Audiology
Editor-in-Chief: J.-M. Aran, Bordeaux

Separatum
Publisher: S. Karger AG, Basel
Printed in Switzerland

Original Paper

Audiology 1995;34:145–159

Multicentric Field Evaluation of a New Speech Coding Strategy for Cochlear Implants

Norbert Dillier[a]
Rolf D. Battmer[b]
Wolfgang H. Döring[c]
Joachim Müller-Deile[d]

[a] ENT Department, University Hospital, Zürich, Switzerland;
[b] HNO-Klinik der Medizinischen Hochschule Hannover,
[c] HNO-Klinik der RWTH, Aachen, und
[d] HNO-Klinik der Universität, Kiel, Deutschland

Abstract

In a multicentric study involving 4 European cochlear implant centers, the speech perception abilities of 20 native German-speaking individuals implanted with the Nucleus 22 Channel Cochlear Implant System when using a new spectral peak (SPEAK) speech coding strategy were investigated. This strategy continuously analyzes the speech signal using 20 digital programmable bandpass filters and presents up to 10 spectral maxima to the 22 implanted electrodes. Each subject's performance on a variety of auditory perceptual tasks was evaluated with the experimental encoder (SPEAK), relative to his or her performance in a reference condition. An ABAB experimental design was used whereby each strategy was reversed and replicated. The reference levels of auditory performance were established using the multipeak (MPEAK) speech-processing strategy of the Nucleus speech processor. Only subjects who achieved open-set monosyllable word recognition in the reference condition were included in this study. Significant differences in group mean scores for most speech recognition subtests were obtained for the SPEAK versus the MPEAK strategy. The largest overall improvements were observed for the sentence tests under noisy conditions.

Key Words
Cochlear implants
Speech coding
Spectral analysis
Speech tests
Multicentric evaluation

Introduction

Cochlear implant coding and processing strategies are designed to overcome the limited channel capacity of the artificially stimulated auditory nerve by selecting the necessary and useful portions of the speech information and to transform acoustic signals into electroneural stimulation according to individual topological requirements for loudness and

Received:
November 29, 1994
Accepted:
April 19, 1995

Dr. N. Dillier
ENT Department
University Hospital
CH–8091 Zürich (Switzerland)

0020–6091/95/
0343–0145$8.00/0

pitch perception while avoiding electrode interaction and current summation effects.

Early attempts to transmit speech information via cochlear implants had used only one [1, 2] or maximally four active electrodes [3, 4], stimulated with analog waveforms which were filtered and amplified replicas of the original speech sounds. The first multichannel device with explicit coding of the most important speech features F_0 (the fundamental voice frequency) and F_2 (the second formant frequency which distinguishes most spoken vowels) used a pulsatile stimulation pattern on 10 bipolar electrode pairs [5].

Subsequent improvements of this system resulted in the currently most widely used device, the Nucleus 22 Channel Cochlear Implant System [6], which can be programmed for a variety of different stimulus conditions and processing strategies [7–10]. The use of pulsatile stimulation schemes for multichannel cochlear implants has now gained wide acceptance due to remarkable improvements in speech recognition for high-rate pulsatile excitation as compared to analog stimulation [11–13].

Two different coding strategies for the Nucleus 22 Channel Cochlear Implant System were considered in this investigation: the multipeak (MPEAK) strategy, which was introduced in 1989 with the Nucleus miniature speech processor (MSP) [14], determines the frequency and amplitude of the first two formants as well as the energy in two higher frequency bands and generates biphasic stimulation pulses on up to 4 out of 22 possible electrodes at a repetition rate dependent on the voice fundamental frequency. The frequency to electrode mapping for the second formant overlaps the mapping of the higher frequency bands.

The spectral peak (SPEAK) strategy, which is based on the Spectral Maxima Sound Processor (SMSP) strategy [15], divides the audio spectrum into 20 frequency bands and continuously selects up to 10 prominent spectral components which generate interleaved stimulation pulses on up to 10 different electrodes within each analysis interval. The rates of stimulation on each channel vary between approximately 180 and 300 Hz depending upon the spectral composition, the sound intensity and the individual's speech processor program. The frequency to electrode mapping follows the normal tonotopic order in the cochlea. Note that this strategy avoids explicit extraction and coding of speech features such as voice fundamental frequency or formant frequencies.

The MPEAK and SPEAK coding strategies were evaluated in a multisite field study involving 8 cochlear implantation centers in English-speaking countries [16] and 8 centers in non-English-speaking countries. The results for 20 subjects from 4 German-speaking cochlear implantation centers are presented in this report.

Methods

Subjects

The selection criteria for participation in this field trial were as follows. Only adult postlingually deafened German-speaking users of the Nucleus Mini-22 Cochlear Implant system were considered as candidates. They should have at least 16 active electrode channels and have had 9 months or more of experience with the MPEAK processing strategy. They should also have used their speech processor on a regular basis of 12 h or more per day, and have reached a monosyllabic word recognition with their MSP of 10% or above. These criteria were established in order to ensure that all tests could be performed by all subjects and that the postimplantation learning curve had already reached a plateau.

The 4 implantation centers, located at the university hospitals of Aachen, Hannover, Kiel and Zürich, agreed upon the same experimental protocol and selected a total of 20 subjects (4, 5, 7 and 4 subjects, respectively). Preliminary results for 16 of the 20 subjects have been reported earlier [17, 18]. Subjects had

to sign an informed consent form and were reimbursed for travel expenses. They were also given a small honorarium for the test sessions at the end of the study.

Test Material

Prerecorded vowel, consonant and monosyllabic word tests [19] as well as two different sentence tests (Innsbruck [20] and Göttingen [21] sentence lists) in quiet and noise were presented from digital audiotape or directly from computer-disk (vowel and consonant tests in Aachen and Kiel, monosyllabic word test in Aachen, Kiel and Zürich) in a sound-treated room via a loudspeaker. The sequence of test lists to be used for each subject was assigned before the start of the experiment. No speechreading tests were included, firstly because the purpose of this study was the comparative evaluation of the auditory performance of the two strategies and secondly because standardized recorded audiovisual test material was not available in German.

The vowel and consonant logatome test comprised 3 randomized repetitions of 8 vowels (/a, o, u, e, i, ä, ö, ü/) in /dV/ context (denoted as VO8) and 3 randomized repetitions of 12 consonants (/p, t, k, b, d, g, m, n, l, r, s, f/) in /aCa/ context (denoted as C12), respectively, spoken by a male voice. The Freiburg monosyllable word test (denoted as 1-Syl) contains 20 lists of 20 German words each and is available on compact disc recording. As mentioned above, some clinics used computerized versions of this test sampled with 16-bit resolution at 22 kHz. Two lists (a total of 40 words) were used for each session. A few lists which were known to produce somewhat unbalanced responses were not used for testing but only as practice lists during the short familiarization after processor switching. Correctly repeated words were scored as percent correct.

