Original Paper

ORL

ORL 2004;66:130–135
DOI: 10.1159/000079332

Received: November 28, 2003
Accepted after revision: April 16, 2004

# Analysis of Ceiling Effects Occurring with Speech Recognition Tests in Adult Cochlear-Implanted Patients

Jan Helms[a] Viktor Weichbold[d] Uwe Baumann[b] Hellmut von Specht[c] Franz Schön[a] Joachim Müller[a] Barbara Esser[b] Michael Ziese[c] Ilona Anderson[e] Patrick D'Haese[e]

[a]HNO-Universitätsklinik Würzburg, Würzburg, [b]HNO-Klinikum Grosshadern, München, [c]Klinik für HNO, Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany; [d]Abteilung für Hör-, Stimm- und Sprachstörungen, HNO-Universitätsklinik Innsbruck, and [e]Clinical Research Department, MED-EL Worldwide Headquarters, Innsbruck, Austria

**Key Words**
Ceiling effect · Cochlear implant · COMBI 40+ · Long-term study · Speech recognition

**Abstract**
This article presents a simple method of analysing speech test scores which are biased through ceiling effects. Eighty postlingually deafened adults implanted with a MED-EL COMBI 40/40+ cochlear implant (CI) were administered a numbers test and a sentence test at initial device activation and at 1, 3, 6, 12 and 24 months thereafter. As a measure for speech recognition performance, the number of patients who scored at the 'ceiling level' (i.e. at least 95% correct answers) was counted at each test interval. Results showed a quick increase in this number soon after device activation as well as a continuous improvement over time (numbers test: 1 month: 51%; 6 months: 73%; 24 months: 88%; sentence test: 1 month: 33%; 6 months: 49%; 24 months: 64%). The new method allows for the detection of speech recognition progress in CI patient samples even at late test intervals, where improvement curves based on averaged scores are usually assuming a flat shape.



## Introduction

Research into long-term effects of cochlear implantation on speech recognition performance of implanted subjects is frequently confronted with the problem that speech test results are distorted by the so-called 'ceiling effect'. A ceiling effect is observed when a considerable number of subjects are scoring at, or close to, the 100% performance level of a test such that the measurement of any further improvement becomes impossible with this test.

When analysing the speech recognition data from a large sample of postlingually deafened adults who received a MED-EL COMBI 40 (C40) or COMBI 40+ (C40+) cochlear implant (CI), it turned out that ceiling

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com


0301–1569/04/0663–0130$21.00/0

Accessible online at:
www.karger.com/orl

Viktor Weichbold, MD, MPhil
Clinical Department of Hearing Voice and Speech Disorders
University Hospital, Anichstrasse 35
AT–6020 Innsbruck (Austria)
Tel. +43 512 504 5438, Fax +43 512 504 3217, E-Mail viktor.weichbold@uklibk.ac.at

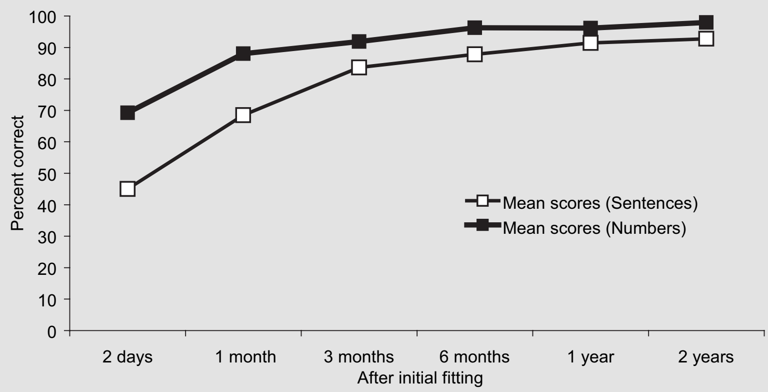


**Fig. 1.** Demonstration of the occurrence of ceiling effects with the numbers test (filled squares) and the sentence test (empty squares). Squares represent the mean speech recognition scores (in percent correct out of 100%) from a sample of 44 adult CI patients.

effects occurred to a considerable extent on practically all measures used. Consequently, test results were biased in the way that the improvement curve over time was too flat and did not reflect the real abilities of the testees (fig. 1). Interestingly, ceiling effects occurred as soon as 1 month after initial device activation on a test that required the recognition of bisyllabic numbers (‘Freiburger Zahlen’) and as soon as 3 months after device activation on a sentence recognition test (‘Innsbrucker Sätze’).

