SPEECH RECOGNITION UNDER CONDITIONS OF FREQUENCY-PLACE COMPRESSION AND EXPANSION

by

Deniz Baskent

A Dissertation Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(BIOMEDICAL ENGINEERING)

AUGUST 2003

Copyright 2003 Deniz Baskent

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3116665

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3116665

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UNIVERSITY OF SOUTHERN CALIFORNIA
THE GRADUATE SCHOOL
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA 90089-1695

This dissertation, written by

Deniz Baskent

under the direction of her *dissertation committee, and approved by all its members, has been presented to and accepted by the Director of Graduate and Professional Programs, in partial fulfillment of the requirements for the degree of*

DOCTOR OF PHILOSOPHY

Joseph Hellige

Director

Date August 12, 2003

Dissertation Committee

Robert Shannon
Chair

[illegible]

[illegible]

[illegible]

Frances Richards

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DEDICATION

for grandma...

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my family for all the unconditional love and endless support they have provided me through the years.

Many thanks go to my academic advisor, Dr. Robert Shannon, for being a great mentor as well as a friend. He showed me how fun science can actually be. And special thanks to Dr. Monita Chatterjee, my virtual advisor, for being so generous in sharing her knowledge and excitement.

Friends are always much appreciated. Thanks to all; Deniz, for being my family in U.S., Tassos, for bringing such joy to my life, old friends from Turkey; Yamac, Aylin, Elif, Senem, Pinar, Guclu, new friends from L.A.; Barbara, Geert, Alan, Alyssa, Marsha, Steve, Rob, Mark, Angelika, Georgios, Lu, Javier, Anna, Ghassan, and friends from House Ear Institute; Sandy and Mark, for just being there to listen, Bear, for constant challenging, Geri, Monica, Sherol, Steve, and Rachel, for her help with my research. Special thanks to Gabriele Larmon for always finding time for me even at the most stressful moments.

I appreciate all the help from the cochlear implant company Med-EL, providing me the software and equipment to run the experiments with patients. Special thanks to Dr. Peter Nopp and Amy Barco for their continuous efforts. Also thanks to Advanced Bionics, particularly to Andy Voelkel, for providing me their research software.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Many thanks to Dr. Dawna Mills from House Ear Clinic for all the help with my research, to Stefan Brill for his input as well as being a great friend, and to Dr. Qian-Jie Fu for letting me use his Tiger Speech Recognition System for sentence recognition tests.

I am feeling particularly lucky for having such great subjects who endured hours of repetitive testing and still could pretend that it is fun. Their efforts are certainly appreciated.

Last but not least, I would like to thank my dissertation committee members, Dr. Francis Richmond, Dr. David D'Argenio, Dr. Gerald Loeb, and Dr. Bartlett Mel, for their valuable comments on my work.

Funding for this research was provided in part by NIDCD grant R01-DC-01526 and contract N01-DC-92100.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

In cochlear implants, the length and the insertion depth of the electrode array determine the cochlear tonotopic range of stimulation, and the speech processor controls the mapping of acoustic frequency information onto this range. Conventional electrode arrays, 16 mm in length, stimulate a cochlear region corresponding to an acoustic frequency range of 500-6000 Hz. However, some implant speech processors map an acoustic frequency range from 150 Hz to 10,000 Hz onto these electrodes. While this mapping preserves the entire range of acoustic frequency information, it also results in a compression of the tonotopic pattern of speech information delivered to the brain. The present study measured the effects of such a compression of frequency-to-cochlear-place mapping on speech recognition, as well as the effects of an expansion. Such an expanded representation of speech in the cochlea might improve speech recognition by improving the relative spatial (tonotopic) resolution, like an "acoustic fovea." Phoneme and sentence recognition scores were measured as a function of compression and expansion with normal-hearing listeners using a noise-band vocoder, and with implant users by changing the programs in their implant processors. The analysis frequency range was either compressed or expanded relative to the cochlear tonotopic range while the tonotopic range was held constant. Speech recognition in matched conditions was generally better than compression and expansion, even when the matched condition eliminated a considerable amount of acoustic information. It was also more beneficial to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

match the frequencies that contribute more to speech information. The results suggest that speech recognition, at least without training, is dependent on the mapping of acoustic frequency information onto the appropriate cochlear place. Further experiments detailed the trade-off between information loss and the accuracy of the cochlear location where the acoustic information is presented. A minor modification of the classic Speech Intelligibility Index model was able to account for the drop in scores due to the mismatch of the frequency to place. Overall results provide guidelines for achieving an optimal frequency-place map for implant users.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1