The Innsbruck sentence test consists of 10 lists of 10 short sentences with an average number of 5 words per sentence spoken by a female voice. The Göttingen sentence test consists of 20 lists of 10 short sentences with an average number of 5 words per sentence spoken by a male voice. Two lists (about 100 words) were used per condition for both the Innsbruck and the Göttingen sentence test and the number of correctly recognized words were scored as percent correct. All sentences were presented only once without repetition during a test session. As 3 conditions had to be tested in 4 sessions, it was inevitable that some lists had to be repeated during the course of the trial. Although a larger number of sentence lists would have been highly desirable, it was decided to use the existing material which had been evaluated in previous studies instead of creating new lists or rerecord a whole new sentence test. The preassignment of lists to subjects, sessions and conditions was therefore carefully done in order to maximize the time interval between repeated lists. The first test condition was speech without noise (denoted as In-Q and Gö-Q, respectively), the second condition speech and noise at 10 dB signal-to-noise ratio (SNR) (In-10 and Gö-10, respectively). The third test condition was dependent on the subject's results with the MPEAK strategy in the second condition (at 10 dB SNR) of the first session (A1). If the score was above 50%, then a SNR of 5 dB was selected for the third condition in all experimental sessions. Otherwise, a SNR of 15 dB was selected for the third condition. As will be shown in more detail in the Results section, many of the subjects (12 out of 20) could be tested at 5 dB SNR with the Innsbruck sentences whereas with the Göttingen sentences, the proportion was approximately reversed (6 out of 20). An acoustic analysis of the test material showed that the Innsbruck sentences were spoken rather slowly, at about 121 syllables per minute, whereas the Göttingen sentences were spoken more than twice as fast, with an average speed of 279 syllables per minute [22]. The Innsbruck lists contained simple context-rich everyday sentences whereas the Göttingen lists were constructed as more complex sentences with less contextual information.

Within each experimental session, subjects were also asked to rate the musical quality of short pieces of music (60 s duration) and to identify musical instruments. These tests were added to the protocol because it was hypothesized that the SPEAK strategy which was not designed specifically for speech recognition would provide a more natural reproduction of non-speech signals than the MPEAK strategy which explicitely attempts to extract speech features such as the voice fundamental and formant frequencies. The first of the 12 recorded pieces (Glenn Miller orchestra) was used for setting the listening volume as well as for practice. The next 5 pieces played were mostly instrumental solos (trumpet, piano, violin, guitar, clarinet), and the last 6 pieces consisted of songs, chorales, symphonic music and instrumentals. The musical quality of all 11 test pieces was to be judged on a subjective scale from 1 (very unnatural) to 7 (very natural). The identification of instruments was requested only for the 5 pieces with instrumental solos. Responses in the same instrument category were counted as correct (for instance, violoncello instead of violin or trombone instead of trumpet).

The presentation level was 70 dB SPL for all recorded speech (sound level meter set to linear fast peak) and music (sound level meter set to linear fast

Table 1. Schedule of audiological evaluations

Session	Week No.	Experimental condition	Evaluation
A1	1	MPEAK	speech and music tests, performance and tinnitus evaluation
B1	7	SPEAK	speech and music tests, performance evaluation
A2	9	MPEAK	speech and music tests
B2	11	SPEAK	speech and music tests, performance and tinnitus evaluation

RMS) material. The input sensitivity control of the subject's speech processor was kept unchanged at the standard everyday listening position during the experiments. For speech tests in noise, the level of the speech signal was constant and the noise level was varied. Two different noise recordings were used for the two different sentence tests. For the Innsbruck sentences, a standardized speech spectrum shaped noise was used [23] which was obtained from a commercial audiometry compact disc recording. The noise for the Göttingen sentence test had been generated by statistically averaging all test items of monosyllable test spoken by the male speaker who had spoken the Göttingen sentences. Thus, the long-term spectrum of the masking noise should perfectly match the long-term spectrum of the sentences for the Göttingen test, whereas for the Innsbruck sentences which were spoken by a female voice, the match would not be perfect.

Experimental Protocol

Each subject's performance with the experimental encoder (SPEAK) was evaluated relative to his or her performance in a reference condition, on a variety of auditory perceptual tasks. The reference levels of auditory performance were established using the MPEAK speech processing strategy of the MSP.

The electronic circuitry for the new encoder was embedded in a body-worn processor case identical to that of the MPS. The experimental processor was labelled MSP+ on the underside of the case and given to the subjects during the trial phases with the SPEAK strategy. Although the MSP+ could have been programmed for the MPEAK as well as the SPEAK strategy, the subjects always used their own MSP during the reference phases. The subject's headset (microphone and transmitter coil) was identical for both the MPEAK and the SPEAK processors. Before the first evaluation session, the T- and C- levels (threshold and comfortable loudness levels) of the MSP map were verified by loudness balancing and adjusted if necessary. The same map was subsequently used to program the SPEAK strategy with default parameters as suggested by the fitting software. Only minor adjustments were made after the initial setup of the SPEAK strategy. The new strategy was often characterized as providing different but more complete sound impressions. Some subjects complained about too much high-frequency emphasis and unnaturally high voice pitch. By lowering the stimulus levels of some basal electrodes, these problems could usually be solved in a short time.

As table 1 indicates, all subjects received both the reference (MPEAK, sessions A1 and A2) and experimental (SPEAK, sessions B1 and B2) conditions twice, in a ABAB paradigm, over an 11-week period. Switching from the MPEAK to the SPEAK processor took place after the test sessions A1 and A2, switching back from SPEAK to MPEAK after sessions B1 and B2. Initial readjustment to SPEAK, if required at all, was carried out in an intermediate session between A1 and B1 in week 2. Thus, the subjects had been using the SPEAK processor for 6 weeks before the first trial sessions (B1) and again two weeks before the second trial session (B2). No formal training sessions were conducted. Some subjects were given the opportunity to keep the MSP+ after the end of the trial in order to collect more data on possible long-term learning effects. Results of these further investigations are not part of this study and will be reported later.

Questionnaires were used to assess effects on tinnitus (at sessions A1 and B2) and subjective performance rating (at sessions A1, B1 and B2).

