure is often accelerated by using standard interpolation techniques to estimate T- and C-levels or by measuring T- and C-levels for a group of electrodes simultaneously. In this chapter we examined to what extent a statistical analysis of many T- and C-profiles, collected over many years, may contribute to a strategic approach to speech processor fitting. An important consideration in this approach was that the T- and C-levels across the electrode array are highly correlated. The lowest correlation coefficient, found between electrodes 22 and 3 located at the two borders of the electrode array, is 0.78 for T-levels and 0.81 for C-levels. This strong correlation suggests that the number of independent variables, and thus the number of measurements that have to be made, is considerably less than 20. Principal components analysis and the more elaborate statistical technique of factor analysis were used to find the independent factors governing the T- and C-profiles.

The results showed that only three independent factors determine the T- and C-profiles: *shift*, - an overall level shift nearly equal for all electrodes of the array, *tilt*, - a change in the slope of the level profile across the array and *curvature*, - a change from concave to convex (from peak- to valley-shaped) profiles. If these profile parameters are incorporated in fitting software in terms of *shift*, *tilt*, and *curvature* controls then one can, by varying these three controls, approach each profile measured psychophysically for individual electrodes with high accuracy. The mismatch error, in terms of the standard deviation of the differences between the original psychophysical measurements and the three parameter approach was only 2-3 CUs for T-levels and about 2 CUs for C-levels. This result was found for profiles measured 6 to 50 months after implantation. Profiles measured shortly after implantation showed a larger error. This implies that the levels adjusted shortly after implantation were more irregular across the electrode array. The finding that they became more regular with successive readjustments suggests that the early adjustments were *too* irregular and that the three-parameter based adjustment might speed up the processor fitting procedure after implantation.

A disadvantage of the initial approach was that the two parameters governing the shape of the profiles, *tilt* and *curvature*, are not related to easily identifiable aspects of sounds. It will be very difficult to find appropriate settings of *tilt* and *curvature* controls by listening to sounds while these controls are manipulated. However, the result could be used in another way. The finding that the T- and C-profiles are governed by only three factors indicates that it suffices to measure T- and C-levels for three electrodes remotely located from one another, after which the three parameter values can be calculated and the T- and C-levels for the other electrodes can be estimated. In the present analysis we followed the approach by way of principal components but essentially regression between levels not measured on those measured is used.

The analysis was continued in order to find parameters that are better related to certain aspects of sounds. In order to optimise this relation, which is of great practical importance, a 25% increase of the mismatch error was allowed (from 2.8 to 3.5 CUs for T-levels and from 2.3 to 2.9 CUs for C-levels). The analysis resulted in two appealing parameters determining the shape of the T- and C-profiles, *bass* and *treble*, and the former *shift* parameter. *Bass* and *treble* controls might be very helpful in adjusting C-levels. Implant recipients may be well able to indicate when sounds become too “boomy” or too “sharp”. Moreover, *bass* and *treble* are controls that could be used by the implant recipients themselves. Adjustment of T-profiles using these parameters is less easy to propose. An implant recipient with little hearing experience will have problems indicating when low- or high-frequency components of sounds become audible. The correlation of about 0.65 between the T- and C-values for *bass* and *treble* suggests that one could use the C-values to set the shape of a T-profile and successively perform a simple threshold measurement using the *shift* control. Alternatively, one could measure the T-levels for two marginally located electrodes and one located in the middle of the array and then estimate the other levels as described above. However, one should bear in mind that adjustment of T-levels is not critical. This becomes evident from the erratic time course of T-level readjustments after implantation. Although the *bass* and *treble* controls may be less suited for threshold measurements, they may be very helpful readjusting T-profiles in relation to speech perception at low sound levels. When there are many fricative misperceptions at low speech levels the *treble* control may provide a practical means to find the best high frequency T-level setting. Likewise, vowel misperceptions indicating weak low formants could be easily corrected with the *bass* controls. The above examples demonstrate how profile parameter control could contribute to effective speech processor fitting. The relative roles of individual electrode level adjustment and profile adjustment should evolve in clinical practice. An important aspect of profile parameter control will be that it enables readjustment of the speech processor in direct relation to speech sounds or other natural sounds, rather than using laboratory type beeps, which are required when adjusting the levels for individual electrodes. This will speed up the fitting procedure, particularly in children.

Another interesting aspect of the profile parameter approach was the question of whether or not the values of these parameters demonstrate some degree of constancy during speech processor readjustments. Analyses not presented here showed that while overall level, and thus the *shift* parameter followed an exponential growth curve (see Chapter 1), the time course of the two profile-shape related parameters was rather erratic. The presentation in this chapter was limited to a comparison between the parameter values of the first complete T- and C-profiles measured shortly after implantation and the most recent ones. The correlation coefficients were somewhat disappointing. For the *bass* and *treble* parameters they were limited to values between 0.4 and 0.5.

Chapter 3

Thresholds of electrically evoked compound action potentials; relation to T- and C-levels

Summary

Objective

In principle, it should be possible to use the objective measures of electrically evoked compound action potentials (ECAPs) as an alternative to the more time- and resource-consuming perceptually-based measures typically used to program speech processors. However, when introducing these objective measures to speech processor programming, it is important to understand the relation between ECAP thresholds and psychophysically measured T- and C-levels. In this chapter, this relation is analysed according to principal components (*e.g.*, *shift*, *tilt*, and *curvature* – see Chapter 2) that characterize the overall level and shape of the ECAP thresholds and T- and C-levels across the electrode array. The principal components are used to find mutually related aspects of these objective and subjective measures.

Conclusions

Principal components analysis of the profiles of ECAP thresholds across the electrode array showed that the ECAP components differed from those of the perceptually based T- and C-level profiles. Moreover, the three major ECAP components explained less of the variance in the ECAP thresholds than the three major T- and C-components did in the T- and C-levels. ECAP thresholds, mostly measured during surgery, were found to be more variable across the electrode array than T- and C-profiles. This was partly due to the ECAP measurement technique. When masker levels were presented at 10 current units (CUs) above probe level, ECAP profiles were less variable than when masker and probe were at the same level. When ECAP thresholds are measured during surgery, electrode impedance priming, by stimulating the electrodes before starting ECAP measurements, may improve ECAP measurement accuracy. Stimulation rate affected ECAP thresholds and T-levels differently. As the stimulation rate was increased from 80 to 250 Hz, ECAP thresholds increased by 5 CU. T-levels decreased by 10 CU when the rate was increased from 250 Hz to 720-1200 Hz. These rate effects should be considered when comparing ECAP thresholds to T-levels.

The comparison between the profiles of the ECAP thresholds and those of the T- and C-levels was based on the principal components of the T- and C-level profiles because the T- and C-levels are commonly used in speech processor programming. A modest correlation was found between the overall level (*i.e., shift*) of the ECAP profile and the overall level of the T-profile (R=0.48) and between the overall level of the ECAP profile and the C-profile (R=0.36). A modest correlation was also found between the slope (*i.e., tilt)* of the ECAP profile and the C-profile (R=0.47). For other components (*i.e.*, *curvature, bass,* and *treble*) correlation coefficients between ECAP profiles and T- and C-profiles were small. In addition to these modest correlations, inter-subject differences in overall T- and C-levels were about 1.7 times larger than those observed with ECAP thresholds.

Finally, previous experiments have shown that speech perception performance was similar between speech processors programmed using ECAP thresholds or conventionally programmed processors using T- and C-levels. The two methods of programming the speech processor may not to be critical to speech perception (in quiet listening conditions), they may be important to other listening tasks such as speech in noise and music perception.

3.1 Introduction

The streamlined fitting procedure described in Chapter 2 proposed speech processor programming using T- and C-profile parameters. The analysis in Chapter 2 showed that it is possible to match conventionally measured T- and C-levels using only three profile parameters: *shift, tilt* and *curvature*. Subsequently, it was found that two other parameters, *i.e. bass* and *treble*, combined with the *shift* parameter, also could provide good matches to the conventionally measured T- and C-levels. *Bass* and *treble* represent more familiar aspects of sounds than *tilt* and *curvature* and will therefore be more effective in a fitting procedure conducted interactively between implant recipient and clinician. However, while these profile parameters can provide good matches, it is possible that the result of profile adjustment deviates from the conventionally measured T- and C-levels. Therefore, one may conclude that profile parameters may be better suited for later speech processor adjustments (particularly for C-levels), after conventionally measured T- and C-levels have been established. Nevertheless, Chapter 2 showed that the profile parameters can be used to streamline the initial stage of the fitting procedure in another fashion. After psychophysically measuring T- and C-levels for one apical, one basal and one middle electrode one can estimate the other T- and C-levels to be expected using the profile parameters.

Another way of streamlining the fitting procedure would be to obtain an initial stimulation profile derived from electrically evoked compound action potential (ECAP) thresholds. ECAP thresholds are objectively measured, requiring no input from the patient. This approach has the important advantage that one can start programming the speech processor immediately with live speech, which would be particularly helpful when fitting children. In this fitting procedure, the clinician would globally increase the stimulation levels from the ECAP threshold profile until obtaining hearing threshold (the alternative T profile) and then further increasing until obtaining a comfortably loud listening level (the alternative C profile). Subsequently, *bass* and *treble* could be adjusted to comfortable loudness. In addition, the clinician may choose to trim the levels of individual electrodes by loudness balancing the T- and C-levels across the electrode array. This chapter examines the relation between the profiles of ECAP thresholds and conventionally measured T- and C-levels to assess how well these profiles match.

3.2 Methods and materials

3.2.1 Methods

ECAP thresholds were measured using the Neural Response Telemetry software provided by Cochlear (NRT^{TM} , V2.04). Stimulation and recording were performed using standard procedures: only one electrode at a time was stimulated using the external ball electrode as the reference electrode (MP1 mode). The neural response was recorded from an electrode located two positions away from the stimulus electrode in the apical direction using the implant housing as the reference electrode (MP2 mode). When stimulating electrodes 21 and 22 the recording electrode was located two electrodes away in the basal direction. The standard masking procedure was used, with a masker-probe interval of 500 µs, sampling delay of 50 µs, and amplifier gain of 60 dB. If amplifier saturation occurred, the delay was increased stepwise to 70, 90, 120 or 140 µs until a satisfactory response was obtained. If increasing the delay to 140 µs did not remove amplifier saturation, the gain was reduced to 40 dB. Impulse duration was set at 25 µs per phase. Stimulation rate was 80 or 250 Hz. Masker level was set at 10 current units (CUs) above the probe level, or equal to the probe level. Before measuring ECAPs during surgery there was no “priming” of the electrodes by electrical stimulation in order to decrease electrode impedance. The effects of stimulation rate and masker level were subjected to statistical analysis.

ECAP thresholds were not calculated according to standard procedure. In the standard procedure the linear part of the ECAP amplitude growth function is extrapolated to zero amplitude. However, determination of a linear segment in a non-linear growth function can be too subjective and, hence, extrapolation of this linear segment to zero is prone to ill-defined results. In the modified procedure, the ECAP threshold was determined by interpolation of the amplitude growth function, defining the threshold as the stimulation level at which an amplitude of 40 µV was obtained. In addition to a better defined threshold this approach has the advantage that the 40-µV threshold is closer to the visual appearance of the response. Note that the value of 40 µV depends on the noise level of the system and is lower for the newer implant devices.

The ECAP profiles were characterized according to principal components analysis (PCA, see Chapter 2). ECAP threshold profiles were compared to T- and C-profiles in terms of principal components. Statistical analyses were performed using the Statistica software package (Statsoft Inc., release 7.1).

3.2.2 Materials

ECAP thresholds for electrodes 3-22 were collected in 118 adults and children. All implants were Nucleus® devices: CI24M (N=12), CI24(CS) (N=88), CI24R(CA) (N=10) and CI24R(ST) (N=8). Most thresholds (N=106) were collected immediately after surgery, with the implant recipients still under anaesthesia; for 12 subjects thresholds were collected later while the subject was awake. The effects of device type, stimulation rate and masker level on ECAP thresholds were determined for the group of 118 subjects. For most of these subjects (N=99, 90 collected immediately after surgery), ECAP thresholds were compared to T- and C-profiles. The post-implantation period for the remaining 19 subjects was too short for T- and C-levels to stabilise. The most recent T- and C-levels were used for comparison. The ECAP stimulation rates were 80 or 250 Hz; the T-C rates 250 or 720-1200 Hz, corresponding to the rates used in SPEAK and ACE strategies. The number of subjects with the combination of ECAP/T-C rates at 80/250 was 19, the number was 45 for 80/720-1200, 2 for 250/250, and 33 for 250/720-1200. Further details of the T and C measurements can be found in Chapter 2.

3.3 Results

3.3.1 Principal components of the ECAP thresholds

Principal components analysis of the ECAP thresholds showed that three components accounted for less variance in these thresholds than the variance in the T- and C-levels that was accounted for by the three T- and C-components in Chapter 2 (90.0% for ECAP; 92.8% for T-levels; 93.7% for C-levels). The weighting factors of the three ECAP components are presented in figure 3.1. They show the familiar *shift* component, a *tilt* compo-

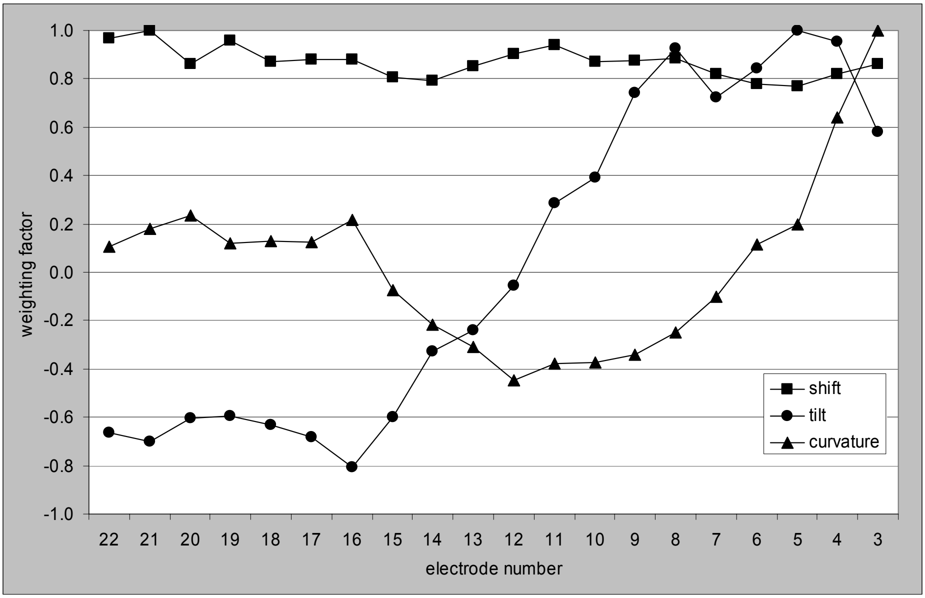

Figure 3.1. Weighting factors for the *shift, tilt* and *curvature* components of ECAP thresholds. The largest weighting factor for each of the three parameters was set to 1.0.

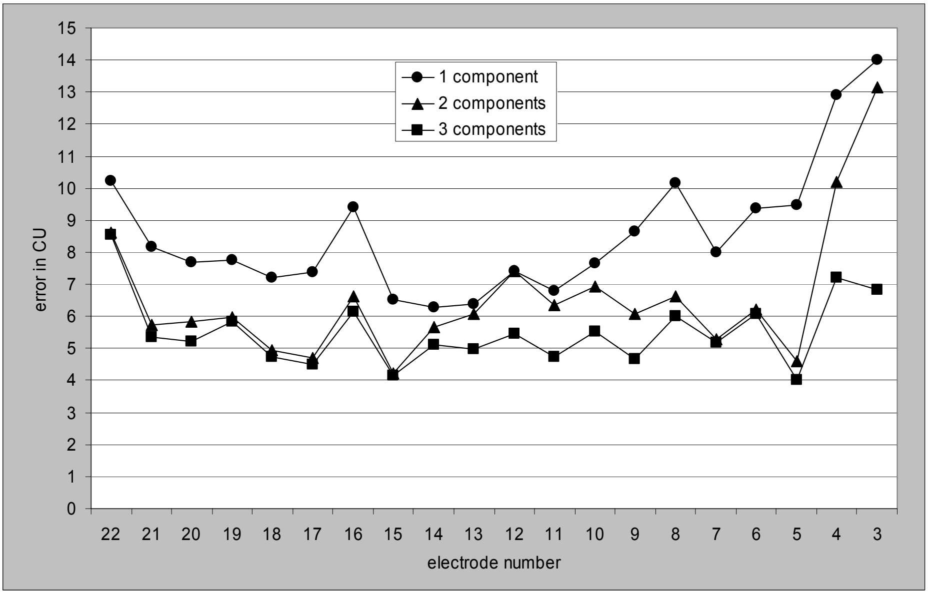

Figure 3.2. Mismatch error in current units between ECAP thresholds and the best matching profiles using 1 component (*shift*), 2 components (*shift* + *tilt*), or 3 components (*shift* + *tilt* + *curvature*) with weighting factors shown in figure 3.1. The error is expressed as the standard deviation of differences per electrode, calculated across 118 implant recipients.

nent with rather flat segments for the basal and apical regions of the electrode array, with a transition in the middle of the array, and a *curvature* component with a flat segment for the apical region of the electrode array, and very high weightings for the basally located electrodes 3 and 4. The latter finding implies that ECAP thresholds for electrodes 3 and 4 differed greatly across implant recipients. The ECAP weightings are markedly different from those found for the perceptually observed T- and C-levels (figure 2.10).

The finding that the amount of variance explained by three components was smaller for the ECAP thresholds than for the T- and C-levels is reflected in the mismatch between the original measurements and the profiles calculated for each implant recipient from the *shift, tilt,* and *curvature* values. The result is presented in figure 3.2. The mismatch error was substantially larger for the ECAP thresholds than for the T- and C-levels (*cf.,* figures 2.5 and 2.6). For three components and most of the electrode range the error was 5-6 CUs in ECAP thresholds whereas it was only 2-3 CUs in T- and C-levels. The ECAP error increased to 7-8 CUs at the margins of the array. The total error in the three component match, across the whole array, was 5.6 CUs for ECAP thresholds while it was 2.8 CUs for T-levels and 2.3 CUs for C-levels. This result suggests that the ECAP profiles may be more variable across the electrode array than the T- and C-profiles. Also, the different weighting factors suggest that the shapes of the ECAP profiles may not be closely related to those of T- and C-profiles.