INTRODUCTION

Speech recognition is adversely affected if the spectral information of speech is presented to an inappropriate cochlear location. For example several studies have shown a reduction in speech recognition when the speech spectrum is shifted up to higher frequencies (e.g. Daniloff *et al.*, 1968; Fu and Shannon, 1999), or if the frequency-to-cochlear place mapping is distorted nonlinearly (Shannon *et al.*, 1998). Changes in speech recognition as a result of such distortions in the frequency-to-cochlear place mapping are of theoretical interest as an indication of the mechanisms by which speech patterns are stored and retrieved in the central nervous system. In addition, understanding the potential effect of an appropriate frequency-place mapping is of critical importance for the design and programming of cochlear implants and hearing aids. These prosthetic devices can stimulate the peripheral auditory system with a tonotopic pattern of information that is distorted relative to the normal acoustic pattern. Such stimulation raises several questions: In the case of hearing loss, can the patient's speech recognition be improved by adjusting the spectral range of speech to match the frequency region of her residual hearing (Braida *et al.*, 1979; Reed *et al.*, 1983)? Or can the resulting frequency-place distortion actually interfere with speech understanding, as shown in some listeners by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Turner *et al.* (1999)? Which of these spectral manipulations are more detrimental and so should be avoided?

Several previous studies have addressed the question of how spectral manipulations affect speech understanding. Fu and Shannon (1999) measured vowel recognition in normal hearing (NH) and cochlear implant (CI) listeners when the acoustic spectral information was mapped to cochlear locations that were shifted apically or basally relative to the acoustic location for that information (for NH listeners) or relative to each listener's clinical frequency-to-electrode map (for CI listeners). They found that vowel recognition was robust to tonotopic shifts up to 3 mm, but dropped significantly for larger shifts. This result matches well with classical studies on frequency shifting (Daniloff *et al.*, 1968; Nagafuchi, 1976; Tiffany and Bennett, 1961).

In a similar study, Dorman *et al.* (1997) measured the effect of a shift in mm between the acoustic frequency range presented and the cochlear range to which it was presented. In the acoustic simulations the analysis filter bands were fixed, and sine wave carriers were shifted in mm along the cochlea relative to the normal acoustic place for that information. Speech recognition performance dropped as the stimulated electrode locations were shifted basally from the normal tonotopic location.

Shannon *et al.* (1998) measured speech recognition under conditions that produced a nonlinear warping of the frequency-place mapping. They used a noise-band vocoder to implement a logarithmic or exponential transformation between acoustic frequency and the normal cochlear place

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for that frequency. Although four spectral channels of information were presented, listeners' performance with the warped mapping dropped to the same level as that seen with a single-channel noise vocoder. This result suggests that nonlinear frequency-place warping can eliminate listeners' ability to utilize spectral cues.

As an extension to previous studies dealing with frequency-place distortions, the present study explored the effects on speech recognition when the acoustic frequency range delivered is larger or smaller than the normal cochlear range.

Note that neither the present study nor the previous studies discussed above address the potential effects of learning. Research by Rosen *et al.* (1999) showed that, following a short training process, listeners could partially adapt to basalward spectral shifts of as much as 6.5 mm. Another study (Fu *et al.*, 2002) showed significant improvement over the first few days by cochlear implant patients using a 2-4 mm apically shifted frequency-place map, but only little change was observed over the following three months. At the end of the three-month training period consonant and HINT sentence recognition scores were comparable to the baseline scores, while vowel and TIMIT sentence recognition scores were still significantly lower than the baseline scores obtained with the patient's own clinical map before the beginning of the test. In the present experiments the emphasis is on speech pattern recognition without any training. The intention was to test the ability of central pattern recognition mechanisms to accommodate alterations in the peripheral pattern of information with no time to adapt.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Frequency compression has historically been used in an attempt to increase the performance of hearing aids. Most hearing aid users have hearing loss at high frequencies with residual hearing at lower frequencies. To make better use of this residual hearing the spectrum of the speech was lowered and compressed so that the entire speech information was delivered to the audible range of the patient. The main techniques used for this purpose were slow playback, frequency shifting, vocoding, and zero-crossing-rate division. In terms of frequency-to-place mapping most of these manipulations consisted of a compressed apical cochlear shift. Braida *et al.* (1979) reviewed frequency compression/shifting studies and concluded that frequency lowering did not result in any substantial improvement in speech recognition, and often decreased the performance compared to simple amplification. Reed *et al.* (1983) evaluated the effect of frequency lowering on consonant recognition in a more systematic way, parametrically varying the frequency compression scheme from linear compression to nonlinear frequency-place warping. The results from this study confirmed that frequency lowering did not improve consonant recognition. Linear frequency compression, where the whole frequency range was compressed, resulted in worse consonant recognition scores than a frequency warping compression in which only higher frequencies were spectrally compressed and lowered.