Statistical Evaluation

It could be expected that the data would contain a lot of variance due to many subject variables which could not be controlled. Thus, a simple factorial analysis of variance was conducted for all subtests to identify the significance of the 3 main factors: (i) Clinic (Aachen, Hannover, Kiel, Zürich), (ii) Strategy/Processor Type (MPEAK or SPEAK), (iii) Phase or Session (A1 and B1 versus A2 and B2). The raw percentage scores were arcsine-square-root transformed to nor-

malize the variances [24]. Average values over sessions, clinics and subjects were all calculated from the transformed values and then inversely transformed to percentage scores for display purposes. Paired samples t-tests were used for group comparisons of the processor effect. Aggregated data were generated for the 4 clinics as well as for the whole set of subjects for all subtests. The rating scales were evaluated using nonparametric statistics (Wilcoxon matched pairs signed rank test). The SPSS for Windows software was used to perform these analyses [25].

Results

Figure 1 shows individual speech test results with MPEAK and SPEAK processors for 3 subjects for all 4 sessions. Subject S6 (fig. 1a) was the top performer in the monosyllable word test and achieved an average score of 85% using the SPEAK strategy (sessions B1 and B2). Subjects S20 (medium performer, fig. 1b) and S17 (bottom performer, fig. 1c) obtained 52.5% and 22.5%, respectively, in the monosyllable word test. Note that S6 was given the 5 dB SNR test condition for both sentence tests as he had obtained greater than 50% correct in session A1 at 10 dB SNR whereas S20 took the Innsbruck sentences at 10 and 5 dB SNR and the Göttingen sentences at 15 and 10 dB SNR. S17, on the other hand, took both sentence tests at 15 and 10 dB SNR. The data for these 3 subjects indicate that there exists a large variation of response patterns which can hardly be interpreted on a subject-by-subject basis. S6, for example, showed poorer performance at the second MPEAK session (A2) than at the first session (A1) while S20 showed both poorer (In-Q, In-10, Gö-Q) as well as improved (In-5) performances, and S17 showed mostly improved performance. The performance with SPEAK, however, seems to be either equal to or better than the performance with MPEAK for all subjects. Ceiling effects are also clearly apparent.

In order to determine the main sources of variance and the significane of these factors, an analysis of variance (simple factor 3-way ANOVA) was performed using the arcsine-transformed scores. It can be seen (table 2) that the factor Clinic accounted for a significant amount of the variance in all but the Göttingen test at 5 dB SNR. The Processor factor accounted for significant amounts of the variance for all but the vowel test (VO8) whereas the Session effect seemed to be negligible except for the Innsbruck test in quiet [note that this value was not significant any more after the data for subject S11 were tentatively excluded from the analysis. This subject had improved his score by 38% from session A1 (42%) to A2 (80%) which was the largest observed difference among all subjects]. None of the factor interactions were significant except for the vowel test, which showed a significant combined effect of Clinic with Processor.

The finding that the Session effect was not relevant for the observed variance can be taken as an indication that none or only negligible learning between the two repeated processor conditions had taken place. Another interpretation which is suggested by the individual data of 3 subjects presented in figure 1 is also possible, namely that the performance with MPEAK had decreased from A1 to A2 and the performance with SPEAK increased from B1 to B2, thereby resulting in an overall insignifcant effect of the Session factor. Examples which would support this interpretation can be found in figure 1a, b while other results of figure 1 suggest that there is no effect at all. To test this hypothesis, separate paired-sample t tests were executed for the two processor types for all speech subtests. While none of the mean differences between session A1 and A2 were significantly different from zero for the MPEAK strategy, it was found that some differences between session B1 and B2 (SPEAK)