3.3.2 Effect of type of implant device, stimulation rate, and masker level on ECAP threshold

A complete three-factor analysis of variance (ANOVA) could not be performed because ECAP thresholds were not measured for all experimental conditions in all subjects. For example, ECAP thresholds were not measured at a rate of 250 Hz with the CI24M device. Also, ECAP thresholds collected at 250 Hz were measured only for the 10 dB masker/probe amplitude offset. In addition, ECAP thresholds were measured only at 250 Hz with the CI24R(CA) device. An ANOVA was first applied to ECAP thresholds collected at the rate of 80 Hz, excluding the CI24R(CA) device. A repeated measures ANOVA (with the data from the 20 electrodes as repeated measure) showed no significant main effect for *masker amplitude offset*, and no interactions between *masker offset* and *electrode number* or *type of implant device*. There was also no significant main effect for *type of implant device*; however, there was a significant interaction between *device type* and *electrode number* ($p<0.001$). The ECAP thresholds for the basal electrodes in the CI24M implant (N=12) were higher than those for the CI24R(CS) device (N=52), as confirmed by Tukey HSD post-hoc analysis of ECAP thresholds per electrode (figure 3.3). The lower basal thresholds found for the CI24R(CS) device are most probably related to the smaller distance between the basal electrodes of this device and the cochlear modiolus. There was no significant difference in thresholds between the CI24R(ST) device and the CI24M, CI24R(CS) devices, possibly because the number of CI24R(ST) measurements at 80 Hz rate was limited to 7. However, no difference is expected between the CI24R(ST) and CI24M devices because both implants use the same stimulus processing scheme and the same electrode array. Any difference that might be present will be related to selection of implant candidates rather than the type of implant device because the CI24R(ST) device is used when one encounters problems inserting the Contour electrode of the CI24R(CS). An ANOVA for the 250 Hz rate condition showed that ECAP thresholds were not significantly different between the CI24R(CS) and CI24R(CA) devices. An ANOVA of the three principal components showed similar results. In particular, ANOVA

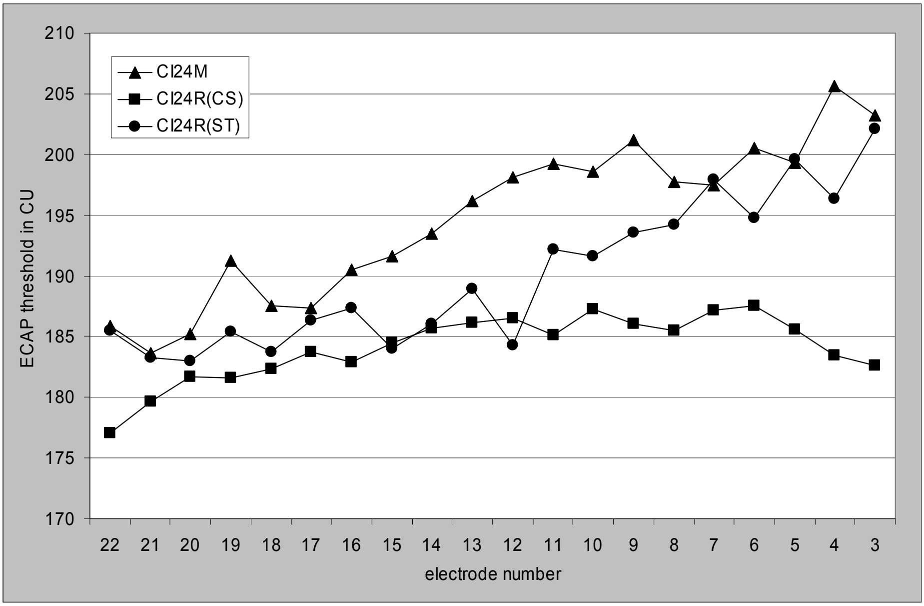

Figure 3.3. Average 40 µV ECAP threshold in current units measured at an impulse repetition rate of 80 Hz for 12 CI24M, 52 CI24R(CS) and 7 CI24R(CS) devices.

showed a significant effect for device type on profile *tilt* (p=0.03), which corresponded to the ECAP thresholds for the basal electrodes found to be higher for the CI24M than for the CI24R(CS) device.

The effect of stimulation rate was tested for the CI24R(CS) and CI24R(ST) devices. An ANOVA (with the 20 electrodes as repeated measure) showed that ECAP thresholds were not significantly affected by *stimulation rate*, and that there was no significant interaction between *stimulation rate* and *electrode number*. However, restricting the analysis to the CI24R(CS) device (N=88), there was a trend toward higher ECAP thresholds at 250 Hz than at 80 Hz (+ 5.7 CUs on average; p=0.08). The *shift* profile parameter was similarly affected by *stimulation rate* but *tilt* and *curvature* were not affected by *rate*. In six implant recipients ECAP thresholds were measured at both 80 and 250 Hz. A within-subject comparison showed a significant increase in ECAP thresholds (+3.4 CUs on average) when the rate was increased from 80 Hz to 250 Hz (p=0.03). While not a definitive sampling, the results suggest that ECAP thresholds may be elevated by about 5 CUs when increasing the stimulation rate from 80 Hz to 250 Hz.

3.3.3 Effect of masker offset level on ECAP profile smoothness

In the previous section ANOVA results showed that masker amplitude offset did not significantly affect overall ECAP thresholds. However, further analysis showed a clear effect of masker offset on the variability (or, *vice versa*, the smoothness) of the thresholds across the electrode array. Maskers presented at 10 CUs above probe level yielded more regular profiles than maskers presented at probe level (0 CU offset). An ANOVA of the three-component mismatch error, calculated per subject across the 20 electrodes, showed

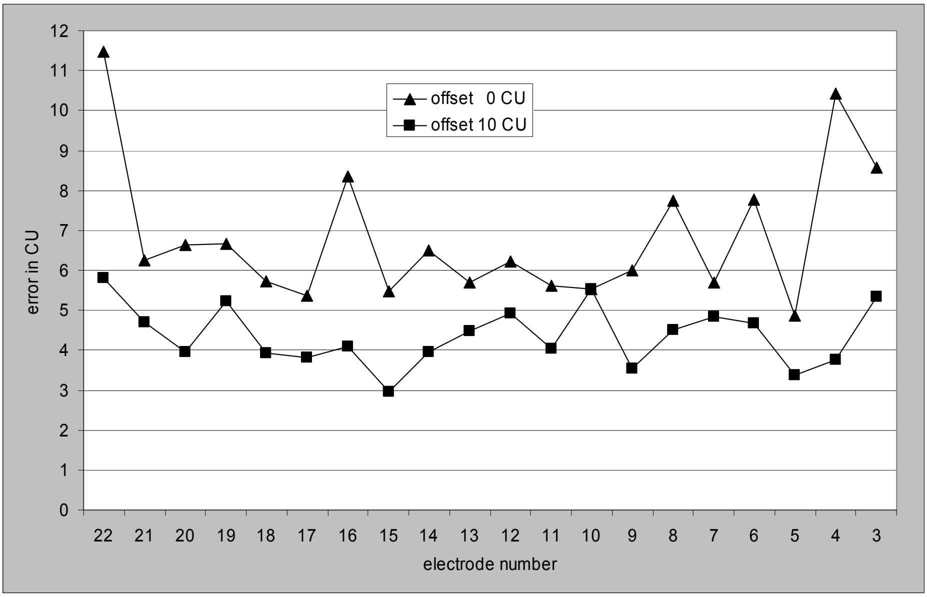

Figure 3.4. Mismatch error in current units between the ECAP thresholds measured with masker offset levels at 10 CUs or 0 CU above probe level and the best matching profiles based on 3 principal components. Error is expressed as the standard deviation of the differences.

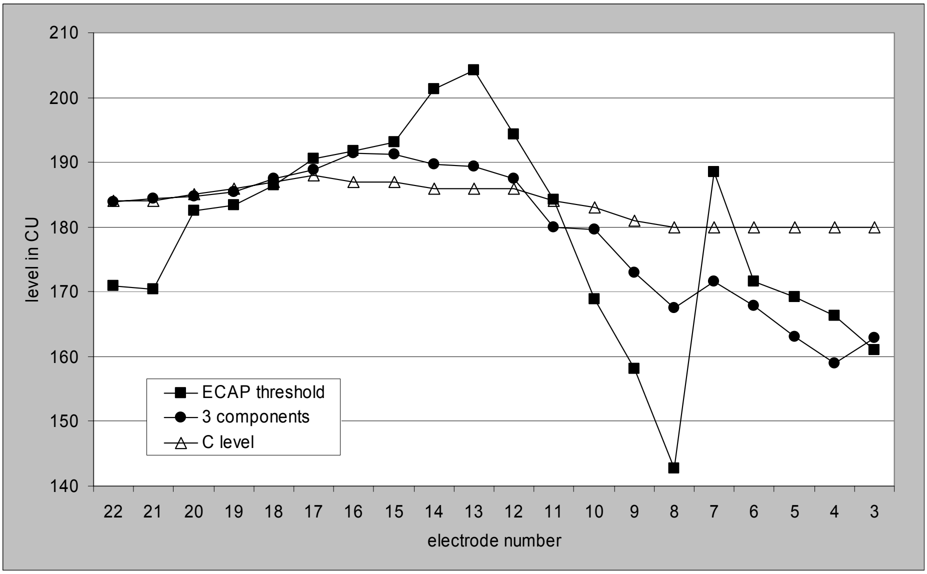

Figure 3.5. Comparison between an irregular ECAP threshold profile measured with 0 CU masker offset, the smoother 3-component profile and the conventionally measured C-level profile.

a significant difference between the 10 CUs (mean error 4.4 CUs) and the 0 CU (mean error 7.0 CUs) masker offsets (p=0.02). Note that the effect of masker offset was analysed for the 80 Hz condition only. However, an ANOVA showed no significant differences in the mismatch errors between the 80 Hz and 250 Hz conditions when measured with a 10 CU masker offset. Moreover, for the 80 Hz rate condition there was no significant effect for *device type* between the CI24M and CI24R(CS) implants on the mismatch error, and no significant interaction between *device type* and *masker offset*. These results strongly suggest that marked variability in ECAP thresholds between successive electrodes may be due to insufficient masking with the 0 CU masker offset. Thus, it is important to ensure that sufficient masking is used when measuring ECAP thresholds if the ECAP profile is to be used in speech processor fitting. The three-component mismatch error for the two masker offsets is shown in figure 3.4.

An example of an irregular ECAP threshold measured with the 0 CU masker offset, along with its three principal components match and the perceptually measured C-levels is presented in figure 3.5. The contours suggest that the three-component approach could be used to smooth the ECAP profile, and that a subsequent *treble* adjustment could yield a good match to the perceptually measured C-profile.

3.3.4 Comparison of ECAP thresholds to T- and C-levels

In order to fairly compare the profile parameters of the ECAP thresholds with those from conventionally measured T- and C-levels, it is necessary to establish a common set of principal components. However, figure 3.1 showed that the weighting factors for the ECAP principal components differ substantially from those for the T- and C-level components. Since T- and C-levels are commonly used for speech processor fitting, the T- and C-components were used for comparison. A larger mismatch error was expected between the measured ECAP thresholds and the best fitting profile when using these three components derived from the T- and C-levels than when using the three ECAP components. The mean errors for the ECAP components shown in figure 3.4 were 4.4 CUs for the 10 CUs offset and 7 CUs for the 0 CU offset. About half of the 99 implant recipients were tested in each masker offset condition; thus the mean error across offset conditions was about 5.8 CUs. The mismatch error between the measured ECAP thresholds and the profiles derived from the three principal T- and C-components *shift, tilt* and *curvature* was 6.2 CUs for the T-levels and 6.3 CUs for the C-levels. For the *shift, bass* and *treble* components these values were 7.4 CUs for the T-levels and 7.2 CUs for the C-levels. Taken together, the results suggest that the mismatch error was not much greater when the *shift, tilt* and *curvature* components derived from T- and C-levels, rather than the ECAP components, were used to fit the ECAP thresholds and that the mismatch error was somewhat greater when the *shift, bass* and *curvature* components were used.

The correlations between ECAP thresholds and T- and C-levels for all profile parameters and for *overall level* and *slope* of the profiles are shown in Table 3.1. In Chapter 2 the mean level of the T profiles measured with high rates (720-1200 Hz) was shown to be 9.6 CUs lower than that measured with a relatively low rate (250 Hz). As described above, ECAP thresholds measured at 250 Hz were found to be ~5.0 CUs higher than those measured at 80 Hz. If the mean values are adjusted for these rate effects, the correlation coefficients for *shift* and *overall level* increase from 0.42 to 0.48 for T-levels, and from 0.33 to 0.36 for C-levels. Note that the coefficients listed in Table 3.2 for *shift* and *overall level* were obtained after this rate adjustment. The correlation coefficients are modest.

parameter	shift	tilt	curvature	bass	treble	overall level	slope
T-levels	0.48	0.32	0.19	0.29	0.11	0.48	0.31
C-levels	0.36	0.47	0.20	0.29	0.15	0.36	0.46

Table 3.1. Coefficients of the correlation between ECAP thresholds and T- and C-levels for 7 parameters characterizing the T- or C-profiles. Coefficients for *shift* and *overall level* are calculated after level adjustments for different stimulation rates.

This was expected for the *shift* and *overall level* parameters in view of previous research. However, they are also modest for the parameters related to profile shape whereas it was hoped that these correlations might come out higher. The largest coefficient is 0.47 for the *tilt* parameter (and the related *slope* parameter) of the C-profiles. The standard deviation of the *tilt* values for the ECAP profiles (across 99 implant recipients) was about 1.4 times larger than that for the C-profiles. Thus, even though the correlation was modest and ECAP profiles may somewhat overestimate the range of *tilts* in C-profiles, this result suggests that the *tilt* in the ECAP profiles may be used to predict the C-profile *tilt*. With regard to *curvature*, *bass* and *treble,* the ECAP thresholds had little in common with the T- or C-profiles.

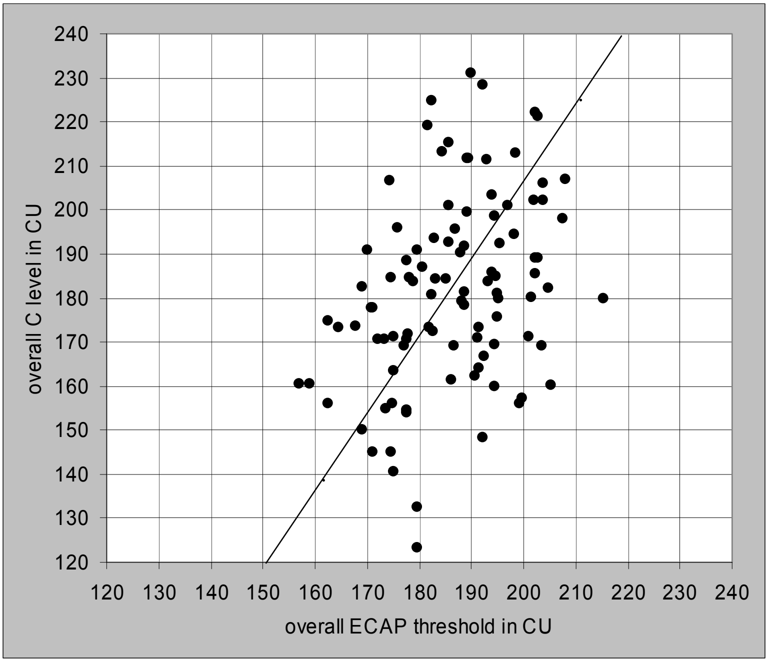

Figure 3.6. Average C-level across 20 electrodes versus the average 40 µV ECAP threshold at 80 Hz stimulation rate. The correlation coefficient is 0.36. The slope of the regression line, minimising the variance perpendicular to the regression, is 1.76.

The mean 40 μV ECAP threshold was calculated for the group of 99 implant recipients and 80 Hz stimulation rate. The thresholds measured at 250 Hz were adjusted by -5 CUs. After these adjustments the mean 80 Hz ECAP threshold was 186 CUs. The mean overall level of the T-levels, adjusted by -9.6 CUs when measured at a rate of 250 Hz rather than 720-1200 Hz, was 136 CUs and the mean overall C-level, independent of stimulation rate, was 182 CUs. It may be tempting to estimate the overall T- and C-levels from the overall ECAP levels, despite the only modest correlations reported above. In that case it should be noted that the standard deviation of the overall ECAP thresholds (after rate adjustments) was 12.3 CUs while it was 20.0 CUs for overall T-levels (after rate adjustments) and 21.6 CUs for overall C-levels. This implies, for example, that when a particular overall ECAP threshold differs by 10 CUs from the group average, the predicted overall T-level differs by 20.0/12.3 = 16.3 CUs from the group T-level average and the predicted overall C-level by 21.6/12.3 = 17.6 CUs from the group C-level average. The relation between the overall ECAP thresholds (adjusted to 80 Hz stimulation rate) and the overall C-levels of the most recent fitting session is shown in figure 3.6.

3.4 Discussion

It is desirable to include objective measures such as ECAP thresholds into the speech processor fitting procedure. Not only would objective measures reduce the time, effort and expense of speech processor programming, but they would also be particularly beneficial in programming the processors of children with little-to-no auditory experience. Until recently, psychophysically measured T- and C-levels are typically used for speech processor programming. Thus, it is important to know how ECAP thresholds relate to perceptually-based T- and C-levels. This was the principle objective of this chapter.

ECAP threshold levels can be compared to T- and C-levels for individual electrodes. However, previous analysis (Chapter 2) has shown that the levels across the array, the level profiles, can be described by only three parameters. One parameter, *shift*, is related to the average stimulation level across electrodes. The two other parameters are related to the shape of the profiles: *tilt* to the slope of the stimulation levels across electrodes and *curvature* to a peak or valley shaped profile. Alternatively, one can use the shape related parameters *bass*, the level of low-frequency (apical) stimulation and *treble*, the level of high-frequency (basal) stimulation. Using these parameters contributes to better understanding the relation between ECAP thresholds and T- and C-levels.

Considering first the ECAP measurements themselves, the results showed that there was a greater mismatch between measured ECAP thresholds and ECAP thresholds predicted by three ECAP profile components than observed between the measured T- and C-levels and their predictions. The greater mismatch implied that the ECAP thresholds across the electrode array are more irregular than the T- and C-levels across the array. The difference between masker and probe level appeared to significantly affect the irregularity; when the masker was presented 10 CUs above probe level ECAP profiles were smoother than when the masker and probe levels were the same. The accuracy of ECAP measurements during surgery may be improved by priming the electrode impedance, *i.e.*, by stimulating all electrodes before measuring ECAPs so that electrode impedance decreases. Also, it was found to be better to define threshold in terms of the stimulation level needed to obtain a certain response amplitude *(e.g.*, 40 μV), than by extrapolating a linear segment of the amplitude growth function to zero amplitude. Finally, ECAP thresholds were found to depend on stimulation rate. ECAP thresholds measured at 250 Hz were ~ 5 CUs higher than those measured at 80 Hz. An opposite effect was found for psychophysically measured T-levels, which were almost 10 CUs lower for high rates (720-1200 Hz) than for low rates (250 Hz). Thus, stimulation rates have to be included in comparisons of ECAP thresholds to T- and C-levels. These various findings suggest that ECAP thresholds must be very carefully measured if they are to be used in speech processor programming.

Principal components analysis showed that the weighting factors of the ECAP components differed substantially from those of the T- and C-components. This result suggested that the match between ECAP thresholds and T- and C-levels will not be perfect. At best, ECAP thresholds were only moderately correlated with T- and C-levels. After correcting for rate effects, the best correlations were found for the *shift* parameter and *overall level* (ECAP thresholds versus T-levels: R=0.48; ECAP thresholds versus C-levels: R=0.36). Along with these only modest correlations, the inter-subject variability in overall T-levels was 1.63 times greater than that in ECAP thresholds; the variability in C-levels was 1.76 times greater. With respect to the parameters related to stimulation profile shape (*tilt, curvature, bass,* and *treble*), only the *tilt* in ECAP thresholds was modestly correlated with the *tilt* in T-levels (R=0.32) and with the *tilt* in C-levels (R=0.47); the other profile parameters showed little correlation. Thus, the *tilt* of the ECAP profile may be useful in adjusting T- and C-levels. However, note that the deviation from average *tilt* in ECAP thresholds tends to overestimate the deviation from average *tilt* in C-levels by a factor of 1.4.

ECAP thresholds have been used to program the speech processor (Smoorenburg et al. 2002, Willeboer et al., 2006). For each patient, using live speech, the amplitude of the ECAP profile was increased from below hearing threshold until obtaining audibility and then further increased until obtaining a comfortably loud listening level. Speech performance using this adjustment procedure was virtually equivalent to that with conventionally programmed speech processors (*i.e.*, adjustment of each electrode to the psychophysically measured T- and C-levels). However, the T-levels using the ECAP-based procedure were much lower than conventionally measured T-levels, and often the shape of the ECAP profile differed substantially from the T- and C-profiles. This raises the question of how important accurate T- and C-levels are for implant recipients' speech performance. The results demonstrate that the overall level and shape of the T and C profiles may not be critical for speech perception (at least for speech in quiet environments). Accurate speech processor programming may be more important for difficult listening tasks (*e.g.*, speech in noisy environments, music perception, etc.). Exploration of the programming requirements for these difficult listening tasks may be facilitated by using the profile shape parameters when searching for the best processor adjustment.

3.5 Complementary reading

Smoorenburg, G.F., Willeboer, C. and Dijk, J.E. van (2002). Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds, Audiol. Neurootol. 7, 335-347.

Willeboer, C., Zanten, G.A. van, 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 and Hearing, 27(6), 789-798.