The occurrence of ceiling effects is contrary to a study that attempts to assess long-term speech recognition improvement in CI patients. With such test scores, the real extent of the improvement cannot be demonstrated. Nevertheless, it is still possible to show that progress of performance is ongoing. It just requires a change of the strategy: instead of using averaged test scores as a measure for the sample’s performance (which is the most common strategy), it is favourable to count the number of subjects who score at the maximum performance level per test interval. If improvement within the sample is continuing, then the number of ‘ceiling scorers’ should steadily increase. Long-term benefit of device use, thus, is indicated through the continuously growing number of patients who attain the maximum performance level with the test. Moreover, with this strategy, it becomes possible to investigate the factors that account for the rapidity of arriving at the ceiling of a test’s difficulty level.

The current article reports on the long-term findings from two speech recognition tests (a numbers test and a sentence test), where many patients are seen to arrive at the maximum scoring level soon after cochlear implantation. It will be shown that, despite the average scores from these tests being unsuitable to trace long-term progress, there is still ongoing improvement that can be quantitatively demonstrated and analysed. In addition to showing the long-term outcome, this paper will investigate which factors account for the observation that some subjects rapidly reach the maximum scoring level, while others need more time to achieve it. Variables that are considered as candidates for such factors are: device type (C40 versus C40+), gender, age at implantation, duration of deafness, side of implantation and initial performance level (test scores at the 2-day interval).

## Method

*Subjects*

Data from 80 postlingually deafened German-speaking adults were included in the current analysis. These patients are part of a larger sample of MED-EL CI recipients from various clinics in Germany, Austria, Switzerland and Italy (German-speaking region), who have been participating in the C40/C40+ surveillance study. Details of this study including the criteria for inclusion/exclusion have been previously reported [1]. As a number of study participants did not meet all scheduled visits to their clinic, complete datasets were not available from all of them. It was decided to include only subjects in the current analysis who have either complete data over a 2-year period or whose data are missing only after they had already arrived at the maximum scoring level (see below for the definition). With this criterion, a sample of exactly 80 patients was recruited.

The patients’ mean age at implantation was 55.4 years, with an age range from 22 to 77 years. Their mean duration of deafness prior to implantation was 7.7 years (1 month to 31 years). Causes of deafness were otosclerosis (12), meningitis (7), skull fracture (4), ototoxic medication (4), Cogan syndrome (2), cholesteatoma (2), Ménière’s disease (1) and unknown in 48 patients.

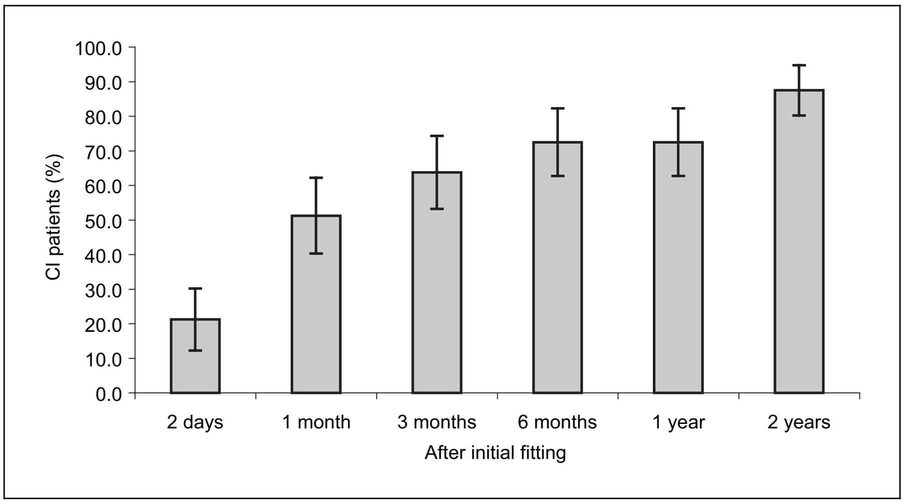


**Fig. 2.** Percentages of CI patients (n = 80) who scored at the maximum performance level (i.e. at least 95% correct answers) with the numbers test. Error bars indicate the 95% confidence intervals.