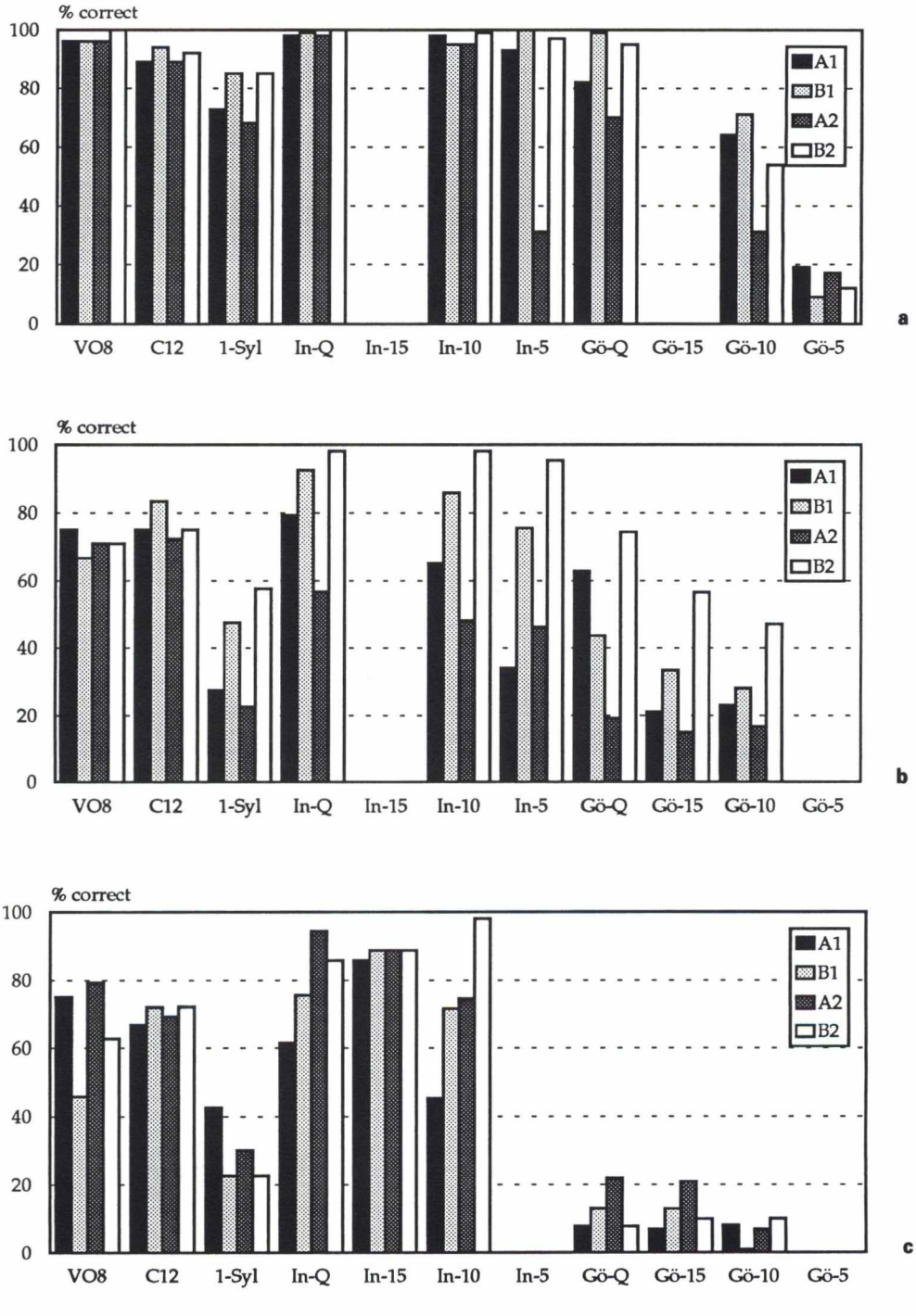

% correct
100
80
60
40
20
0
A1
B1
A2
B2
VO8
C12
1-Syl
In-Q
In-15
In-10
In-5
Gö-Q
Gö-15
Gö-10
Gö-5
a
% correct
100
80
60
40
20
0
A1
B1
A2
B2
VO8
C12
1-Syl
In-Q
In-15
In-10
In-5
Gö-Q
Gö-15
Gö-10
Gö-5
b
% correct
100
80
60
40
20
0
A1
B1
A2
B2
VO8
C12
1-Syl
In-Q
In-15
In-10
In-5
Gö-Q
Gö-15
Gö-10
Gö-5
c

Table 2. Simple-factor 3-way ANOVA for all speech subtests

Subtest	Clinic	Processor	Session	Clinic* Processor
VO8	0.000 ***	0.804 (n.s.)	0.146 (n.s.)	0.000 ***
C12	0.000 ***	0.038 *	0.232 (n.s.)	0.776 (n.s.)
1-SYL	0.000 ***	0.037 *	0.242 (n.s.)	0.730 (n.s.)
In-Q	0.012 *	0.000 ***	0.040 *	0.727 (n.s.)
In-10	0.000 ***	0.000 ***	0.472 (n.s.)	0.054 (n.s.)
In-5	0.001 ***	0.000 ***	0.615 (n.s.)	0.195 (n.s.)
Gö-Q	0.000 ***	0.047 *	0.969 (n.s.)	0.682 (n.s.)
Gö-15	0.000 ***	0.014 *	0.765 (n.s.)	0.055 (n.s.)
Gö-10	0.000 ***	0.008 **	0.558 (n.s.)	0.234 (n.s.)
Gö-5	0.116 (n.s.)	0.022 *	0.212 (n.s.)	0.080 (n.s.)

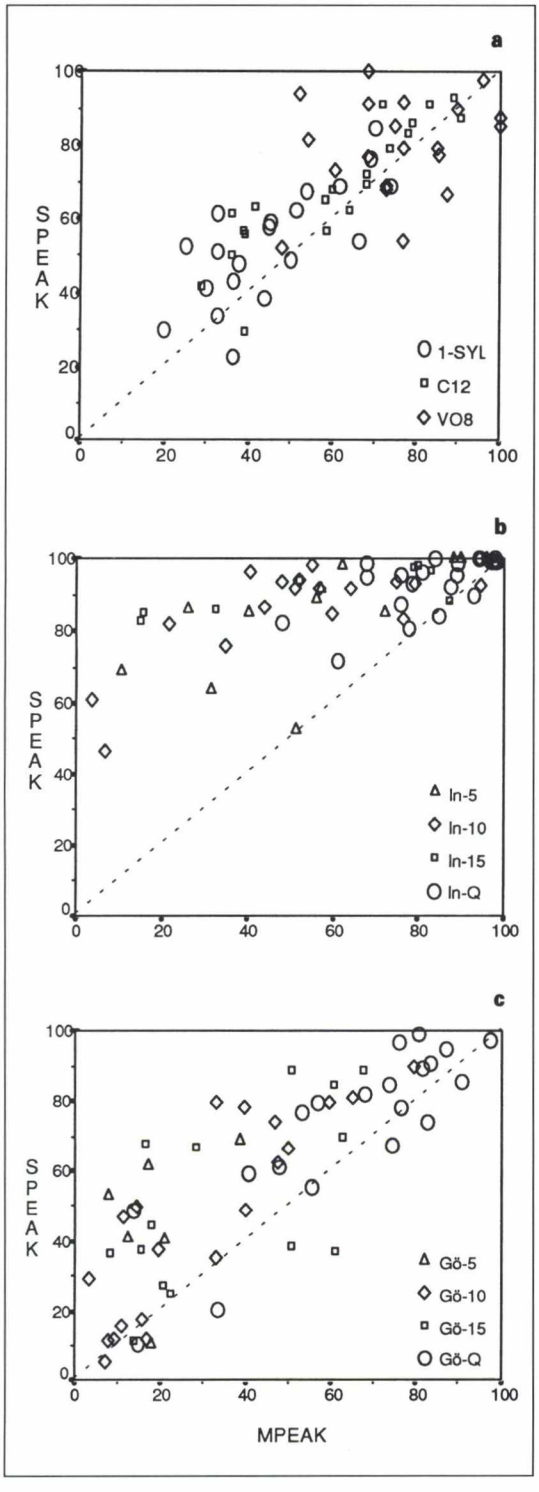
F values for the primary factors Clinic, Processor and Session and the interaction of the factors Clinic and Processor. Significance of F: n.s.: ≥ 0.05, * <0.05, ** <0.01, *** <0.005.

were indeed significant at a level of 5% (VO8, C12, 1-SYL and Gö-10) or even 0.5% (In-Q). This means that the performance with the MPEAK strategy did not change significantly over the time of the study whereas the performance with the SPEAK strategy was improved over time and demonstrated a learning effect.

For the remaining analysis, the data for the two sessions will be aggregated in order to emphasize the main factors of variance. Figure 2 shows scatter plots for all 11 subtests and all 20 subjects. The data cover a wide range of values which is well reflected in the range for the monosyllable word test (20–74% for MPEAK and 22.5–85% for SPEAK) or the Innsbruck sentence test in 10 dB SNR (4–96% for MPEAK and 46–100% for SPEAK). Figure 3 displays the results for the two sentence tests at 10 dB SNR, ordered according to the average score with the SPEAK processor. The numbers 1–20 on the abscissa are the subject identifications. It can be seen that most subjects reached 80% or above with the SPEAK strategy, with differences between processors being as large as 60%. The largest difference in the Göttingen sentence test was 45%. Note that all differences, with one or two minor exceptions, were positive towards the SPEAK strategy.