Chapter 4

Development of speech perception and language acquisition in children over time; comparison of performance measures

Summary

Objective

Performance of young cochlear implant recipients is measured primarily in terms of speech perception and language acquisition. These measures are used to quantify the results of cochlear implantation, for example in cost-benefit analysis, and to analyse the factors that affect these results, which may contribute to improving performance and performance prognosis. A variety of tests have been applied. The question arises whether or not we can reduce the number of tests to a small set covering the essential aspects of speech and language performance after cochlear implantation. This chapter presents the results of five tests applied over a period of four years after implantation. The presentation focuses on the development of performance and a comparison of the five tests.

Conclusions

Four years after implantation the average speech perception score in children, 5 to 8 years of age, counting the number of phonemes in CVC words responded correctly, was 55% in auditory-only (A) mode and 75% in auditory-visual (AV) mode. In children, congenitally deaf or deafened before the age of 2, these scores reach 50-95% in A mode and 65-100% in AV mode if they are implanted within two years of the onset of deafness.

The results for receptive language and speech production tests are less promising. Almost no child shows progression to normal after implantation. On the contrary, most children show an increase in the scores over the three year post-implantation period smaller than the increase in the standard score for normal-hearing children: 77% of the normal increase for the Peabody Picture Vocabulary Test and 47% for both the Reynell Verbal Comprehension Test and the Schlichting sentence production test.

Speech perception performance appears to be no good indicator of language acquisition. The phoneme scores collected at 6 and 12 months post implantation are not substantially correlated with the increase over time in the scores of the three receptive language and speech production tests (R = 0.18-0.36). Therefore implant recipient performance should be evaluated with at least a speech perception test and a receptive language test.

With respect to the language oriented tests the present evaluation points at the Reynell test as the best complement to a speech perception test, while the Schlichting speech production test could be viewed as a valuable addition to the Reynell test. The Erber Word Comprehension Test, also included in the present evaluation, appears to have limited applicability.

4.1 Introduction

Speech and language performance can be measured with tests ranging from a simple word reception task to complicated analyses of spoken language. One may simply ask a child to repeat a word. The child may repeat the word without knowing its meaning; speech reception without understanding. In this test the child's pronunciation of the repeated word may be poorly intelligible, degrading the accuracy of the score. Alternatively, one may ask the child to point at a picture related to the spoken word. This implies word comprehension. The child is unable to point at the picture when it does not know the meaning of the word but the intelligibility of the child's speech does not affect the score. The complexity of the test may be increased by including combinations of words (point at the red car rather than point at the car) or more complex instructions. Finally, one may switch from these receptive language tests to analysis of spoken language. In order to be able to conduct longitudinal studies over at least four years it is necessary to include different levels of complexity in a single test to avoid floor and ceiling effects. Higher scores in these tests will then reflect positive responses to more complex test items. The design of these tests is rather complicated: complex items should not frustrate children who cannot handle these items. The five tests included in this chapter cover the aspects mentioned above. They are described below in the Materials and Methods section. Three tests are versions of American-English tests translated into Dutch, two tests are developed in the Dutch language.

The present results are based upon a retrospective analysis of clinical data collected over about a ten year period. The analysis is not based upon a prospective study with a well balanced design. This approach has certain advantages and disadvantages. An advantage is that the analysis reflects the results of everyday clinical practice whereas in experimental settings performance may be different because of extra attention being paid to child guidance and testing procedures. A disadvantage is that the data sets are incomplete, which introduces a considerable risk of biased results due to confounded variables. Moreover, missing data may imply a selection because, for example, the examiner anticipated a 0 or 100% score and therefore did not conduct the test. In view of these disadvantages the present data sets have been inspected very carefully and, in case of doubt, results of analyses have been rejected.

4.2 Materials and methods

4.2.1 The five speech and language tests are:

(1) The Bosman-Smoorenburg CVC word reception test
Subjects have to repeat orally a word consisting of the sequence consonant-vowel-consonant (CVC). The test is derived from the standard audiological speech reception test in the Dutch language. Each list consists of 11 CVC words. Only those words of the original standard test repeated well by children are included in the present test. The responses are scored in terms of number of phonemes correct (33 in total). The test has been developed in Dutch. The Dutch language contains a large enough number of CVC words. Statistical analysis of results collected previously has shown that the redundancy in these words is small; on average the CVC words contain 2.4 independent phonemes per word. Hence, scoring phonemes correct rather than words correct improves the accuracy of the score. Oral repetition of the words implies that the score depends on the qual-

ity of the child's articulation. Thus, for previously deaf children the test is not just a speech reception test. Further, the child does not need to know the meaning of the word in order to be able to respond. However, knowing its meaning might affect the score. The test is administrated in auditory only, visual, and auditory-visual mode using recorded materials. With the youngest children live presentation was sometimes necessary in order to keep their attention. Presenting three lists takes about 5 minutes. Starting with a training list is recommended.

(2) The Erber word comprehension test
This test is a Dutch version of a test originally developed in the United States. Subjects have to respond to a word, spoken by the examiner, by pointing at one picture out of a set of 12 presented. The list consists of four types of words: monosyllables, trochees, spondees and three-syllable words. The response can be scored in terms of the type of word correct (1 of 4) or meaning correct (1 of 12). In our experience the score in terms of type of word correct was unreliable. Recognizing the type of word in terms of number of syllables and stressed syllable requires analytical power that appeared to make the test less suited for children less than 5 years of age. Therefore, the results presented here are limited to the score in terms of word comprehension. The total score is based upon 24 presentations. The test is administrated in auditory-only and combined auditory-visual mode. It requires rehearsal of the materials before commencing the test. The extent of rehearsal may affect the test result. This test does not depend on the child's articulation. Without rehearsal it takes about 10 minutes.

(3) The Peabody Picture Vocabulary Test
In this receptive language test subjects have to respond by pointing at one of four pictures in response to a word spoken by the examiner. The test focuses on vocabulary size with words of somewhat increasing complexity. It is administrated auditory-visually. The test is a Dutch version of a test originally developed in American-English. The response can be expressed in terms of equivalent age: the average age of a child with normal hearing yielding the same score. For children with normal hearing the test covers the age range from 2;05 to 8;01 years. A newer version of this test has been in use since 2004. The data presented here were acquired using the older version. The test takes 5 to 20 minutes, depending on the level of complexity reached when the test is terminated.

(4) The Reynell Verbal Comprehension Test
This test comes with a kit of objects. The child has to respond to instructions of increasing complexity using these objects. Thus, the test includes vocabulary and also sentence/language comprehension. The test is administrated auditory-visually. It is a Dutch version of a test originally developed in American-English: the Verbal Comprehension part of the Reynell Developmental Language Scales. The Dutch version has been validated independently of the original version. The test results can be expressed in equivalent age. For children with normal hearing the test covers the age range from 1;02 to 6;02 years. All scores pertain to a new version of the test in use since May 1997. Some data collected earlier were transformed to scores for the new test via the equivalent age parameter. The test takes 5 to 30 minutes, depending on the level of complexity reached when the test is terminated.

(5) *The Schlichting sentence production test*
In this test the examiner elicits responses of the child to single words encouraging the child to produce sentences. Their responses are analysed in terms of sentence develop-

ment: syntax and grammar. The test has been developed in the Dutch language. The results can be expressed in terms of equivalent age. For children with normal hearing the test covers the age range from 1;09 to 6;02 years. The test takes 5 to 30 minutes, depending on the level of complexity reached when the test is terminated. The test result can be affected by poor articulation.

4.2.2 Subjects and implants

The results presented are from about 100 children implanted between mid-1994 and mid-2002. The average age at implantation was 53 months with a standard deviation of 29 months (range 13-143 months). 61% of children was congenitally deaf, 12% had some useful hearing at the time of implantation, 27% deafened at 24 months on average with a standard deviation of 22 months (range 1-88 months). The average duration of deafness, from onset of deafness to the time of implantation, was 44 months with a standard deviation of 31 months (range 3-143 months).

All subjects received a Nucleus® implant: the CI22M with the Spectra™ processor, the CI24M with the Sprint™ processor, and the CI24R(CS) implant with the "Contour" electrode and the Sprint processor. Most children with the latter implant received the Esprit™ processor after the Sprint processor but not during the period analysed. Subsequent implantations using the CI24R(CA) implant with the "Contour Advance" electrode have been excluded from the analysis because most data did not cover the 3 to 4 year post-implantation period considered in this analysis. Moreover, this population of implant recipients differed from the previous set of subjects in the sense that we gradually widened the intake criterion from deaf to profoundly hearing impaired.

4.3 Results

4.3.1 Results from the Bosman-Smoorenburg CVC word reception test

Figure 4.1 shows the average scores for 13 subjects that were all tested 6, 12, 24, 36, and 48 months after implantation. The age at implantation of these children lies between 5 and 8;06 years. The auditory (A) and auditory visual scores (AV) clearly increased over time, even during the fourth year. The scores for speech reading alone (V) showed some increase over time.

To a first approximation the AV scores may be considered to be the combination of the A and V scores in probabilistic terms. In other words: the probability of incorrect AV scores equals the probability that both the A and the V scores were incorrect. At 48 months, for example, the average A score equalled 0.45 and the average V score equalled 0.48. Thus the probability of an incorrect AV score would be (1-0.45) * (1-0.48) = 0.29. Hence, the probability of a correct AV score becomes 0.71 while the average score was 0.73.

The difference between the average AV and A scores in figure 4.1 depended on the age at implantation. A separate analysis of the larger set of data, collected 12 months post implantation, showed that up to an age at implantation of 4 years this difference was much smaller than shown in figure 4.1; speech reading contributed less.

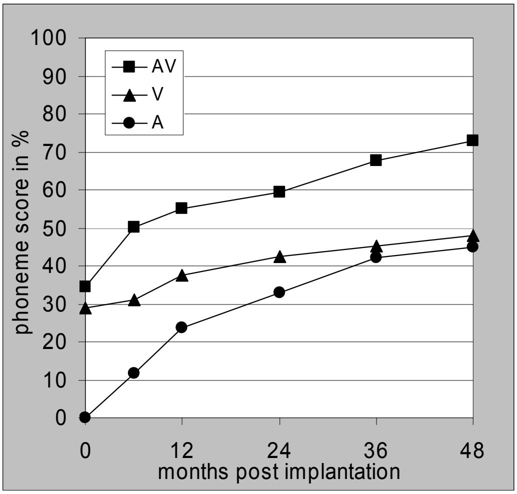

Figure 4.1. Average phoneme scores for CVC words presented in auditory-only (A), Visual (V), and auditory-visual (AV) mode to 13 CI recipients all tested 6, 12, 24, 36, and 48 months post implantation.

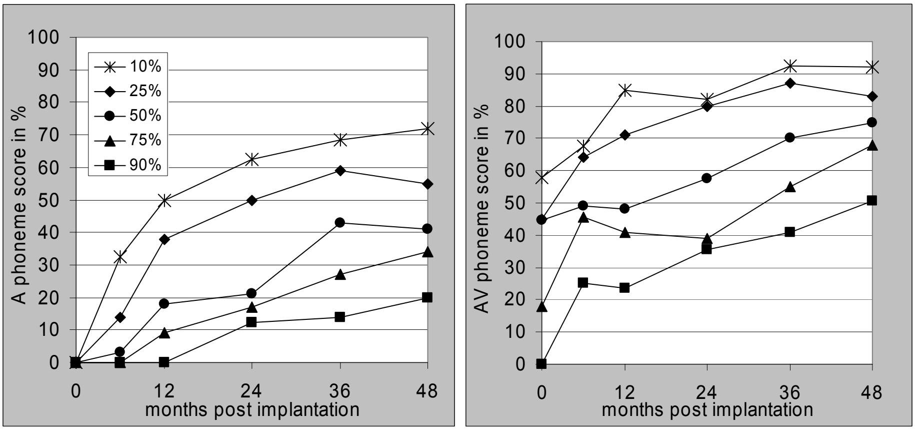

Figure 4.2. Distribution of post-implantation phoneme scores for CVC words presented in auditory-only mode (left panel) and auditory-visual-mode (right panel) across 13 CI recipients. The percentiles indicate the fraction of results with scores above the respective curves.

Clinically, the average scores are of limited importance. It is interesting to note how the scores were distributed across individuals. Therefore, figure 4.2 presents the 90th, 75th, 50th, 25th and 10th percentiles of the A and AV scores. These figures show that there were large interindividual differences. This result is of great concern to clinicians. Whereas after four years of implant use some subjects reached scores of 70% of phonemes correct with auditory presentation only, others reached only 20%. In 60 children who did not have data at all sample times we found, after four years, 10% of children with scores above 80% and 10% below 25%.

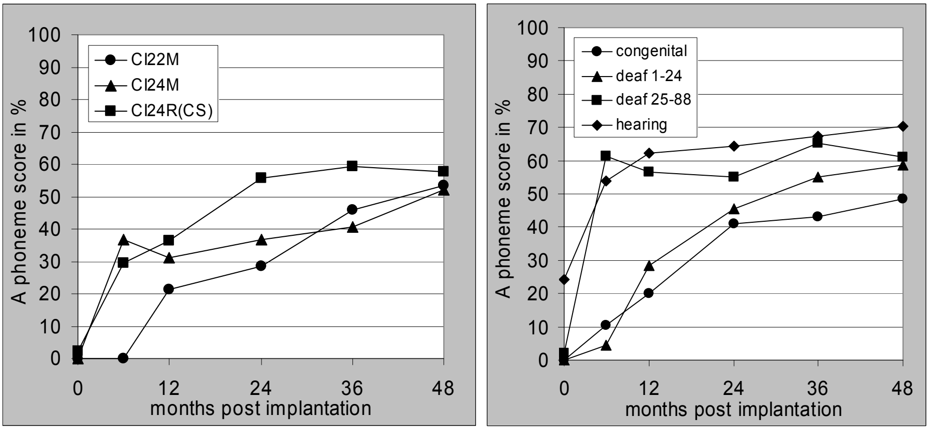

Figure 4.3, left panel. Post-implantation phoneme scores for CVC words presented in auditory-only mode with three types of Nucleus devices: CI22M (N=6), CI24M (N=25) and the Nucleus "Contour" CI24R(CS) (N=30). Right panel. Post-implantation phoneme scores for CVC words presented in auditory-only mode in four groups of implant recipients with different onsets of deafness: congenital (N=36), deafened between 1 and 24 months after birth (N=10), deafened between 25 and 88 months after birth (N=5) and those with some useful hearing at the time of implantation (N=12). Results for all subjects with at least four post-implantation data points.

The large interindividual differences may be due to different types of implanted devices, differences in age at onset of deafness and differences in duration of deafness, from its onset to the time of implantation. These factors are analysed below.

Figure 4.3, left panel, shows the average scores collected in auditory-only mode over the four year period for the three types of implant devices. The results of figure 4.3 are from data sets larger than those of figures 4.1 and 4.2. They were not necessarily complete over the 48 month period for each individual. The newer implants, the CI24M (N=25) and particularly the CI24R(CS) (N=30) show a more rapid increase of the phoneme scores for auditory-only presentation than the CI22M (N=6). However, after four years there was no significant difference between the scores for the three types of implants. The mean scores after four years for this larger group of 61 subjects were 55% in A mode and 77% in AV mode. The initial results for the CI22M may be related to the duration of deafness in the first group of 6 implant recipients; this duration being somewhat longer in the first group than in the two other groups.

With respect to the effect of age at onset of deafness four groups were distinguished: congenitally deaf (N=36), those who became deaf at age 0-2 years (N=10), deafened at age 2-7 years (N=5, 88 months being the highest onset age) and children with some useful hearing, still using hearing aids, at the time of implantation (N=12). The mean A scores for these four groups are presented in figure 4.3, right panel. After four years of implant use mean phoneme scores for the congenitally deaf reached 50% while the scores for those who had some useful acoustic hearing were 70%. The first two years showed the largest increase in scores. Note, however, that these scores might have been affected by articulation; articulation by children in the "hearing" and "deaf 25-88"group is likely to be better than that in the congenitally deaf group, particularly in the first years after implantation.

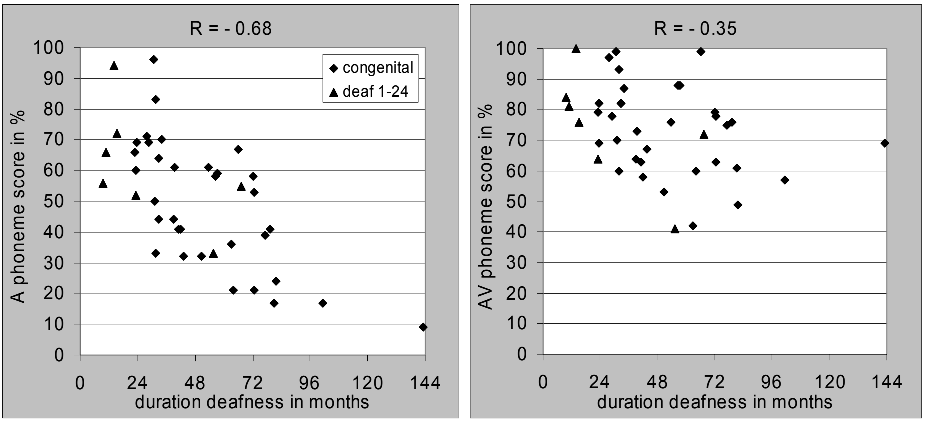

Figure 4.4. Phoneme scores four years after implantation for CVC words presented in auditory-only mode (left panel) and auditory-visual mode (right panel) as a function of the duration of deafness, from its onset to the time of implantation, for 33 congenitally deaf subjects and 7 subjects with onset of deafness between 1-24 months of age. The correlation coefficients, R, are presented on top.

The effect of duration of deafness, from its onset to the time of implantation, could not be analysed in general, because this analysis should also include the age at which the child became deaf. The data set was too restricted to perform a two-variable analysis. However, an interesting result was found when the analysis was limited to the large set of congenitally deaf children (N=33). Figure 4.4 shows the phoneme scores for the auditory and auditory-visual conditions, measured four years after implantation as a function of the duration of deafness (or equally the age of implantation) up to 12 years (144 months). Including the data for children (N=7) that became deaf at age 0-2 years did not change the overall picture (in terms of duration of deafness, not age of implantation). The figures show a clear correlation. The scores decrease markedly with increasing duration of deafness before implantation. These data are in line with results published elsewhere, most of which, however, covered a shorter period after implantation. The present data stress the importance of early implantation because, even four years after implantation, the effect of duration of deafness before implantation is strongly present. We may assume that this effect concerns primarily speech reception. After four years of implantation the dependence of these scores on articulation quality will have diminished.

The auditory-visual scores in figure 4.4 may be considered as being the combination of the A and V scores in probabilistic terms as was discussed in relation to figure 4.1. Calculating the probability of correct responses in AV mode from the A and V scores for 40 individuals and comparing these calculated values with the measured scores yielded a correlation coefficient of 0.78.

In our experience this word reception test can be conducted very well when examining young implant recipients. The range of scores is such that the test easily accommodates the development of word reception over a four year post-implantation period without risking floor and ceiling effects, both in auditory and auditory-visual mode. It was possible to conduct this test with most children, starting at age three to four, without serious limitations caused by poor articulation by the child.

4.3.2 Results from the Erber word comprehension test

Figure 4.5 shows the average scores for 38 subjects, all measured at 6, 12, 24, 36, and 48 months. The initial auditory-only (A) and auditory-visual scores (AV) were close to the phoneme scores presented above but after four years post-implantation one notes higher A and AV scores, 73 and 85% respectively.

Figure 4.6 presents the 90th, 75th, 50th, 25th and 10th percentiles of the A and AV scores. These figures show large interindividual differences, differences even larger than were found for the phoneme scores presented above. While after two years of implant use some subjects reached scores of 100% correct with auditory-only presentation, one notes that others stayed at 0% over 2 years and reached only 30% after four years, even in AV mode. The range of Erber scores was markedly larger than the range of phoneme scores. The data showed many more floor and ceiling scores. This large range of scores limits considerably the suitability of the Erber word comprehension test for longitudinal studies.