The patients participated in 6 test sessions: 2 days after initial device activation and then 1 month, 3 months, 6 months, 1 year and 2 years thereafter. Pre-operatively, the performance was evaluated with an open-set monosyllabic word test in the best-aided condition, but scores were generally at the 0% level. Hence, pre-operative results were not included in the current analysis.

*Devices*

All subjects had received either a MED-EL C40 or a C40+ unilaterally. Both devices implement the continuous interleaved sampling strategy; however, the C40 features 8 channels and an overall stimulation rate of 12,120 pulses per second (pps), while the C40+ features 12 channels and an overall stimulation rate of 18,180 pps. Stimulation rate per channel, if all channels are activated, is 1,515 pps for both devices. Further information on technological features of the devices is provided in Zierhofer et al. [2, 3].

*Test Materials*

Tests for assessment of speech perception were a numbers test (Freiburger Zahlen) and a sentence test (Innsbrucker Satztest). The numbers test contains various lists of 10 two-digit numbers, where the patient's score is determined by the percentage of correctly repeated numbers. The items are presented from a CD (WESTRA CD 1, DIN 45621) at 70 dB SPL. The sentence test consists of various lists of 10 everyday sentences containing 3–8 words each. The sentences are presented by live voice (with the speaker's face hidden) at a normal conversational loudness level and at normal conversational speed. The score is computed through the percentage of correctly repeated words. All speech tests were presented in quiet. Additional information on the two tests is provided in the articles of Helms et al. [1] and Gstöttner et al. [4].

*Definition of Maximum Scoring Level*

As a measure for the speech recognition performance the number of subjects who reached the maximum scoring level was taken. Thereby, achievement of the maximum scoring level was defined by two criteria:

(a) reaching at least 95% correct answers and

(b) repeat scoring of at least 95% correct answers at the subsequent test interval.

Hence, a subject was rated to be at the maximum performance level at a particular test interval, when he/she scored 95% or higher and when he/she also scored 95% or higher at the following test interval.

*Data Analysis*

With the above measure, statistical analysis is confined to counting the number of subjects who meet the criteria of performing at the maximum scoring level. For the purpose of comparison between test intervals, the 95% confidence interval was calculated at each test interval, i.e. the range of values where the true number of maximum performers is contained with a 95% probability. If the number of maximum performers at a particular test interval falls outside of the 95% confidence interval of a preceding test interval, then a significant difference between the two intervals can be assumed.

To examine which factors influence the rapidity of reaching the ceiling level, a Pearson correlation was computed between the factors under consideration and the inverse ranks of the test interval, at which the individual reached the ceiling level (i.e., never = 0, 2 years = 1, 1 year = 2, 6 months = 3, 3 months = 4, 1 month = 5, 2 days = 6). Bivariate factors (device type, implantation side and gender) were coded as dummy variables (i.e., through either 1 or 0).