Table 3 lists group mean differences with two-tailed statistical significances of a paired-sample t test and the number of subjects with improved, unchanged and decreased performance for the 11 subtests. Differences of less than 5% were arbitrarily counted as unchanged performance in table 3. Note that ceiling effects were responsible for some of the small differences found. The mean values

Fig. 1. Individual speech test results with MPEAK and SPEAK processors for 3 subjects. **a** Results for subject S6 (top performer in monosyllabic word test). **b** Results for subject S20 (medium performer in monosyllabic word test). **c** Results for subject S17 (bottom performer in monosyllabic word test).



with asterisks indicating statistically significant differences are shown in figure 4. Note that all differences were positive, i.e. the mean results with SPEAK were all higher than with MPEAK. The largest differences were observed for the Innsbruck sentence test (34, 32 and 34% mean difference for a signal-to-noise ratio of 15, 10 and 5 dB, respectively, and 10% for the quiet condition). Mean differences above 10% were also observed for the Göttingen sentences presented in noise (19, 18, 28% for SNR of 15, 10 and 5 dB, respectively) whereas the difference for speech without noise was 10%. All differences between the MPEAK and SPEAK conditions were significant at a level of 0.05 or better (paired-sample t test) with the exception of the vowel test (VO8). It should be noted at this point that the MPEAK results for the vowel test are already remarkable, which may be explained by the explicit formant extraction of this strategy which allows near-perfect identification of the 8 long vowels used in this study.

A second reason for the lack of significant benefit of the SPEAK strategy for vowel identification may be found in the interaction with the Clinic factor as revealed in the ANOVA (table 2). The mean differences have therefore been calculated separately for the 4 clinics. Figure 5 shows mean results of the 4 evaluation centers for 4 subtests. Significant differences as determined by a Wilcoxon matched-pairs signed rank test are indicated by asterisks. The only 2 cases with a negative difference between SPEAK and MPEAK were the vowel tests for centers C1 and C4 (fig. 5a). The difference for center C3, on the

Fig. 2. Scatter plots of speech test results with SPEAK vs. MPEAK strategy for 20 subjects and 11 subtests. For the abbreviations refer to text. **a** Word and logatome tests. **b** Innsbruck sentence test in quiet and noise. **c** Göttingen sentence test in quiet and noise.

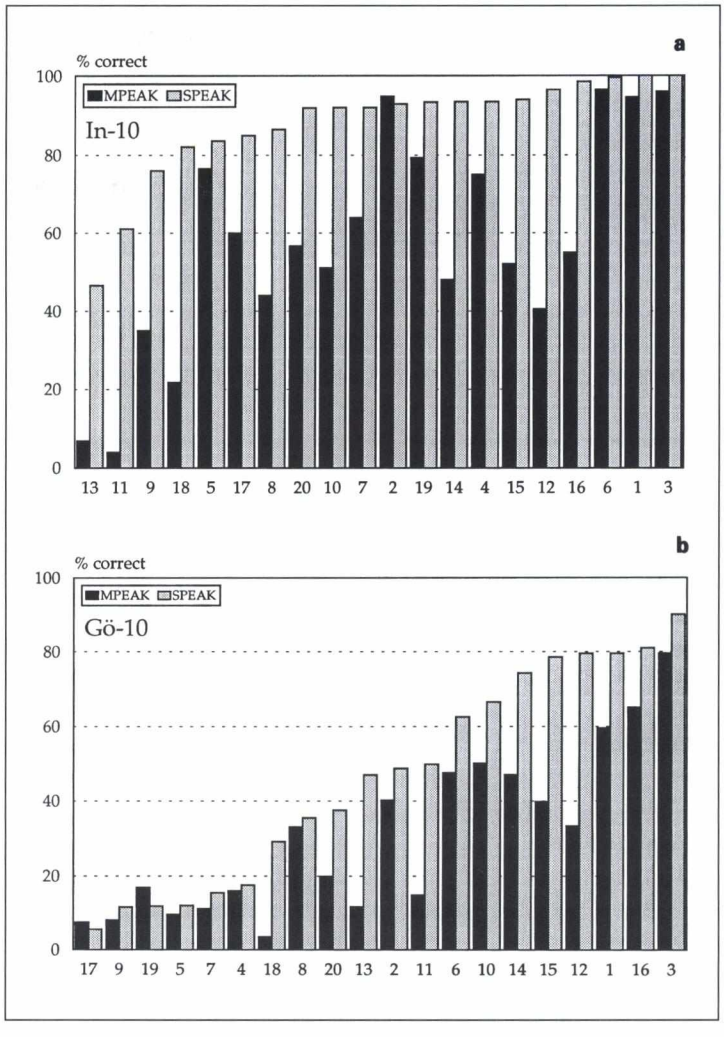


Fig. 3. Sentence tests in noise (10 dB SNR). Numbers 1–20 are the subject identifications. Pairs of data bars are sorted by SPEAK results from low to high. **a** Innsbruck sentences. **b** Göttingen sentences.

other hand, was significant for the vowel test as well as for all other tests. The sentence test results for the 10 dB SNR condition were also significantly different for center C2 whereas the other differences were not statistically different. Note that the mean values for C3 were calculated for 7 subjects whereas only 5 (C2) or 4 (C1 and C4) subjects were aggregated for the other centers. It should be noted that the response pattern across centers varied depending on the test. Subjects at C4 for example performed quite well on the Innsbruck test while their performance with the Göttingen sentences was rather poor. Subjects at C3, on the other hand, did less well with the Innsbruck sentences but scored much better on the Göttingen sentences. These differences correspond well with subjective comments addressing geographical and dialectic speaker and listener idiosyncrasies.

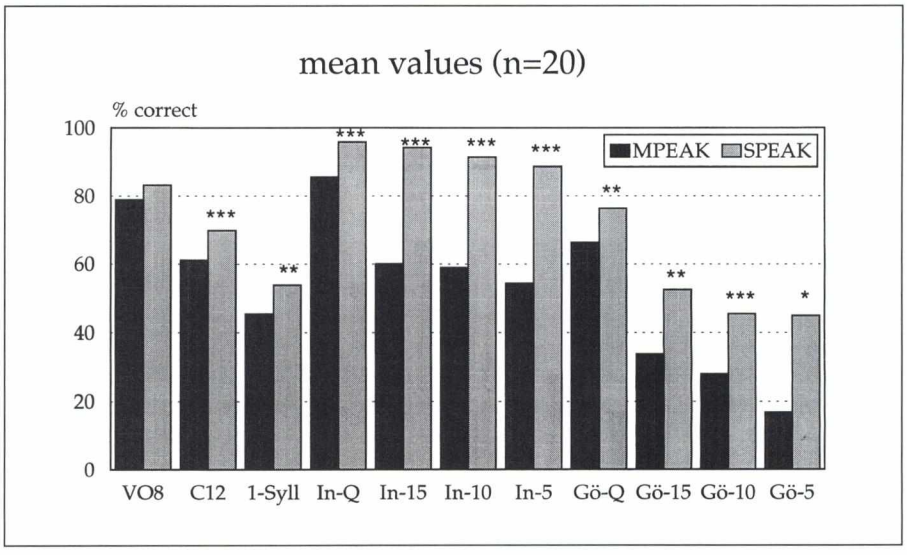


Fig. 4. Group mean results with MPEAK and SPEAK for 20 subjects and 11 subtests. Significant differences are indicated by asterisks.