Another limitation of the Erber test emerged clearly when the effect of duration of deafness, from its onset to the time of implantation, was considered. Figure 4.7 shows the auditory-only data collected one and four years after implantation. Many scores collected one year after implantation (left panel of figure 4.7) in subjects with a period of up to 60 months between onset of deafness and implantation equalled zero. This did not occur in the phoneme scores of the previous test after one year (not shown). The present result suggests that poor word comprehension may have affected the scores of the Erber test substantially up to 5 years after implantation. Figure 4.7, right panel, shows the results four years after implantation. The zero scores have vanished suggesting that word comprehension has improved. However, ceiling effects now limit the applicability of the test. The 100% scores reduce the correlation coefficient (R = -0.51), which is lower than the one found for the phoneme scores presented in figure 4.4 (R = -0.68).

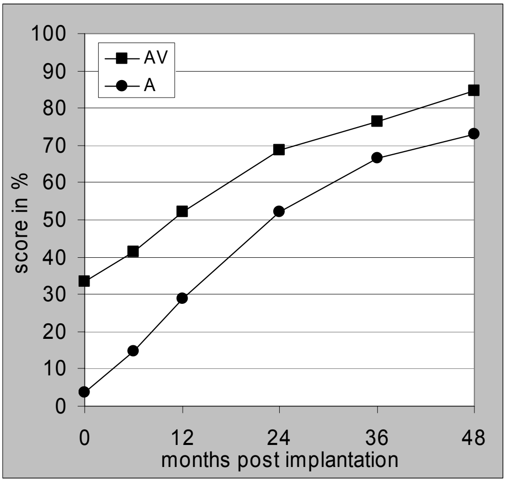

Figure 4.5. Auditory-only (A) and auditory-visual scores (AV) found for the Erber word comprehension test. Results from 38 subjects all tested 6, 12, 24, 36, and 48 months post implantation.

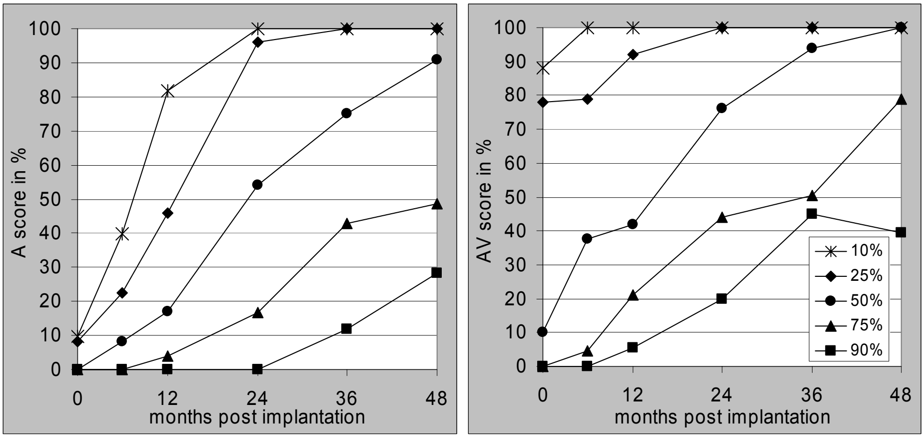

Figure 4.6. Distribution of post-implantation Erber scores for words presented in auditory-only mode (left panel) and auditory-visual-mode (right panel) across 38 subjects. The percentiles indicate the fraction of results with scores above the respective curves.

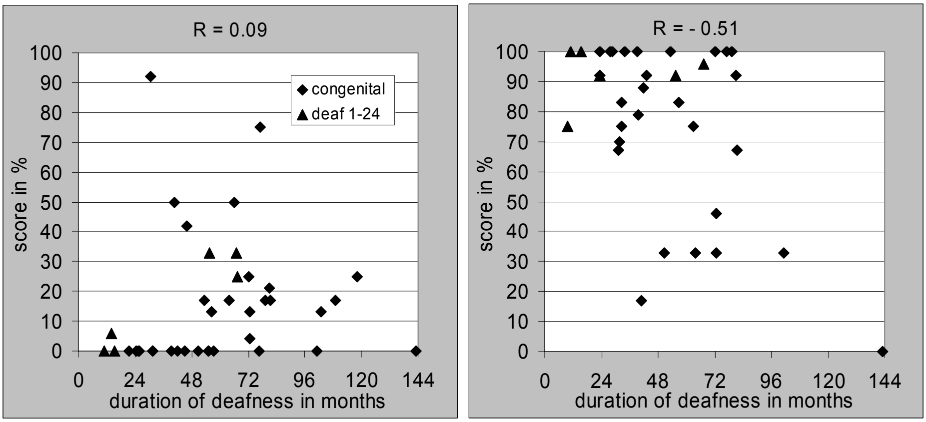

Figure 4.7. One year (left panel) and four year (right panel) post-implantation Erber scores for tests presented in auditory-only mode as a function of the duration of deafness, from its onset to the time of implantation. Data for 32 congenitally deaf subjects (left panel), 28 congenitally deaf subjects (right panel) and 6 subjects with onset of deafness between 1-24 months of age (both panels). The correlation coefficients, R, are presented on top.

The above results show that the Erber word comprehension test has a smaller application range than the CVC word reception test. It is less suited for assessing performance over periods of more than two years. For shorter periods it may offer somewhat better resolution but the present data are insufficient to perform an analysis concerning the question of whether or not accuracy, in a limited time period available for testing, is better with the Erber word comprehension test than with the CVC word reception test. Finally, the Erber test is less suited for children under 5 years of age because poor word comprehension may lead to floor effects (figure 4.7, left panel).

4.3.3 Results from the Peabody Picture Vocabulary Test

The Peabody Picture Vocabulary Test (PPVT) primarily measures receptive language. Since language acquisition is intimately coupled with age and individual development figure 4.8 presents the individual data for 27 children as a function of age. The children included are those with test results at all post-implantation intervals: 6, 12, 24, 36 months. The PPVT test offers the opportunity to convert the raw scores to equivalent age. However, this scale has a lower limit of 29 months and an upper limit of 97 months whereas the scores extend beyond these limits. Therefore, the raw scores are presented in figure 4.8 and the conversion curve is added to the data for comparison. In addition, the raw scores were averaged as a function of post implantation time interval. In view of the limitations in the conversion from raw scores to equivalent age, averaging the raw scores rather than averaging equivalent age was the preferred approach. Less raw than converted data suffered from floor and ceiling effects. Figure 4.8 shows that the raw scores of some individuals increased even faster than the standard scores. However, the curves of most individuals show slopes less steep than the slope of the standard scores. The averaged post-implantation scores increased by 27.8 units over a 36 months period whereas the standard score increases by about 36 units, a ratio of 0.77. Calculating the slopes of the individual curves from linear regression over the 36 months period yielded average slopes and standard deviations (the spread in the data, not the error in the mean) of 0.67 ± 0.28 units per month for congenitally deaf children (N=16), 0.72 ± 0.31 for those deaf at ages 1 -24 months (N=7), and 1.10 ± 0.72 for those with some useful hearing at the time of implantation (N=3). (In this case there were no data for the group deafened at 25-88 months.) The slope of the standard curve is about 1.0 unit per month. Thus, on average the increase in the vocabulary of 23 children deaf at birth or in the first two years remained delayed with respect to normal-hearing children although the implanted children started with a considerable disadvantage. However, 8 of these 23 children show slopes in figure 4.8 between 0.9 and 1.1, an increase in parallel with the standard curve. No child shows a steeper slope, which would be indicative of progression toward normal. The results for the 3 children with some useful hearing at the time of implantation show large interindividual differences. The individual slopes are 0.46, 0.95, and1.88. Hence one child showed progression toward normal.

Analysing performance with the CVC and Erber tests, there appeared to be a large negative effect of the duration of deafness before implantation in the congenitally and early deaf children (figures 4.4 and 4.7). Thus, vocabulary size may also be negatively affected by the duration of deafness. However, vocabulary size will also increase with age on the basis of visual input. Hence, it is interesting to examine whether or not the detrimental effect of duration of deafness, particularly on speech reception, carries over on vocabulary size. Figure 4.9 presents the raw scores found for the congenitally deaf children (N=14) and for those deaf 1-24 months after birth (N=6) collected four years after implantation as a function of their age at the time of data collection. There appears to be a small negative correlation of R = -0.24 for these data. Figure 4.9 clearly shows the delay of these children with respect to the standard reference curve. Four years after implantation the delay is about 9 months up to 72 months. Rather than considering the raw scores four years after implantation one may also pose the question of whether or not the increase in the raw scores over time, of these congenitally and early deaf children, derived from linear regression analysis, depends on the age at implantation. The answer was negative: the correlation coefficient was only 0.007.

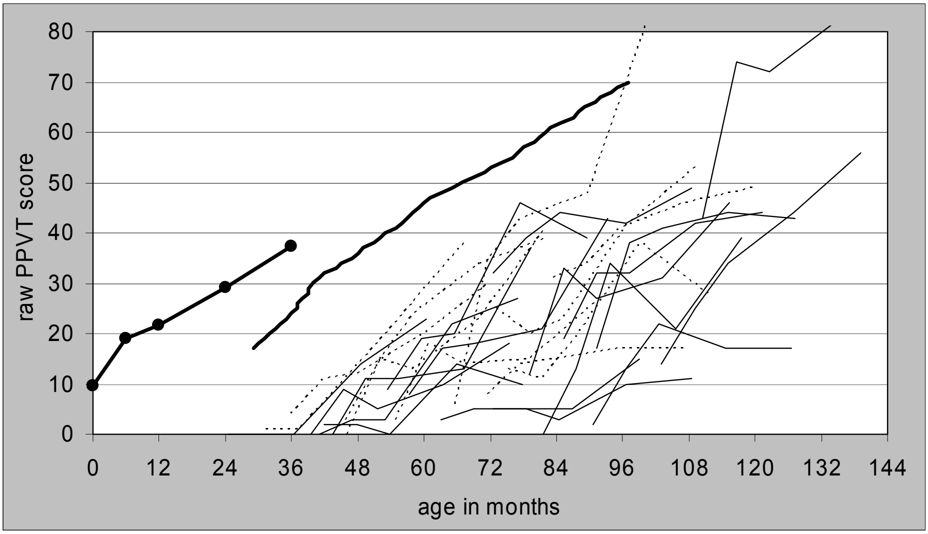

Figure 4.8. Individual raw scores from the Peabody Picture Vocabulary Test (PPVT score) as a function of age measured pre-implantation (left end of each individual curve) and at 6, 12, 24, and 36 months post-implantation. Solid curves from 16 congenitally deaf children, dotted curves from 11 post-natal deaf children. The mean raw scores as a function of the post-implantation period, not age, are presented on the left hand side (bold curve with data points, at 0 months the pre-implantation average). The bold curve without data points presents the standard score for normal-hearing subjects as a function of true age. This curve can be used to convert raw scores into equivalent age.

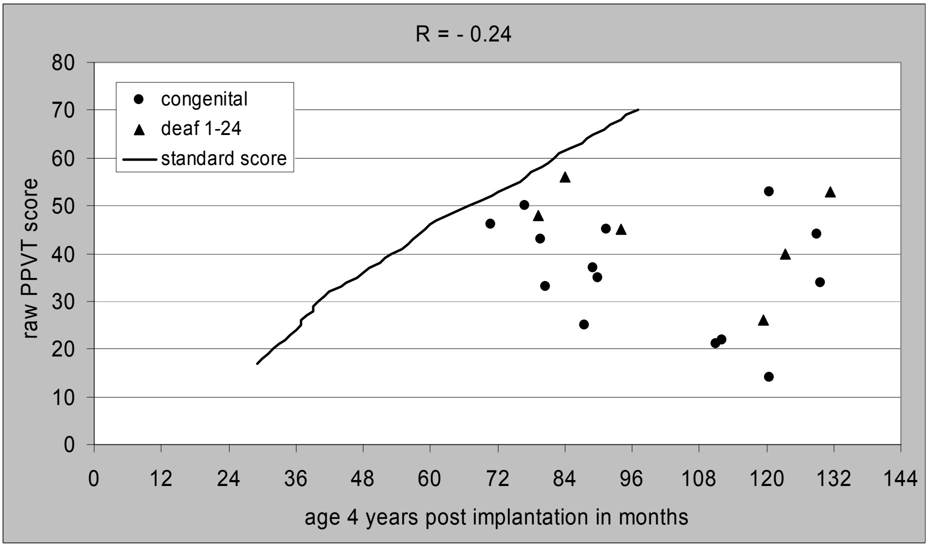

Figure 4.9. Individual raw scores from the Peabody Picture Vocabulary Test (PPVT score) as a function of true age 4 years after implantation. Data for congenitally deaf children (N=14) and children with onset of deafness at 1-24 months of age (N=6).

Figure 4.8 shows that the raw scores over time may not increase monotonically whereas one may assume that size of vocabulary does. The jumpiness in the data probably reflects variability in motivation, concentration and mood of the child from one test session to the next. In order to characterize this variability to a first order we used the variance not explained by linear regression of the individual curves in figure 4.8. The standard reference curve suggested that the increase of the raw scores over time is largely linear. The average unexplained variance in the data from the Peabody test was 22%. The results from the Erber test also showed some jumpiness. However, the measure based upon linear regression introduced here could not be used for those data because of the prominent floor and ceiling effects in the data. The scores from the Bosman-Smoorenburg speech reception test increased monotonically over time.

4.3.4 Results from the Reynell Verbal Comprehension Test

The results of the Reynell test had to be evaluated, like those of the Peabody test, with respect to the age of the child. Figure 4.10 presents the individual data for 31 children with test results pre implantation and post implantation at 6, 12, 24, and 36 months. Similarly to the results for the Peabody test, figure 4.10 presents the raw data in conjunction with the standard score. The standard score increases linearly with age until a score of about 65, above which the score levels off. The average scores as a function of post implantation time (including the average pre-implantation score: 0 months) were added to figure 4.10. However, in view of the nonlinear growth of the scores we excluded from averaging the two curves from individuals with the highest scores. These two (dotted) curves represent results for two children with some useful hearing at their time of implantation. The average increase in the raw scores post implantation was 29 units in 36 months (0.8 unit per month) whereas the standard score increases 29 units in 17 months (1.7 units per month). Hence, the ratio of the average increase found for the implant recipients to the increase in the standard score equalled 0.47. Calculating the slopes of the individual curves (excluding the two curves mentioned before) from linear regression over the 36 months period yielded average slopes and standard deviations of 0.77 ±0.39 units per month for the congenitally deaf children (N=19), 0.95 ± 0.26 for those deaf at ages 1 -24 months (N=5), and 0.90 ± 0.51 for those with some useful hearing at the time of implantation (N=4), while the slope of the standard curve was about 1.7 units per month.

Thus, on average, the increase in verbal comprehension of nearly all children stayed behind with respect to the increase found in normal-hearing children although the implanted children started with a considerable disadvantage. The results of only two children showed an increase of 1.5 units per month, close to that of 1.7 for the standard score. All other results were lower.

Examining the results from the congenitally and early deaf children (N=34) four years after implantation, as was done with the previous test in figure 4.9, the results looked better. Figure 4.11 shows that some children did reach the maximum score of 70 to 80 units corresponding to the performance of normal-hearing children at an age of about 72 months. A few, implanted at about 24 months, even reached the maximum score at the same age as normal-hearing children. However, the data of figure 4.11 also show that many other children, deafened at the same age, scored much lower. After four years of

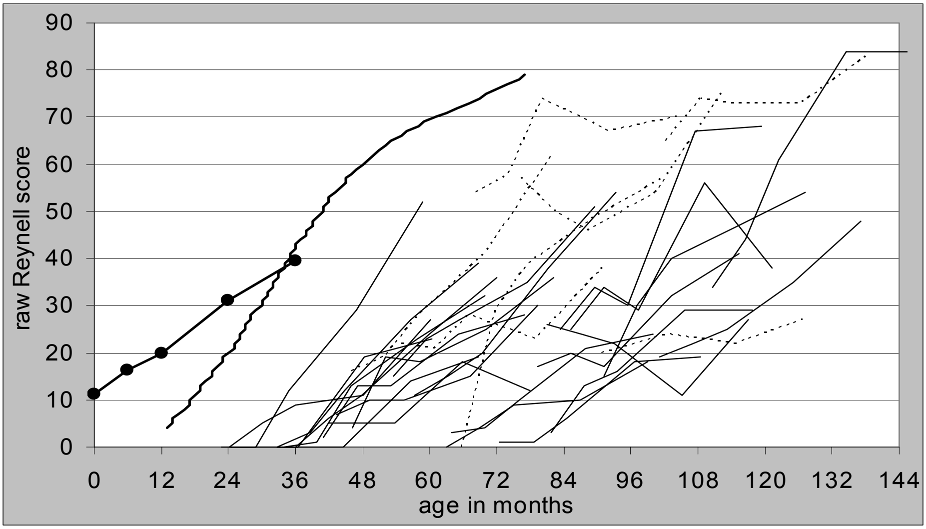

Figure 4.10. Individual raw scores from the Reynell Verbal Comprehension Test as a function of age measured pre implantation (left end of each individual curve) and at 6, 12, 24, and 36 months post implantation. Solid curves from 19 congenitally deaf children, dotted curves from 12 post-natal deaf children. The average raw scores as a function of the post-implantation period (including the pre-implantation average at 0 months) are presented on the left hand side (bold curve with data points). The two dotted curves from individuals with the highest scores are excluded from this average. The bold curve without data points presents the standard score for normal-hearing subjects as a function of true age. This curve can be used to convert raw scores into equivalent age.

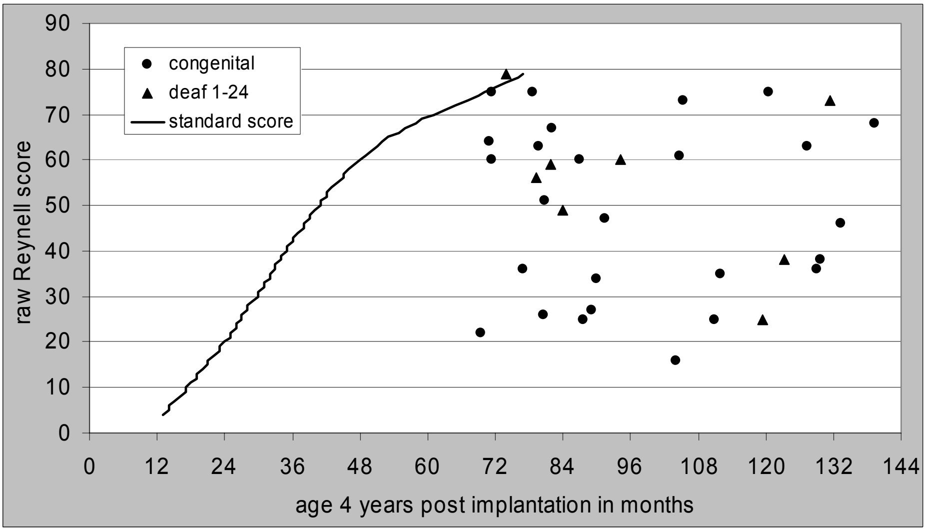

Figure 4.11. Individual raw scores from the Reynell Verbal Comprehension Test as a function of age, 4 years after implantation. Data for congenitally deaf children (N=26) and children with onset of deafness at 1-24 months of age (N=8).

implant use we found congenitally deaf children implanted between about 24 and 36 months with raw scores of between 20 and 30, implying a delay in verbal comprehension of about 48 to 60 months with respect to their normal-hearing counterparts. The spread in these scores was very large.

In figure 4.11 there appears to be no correlation between the raw scores and age. Thus, duration of deafness, from its onset to the time of implantation, did not affect the scores four years after implantation. (There was a small correlation in figure 4.9 in the Peabody Picture Vocabulary Test.) Also, there was no correlation between the age of implantation and the increase in the scores (the slope) over a three year period (R = -0.01).

The jumpiness of the Reynell data, expressed as before in the total variance unexplained in linear regression analysis of the individual curves, was 17%. Again, we excluded the two (dotted) curves with the highest scores because they were not in the linear part of the standard curve. Considering that two other curves showing saturation were included, which somewhat increased the unexplained variance, we may conclude that the jumpiness in these data is smaller than the value (22%) found in the Peabody test.