## Results

Figures 2 and 3 illustrate the percentages (including 95% confidence intervals) of maximal scoring individuals per test interval with the numbers test and the sentence test, respectively. Prior to implantation, none of the subjects had scored at this level. After device activation, the number of maximum performers increased up to 2 years, where the largest gains occurred within the first 3 months

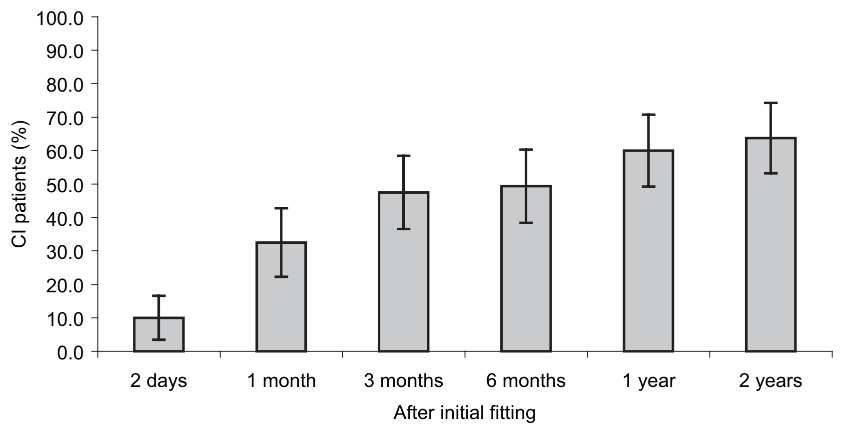


**Fig. 3.** Percentages of CI patients (n = 80) who scored at the maximum performance level (i.e. at least 95% correct answers) with the sentence test. Error bars indicate the 95% confidence intervals.

after CI activation. With the numbers test, 21% of the subjects performed at the ceiling level already after a few days of device experience. After 1 month, more than half of them (51%) achieved the maximum score. This number rose up to 73% at 6 months and to 88% at 2 years after implantation. Figure 2 also shows that significant improvement occurred up to 2 years: the number of ceiling level performers at this test interval was beyond the upper boundary of the 1-year 95% confidence interval. Noteworthy, the averaged test scores did not reflect this improvement, as they were hardly changing any more after the 6-month interval (fig. 1).

With the sentence test, the number of maximum performers was generally a bit lower than with the numbers test, but the course of improvement was quite similar. At the initial test interval, 10% of the subjects were maximum scorers. After 3 months, 48% scored at the maximum level and this number rose to 62% at the 2-year post-implantation interval. Again, it becomes apparent that the number of ceiling performers was steadily growing up to 2 years and that significant improvement was still occurring between the 6-month and the 2-year interval.

Analysis into variables that influence the rapidity of reaching the ceiling did not yield consistent results. Correlation coefficients between the variables and arrival at the 100% performance level are included in table 1. Only one factor was found to have a strong predictive value: the initial test score. Both with the numbers test and with the sentence test, subjects showed the fastest achievement of ceiling if their scores were already high immediately after initial device fitting. Device type (C40+ versus C40) turned out to be another potential factor, but the results were not fully clear. The C40+ yielded significantly quicker progress than did the C40 with the numbers test (p = 0.019), but with the sentence test, the correlation was slightly above the significance threshold (p = 0.065). Nevertheless, this finding might be a hint that the C40+ is allowing for quicker results than is the C40 system. Not unexpectedly, age at implantation, side of implantation and gender of the patient did not turn out to be crucial. Interestingly, also duration of deafness was not found to have an impact on the rapidity of reaching the top of test performance.

**Table 1.** Correlation (Pearson's r) between rapidity of reaching the test 'ceiling' and presumed influential factors

| Factor | Numbers test | Sentence test |
|---|---|---|
| Duration of deafness | 0.07 (0.57) | 0.09 (0.41) |
| Age at implantation | 0.05 (0.66) | 0.01 (0.96) |
| Device type | 0.27 (0.019) | 0.22 (0.065) |
| Gender | 0.09 (0.42) | 0.11 (0.32) |
| Side of implantation | 0.02 (0.87) | 0.06 (0.59) |
| Initial test scores | 0.36 (0.0034) | 0.41 (<0.001) |

Figures in parentheses indicate p values.