Table 3. Group mean comparison for all speech subtests

Subtests	SPEAK-MPEAK (paired differences)				(+)	(0)	(–)
	% diff.	t value	d.f.	2-tail sign.			
VO8	4.3	0.90	19	0.378 (n.s.)	8	6	6
C12	8.8	4.13	19	0.001 ***	13	6	1
1-SYL	8.4	3.15	19	0.005 **	14	2	4
In-Q	10.2	5.11	19	0.000 ***	12	8	0
In-15	34.1	5.20	8	0.001 ***	7	1	0
In-10	32.5	8.21	19	0.000 ***	17	3	0
In-5	34.3	5.49	10	0.000 ***	10	2	0
Gö-Q	10.0	2.95	19	0.008 **	12	4	4
Gö-15	18.7	3.09	13	0.009 **	10	2	2
Gö-10	17.5	4.72	19	0.000 ***	13	6	1
Gö-5	28.2	3.20	5	0.024 *	5	0	1

Percentage differences of mean scores (SPEAK-MPEAK) with t values, degrees of freedom (d.f.) and two-tailed statistical significances of paired sample t tests, number of subjects with improved (+), unchanged (0) and decreased (–) performance.

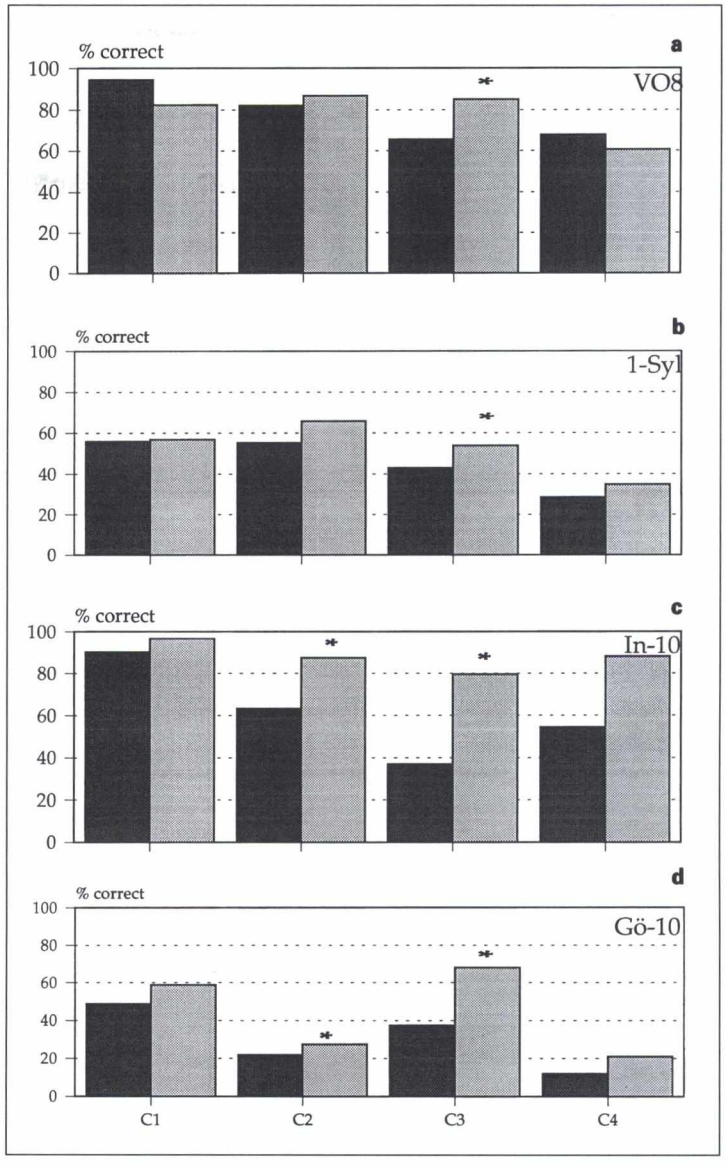


Fig. 5. Group results of the 4 evaluation centers for 4 subtests. **a** Vowel test VO8. **b** Monosyllabic word test. **c** Innsbruck sentence test (at 10 dB SNR). **d** Göttingen sentence test (at 10 dB SNR).

Figure 6 shows the results of the subjective performance rating through questionnaires (fig. 6a) and of the music quality rating (fig. 6b). Only the last of the three performance evaluation questionnaires was formally evaluated with 25 questions comparing processor performance and overall impressions at home and at work in quiet and noisy environments. A clear preference for SPEAK (69%) versus MPEAK (8%) was obtained. This corresponded with the evaluation of the music quality rating which was carried out during all 4 sessions. The 11 music pieces had to be judged on a subjective scale which