4.3.5 Results from the Schlichting sentence production test

Utterances elicited by children are analysed in the Schlichting test in terms of sentence development: syntax and grammar. Like the previous receptive language tests this test is intrinsically age related. Therefore, the results are presented in figure 4.12 in a fashion similar to that of figures 4.8 and 4.10. The results presented are all from subjects with measurements pre implantation and at 6, 12, 24, and 36 months post implantation, N=58. The raw data show the same trend as in the previous figures. The individual results showed no progression to normal over time. On the contrary, within the 36 month post-implantation period most results stayed behind with respect to the increase in the standard score. The average increase within this period (left hand curve with data points) equalled 8.6 units whereas the standard score increased by 18.5 units over the same period, a ratio of 0.47. The outstanding individual result (dotted curve) near the standard curve in figure 4.12 concerns an implant recipient who became deaf at 15 months and was implanted at 26 months. This individual result corresponds to the upper data point (triangle) in figure 4.11. Calculating the slopes of the individual curves from linear regression over the 36 months period yielded average slopes and standard deviations of 0.22 ± 0.14 units per month for congenitally deaf children (N=38), 0.27 ± 0.22 for those deafened at ages 1 -24 months (N=10), and 0.32 ± 0.14 for those with some useful hearing at the time of implantation (N=8), while the slope of the standard curve was about 0.51 units per month. (In this case there were only 2 data points for the group deaf at 25-88 months.) Thus, also with this test the increase in performance of most children was smaller than the increase in normal-hearing children, although the implanted children started with a considerable delay. Implant recipients with some useful hearing did somewhat better than those deafened in the first two years (ratios of 0.63 and 0.53 with respect to the standard score, respectively). The latter, in turn, did better than the congenitally deaf (ratio of 0.43). Only 5 of 58 children from the three deafness categories showed an increase of 0.46 units or more per month, close to that of 0.51 for the standard score. All other results were lower.

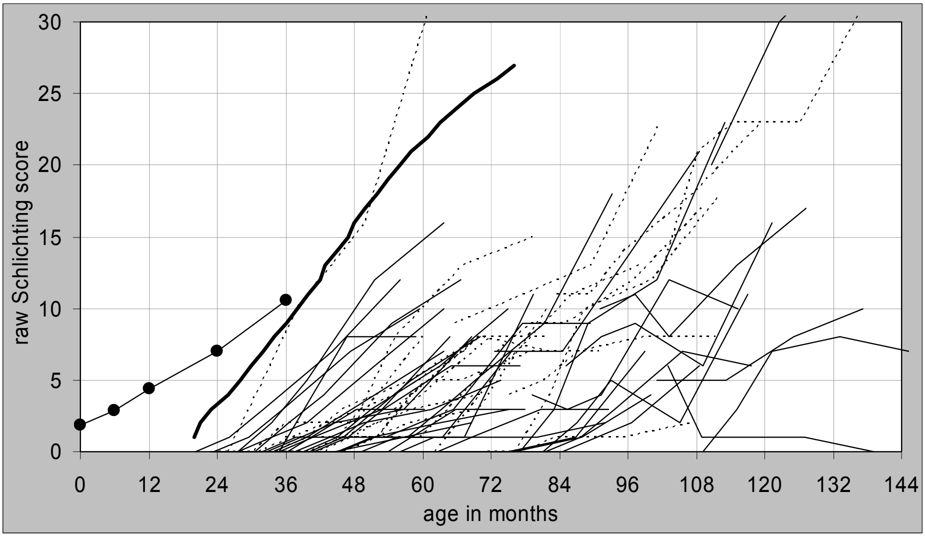

Figure 4.12. Individual raw scores from the Schlichting sentence production test as a function of age measured pre implantation (left end of each individual curve) and at 6, 12, 24, and 36 months post implantation. Solid curves from 38 congenitally deaf children, dotted curves from 20 post-natal deaf children. The average raw scores as a function of the post implantation period (including the pre-implantation average at 0 months) are presented on the left hand side (bold curve with data points). The bold curve without data points presents the standard score as a function of true age. This curve can be used to convert raw scores into equivalent age.

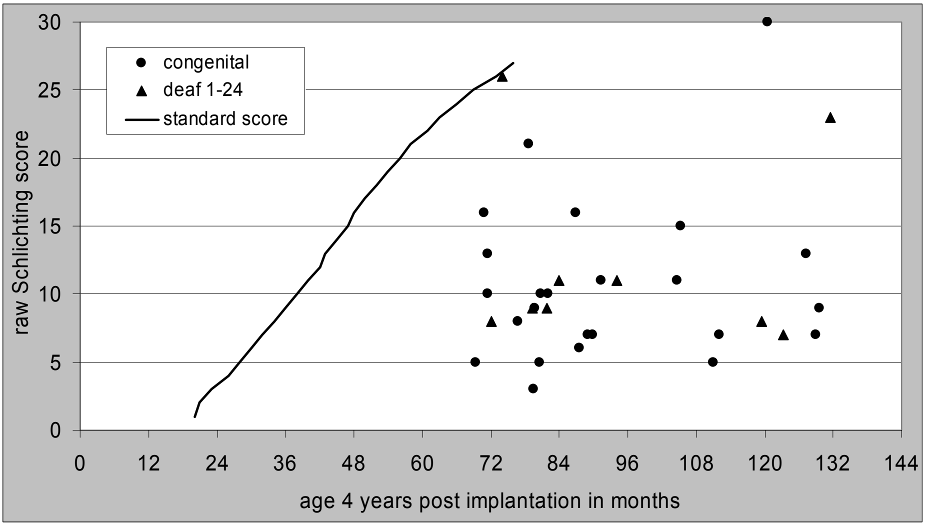

Figure 4.13. Individual raw scores from the Schlichting sentence production test as a function of age 4 years after implantation. Data for congenitally deaf children (N=24) and children with onset of deafness at 1-24 months of age (N=9).

The results from congenitally and early deaf children (N=33) four years after implantation are presented in figure 4.13. Only two children reached the maximum score found for normal-hearing children at age 72 months. One of these two children was the well performing child previously identified (triangular data point). In contrast to figures 4.11 and 4.9, however, figure 4.13 shows that sentence production performance of many children stayed low at about 9 units, corresponding to the performance of normal-hearing children at an age of 36 months. Sentence production performance showed more delay than the results from the two previous receptive language tests. Also, there was no effect of age of implantation on the scores from these congenitally and early deafened children. Moreover, the increase in the scores over the 36 months period, derived from linear regression analysis, showed no correlation (R = -0.003) with age at implantation. The jumpiness in the present individual curves in terms of unexplained variance in linear regression was calculated at 13%, better than the result in the two previous tests.

4.4. Discussion

4.4.1 Post implantation performance

Speech reception developed well in children who received cochlear implants. After 4 years the mean phoneme score for CVC words reached 55% in auditory-only mode and 75% in auditory-visual mode (N=61). One should bear in mind that these were scores for individual phonemes in CVC words with little redundancy. Statistically, the number of independent phonemes per word is 2.4, close to the number of phonemes per word of 3. Every-day speech is much more redundant. The data showed considerable interindividual spread. Some spread may be related to the type of implant. Figure 4.3 suggested, however, that the newer implants may yield higher scores in the first two years after implantation, but four years after implantation the influence of the type of implant may be rather limited. The age at which the child became deaf was more important. As one may expect the results were best for those children with some useful hearing at the time of implantation while the lowest scores were found for the congenitally deaf. Yet, even for the congenitally deaf children, and children with onset of deafness in the first two years of their life, we found CVC phoneme scores four years after implantation between 50 and 95% in auditory-only mode and between 65 and 100% in auditory-visual mode when the duration of deafness, from its onset to the time of implantation, was limited to two years. For longer durations of deafness before implantation, the data showed a marked decrease in performance. The same effect occurred in the Erber word comprehension scores, although less clear because these scores suffered from floor and ceiling effects (figure 4.6). These results stress the importance of early implantation. Even four years after implantation the effect of duration of deafness before implantation was strongly present.

The negative effect of an increasing duration of deafness before the time of implantation on the performance of congenitally and early deaf children in the previous tests was less apparent in the Peabody and Reynell receptive language tests and in the Schlichting speech production test. These tests are intrinsically age dependent. With increasing duration of deafness speech reception may decrease, but increasing visual input may reduce its negative effect on language acquisition. In agreement with this view, figure 4.9 showed a very small (negative) correlation between the Peabody scores collected four years after implantation in the congenitally and early deaf children. Moreover, figures 4.11 and 4.13 showed no such correlation in the Reynell and Schlichting scores.

The results from the latter three language oriented tests were, in contrast to the first two tests, rather deceptive. Nearly all individual scores in figures 4.8, 4.10, and 4.12 did not show a progression to normal. On the contrary, most individual curves showed an increase in score less than the increase in the standard score, although these children started with a considerable delay. The increase in scores was quantified by the slope calculated from linear regression analysis. In the three tests these slopes were independent of age. Thus, the data could be summarized in the three figures by calculating the mean score as a function of post-implantation time. The ratio of the slopes found for the implant recipients to the slope in the standard score was 0.77 for the Peabody test, 0.47 for the Reynell test, and 0.47 for the Schlichting test. These three ratios varied from 0.67, 0.45, and 0.43 for the congenitally deaf to 1.10, 0.53, and 0.63 for the implant recipients with some useful hearing (hearing aid users) at the time of implantation. These results strongly suggest that guidance programs and tests for young implant recipients should not focus on hearing only but language acquisition should receive ample attention. One should bear in mind, however, that a number of these children have multiple handicaps, which delay language development.

Frequently the question is asked whether or not one can predict receptive language performance, in particular the increase in performance to be expected, from the speech reception results. The correlation between the CVC scores over time and the Reynell verbal comprehension scores may yield some insight into this question. Table 4.1 shows that the correlation between the CVC scores, both those measured in auditory-only and those measured in auditory-visual mode, 6 and 12 months post implantation, and the slope in

A	pre	6m	12m	24m	36m	48m
pre	-	0.51	0.45	0.31	0.09	0.03
6m	-	0.47	0.38	0.36	0.17	0.16
12m	-	0.54	0.52	0.40	0.16	0.34
24m	-	0.21	0.37	0.47	0.36	0.35
36m	-	0.17	0.55	0.57	0.47	0.45
48m	-	-0.02	0.52	0.62	0.55	0.45
slope	-	0.18	0.22	0.59	0.58	0.58

AV	pre	6m	12m	24m	36m	48m
pre	0.74	0.74	0.67	0.37	0.36	0.31
6m	0.88	0.84	0.78	0.48	0.46	0.41
12m	0.37	0.84	0.79	0.61	0.44	0.52
24m	0.83	0.80	0.65	0.62	0.56	0.60
36m	0.68	0.77	0.68	0.70	0.68	0.66
48m	0.63	0.39	0.48	0.68	0.77	0.64
slope	0.30	0.26	0.36	0.44	0.54	0.49

Table 4.1. Coefficients of correlation between CVC scores (by row), measured in auditory-only mode (A, upper panel) and auditory-visual mode (AV, lower panel), and the raw Reynell scores (by column). Scores collected pre implantation and 6, 12, 24, 36, and 48 months post implantation. The lower row of each panel presents the correlation between the CVC scores at 0-48 months and the slope of the increase in the Reynell scores over a 36 month post-implantation period. The coefficients for the pre-implantation condition in auditory-only mode are not presented because almost all CVC scores equalled 0.

the Reynell scores was low, from 0.18 to 0.36 (N=7-16). Thus, the CVC scores offered no predictive power. The highest correlation coefficients in Table 4.1 (about 0.8, N=15-25) were found for the auditory-visual CVC scores and Reynell scores up to about one year post implantation. This suggests that the Reynell scores collected within one year after implantation were influenced by auditory-visual speech reception.

4.4.2 Comparison of the five tests

The Bosman-Smoorenburg speech reception test using CVC words and scored in terms of phonemes repeated correctly, showed a wide application range. Thus, it can be used in auditory-only, visual-only and auditory-visual mode over a period of four years post implantation without risking floor and ceiling effects (figure 4.2). Presentation of three lists (33 words) could be completed in about 5 minutes while the inaccuracy calculated from the binomial distribution, taking into account 2.4 independent phonemes per word, was at most 5.6 percentage points. The individual results over time showed little deviation from the expected monotonic increase (little jumpiness), suggesting that the cooperative mood of the child had little effect on the scores. Detailed analysis (figure 4.3 and particularly figure 4.4) showed good discriminative power. The a-priori drawback of oral responses that may be poorly intelligible did not appear to be a real disadvantage in clinical practice. In fact, one may consider this to be an essential aspect of the score when one takes the test as a measure of communication (hearing and being heard) rather than an analytical test of speech reception as such. The Erber Word Comprehension test appeared to be limited in its application range with a large probability of encountering floor and ceiling effects. It is useful only when one wants to focus on progress in performance over a period of not more than two years. In addition, it is less suited for children below 5 years of age.

The two receptive language tests and the speech production test constitute an important addition to the CVC speech reception test. The Peabody Picture Vocabulary Test showed an increase in the score over time of about 77% of the increase in the standard score and it showed some negative effect of age of implantation on the four years post-implantation scores for the congenitally and early deafened children. However, the latter effect was more obvious in the CVC scores, particularly in those collected in auditory-only mode (R=-0.24 in PPVT versus R=-0.68 in CVC). The error (jumpiness) in the individual curves estimated from linear regression of the increase in scores over time was 22%. The increase in the scores of the Reynell verbal comprehension test and the Schlichting speech production test was in both tests only 47% of the increase in the standard score. Thus, these scores reflect larger discrepancies compared with those from normal-hearing children. Finally, the error in these two tests was respectively 17 and 13%, smaller than the error of 22% in the Peabody test. Therefore, the Reynell test appears to be a better complement to the CVC speech reception test than the Peabody test.

If one seeks to reduce the number of tests one may conclude that the CVC word reception test, scored in phonemes and the Reynell Verbal Comprehension Test are the most relevant indicators of performance, while the Schlichting speech production test could be viewed as a valuable addition. The Schlichting speech production test yielded lower scores with respect to those from normal-hearing children than the Reynell test. In this respect it is an important complement to the Reynell test, in particular for children with relatively good performance. Moreover, the increase in scores with age is steadier in the Schlichting test than in the two other language tests.

Chapter 5

Development of speech perception in adults over time; comparison of performance measures

Summary

Objective

Performance of adult cochlear implant recipients is typically expressed in terms of speech perception. Other aspects like music perception and speech quality (as it is affected by hearing one's own voice via the implant) receive some attention but speech perception is the primary choice when the results of cochlear implantation in adults are evaluated in terms of recipient performance or cost-benefit analysis of the intervention. Speech reception can be measured in various ways: from an analytical approach at word level, scoring individual phonemes perceived correctly, to an approach reflecting everyday situations in which the clinician measures how the listener keeps track of connected discourse. In addition, one may include or exclude the contribution of speech reading and the detrimental effect of interfering environmental noise. Finally, one may consider including the contribution of acoustic hearing aids, which are sometimes used in conjunction with the cochlear implant. This chapter presents the results of three tests, covering the range from presentation of isolated words to connected discourse, over a period after implantation of up to 10 years. It includes the effect of lip reading and the role of acoustic hearing aids, but not the effect of interfering noise. In addition, this chapter presents measurement accuracy of the three tests in relation to the time it takes to complete the tests.

Conclusions

In adults speech perception performance improved over a period of up to 2-3 years after implantation, considering the number of phonemes-in-words and the number of syllables-in-sentences responded correctly. After 2 years, performance stabilised at an average of 50 to 70% phonemes perceived correctly in auditory-only mode, 60 to more than 90% syllables perceived correctly in auditory-only mode and 85 to 100% syllables perceived correctly in combined auditory-visual mode, depending on the type of implant and excluding the first series with the CI22M implant. Connected discourse performance stabilised already one year after implantation. After one year 45 to 70 words per minute were perceived correctly in auditory-only mode while the scores in auditory-visual mode

reached 70 to 85 words per minute. Normal-hearing people score above 80 words per minute, on average about 100 words per minute.

The scores differed seriously among implant recipients. Two years after implantation 10% of the Nucleus CI24M and CI24R(CS) users showed syllable-in-sentence scores in auditory-only mode below 15% whereas another 10% of users scored above 95%.

The number of adult implant recipients who continued to use their acoustic hearing aid in the contralateral ear immediately after implantation decreased by 50 % over a three year post-implantation period. After three years this percentage did not change over another two years. The average phoneme score in auditory-only mode for the acoustic hearing aid with the implant switched off was 12%. The acoustic hearing aid contributed about 5% to the total score for the cochlear implant and the acoustic hearing aid together.

Considering the accuracy of measurement reached in a given duration of a measurement session, the phoneme-in-word score does substantially better than the syllable-in-sentence score and the connected discourse score. The CVC word test would be the primary choice if one wishes to reduce the number of tests. The connected discourse test could be viewed as a valuable addition.

5.1 Introduction

Speech perception by adult implant recipients can be measured with tests ranging from a simple word reception task to assessing the perception of connected discourse. Using simple words in a test implies that one reduces the influence of language skills, focussing on perception of speech elements. This may contribute to, for example, fitting the speech processor. Analysing the mistakes in perceived speech elements, or phoneme confusions, may provide insight into the question of whether or not the frequency response of the speech processor has been tuned optimally in relation to the individual's sensitivity to electrical stimulation across the electrode array. If one aims at measuring speech performance of an implant recipient in relation to everyday situations one may choose to assess how well the recipient is following connected discourse. However, the choice of test could also be determined by quite different considerations. If one wishes to follow the development of speech perception performance after implantation over time it is important that the test has high measurement accuracy within a reasonable duration of measurement. From a statistical point of view this means that the test should contain a high number of independent elements. Moreover, the test should be insensitive to the performance level of the implant recipient determined by factors other than auditory or auditory-visual speech perception; factors such as changes in concentration and motivation. This chapter presents the results of three tests at the word, sentence and connected discourse level that provide a broad scope of the development of speech perception after cochlear implantation, while they can also be compared in terms of accuracy. The three tests have been developed in the Dutch language. The word and sentence tests are based on standard speech-audiometry tests in the Netherlands.

The chapter presents a retrospective analysis of clinical data, collected over about a ten year period. The analysis does not concern a prospective study with a balanced design. Such an analysis has certain advantages and disadvantages. An advantage is that the analysis reflects the results of everyday clinical practice whereas in experimental settings performance may be different if extra attention is paid to implant recipient guidance and testing procedures. A disadvantage is that the data sets are incomplete, which introduces a considerable risk of biased results due to confounded variables. Moreover, missing data may imply a selection if, for example, the examiner anticipated a 0 or 100% score and therefore did not conduct the test. In view of these disadvantages the present data sets have been inspected very carefully and, in case of doubt, results of analyses have been rejected.

5.2 Materials and methods

5.2.1 The three speech tests are:

(1) The Bosman-Smoorenburg CVC word reception test

Subjects have to repeat orally a word consisting of the sequence consonant-vowel-consonant (CVC). The test used here is a precursor of the standard audiological speech reception test in the Dutch language. Each list consists of 12 CVC words. The responses are scored in terms of the number of phonemes repeated correctly (36 in total). The test has been developed in Dutch. The Dutch language contains a large enough number of CVC words. Statistical analysis of results collected previously has shown that the redundancy in these words is small; on average the CVC words contain 2.4 independent phonemes per word. Hence, scoring phonemes perceived correctly rather than words perceived correctly improves the accuracy of the score. Oral repetition of the words implies that the score depends on the quality of the implant recipient's articulation. Thus, the test is not just a speech reception test. The test is administrated in auditory-only mode using recorded materials. Presenting one list takes about 1 minute. The number of lists presented in one session varied between 4 and 16, on average it was 10. If the implant recipient also used an acoustic hearing aid we determined scores for the following conditions: implant only, hearing aid only and implant plus hearing aid.