## Discussion

Long-term results of speech recognition in adult CI patients are not yet firmly established. All studies agree that the largest gains in speech recognition are generally made within the first 6–9 months of CI use [1, 4–10], but it is still doubtful how long and to what extent the improvement is continuing after that period. There is evidence that it goes on up to 36 months, with the rate of improvement slowing down after 1 year [8]. Some authors claimed that long-term improvement would occur only with some tests and not with others [4, 8]. A crucial problem of the assessment of long-term development is that, in the course of time, a growing part of the patient sample is reaching the ceiling of the test's difficulty level. As soon as this occurs, averaged test scores do not reflect the patients' real skills, but spuriously indicate that there would be only little or even no progress any more in the sample. Thus, the slowing down of the improvement rate observed in long-term studies of CI-aided speech perception may just be an artefact coming from the fact that too many patients have already reached the top of the test's difficulty level.

The occurrence of 'ceiling scores' is almost unavoidable in studies whose observation period spans over several years, especially, when they address quickly developing abilities. There are hardly any tests available which cover the full performance range of these abilities, beginning with very elementary reactions, continuing with more and more complex response patterns and ending up with highly sophisticated behaviours. Thus, one usually has to start with some low-difficulty test (e.g. test A) and to switch over to a more difficult one, when the patient has reached the maximum performance level of A. While the patient is not tested any more with A, he/she still remains included in the sample and is given a 100% score on this test at every subsequent interval. This is a quite reasonable solution, but when the number of maximum performers grows, the improvement rate gets tangibly biased through the ceiling effect. It should be clear at this point that excluding the maximum performers from the sample is an unacceptable solution. This would have two detrimental effects on the study: first, the size of the patient group tested with A would continuously decrease, until only a few subjects remain; second, the growth curve of performance with A will assume an arbitrary shape as the best performers are regularly taken away, while the worst ones are still being tested. Hence, the 100% scores must remain included in the dataset as long as the test is administered to patients. With this is mind, the question arises how data, despite containing such scores, can still be used for a realistic analysis of long-term progress.

In this article, we described an easy method of analysing speech recognition data that were affected by the ceiling effect. The aim was to show that these data could still be used for documentation and quantification of the benefits of cochlear implantation. This was achieved by a simple method: by counting the number of 'ceiling performers' per interval and computing 95% confidence intervals, which allows us to estimate whether or not the numbers of 'ceiling performers' at two distinct test intervals differ significantly. While averaged scores from speech tests are frequently not significant any more after 3 or 6 months of CI experience, the number of 'ceiling performers' may be much more sensitive to processes of ongoing improvement in the sample up to several years after implantation.

When applying this method to our data, we found that a lot of CI patients reached the ceiling performance level quite soon. With the numbers test, half of the patients achieved the maximum score after only 1 month of CI experience. With the sentence test, nearly 50% of the sample was capable of fully understanding simple sentences after 3 months of CI use. This finding is emphasizing the need for data analysis procedures which are robust against ceiling effects. In addition, this finding confirms the results from previous studies by showing that substantial progress of listening skills occurs quickly after implantation. Clearly, neither recognition of numbers nor understanding of simple sentences in quiet are significant measures of everyday communication skills; yet, they are measures of basic listening abilities, which indicate that the cochlear-implanted subjects have attained some elementary degree of speech understanding. While the acquisition of more advanced listening skills may proceed over a longer period, the gain of fundamental skills is obviously achieved within a few months after CI activation. As these skills are certainly helpful to the patients in their everyday activities, one can claim that cochlear implantation is quickly contributing to the enhancement of the quality of life of many deaf subjects.

As for the assessment of speech recognition with low-difficulty tests beyond intervals of 6 months or 9 months, the use of averaged scores was shown to be little expressive (fig. 1). However, when counting subjects who perform at a test's ceiling level per interval, significant progress in the sample could be tracked up to 2 years. With the numbers test, a significant increase in the number of 'ceiling performers' was found to occur between the 1-year and 2-year interval and with the sentence test, such an increase occurred between 6 months and 2 years. Also, it can be seen from figures 2 and 3 that the improvement curve at late intervals, when based on the number of 'ceiling performers', is not so flat like the one which is based on averaged scores (fig. 1). This observation indicates once again that progress occurs in the sample even at later intervals, but that averaged measures are not an appropriate means to detect it, when the data contain too many 100% scores.