These studies provide insight into the mechanisms used by the central nervous system for storing and retrieving tonotopic patterns of speech. If speech patterns were stored in a "positionally relocatable" fashion, then a tonotopic shift that maintained the overall spatial distribution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

should still be intelligible. This is clearly not the case, because frequency shifting usually reduces speech recognition. If only the relative order of spectral features were important, then monotonic alterations in the tonotopic pattern would still be intelligible. This is also not the case, because frequency compression reduces speech recognition. The present experiment was designed to further quantify the importance of linear compression or expansion of the tonotopic pattern of information (in cochlear mm). If the central pattern recognition system stores and retrieves information in terms of the relative tonotopic pattern, then it might be able to tolerate a substantial amount of linear compression or expansion.

These issues are not only noteworthy in terms of understanding the relative importance of peripheral vs. central pattern recognition for speech, but are of critical importance for the design and fitting of cochlear implants and hearing aids. In a cochlear implant, the electrode array is typically inserted into the scala tympani, reaching a depth of 20-30 mm inside the round window. The average insertion depth from 20 Nucleus implant patients was estimated to be 20 mm by Ketten *et al.* (1998). However newer electrode designs are intended to achieve array insertions as deep as 30 mm (Gstoettner *et al.,* 1999). The active stimulation range is typically 16 mm in length for Clarion I and Clarion II, and 16.5 mm for Nucleus 22 and Nucleus 24 devices. According to Greenwood's (1990) frequency-to-place equation, and assuming a 35 mm cochlear length in humans, this stimulated cochlear region corresponds to an acoustic frequency range of 500-6000 Hz in humans for a 25 mm insertion depth, and an acoustic frequency range of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1-12 kHz for a 20 mm insertion depth of the electrode array. Present cochlear implants offer only a limited choice of analysis filters, which cannot be changed individually to match a given patient's electrode location. Most commercial implant speech processors assign a wider fixed acoustic frequency range to this limited cochlear region regardless of the length or the insertion depth of the electrode array. For example Clarion II assigns an acoustic range of 350-8000 Hz (Advanced Bionics Corporation, 2001) and the default frequency allocation of the Nucleus-22 implant (SPEAK strategy Table 9) assigns a frequency range of 150 Hz-10 kHz to the electrodes (Cochlear Corporation, 1995). The latter acoustic range would normally cover a 25 mm range in the cochlea, specifically from 5 mm to 30 mm from the round window, rather than the 16.5 mm covered by the electrode array. Thus, mapping the larger acoustic frequency range onto the electrode array results in a compression of the frequency-to-place mapping. In some cochlear implant patients there may also be a tonotopic shift due to the discrepancy between the actual electrode location and the acoustic information assigned. The present experiment evaluated the effect of frequency-place compression on speech recognition in normal-hearing listeners in conditions that simulated two implant electrode insertion depths.

In addition to compression, frequency-place expansion was also evaluated. In this condition, the mid-frequency region was expanded in terms of its representation in the cochlea, effectively increasing the sensory resolution within this frequency range. This frequency-place expansion is analogous to the "acoustic fovea" in bats or cetaceans, where a large portion

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the cochlea is devoted to the small frequency region used for echolocation (e.g., Echteler *et al.,* 1994). While this type of expansion results in the loss of some acoustic information, the most critical spectral information is presented to a larger cochlear region, resulting in better neural resolution (increased mm/Hz) within that range. Also with developing technology longer electrode arrays are available, such as Med-El Combi 40+ electrode array that covers a cochlear length of 26.4 mm with 12 electrodes. A full insertion of 31 mm can be achieved with this electrode array which results in the stimulation of the cochlear region from around 4.6 mm to 31 mm from the round window. With Greenwood mapping function this region is found to respond best to acoustic frequencies between 150 Hz and 11 kHz. Yet the widest frequency ranges that can be assigned onto these electrodes are 200 Hz-5.5 kHz with the body worn processor (CISPRO) and 200 Hz - 8.5 kHz with the behind the ear processor (TEMPO+). Assigning this smaller frequency range onto the wider range covered by this long electrode array results in expansion in the frequency-place mapping.