(2) The sentence reception test

Subjects have to repeat orally a sentence consisting of 8 or 9 syllables. The test is based upon a standard audiological speech reception test in the Dutch language introduced by Plomp and Mimpen. In its original form it applies an adaptive procedure, changing sentence presentation level, and it is primarily meant to be presented in noise. However, when presented to profoundly hearing impaired persons and early cochlear implant users the adaptive procedure may fail. Therefore, the test is administrated at fixed levels. In addition, scoring is conducted in terms of the number of syllables responded correctly, whereas the original adaptive procedure is based on whether or not sentences were repeated completely correctly. Previous research demonstrated that the sentences contain about three statistically independent elements. Thus, scoring in terms of syllables correct, rather than whole sentence correct, improves the efficacy of the test in terms of accuracy reached within a given test duration. Finally, the test is presented without background noise.

The test was recorded on video tape with frontal presentation of the speaker's face. It was presented in auditory-only, visual-only and auditory-visual mode. The implant recipient was allowed to use a contralaterally placed acoustic hearing aid in conjunction with the implant if they normally used their hearing aid. A test consisted of lists of 13 sentences. There was no contextual relation among the sentences within a list. The first 3 sentences were not scored but used to acclimatise. Thus, the score was based on 10 sentences with, on average, 85 syllables. In each condition the test was repeated once, collecting the scores for test and retest individually. The test-retest scores allowed for a calculation of measurement accuracy. The presentation of one list took about 2.5 to 3 minutes.

(3) The connected discourse tracking test

This test is not based upon a standard test in the Dutch language but it was introduced when starting the cochlear implant project. It was modelled after examples from abroad. A simple text of 350 to 400 words is presented live to the implant recipient. The subject has to repeat the text sentence by sentence. When the response contains an error the sentence is repeated by the examiner, but no more than twice, after which, if the response is still in error, the sentence is rephrased until a correct response is received. The test is scored in terms of words per minute perceived correctly. The scores for normal-hearing people range from 80 to sometimes more than 120 words per minute, depending somewhat on the examiner. 100 words per minute may be adopted as a standard reference for normal-hearing people. When testing implant recipients the test took, on average, 6.5 minutes. The test was administrated in auditory-only, visual-only and auditory-visual mode. Implant recipients were allowed to use a contralaterally placed acoustic hearing aid in conjunction with the implant if this is what they normally did.

5.2.2 Subjects and implants

The results presented are from 160 adult implant recipients who received their cochlear implant between 1992 and 2004. From 1992 to 1995 subjects received the Nucleus® CI22M device (N=17) with the MSP™ or Spectra™ speech processor. From 1997 to 2001 they received the Nucleus CI24M device (N=55) with the Sprint processor up to 2000 and the Esprit™ processor thereafter. After mid 2002 the Sprint processors were gradually replaced by Esprit processors, if preferred by the implant recipient. From 2000 to 2004 subjects received the Nucleus CI24R(CS) device (N=64) supplied with the "Contour" electrode and with the Sprint processor up to 2001 and the Esprit(3G) processor thereafter. Also with this implant Sprint processors were gradually replaced by Esprit(3G) processors after mid 2002, if preferred by the user. From mid 2003 to 2004 subjects received the Nucleus CI24R(CA) (N=6) device with the "Advanced Contour" electrode and the Esprit3G processor. Cases with CI24R(ST) devices (N=5) have been excluded from this analysis because the choice of this implant implies a particular medical indication rendering these data unsuited for statistical analysis. Likewise, two cases with split electrodes have been excluded. The scores found for these 7 excluded cases in auditory-only mode varied from 10 to 70% correct. In addition to the Nucleus implants there was a group of subjects who received, from 1995 to 1997, a Med-El COMBI 40® implant (N=11) with the C40+ speech processor. Starting in 2003 the C40+ processor was replaced by the TEMPO+ processor.

Most electrode arrays reached full insertion depth. Four electrodes of the CI24R(CS) implant reached 25 or 26 mm rather than the full insertion depth of about 32 mm. This did not affect the speech performance scores. Below, no distinction between these four cases and the cases with full insertion are made.

Onset of deafness in the best ear occurred in this cohort of implant recipients at a mean age of 37 years. The distribution is skewed with a standard deviation of 20 years, a maximum age of 73 years, 25 subjects deafened at less than 10 years of age, and one congenitally deaf subject. The duration of deafness, from its onset to the time of implantation, was 15 years on average. Also this distribution was skewed with a standard deviation of 16 years, a maximum duration of 60 years, and a minimum duration of 0.5 year.

5.2.3 Other factors included in the analysis

The analysis also addresses, in addition to the longitudinal behaviour of the speech perception scores, the effects on the scores of aetiology of deafness, age at which deafness occurred, and duration of deafness, from its onset to the time of implantation. The Statistica software package was used for statistical analyses (Statsoft Inc., release 7.1).

5.3 Results

5.3.1 Results from the Bosman-Smoorenburg CVC word reception test

Figure 5.1 shows the average CVC scores collected in auditory-only mode for five types of implant devices over a period from 1 year for the Nucleus CI24R(CA) device up to 10 years for the Nucleus CI22M device. For each type of device subjects were included only if their individual data covered all observation times given in figure 5.1 for that implant.

Thus, the curve given for each implant is not affected by subjects dropping out during the period covered by that curve: those subjects were completely excluded from this analysis. Except for the CI22M implant, the data show a gradual increase of the CVC scores during the first two to three years after which the scores tended to stabilise. The CI22M users were called back after 6 years, a period in which they were not measured. The higher scores at 84 months and over can be ascribed to replacement of the MSP speech processors by newer types. Similarly, the increase at 84 months in the COMBI 40 results can be ascribed to changing the processor from the C40+ to the TEMPO+.

Figure 5.1 shows the scores found for the various types of Nucleus devices; the newer types show better performance. The low scores for the first type of device, the CI22M, may not only be related to the technological state of the art at that time but to some extent also to the selection of implant candidates between 1992 and1995. None of those candidates used acoustic hearing aids; they were selected for complete deafness while later on this intake criterion was somewhat released. The scores for the COMBI 40 are relatively high considering the time this device became available. Analysis of variance of the data at 3 and 12 months post implantation showed that the CI22M scores were significantly ($p<0.0001$) lower than the scores for all other types ofdevices and that the CI24M scores were significantly ($p<0.02$) lower than the scores for the CI24R(CS) and CI24R(CA) devices. Although the results in figure 5.1 might suggest that the scores found for the COMBI 40 are also higher than those for the CI24M it appears that the difference between these scores is too small to attain statistical significance ($p=0.15$ at 3 months and $p=0.6$ at 12 months) in view of the small number of subjects in the COMBI 40 group ($N=11$).

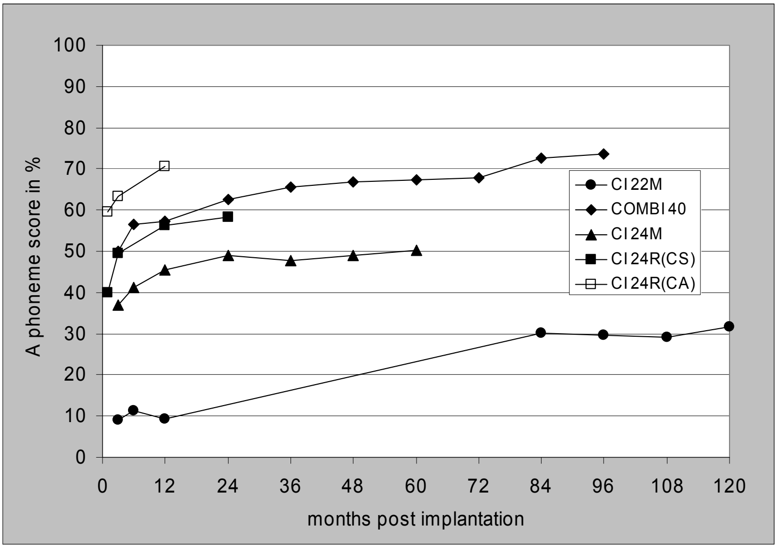

Figure 5.1. Time course after implantation of the average phoneme scores for CVC words presented in auditory-only (A) mode shown for five types of implant devices. Number of subjects per type of device: CI22M, N=15; COMBI 40, N=11; CI24M, N=49; CI24R(CS), N=64, CI24R(CA), N=6.

Clinically, the average scores are of limited importance. It is important to know how individuals are performing. Therefore, figure 5.2 presents the distribution of the scores across two groups of subjects; those with the CI24M and CI24R(CS) devices. These were the largest groups, respectively N=49 and N=64. The data are presented in terms of the scores exceeded by 10 to 90% of the recipients in an implant group. The interindividual differences appear to be dramatic. Whereas 10% of implant recipients may score above 80% we note that another 10% of recipients scores 20% or less, even with the newer CI24R(CS) device. Further, the CI24M results show that there is no improvement after 24 months. On the contrary, the lower decile suggests that after 24 months the scores may decrease. Again, these results are derived from complete data sets for each individual across the whole post-implantation period indicated in the figure, avoiding artefacts that may be introduced by drop-outs during the observation period.

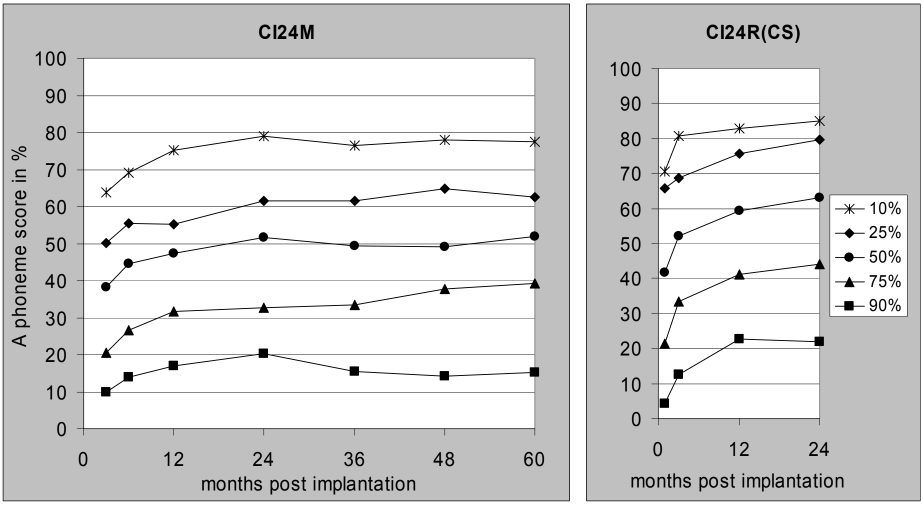

Figure 5.2. Distribution of post-implantation phoneme scores for CVC words presented in auditory-only mode, shown for the CI24M (N=49) and the CI24R(CS) device (N=64).The percentiles indicate the fraction of results with scores above the respective curves.

Figure 5.3 shows the phoneme scores for implant recipients who continued to use an acoustic hearing aid contralaterally after cochlear implantation, for whom we had pre-implantation hearing aid scores, and who could be followed for 60 months post implantation. The group (N=25) consisted primarily of CI24M users. The upper curve shows that after three years the number of these implant recipients using an acoustic hearing aid decreased by 50%. Beyond three years there is not much change. At each observation time the scores represent results from acoustic hearing aid users only. Hence, the curves represent a number of subjects decreasing with the post-implantation period. Figure 5.3 shows that the CVC scores for the acoustic hearing aid alone (implant off, 12%) did not increase with time. On average, an acoustic hearing aid score of 12% seemed to be enough to keep the hearing aid. With this score of 12% the acoustic hearing aid contributed about 5% to the CVC scores found for the implant alone. The implant recipients still using their acoustic hearing aids 60 months after implantation may be the ones who had the highest acoustic scores shortly after implantation. After 60 months their acoustic scores may have de-

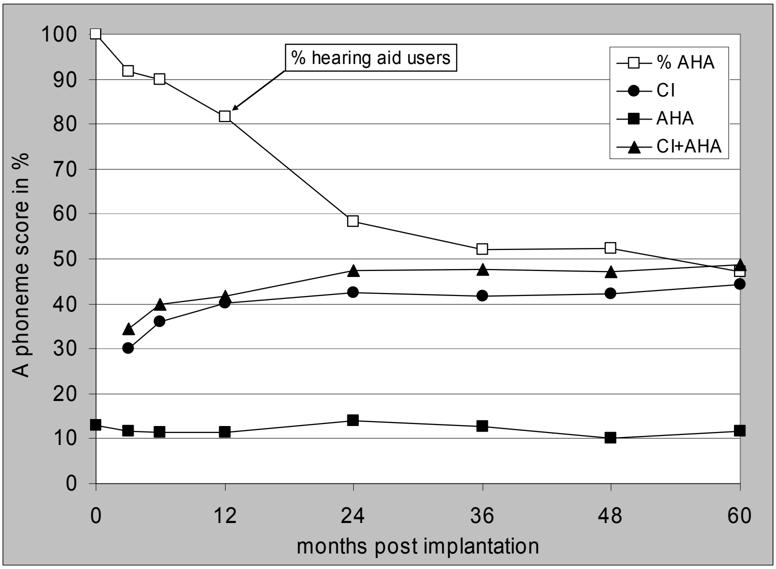

Figure 5.3. Time course of the mean phoneme scores for CVC words presented in auditory-only mode from implant recipients who, after implantation, continued to use a hearing aid in the contralateral ear. Results for all types of devices, mainly the CI24M. The upper curve (% Acoustic Hearing Aid, % AHA) shows the decrease in the number of hearing aid users in % with respect to 25 users immediately after implantation. The other curves present the average phoneme scores for the implant alone (CI), the acoustic hearing aid alone (implant switched off, AHA) and both aids switched on (CI+AHA). Note that the scores are mean values for a number of subjects decreasing with post-implantation time.

creased to about 12% while the other recipients with lower initial acoustic scores may have stopped using their acoustic hearing aids after 60 months.

The effect of the aetiology of deafness on the CVC scores collected 3 and 12 months post-implantation was analysed using six aetiological categories: Cogan's disease, infection excluding meningitis, meningitis, progressive-hereditary, skull trauma, and unknown origin. Anatomical malformations, noise trauma, otosclerosis, and sudden deafness were categories also identified but not included in the analysis because the number of cases was too small and the few cases present were not well distributed over types of implant devices and acoustic hearing aid usage, which implied the risk of confounding interaction. In order to avoid confounding interactions, analysis of variance (ANOVA) was conducted including *type of device* and *acoustic hearing aid usage* as factors in the analysis. The analysis was conducted for two groups with a sufficient number of subjects: those using the CI24M and CI24R(CS) devices, each containing about 50 subjects. ANOVA showed a significant effect of *aetiology of deafness* on the scores ($p<0.05$). The results for recipients deafened by Cogan's disease tended to be higher than average, those for recipients deafened by meningitis tended to be lower than average.

The effect of *age at onset of deafness* and *duration of deafness*, (duration from onset of deafness to the time of implantation) on the CVC scores collected 3 and 12 months post-implantation was analysed applying the multiple regression technique on the data of the two groups mentioned above. Analysis of variance is considered to be less suited because the two independent variables are not categorical but continuous in nature. The partial correlation between each of these two variables and the CVC score appeared to be small and statistically insignificant ($p>0.3$).

Measurement error is estimated from the differences between the CI24M scores collected at 36 and 48 months and between 48 and 60 months and from the differences between the CI24R(CS) scores collected at 12 and 24 months. Only scores between 40 and 60% are considered because these scores will show the largest error. This can be understood conceptually because the scores will follow a binomial distribution. Additionally, it will be clear that scores at 0 or 100% will show virtually no error (see also figure 5.6 and its explanation). However, comparing the results a year apart from one another is not truly a test-retest situation. Performance may have changed over that year. The three sets of data showed an average increase of 0.6, 1.0 and 1.6 percentage points over one year, respectively. Measurement error is expressed in the standard deviation; error variance in the square of the standard deviation. The error variance is given by half the mean squares of the differences between the first and second measurement (MSD/2) if there would have been no average increase due to a change in performance. The change in performance is taken into account by reducing MSD/2 by half the square of the average increase. However, this is correct only if it is assumed that all individuals had the same increase in score; but the difference in performance may vary from one person to the next. Theoretical calculations showed that the final result depended little on the assumption concerning the individual increase in performance. When, for example, it was assumed that these increases are uniformly distributed across subjects, between no increase and twice the average increase (so that the distribution average attained the value of the measured average increase), rather than being the same for each subject, it could be shown mathematically that the proper estimate of the error variance is obtained by subtracting 2/3 of the square of the average increase, rather than half the square, from MSD/2. Accepting the latter assumption, estimates of measurement error of 3.8, 4.3, and 6.5 percentage points were found for the three sets of data. Combining the three estimates yielded a measurement error of 5.1 percentage points. This error was related to the presentation of 10 word lists on average in about 10 minutes.

The measurement error of 5.1 percentage points is greater than the theoretical error derived from the binomial distribution assuming 2.4 independent elements per word; a result taken from previous research on speech perception by the hearing impaired. The theoretical error was calculated at 3.0 percentage points for 10 lists. Thus, one has to conclude that, in addition to the intrinsic theoretical error, taking into account a correction for a real increase in performance after a year, there also was a fluctuation in performance from one fitting session to the next. This fluctuation in performance, revealed in the present data, was calculated at 4.1 percentage points, assuming that there was no correlation between measurement error and fluctuation in performance so that the variances of the two sources of error could be added.

5.3.2 Results from the sentence reception test

Figure 5.4, upper panel, shows the average syllable scores collected in auditory-only mode for five types of implants over a period from 1 year for the Nucleus CI24R(CA)

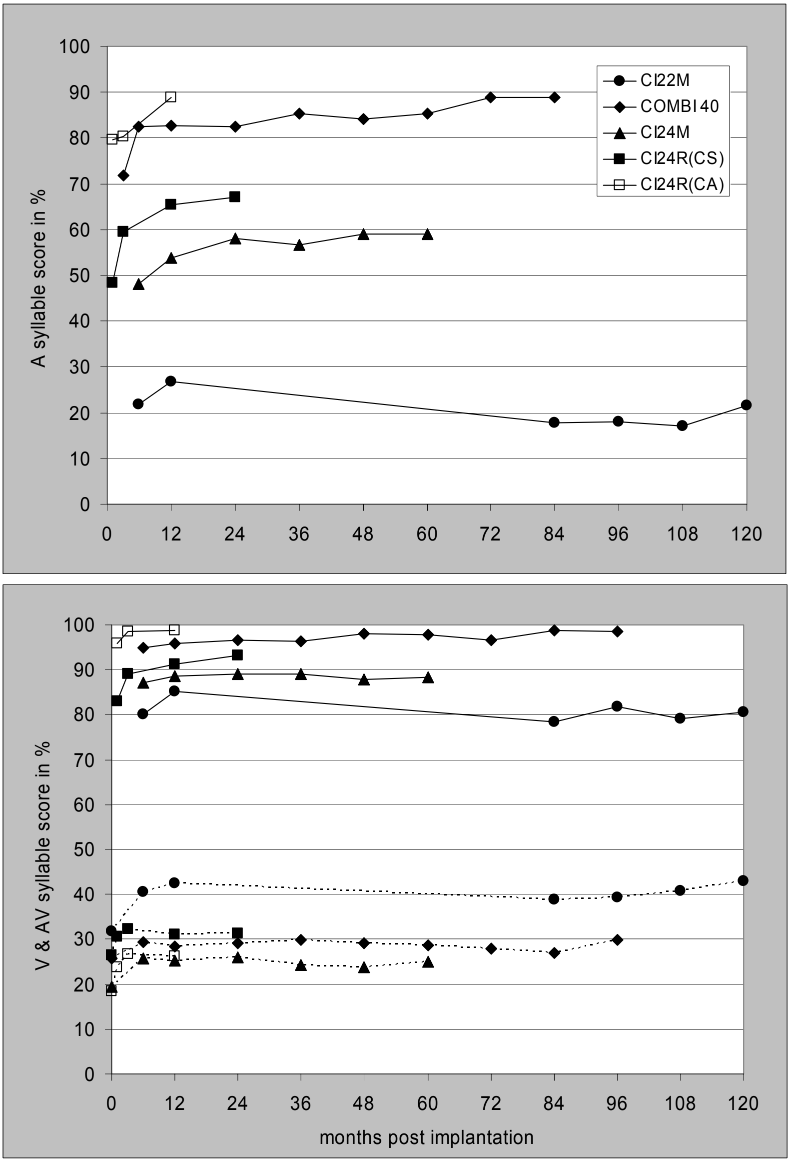

Figure 5.4. Upper panel: Time course after implantation of the average syllable scores for sentences presented in auditory-only mode (A-score). Lower panel: data collected in visual-only mode (speech reading, lower set, dotted curves, V score) and in auditory-visual mode (upper set, solid curves, AV score). Results for five types of implant devices. Number of subjects per device: CI22M, N=17; COMBI 40, N=11; CI24M, N=50; CI24R(CS), N=64, CI24R(CA), N=6.

device up to 10 years for the Nucleus CI22M device. Figure 5.4, lower panel, shows the results collected in visual-only (V, speech reading) and auditory-visual (AV) mode. For each type of implant and each mode subjects were included only if data for all observation times indicated in the respective curves were present. Thus, each individual curve is not affected by subjects dropping out after a certain number of months; those subjects were excluded from this analysis. The visual-only data show that after 6 months the CI22M users had better speech reading, about 40% syllable score, than the other groups with scores between 25 and 30%. Before, we mentioned that the CI22M group differed from the other groups in the sense that the selection of these early implant recipients was based upon complete deafness. For all types of implant devices speech reading improved up to one year after implantation. Also, speech reading contributed substantially to the syllable score when using the cochlear implant. The scores found in auditory-only mode ranged from 20 to 90% while the scores in auditory-visual mode reached values from 80 to almost 100%. The contribution from speech reading was particularly large for the CI22M group. The increase in score to about 80% in auditory-visual mode was more than the sum of the separate scores for the auditory-only and visual-only modes (20+40%) and more than the score calculated on a probabilistic basis, being 52% (see Ch. 4.3.1).