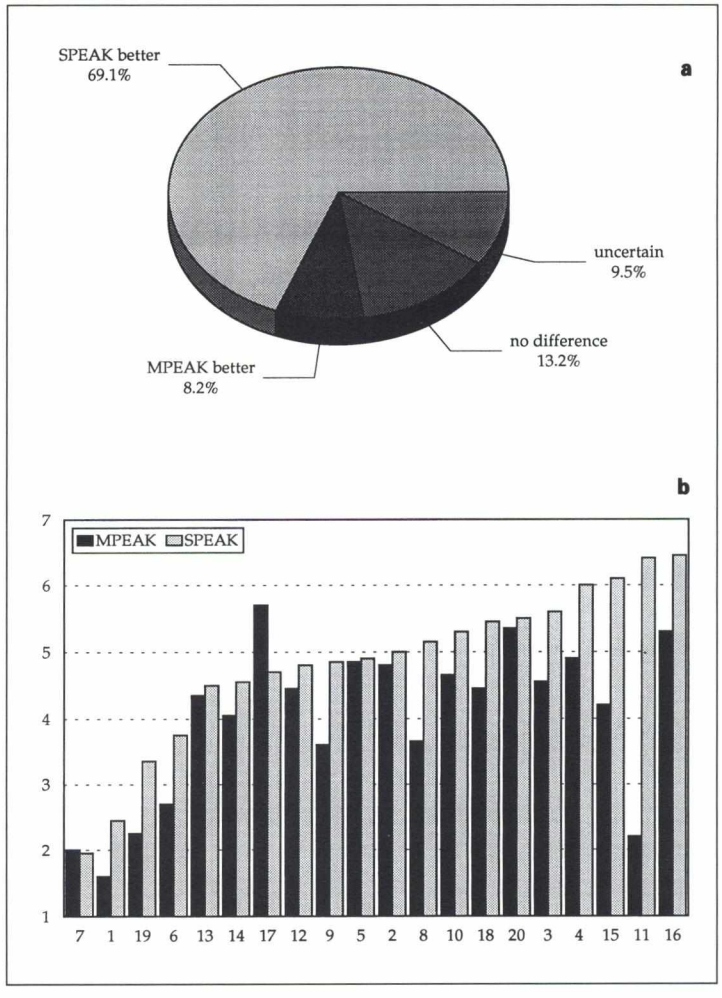


Fig. 6. Subjective performance rating through questionnaires (**a**) and results of the music quality rating (**b**). The subjective scale ranged from 1 (very unnatural) to 7 (very natural).

ranged from 1 (very unnatural) to 7 (very natural). Scores for sessions A1 and A2 as well as B1 and B2 were averaged and compared with a Wilcoxon matched-pairs signed rank test. As shown in figure 6b, the MPEAK scores ranged from 1.6 to 5.7 (mean = 3.98) whereas the SPEAK scores ranged from 1.95 to 6.45 (mean = 4.84) and the differences were significant at a level of $p < 0.0005$ ($z = 3.49$). The evaluation of the musical instrument identification test (5 main instruments in 5 musical pieces) showed that 10 of the 20 subjects could identify more instruments using the SPEAK strategy, while 4 others had obtained higher scores with MPEAK and the remaining 6 did not show any difference in identification scores. The difference between the two strategies was statistically significant ($p < 0.05$).

The tinnitus questionnaire consisted of 12 questions which had to be answered at sessions A1 and B2. The first question asked about the occurrence of tinnitus while wear-

ing the speech processor, the second about tinnitus when not wearing the speech processor. Four subjects did not perceive tinnitus at all and therefore did not answer the remaining questions. The severity of tinnitus was rated from 0 (never) to 3 (always) with intermediate levels 1 (sometimes) and 2 (often). The cumulated scores were 20 and 17 for tinnitus with MPEAK versus SPEAK and 27 versus 22 for tinnitus without processor at the first session (A1) versus the second (B2). These differences were not statistically significant (χ^2 test). The remaining questions (MPEAK/SPEAK scores are given in parentheses) dealed with the ability to concentrate (7/6), sleeping problems (4/5), relationship with stress (3/1) and drugs (3/1) and recognition of speech (8/4). Three subjects reported that the tinnitus was stronger when they were using the MPEAK processor. All other subjects did not perceive any change in the intensity of their tinnitus when switching processors.

Discussion

The main outcome of this study is the marked difference in performance between the two processing strategies investigated. As shown above, the mean scores calculated for the whole group were superior at a significance level of 0.05 or better for the SPEAK processor in all tests except the vowel logatome identification test. As mentioned above, the performance with MPEAK for vowel identification is already high. In addition, it should be noted that there were only 24 items in each test, which precludes further analysis.

The differences in performance for the four evaluating centers were somewhat surprising. Some possibilities for systematic differences between test sites are variations in the experimental setup, differences in sound delivery systems, test application and scoring of results. All of these factors have been carefully checked and excluded by a rigorous experimental protocol. Thus, the most probable cause for the observed differences are the subjects themselves (linguistic background and regional origin) and possibly the distribution of subjects at the four centers and may be the amount or type of rehabilitation programs at these centers. Note however that the small number of subjects in each group makes the statistical analysis of these data highly susceptible to outliers and individual values.

Learning did not occur in the way it was originally hypothesized. The analysis of variance showed that the session effect was not significant. However, separate analyses for MPEAK and SPEAK revealed that the scores with MPEAK remained stable during the study period whereas the SPEAK scores generally increased in the second phase. Thus it can be expected that the performance with SPEAK may still increase further after some more time. This has indeed been observed in some cases where subjects had been retested a few months after the end of the field trial. On the other hand, it can be assumed from anecdotal comments that most subjects did notice a more or less immediate benefit when they switched from the MPEAK to the SPEAK strategy and that the time for familiarization was generally rather short. Apart from sometimes annoying high-pitched sensations which usually disappeared or became unnoticeable after a few days, the sound quality and the overall performance was also judged to be superior by most subjects. The ability to identify musical tunes and instruments was noted as an important aspect of the new processor by a majority of the participating CI users.

Concerns about the use of higher stimulus rates and possible effects on provoking or enhancing tinnitus could not be corroborated in this investigation. If a change in the intensity of the tinnitus was noted at all (3 cases), it

was reported to have occurred during the MPEAK sessions. However, this result does not answer all questions about long-term effects of high rate stimulation and more data will have to be collected on this issue [26].

Finally, it may be asked whether poor performers would also benefit from the new strategy or whether the observed improvements were somehow related to the selection criteria which excluded subjects without open set speech understanding. Generally, this question cannot be answered based on the experimental data of this field trial. There are some indications (fig. 1, 3) that poor performers may improve their scores substantially. There are, however, examples which indicate no change for poor performers as well. It should be noted that poor performers may not have all 22 electrodes functioning with average dynamic ranges and that sometimes, threshold values would be increased. In order to avoid excessive current amplitudes and still reach suprathreshold stimulation, the average pulse widths would have to be increased which would reduce the overall stimulus rate or the number of possible frequency bands per stimulation cycle and therefore result in a decreased spectral resolution.

Conclusions

In summary, considerable individual improvements in speech perception were noted for most subjects on at least one measure when using the SPEAK strategy. The performance and benefit reported by the 20 deaf participants of this study indicate that the new coding strategy enables almost all subjects to recognize speech as good as or significantly better than with the MPEAK strategy that they had used for at least 9 months prior to this trial. Differences between the 4 centers were observed which were mainly related to geographical speaker and listener variables. Additional tests and the analysis of the performance questionnaires revealed generally more pleasant music listening, improved identification of musical instruments and mainly very positive qualitative judgments of the sound quality with the SPEAK coding strategy as compared to the MPEAK strategy.