When looking for factors that influence the speed of speech perception progress, an interesting result was obtained. No evidence was found that duration of deafness prior to implantation predicted how quickly a patient arrived at the ceiling. This is surprising, because duration of deafness has been overwhelmingly demonstrated to predict postimplantation outcome [11, 12]. On the other hand, our data suggest that the device type may be an influencing factor. It seems that patients with the C40+ are likely to attain the ceiling level faster than do patients with the C40. Perhaps, this is due to the higher stimulation rate the C40+ system is equipped with and to the different electrode design: as the C40+ electrode is longer than the C40 electrode, it covers a larger part of the cochlea and allows for a better topological mapping and for a more 'natural' sound sensation. This hypothesis, however, demands a specific investigation which directly addresses the effects of stimulation rate and electrode insertion depth on the improvement rate of speech recognition.

How long is improvement of speech recognition continuing after implant activation? Gstöttner et al. [4] and Tyler et al. [8] presented data which suggest that improvement occurs up to 36 months after implantation, at least on some measures they applied within their studies. Findings from the present analysis are essentially in line with their reports. When defining 'improvement' in the way that an increasing number of subjects are scoring at the 100% performance level of a test, then it was observed to happen up to 2 years after implantation, both with the numbers and the sentence test. But these tests only cover basic listening abilities, so it can be assumed that higher sophisticated hearing skills will continue to develop for several years. To demonstrate this, more difficult tests need to be administered as well as ceiling effects must be excluded. It is not before these two requirements are met that we will be able to draw a realistic picture of the long-term development of CI patients' speech understanding.

## References


1 Helms J, Müller J, Schön F, Moser L, Arnold W, Janssen T, et al: Evaluation of performance with the COMBI 40 cochlear implant in adults: A multicentric clinical study. ORL 1997;59: 23–35.
2 Zierhofer CM, Hochmair-Desoyer IJ, Hochmair ES: Electronic design of a cochlear implant for multichannel high-rate pulsatile stimulation strategies. IEEE Trans Rehabil Eng 1995;3:112–116.
3 Zierhofer CM, Hochmair IJ, Hochmair ES: The advanced COMBI 40+ cochlear implant. Am J Otol 1997;18(6 suppl):S37–S38.
4 Gstöttner W, Adunka O, Hamzavi J, Lautischer M, Baumgartner WD: Sprachdiskrimination bei postlingual ertaubten cochlear-implantierten Patienten. Wien Klin Wochenschr 2000;112:487–491.
5 Spivak LG, Waltzman SB: Performance of cochlear implant subjects as a function of time. J Speech Hear Res 1990;33:511–519.
6 Tye-Murray N, Tyler RS, Wordworth GG, Gantz BJ: Performance over time with a nucleus or ineraid cochlear implant. Ear Hear 1992;13:200–209.
7 Cohen NL, Waltzman S, Fisher SG, the Department of Veterans Affairs Cochlear Implant Study Group: A prospective randomized study of cochlear implants. N Engl J Med 1993;328: 233–237.
8 Tyler RS, Parkinson AJ,Woodworth GG, Lowder MW, Gantz BJ: Performance over time of adult patients using the ineraid or nucleus cochlear implant. J Acoust Soc Am 1997;102:508–522.
9 Välimaa TT, Sorri MJ: Speech perception after multichannel cochlear implantation in Finnish-speaking postlingually deafened adults. Scand Audiol 2000;29:276–283.
10 Anderson I, Weichbold V, D'Haese P: Recent results with the MED-EL COMBI 40+ cochlear implant and TEMPO+ behind-the-ear processor. Ear Nose Throat J 2002;81:229–233.
11 Van Dijk JE, Langereis MC, Mens LHM, Brokx JPL, Smoorenburg GF: Predictors of cochlear implant performance. Audiology 1999;38:109–116.
12 Shea JJ, Domico EH, Orchik DJ: Speech recognition ability as a function of duration of deafness in multichannel cochlear implant patients. Laryngoscope 1990;100:223–226.