The upper panel of figure 5.4 shows, similarly to the CVC scores in figure 5.1, that the present syllable scores collected in auditory-only (A) mode stabilised within a period of two to perhaps three years after implantation. The COMBI 40 showed a small increase after 72 months, an increase similar to the one found for the CVC scores. This increase was attributed to the change of processor from the C40+ to the TEMPO+. The CVC scores for the CI22M showed a clear increase 84 months after implantation, which was attributed to replacing the MSP processor by a newer type. However, this change is not reflected in the present syllable scores. Language-based skills may have dominated these scores.

The auditory-only scores found for the different types of implant devices showed the same trend as those found for the CVC scores (figure 5.1). Analysis of variance of the data collected at 12 months post implantation showed that the CI22M scores were significantly ($p<0.001$) lower than the scores for all other types of implants except the CI24M. The scores for the COMBI 40 were significantly ($p<0.05$) higher than the scores for the CI24M. (The latter difference did not reach significance for the CVC word scores.) All other differences were insignificant.

Figure 5.5 presents, similarly to figure 5.2 for CVC phoneme scores, the distribution of the syllable scores across the two groups of subjects using the CI24M (N=50) and CI24R(CS) (N=64) devices. The interindividual differences were even larger than those found for the CVC phoneme scores: 10% of the implant recipients scored above 95% whereas another 10% of recipients scored less than 5% with the CI24M, and less than 13% with the newer CI24R(CS) device. Further, the CI24M results show again that there is no improvement after 24 months. On the contrary, the lower decile curve in figure 5.5 suggests that after 24 months scores for these subjects may decrease to virtually zero. Again, these results are derived from complete data sets for each individual across the whole post-implantation period indicated in the figure, avoiding artefacts that may be introduced by drop-outs during the observation period.

The effect of the aetiology of deafness on the syllable-in-sentence scores collected 12 months after implantation in auditory-only mode was analysed using the six aetiological

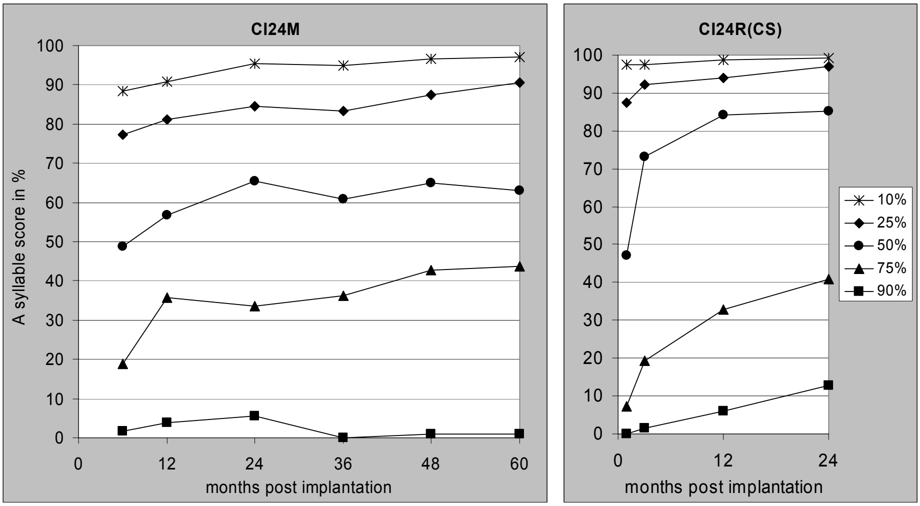

Figure 5.5. Distribution of post-implantation syllable scores for sentences presented in auditory-only mode, shown for the CI24M (N=50) and the CI24R(CS) device (N= 64).The percentiles indicate the fraction of results with scores above the respective curves.

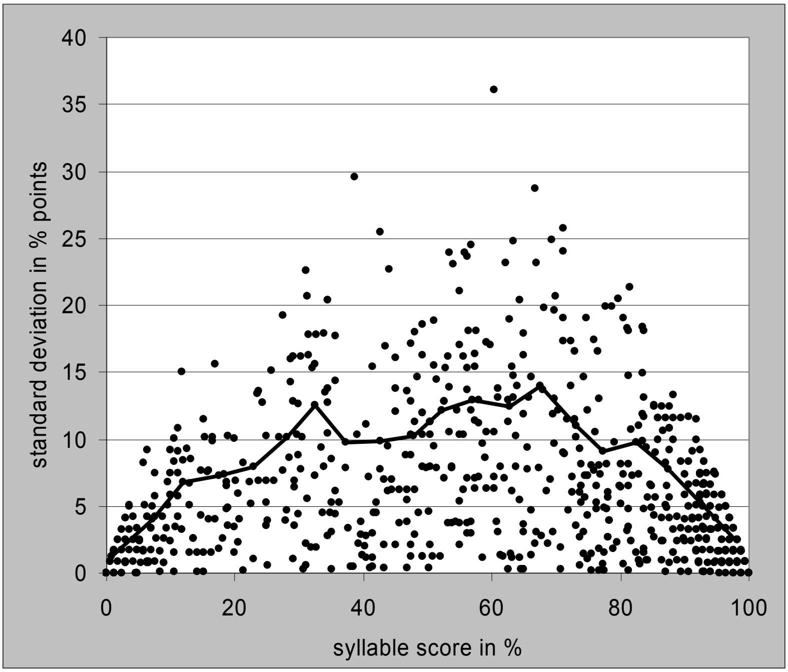

Figure 5.6. Measurement error (standard deviation) for individual syllable scores collected in auditory-only mode derived from test-retest measurements within one session. The solid line presents the total standard deviation per group of syllable scores in steps of 5%.

categories listed in the previous section. Again, the analysis was conducted for the two implant groups with a sufficient number of subjects: those using the CI24M and CI24R(CS) devices, each containing about 50 subjects. ANOVA was conducted including *type of implant* and *acoustic hearing aid usage* as factors in the analysis in order to avoid confounding interactions. ANOVA showed no significant effect on the syllable scores from any individual aetiology of deafness (p>0.3). The effect of *age at onset of deafness*, and *duration of deafness*, from its onset to the time of implantation, on the syllable scores collected 12 months after implantation in auditory-only mode was analysed applying multiple regression analysis (see Sec. 5.3.1). The partial correlation between each of these two variables and the syllable score was statistically insignificant (p>0.8).

Measurement error of the syllable score for sentences could be estimated directly from the test-retest data collected within one session. The error in the auditory-only condition was 11.4 percentage points in the 40 to 60% range. The difference between the average value of the first and second scores was so small that a potential contribution from a learning effect within one session (or vice versa from a fatigue phenomenon) to the calculated error could be considered negligible. The distribution of the errors is illustrated in figure 5.6. Note that the error is largest for the scores near 50% and that it decreases to zero at scores of 0 and 100%, as was mentioned before in Sec. 5.3.1. Assuming that the scores follow a binomial distribution it was possible to calculate the number of independent elements in the sentences. In the present case this number appeared to be 2 (as compared to 8 or 9 syllables in the sentences), whereas a previous analysis conducted for normal-hearing people resulted in an estimate of about 3 independent elements. Interestingly, the measurement error found for the visual-only condition also yielded an estimate of 2 independent elements, whereas the error found for the auditory-visual mode yielded 3.4 independent elements. This suggests that contextual cues within the sentence played a smaller part in the combined auditory-visual presentation. Comparing the differences in the scores collected between 36 and 48 and between 48 and 60 months for the CI24M device and between 12 and 24 months for the CI24R(CS) device, and following the rational presented in the previous section, yielded an error estimate of 9.4 percentage points. This result should be compared with the measurement error of 11.4 percentage points, calculated from the test-retest data, after dividing this latter percentage by $\sqrt{2}$ because the scores at each observation time were based on two measurements. Thus, the error between sessions at different observation times was 9.4 percentage points while the measurement error within a session equalled $11.4/\sqrt{2} = 8.1$ percentage points. Similarly to the analysis in the previous section this difference suggests that there was a fluctuation in performance from one session to the next. This fluctuation was calculated at 4.8 percentage points. (The fluctuation for CVC words was estimated at 4.1 percentage points.) The error between sessions of 9.4 percentage points, including the fluctuations in performance, was related to the presentation of two sentence lists per session, taking about 5 to 6 minutes.

5.3.3 Results from the connected discourse tracking test

Figure 5.7, upper panel, shows the tracking speed for connected discourse presented in auditory-only mode for five types of implant devices over a period of one year. Figure 5.7, lower panel, shows the results collected in visual-only (V, speech reading) and auditory-visual (AV) mode. For each type of implant device and each presentation mode subjects were included only if data for all observation times indicated in the respective curves were present. The results show that tracking speed tended to stabilise at or even before 12 months. Thus, it seems that these measurements do not have to be continued for more

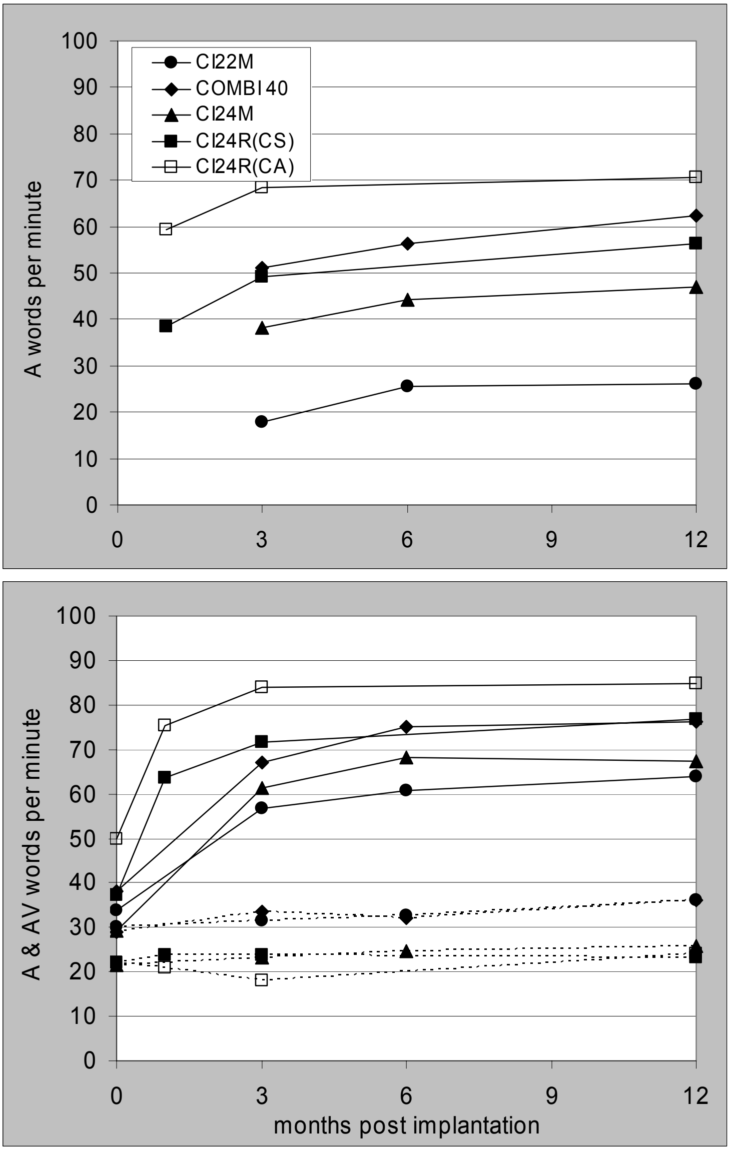

Figure 5.7. Upper panel: Time course after implantation of the average scores in words per minute for connected discourse tracking in auditory-only (A) mode. Lower panel: data collected in visual-only mode (V, speech reading, lower set, dashed curves) and in auditory-visual mode (AV, upper set, solid curves). Results for five types of implant devices. Number of subjects per device: CI22M, N=21; COMBI 40, N=12; CI24M, N=54; CI24R(CS), N=67, CI24R(CA), N=6.

than one year after implantation. Tracking speed stabilises markedly earlier than phoneme scores for CVC words and syllable scores for sentences do. Thus, the latter part of the increase in phoneme and syllable recognition does not seem to contribute to the speed by which connected discourse can be followed. Language-based skills may dominate tracking speed after one year.

The average speech reading speed (visual-only mode) was about 30 words per minute for the devices first implanted (CI22M and COMBI 40) down to about 20 words per minute for devices implanted more recently. In auditory-only mode there were marked differences per device. At 6 to 12 months the tracking speed was 25 words per minute for the CI22M device, while it was 70 words per minute for the CI24R(CA). Adding visual input to the auditory input enhanced the CI22M scores by 35 words per minute and the CI24R(CA) scores by 15 words per minute, with intermediate results for the other devices. In the previous section we found an exceptionally large contribution from visual input to the syllable-in-sentence scores for the CI22M device. Here, this effect does not seem to be present in the tracking speed found for this device. However, it is not clear how one should "add" tracking speeds. The auditory-visual scores showed that tracking speed approached the result found for normal-hearing people (80 words and up) when the newest device in the set studied was used, the CI24R(CA), whereas the scores for the other devices did not reach normal tracking speed. The differences between the results for different implants followed the same pattern as reported before for the word and sentence tests. Analysis of variance of the auditory-only data collected 12 months post implantation showed that the CI22M scores were significantly ($p<0.02$) lower than those for the COMBI 40, CI24R(CS), and CI24(CA). The other differences were insignificant ($p>0.1$).

Figure 5.8 presents the distribution of the connected discourse tracking scores collected in auditory-only mode for the CI24M (N=54) and CI24R(CS) (N=67) implants, similarly to figures 5.2 and 5.5. The tracking scores also show serious differences in performance.

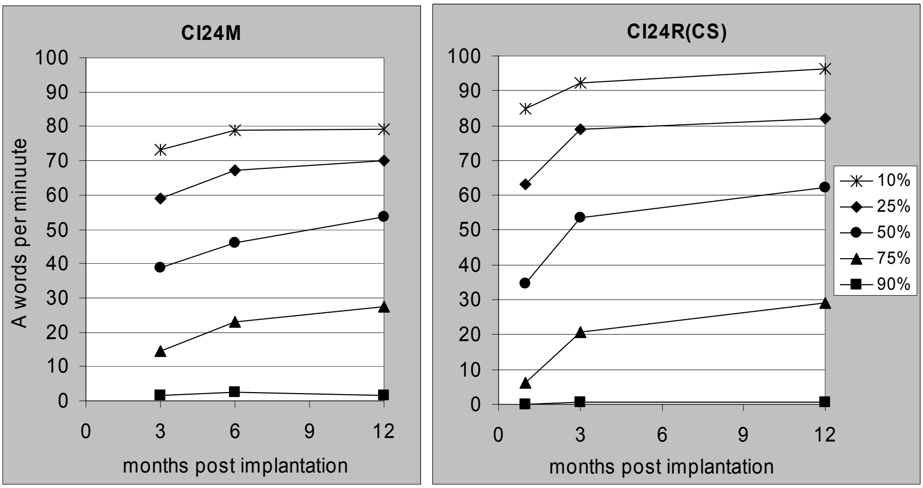

Figure 5.8. Distribution of post-implantation tracking scores for connected discourse presented in auditory-only mode, shown for the CI24M (N=54) and the CI24R(CS) device (N= 67). The percentiles indicate the fraction of results with scores above the respective curves.

While 10% of the CI24M users and 25% of the CI24R(CS) users reached the performance of normal-hearing people, another 25% stayed below 30 words per minute and 10% clearly had problems completing the task at all with both the CI24M and CI24R(CS) devices. After 12 months the lower quartile in figure 5.8 may still show some increase in performance but the lowest decile does not show any upward trend. Again, these results are derived from complete data sets for each individual across the whole post-implantation period indicated in the figure, avoiding artefacts that may be introduced by drop-outs during the observation period.

The effect of the aetiology of deafness on tracking speed measured 12 months after implantation in auditory-only mode was analysed using the six aetiological categories listed in section 5.3.1. Again the analysis was conducted for the two implant groups with a sufficient number of subjects: those using the CI24M and CI24R(CS) devices, each containing about 50 subjects. ANOVA was performed including *type of implant* and *acoustic hearing aid usage* as factors in the analysis in order to avoid confounding interactions. ANOVA showed a significant effect of *aetiology of deafness* on the scores ($p<0.05$). The results for recipients deafened by Cogan's disease were clearly higher than average. The tendency of lower scores in recipients deafened by meningitis, found for the phoneme-in-word scores (Sec. 5.3.1), was less clear in this speech tracking task. The effect of *age at onset of deafness* and *duration of deafness*, from its onset to the time of implantation, on the tracking speed measured 12 months after implantation in auditory-only mode was analysed applying multiple regression analysis (see Section 5.3.1). The partial correlation between each of these two variables and tracking speed was statistically insignificant ($p>0.2$).

Measurement error was estimated in the same fashion as discussed for the phoneme-in-word scores. The differences between the scores collected 6 and 12 months after implantation for the CI24M device (N=45) and between the scores collected 3 and 12 months after implantation for the CI24R(CS) device (N=57) were used. Clearly, these scores did not completely stabilise in these periods. The average difference between the scores at the two sample times were 3.2 words per minute for the CI24M device and 8.1 words per minute for the CI24R(CS) device. The standard deviations calculated for the individual differences were 9.0 and 12.5 words per minute, for the CI24M and CI24R(CS) devices, respectively. After eliminating the change in performance between the two sessions, given by the average differences, from these calculated standard deviations (see Sec. 5.3.1) the error estimates became 8.6 and 10.6 words per minute, respectively. Combining the two estimates yielded a measurement error of 9.8 words per minute. The average tracking speed was 58 words per minute. Thus, the relative measurement error was about 17%. Rather than taking the relative error one might prefer to consider the error with respect to the range of tracking speeds covered by the data, because this gives some indication of the sensitivity of the test. The tracking speed covered the range from 0 to about 100 words per minute. Thus, in terms of numbers, the range of the tracking speeds (0-100) is about equal to the range of the phoneme and syllable scores (0-100%), which implies that one may compare the accuracy of the three tests in terms of relative measurement error. The error of 17% was related to an average duration of the connected discourse tracking test of 6.5 minutes.

5.4 Discussion

5.4.1 Performance after implantation

The phoneme scores collected with CVC words and the syllable scores collected with sentence materials showed an increase in performance over a two to three year period after implantation. This was seen in the average scores (figures 5.1 and 5.4) and over the full range of percentiles (figures 5.2 and 5.5). Thus, cost-benefit analyses should take into account this period in which hearing develops. The connected discourse tracking scores stabilised already within 12 months. Thus, the increase in the phoneme and syllable scores still present after one year was not clearly reflected in the tracking speed of connected discourse. This suggests that language based skills may dominate tracking speed one year after implantation and beyond.