Résumé

Dans une étude comprenant 4 centres européens d'implantation cochléaire, 20 individus de langue allemande munis du système Nucleus 22 ont été examinés par rapport à leur capacité de perception de la langue avec une nouvelle stratégie de codage de l'information spectrale de la parole (SPEAK). Ce procédé analyse continuellement le signal sonore en 20 filtres passe-bande programmables et offre jusqu'à 10 pointes spectrales aux 22 électrodes implantées. La performance de chaque sujet portant un boitier processeur expérimental a été évaluée en relation avec la performance dans une condition de comparaison à l'aide de tâches auditives multiples, la condition de comparison étant la stratégie de codage multi-pic (MPEAK) du processeur vocal Nucleus. Les deux stratégies ont été changées et répétées d'après un schéma expérimental ABAB. Seuls des sujets capables de reconnaître des mots monosyllabiques sans lecture labiale étaient inclus dans cette étude. Des différences significatives ont été observées dans les résultats dans la plupart des sous-tests lors de l'usage de la stratégie SPEAK en comparaison avec la stratégie MPEAK. La plus grande amélioraton a été obtenue dans des tests de phrases dans le bruit.

Acknowledgements

The authors would like to thank their colleagues, especially Dr. W.K. Lai, M. Olesch, A. Krings, P. Köppen for assistance with this study. Special thanks are also due to all the subjects who participated in this research. The support of many staff members of Cochlear AG, Basel, in the planning and evaluation phase is also greatly acknowledged. The contributions of Dr. E.L. von Wallenberg, J. Reid and C. Everingham have been particularly valuable. The study was funded by Cochlear AG, Basel, and the participating implantation centers.

References

1 House WF, Urban JC: Long term results of electrode implantation and electrical stimulation of the cochlea in man. Ann Otol Rhinol Laryngol 1973;82:504–510.

2 Burian K, Hochmair ES, Hochmair-Desoyer IJ, Lessel MR: Designing of and experience with multichannel cochlear implants. Acta Otolaryngol (Stockh) 1979;87:190–195.

3 Eddington DK: Speech discrimination in deaf subjects with cochlear implants. J Acoust Soc Am 1980;68: 885–891.

4 Michelson RP, Schindler RA: Multichannel cochlear implant: Preliminary results in man. Laryngoscope 1981;91:38–41.

5 Tong YC, Clark GM, Seligman PM, Patrick JF: Speech processing for a multiple-electrode cochlear implant hearing prosthesis. J Acoust Soc Am 1980;68:1897–1899.

6 Patrick JF, Clark GM: The Nucleus 22-Channel Cochlear Implant System. Ear Hear 1991;12(suppl to No 4):3S–9S.

7 Tong YC, Lim HH, Clark GM: Synthetic vowel studies on cochlear implant patients. J Acoust Soc Am 1988;84:876–887.

8 McKay C, McDermott HJ, Vandali A, Clark GM: Preliminary results with a six spectral maxima sound processor for the University of Melbourne/Nucleus multiple-electrode cochlear implant. J Otolaryng Soc Austr 1991;6:354–359.

9 Dillier N, Lai WK, Bögli H: A high spectral transmission coding strategy for a multi-electrode cochlear implant. Advances in Cochlear Implants. Proc 3rd Int Cochlear Implant Conf, Innsbruck 1994, pp 152–157.

10 Dillier N, Bögli H, Spillmann T: Speech encoding strategies for multielectrode cochlear implants. A digital signal processor approach. Prog Brain Res 1993;97:301–311.

11 Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M: Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. J Rehabil Res Dev 1993; 30:110–116.

12 Schindler RA, Kessler DK, Haggerty HS: Clarion cochlear implant: Phase I investigational results (corrected; erratum to be published). Am J Otol 1993;14:263–272.

13 Zierhofer C, Peter O, Brill S, et al: A multichannel cochlear implant system for high-rate pulsatile stimulation strategies. Advances in Cochlear Implants. Proc 3rd Int Cochlear Implant Conf, Innsbruck 1994, pp 204–207.

14 Patrick JF, Seligman PM, Money DK, Kuzma JA: Engineering; in Clark GM, Tong YT, Patrick JF (eds): Cochlear Prostheses. Edinburgh, Churchill Livingstone, 1990, pp 99–124.

15 McDermott HJ, McKay C, Vandali AE: A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. J Acoust Soc Am 1992;91:3367–3371.

16 Skinner MW, Clark GM, Whitford LA, et al: Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. Am J Otol 1994;15(suppl): 15–27.

17 Dillier N, Battmer RD, Döring WH, Müller-Deile J: Multicentric field evaluation of a new spectral peak coding strategy with native German speaking cochlear implant users (abstract). Second Eur Symp on Paediatric Cochlear Implantation, 1994; La Grande Motte, May 26–28, 1994.

18 Dillier N, Battmer RD, Döring WH, Müller-Deile J: Multicentric field evaluation of a spectral peak (SPEAK) cochlear implant speech coding strategy. Proc 5th Int Symp on Cochlear Implants. Surgical Workshop and 1st Int Auditory Brainstem Satellite Symp, Freiburg, June 24–26, 1994, in press.

19 Keller F: Verschiedene Aussprachen des Sprachverständnistests nach DIN 45621 (Freiburger Test). Biomed Tech (Berlin) 1977;22:292–298.

20 Hochmaier-Desoyer IJ, Hochmair ES, Fischer RE, Burian K: Cochlear prostheses in use: Recent speech comprehension results. Arch Otorhinolaryngol 1980;229:81–98.

21 Wesselkamp M, Kliem K, Kollmeier B: Erstellung eines optimierten Satztests in deutscher Sprache; in Kollmaier B (ed): Moderne Verfahren der Sprachaudiometrie. Heidelberg, Median-Verlag von Killisch-Horn, 1992, pp 330–343.

22 Müller-Deile J: Logopädische Bemerkungen zum Satztest, unpubl, 1994.

23 Niemeyer W: Sprachaudiometrie mit Sätzen. I. Grundlagen und Testmaterial einer Diagnostik des Gesamtsprachverständnisses. HNO 1967;15:335–343.

24 Thornton ARD, Raffin MJM: Speech-discrimination scores modeled as a binomial variable. J Speech Hear Res 1978;21:507–518.

25 Norusis MJ: SPSS for Windows Professional Statistics Release 5. Chicago, SPSS Inc. 1992.

26 Ni D, Seldon HL, Shepherd RK, Clark GM: Effect of chronic electrical stimulation on cochlear nucleus neuron size in normal hearing kittens (review). Acta Otolaryngol (Stockh) 1993;113:489–497.