The three types of measurements showed a systematic increase in score with the development of implant technology in the Nucleus devices. Although the differences collected with a particular test did not all reach statistical significance, the consistent trend across the three tests does suggest that this increase truly reflects an improvement in performance in line with technological development. The low scores for the device marketed first, the CI22M, may partly be due to candidate selection because at that time candidates received an implant only if they were absolutely deaf, and were not using a hearing aid. The COMBI 40 showed scores which were relatively high for the time that device was marketed.

The average scores found in children (figure 4.1) were very close to those reported here for adults. The average score for the CI24M device, four years after implantation, was about 50% for both groups of implant recipients; the average score for the CI24R(CS) device, after 2 years, was almost 60% in both groups. It is interesting to note that after four years the average score for the CI22M reached about 50 % in children, whereas it remained at about 30% up to ten years after implantation in adults. This may be due to the initial selection of adult candidates mentioned above. However, it may also be related to children having less hearing experience at the time of implantation and (therefore) being more flexible in learning to use the simpler sound processing strategies of the early speech processors.

Although the increase in scores with technological development is encouraging one faces a serious problem when the distribution of the scores across the implanted population is examined (figures 5.2, 5.5, and 5.8). The interindividual differences are very large. It is almost impossible to give an implant candidate a proper prognosis of performance after implantation. Several dedicated studies provide evidence that factors such as aetiology, age at onset of deafness, and duration of deafness, from its onset to the time of implantation, affect the success of implantation. However, the present retrospective analysis of clinical data showed that the effect of aetiology of deafness was small and that the effects of age at onset of deafness and duration of deafness did not reach statistical significance. The large interindividual differences found in this analysis could not be attributed to these three factors. These large differences are found in spite of implant candidate selection by a multidisciplinary team, dedicated to cochlear implantation, and operating according to a meticulously drawn selection protocol. The present result suggests that more attention should be paid to improving the poor performers.

Looking at the scores collected in auditory-visual mode it was clear that, on average, implant recipients still performed substantially better when they could also see the speaker. In that case the syllable-in-sentence score reached values from 80 to almost 100% and tracking speed reached values from 60 to 85 words per minute. The latter value touched the scores found for normal-hearing people.

A number of implant recipients continued to use their acoustic hearing aid after cochlear implantation in the contralateral ear. The results for those who could be followed up to 60 months post implantation, and who were subjected to the word test in acoustic-hearing-aid-only and acoustic-hearing-aid-with-implant conditions, showed that the number of acoustic hearing aid users had decreased by 50% three years after implantation. Between three and five years there was not much change in hearing aid usage within this group of implant recipients. On average, the contribution from the acoustic hearing aid to the implant score was limited to a 5% increase in phonemes perceived correctly while this score was about 12% when the acoustic hearing aid was switched on and the implant switched off.

5.4.2 Comparison of the three tests

The Bosman-Smoorenburg CVC word reception test did not show floor and ceiling effects in the upper and lower deciles of the phoneme scores collected in auditory-only mode for both the CI24M and CI24R(CS) devices (figure 5.2). The risk of floor and ceiling effects was somewhat larger for the syllable-in-sentence test (figure 5.5). Thus, the CVC test has somewhat greater applicability than the sentence test, particularly when the auditory-visual presentation mode is included. The connected discourse tracking test did not suffer from ceiling effects but at least 10% of the implant recipients faced problems in performing the task at all.

Although the tests reflect different aspects of speech perception by implant recipients one might wish to reduce the number of tests. The scores in auditory-only mode were highly correlated: R=0.92 between sentences and connected discourse, R=0.85 between CVC words and connected discourse and R=0.87 between CVC words and sentences. Thus, the CVC word test is the most independent one of the three tests. When reducing the number of tests, measurement accuracy could be important in deciding which test to select. For example, measurement accuracy is important when one wishes to examine the change in performance over time. Accuracy of the connected discourse tracking task was 9.8 words per second at an average speed of 58 words per minute; a relative error of 17% related to a test duration of 6.5 minutes. The syllable-in-sentence score reached an accuracy of 9.4 percentage points in about the same time, 5 to 6 minutes. This error occurs at a 50% score; hence the relative error is 19%, a value similar to the previous one. The accuracy of the phoneme-in-CVC-word test was 5.1 percentage points in 10 minutes, consisting of an intrinsic error of 3.0 percentage points and an error ascribed to fluctuations in performance level between sessions of 4.1 percentage points. In 6 minutes, presenting 6 rather than 10 lists, we may assume that the intrinsic measurement error increases by $\sqrt{}10/6$. Also, we may assume that the error due to performance fluctuations between sessions remains the same. Combining the two sources of error, adding the variances, the total error for 6 minutes measuring time was estimated at 5.6 percentage points. Thus, the relative error became 11%. The latter error was considerably smaller than the error found for the two other tests. This suggests that the phoneme-in-word test would be the primary choice if measurement error is of primary concern. The higher accuracy of the CVC word test is also reflected in the results of analysis of variance. The number of significant differences in the scores for different types of implants and different aetiologies of deafness was higher for the CVC word test than for the other two tests (although measurement duration was somewhat longer for this test). Finally, the CVC test showed a smaller risk of floor and ceiling effects. Thus, the CVC word test would be the primary choice if one wishes to reduce the number of tests. The connected discourse test could be viewed as a valuable addition.

The present analysis concerns speech perception tests that were in use already in 1992. It did not include a test addressing speech perception in noisy environments, a test condition that has become more important for newer types of implant devices and speech processing techniques.

Chapter 6

Effect of cochlear implantation on speech production; vowel quality

Summary

Objective

In addition to speech perception performance, it is also important to evaluate the quality of speech production. Clear speech implies that the implant recipient will be understood without demanding too much effort on the part of the listener. If speech production is clear, implant recipient may more easily engage group conversations, which allow opportunities to improve speech and language skills. One way to study the quality of speech production is by measuring differences in vowel formant frequencies, as well as the variability in formant frequencies within repeated vowels. These physical measures may indicate the likelihood that listeners will understand or confuse vowels uttered by cochlear implant recipients. Previously we have reported results for adults. This chapter presents complementary data for children.

Conclusions

In children, between 5 and 9 years of age, contrast between 11 Dutch vowels (no diphthongs) plotted as a function of the frequencies of the first and second formant, F1 and F2, increases substantially during the first two years after implantation. After two years of implant use, relative positions of F1 and F2 are close to vowels produced by normal-hearing persons. Moreover, the variability in formant frequencies for several utterances of the same vowel (pronounced by an individual implant recipient) decreases already during the first year of implant use, and approaches near-normal variability after two years. The larger vowel contrasts (*i.e.*, larger differences between individual vowel formant frequencies), and the reduced variability in the formant frequencies within each vowel, suggest that after an appropriate adaptation period, speech production by cochlear implant recipients is likely to be better understood.

6.1 Introduction

Speech recognition performance is the dominant concern when evaluating the benefits of cochlear implantation. However, the quality of implant recipients' speech production should also be considered. Poor speech quality is often associated with profound hearing loss and deafness, particularly in children. After implantation the recipients are better able to hear their own voice, and thereby produce better voicing and articulation, resulting in better speech production quality. Intelligibility of implant recipients' speech is not only important on the part of the listener, but it may also contribute to improving implant recipients' overall language and speech perception skills. With clear, well-understood speech implant recipients may engage more easily group conversations; active participation in conversations may provide the training needed to improve speech perception, speech production and language acquisition with the cochlear implant.

The quality of speech production can be assessed in many ways. Previously, we reported on the effects of cochlear implantation on vowel production, including vowel quality, vowel intelligibility, voice fundamental frequency and nasality (Smoorenburg et al., 1994, Langereis et al., 1994, 1995, 1997a, 1997b, 1997c, 1998, 1999). These evaluations were made with adult implant recipients. Here, we complement these reports with physical measurements of vowels produced by implanted children.

6.2 Methods and materials

Six pediatric implant subjects (3 male, 3 female; 5 - 9 years of age) participated in the study. Three were born deaf, and the other three were deafened after a meningitis infection between ages 1 year; 4 months, 2 years, and 2 years;7 months. Pre-implantation auditory thresholds were higher than 120 dB HL at most frequencies; low frequency thresholds were higher than 105 dB HL. All subjects used the Nucleus CI22 implant and MSP processor.

Speech samples were recorded prior to implantation and 1-2 years post-implantation. Subjects were asked to produce three utterances each for 11 Dutch monophthongs (/u/, /o/, /ɔ/, /ɑ/, /a/, /ɛ/, /e/, /I/, /i/, /y/, /Λ /); notation of the International Phonetic Association) in a CVC context. In Dutch, these vowels occur in the words (hoed, boot, bot, kat, kaas, mes, peer, kin, mier, muur, and mus); in British (UK) or in American-English (US) in the words (UK booed, US bode ,UK bought, UK father, US buy /ai/, UK bed, US bayed, UK bid, UK bead, - , UK bud). Subjects produced these speech samples after being asked to name an object presented to them.

Vowel contrasts and vowel variability were analysed in terms of the frequencies of the first and second formant, F1 and F2. The vowels were segmented from the utterances and F1 and F2 were extracted using Linear Predictive Coding (LPC) analysis. F1 and F2 frequencies for all subjects were scaled by a multiplicative factor to the group averages, thereby eliminating effects of less relevant inter-subject differences (e.g., differences in vocal cavity among the children).

6.3 Results

6.3.1 Vowel contrast

Figure 6.1 shows mean (across subjects and utterances) F2 frequency as a function of F1 frequency for 11 Dutch vowels; the thin characters and lines show speech production prior to implantation and the bold characters and lines show production 2 years post-implantation Note that the vowel space has expanded significantly after two years of experience with the implant. Post-implantation data show that for the upper branch of the vowel triangle (from /a/ to /i/), F1 frequencies are considerably lower than before implantation. For the lower branch (from /a/ to /u/), both F1 and F2 frequencies are considerably lower than before implantation. The vowels /y/ and /Λ / also have become more distinct. After two years of implant use, F1 and F2 frequencies approach those of normal-hearing paediatric talkers. These results suggest that vowels produced by these children will be better discriminated by the listener; there will be less vowel confusions.

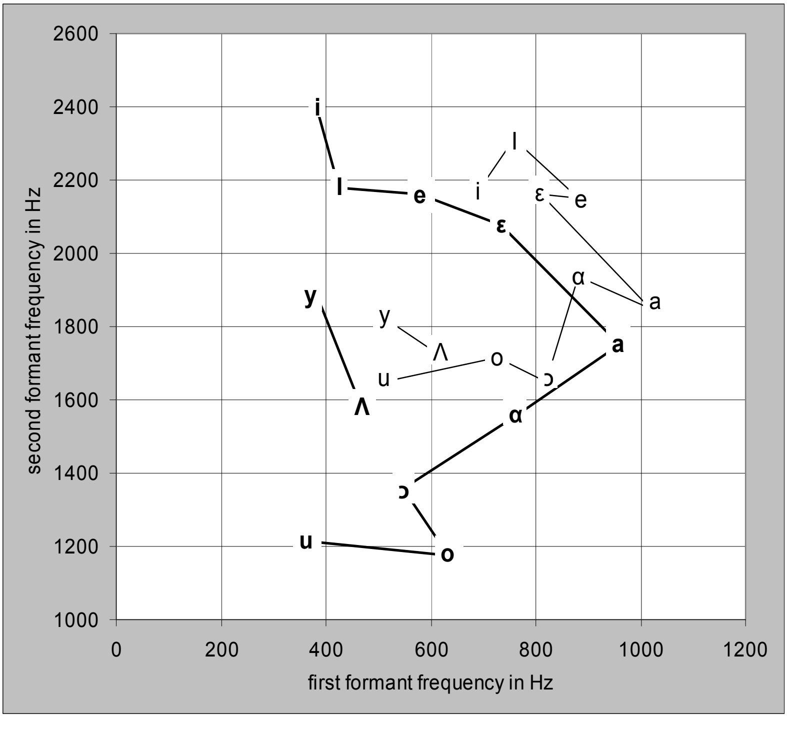

Figure 6.1. Group mean frequencies of the first and second formants, F1 and F2, for 11 Dutch vowels uttered by 6 children before (thin characters and thin lines) and two years after implantation (bold characters and bold lines). Vowels along the branches of the vowel triangle (with corner points /u/, /a/, /i/) and the vowels /y/ and /Λ / are connected by lines to facilitate comparison between pre- and post-implantation speech production. The vowels (/u/, /o/, /ɔ/, /ɑ/, /a/, /ɛ/, /e/, /I/, /i/, /y/, /Λ /) were produced in a CVC context, corresponding to the Dutch words (hoed, boot, bot, kat, kaas, mes, peer, kin, mier, muur, mus).

6.3.2 Vowel variability

Besides changes in vowel contrast it is also important to evaluate the variability in speech production, before and after implantation. If production of a vowel changes with each utterance, then vowel confusions are likely to occur and speech will be more difficult to understand. The increase in the differences between formant frequencies after implantation (figure 6.1) may reduce the number of vowel confusions but if the variability in formant frequencies would increase with the increase in the differences between formant frequencies then there may be no reduction in vowel confusions. Therefore, we calculated formant frequency variability for the three utterances of each vowel. The variance per vowel was then summed across vowels.

Figure 6.2 shows the variability in F1 (top panel) and F2 (bottom panel) for each subject, before, one year and two years after cochlear implantation. The variability is shown in terms of the standard deviation (*i.e.*, the square root of the mean variance across all vowels), in Hz. Significant reductions in variability between pre- and post-implantation data are indicated by * ($p<0.05$) or ** ($p<0.01$). Note that there was no subject with a significant increase in variability after implantation; for four subjects F1 variability was significantly reduced. Subject 4 shows a non-significant increase in F1 variability, but note that F1 variability before implantation was already relatively small for this subject. Subject 5 shows no change in first formant variability after one year although the pre-implantation variability for this subject was large. The recordings for this subject made two years after implantation were incomplete. F2 variability was significantly reduced for half the subjects after implantation, while the other subjects showed a similar trend.

In summary, there is a general trend toward smaller variability in formant frequencies after implantation, approaching the variability observed with normal-hearing talkers. Along with the improved vowel contrasts, the reduced variability in speech production suggests that as children gain more experience with their implant device, their speech becomes clearer and easier to understand.

6.4 Discussion

Because of the limited number of subjects, the variability in formant frequency was summed across all vowels. However, the variability may significantly differ from one vowel to the next. For example, there was greater variability in F1 frequencies across utterances for vowels /I/ and /y/. For the vowel /y/, this variability may be less critical, given its position in the vowel contrast map shown in figure 6.1. However, the cross-utterance variability in /I/ could result in perceptual confusion with /e/.

Shortly after implantation, it is possible that quality of vowel production by paediatric implant recipients may decrease, as children need adequate time to adapt to the new electrically stimulated auditory patterns. Also, given its deafness, the child might have been taught to produce a set of vowels optimally distinguishable, but with formant frequencies somewhat different from the normal ones. After implantation, children may learn to correct their vowel production because they hear both; their vowels and those from others. However, such a possible decrease in vowel quality did not occur in a statistically significant way in our data one year after implantation.

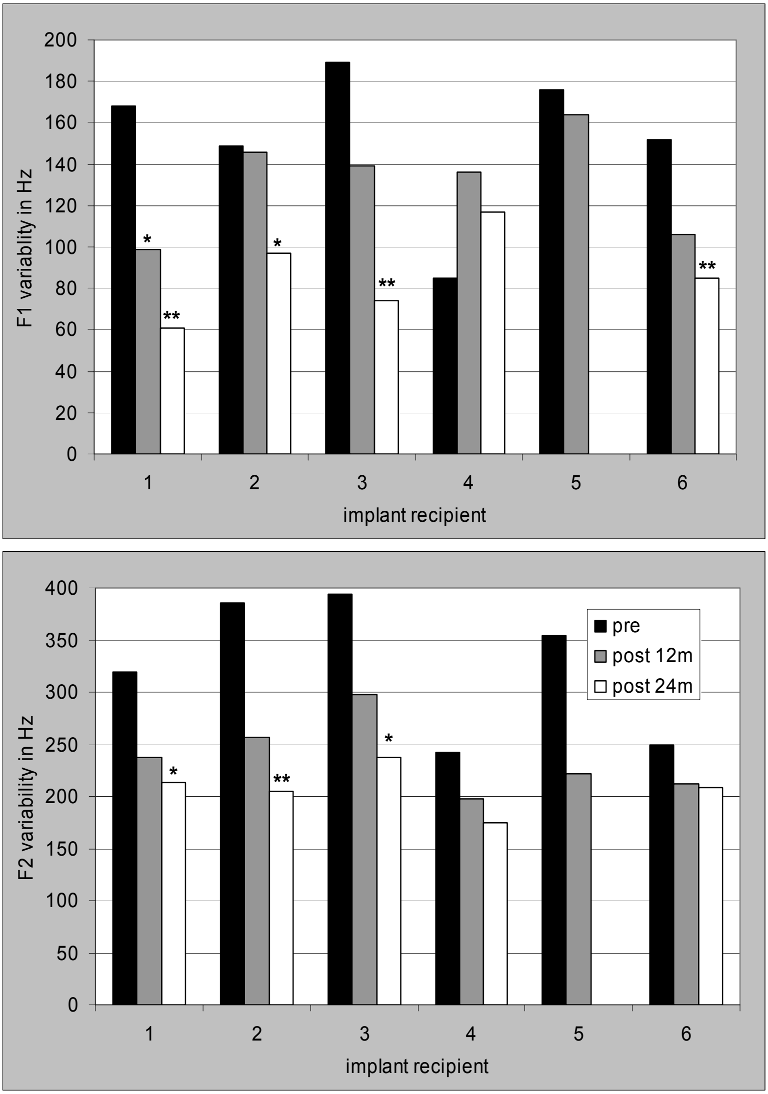

Figure 6.2. Mean variability across 11 Dutch vowels in first formant frequency (F1, upper panel) and second formant frequency (F2, lower panel), for three utterances of each vowel. Individual data are shown for speech production pre-implantation, one and two years after implantation. Significant differences in variability after implantation are indicated by *, $p<0.05$, and **, $p<0.01$. Variability is expressed in the standard deviation of the differences (not the standard error of the mean). Data for recipient 5, two years after implantation, were incomplete.

6.5 Complementary reading

Smoorenburg, Guido F.,Huiskamp, T., Langereis, M. and Bosman, A.J. (1994). Effects of cochlear implants on voice quality and speech production, Advances in Cochlear Implants (L.J. Hochmair-Desoyer and E.S. Hochmair, Eds.), Manz, Wien, 374-379.

Langereis, M.C., Bosman, A.J., van Olphen, A.F. and Smoorenburg, G.F. (1995). Changes in vowel quality in adult cochlear implant users, Ann Otol Rhinol Laryngol 104, Suppl 166, 387-390.

Langereis, M.C. (1997a). Effects of Cochlear Implantation on Speech Production, Dissertation Utrecht, ISBN:90-75188-11-0

Langereis, M.C., Bosman, A.J., van Olphen, A.F. and Smoorenburg. G.F. (1997b). Changes in vowel quality in postlingually deafened cochlear implant users, Audiology 36(5), 279-297.

Langereis, M.C., Dejonckere, Ph., van Olphen, A.F. and Smoorenburg, G.F. (1997c). Effect of cochlear implantation on nasality in post-lingually deafened adults, Folia Phoniatrica et Logopaedica 49, 308-314.

Langereis, M.C., Bosman, A.J., van Olphen, A.F. and Smoorenburg, G.F. (1998). Effects of cochlear implantation on voice fundamental frequency in post-linguistically deafened adults, Audiology 37, 219-230.

Langereis, M.C., Bosman, A.J., van Olphen, A.F. and Smoorenburg, G.F. (1999). Intelligibility of vowels produced by post-linguistically deafened cochlear implant users, Audiology 38, 206-224.