Several studies have attempted to modify cochlear implant users' speech recognition performance by matching the apical end to the characteristic frequency of the electrode array. Whitford *et al.* (1993) showed significant improvement in open-set sentence recognition tests with implant users, even when background noise was added. The subjects wore the modified maps from a few days to a week. Eyles *et al.* (1996) had 11 implant users who were all deafened postlingually and had electrode arrays that were located at various insertion depths from the round window. The location

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the electrode array was estimated by postoperative radiographs for each patient. The range for the insertion depth was found to be from 14 mm to 21.5 mm from the round window. After wearing the modified maps for three months there was significant improvement in speech recognition scores when the sound stimuli were presented with lipreading combined. The effect was more prominent for patients with shallower insertion.

Both studies had only focused on matching the map to the characteristic frequency of the cochlea at the insertion depth, yet they did not explore the idea of matching the whole frequency range to the stimulation range. The present study quantifies effects of such spectral mismatch conditions by methodically changing the amount of the acoustic information assigned to the electrodes as well as exploring the effects of number of channels, insertion depth of the electrode array, and the ambiguity in the actually location of the electrode array, both with simulations and with implant users.

In following experiments, we matched the basal end of stimulation region while changing the frequency range assigned on the apical end only, both in simulations and implants. In addition, we also changed the frequencies assigned on the basal end only while the apical end was always matched. In each experiment perfect matching condition was always when the whole frequency range assigned matched the whole tonotopic location of stimulation region. Most vowel information is located towards the apical end of the stimulation region. Because vowels are more sensitive to such spectral changes compared to consonants a large improvement was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

predicted by just matching the apical end. Even though consonants are higher frequency they still fall more towards the midline of stimulation region in this study. Also they can still be recognized by cues carried by temporal characteristics, even in adverse situations such as a single electrode processor. Therefore a smaller effect was expected to be observed with consonants.

In the simulations all parameters can easily be controlled with the noiseband vocoder software. However with actual cochlear implant patients there are many unknown factors, such as the exact insertion depth of the electrode array, the proximity of the electrodes to the spiral ganglia, the nerve survival pattern, and the possibility of an abnormal anatomy of the cochlea.

The biggest problem for this study among those factors is that the exact location of the electrode array inside the cochlea is unknown. A great deal of work has been done trying to estimate the electrode location by combining mathematical tools and cochlea models with radiographs (Marsh *et al.*, 1993, Cohen *et al.*, 1996a) or CT scans (Ketten *et al.*, 1998, Skinner *et al.*, 2003). For example, the latter studies observed that the surgeons' estimate of the insertion depth of the array was 1-2 mm longer on average than the actual insertion depth.

Even though such imaging techniques might provide a more accurate estimation for the location of the electrode array it might be too time consuming and expensive to go through this process for every single patient who receives an implant. Therefore we also designed an experiment to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

assess how much the effect of frequency-place compression and expansion on speech perception changes if the electrodes are shifted from their estimated location. Such a design will represent the realistic listening conditions of an implant user where only the appropriate insertion depth of her electrode array (but not the exact location) is known.

In the last part of the study Combi 40+ (by Med-El) users and one Clarion II (by Advanced Bionics) user were tested with similar conditions. Combi 40+ users all were reported to have full insertions by their surgeons and all had long electrode arrays on soft carriers designed specifically for deep insertion. On the other hand, the Clarion II user had a much shorter array that was inserted with an electrode positioner system where the positioner keeps his array close to the modiolus. Most recent electrode designs are developed with the motivation of achieving a good insertion depth as well as a selective stimulation of spiral ganglia by having the electrodes closer to the inner wall or having more directional current spread sources. In both designs, even though the exact insertion depth, the lateral position of the array, or the individual cochlear length were not known, just assuming realistic values for these parameters similar results to simulations were observed. One reason is that the spectral pattern recognition in CNS has a tolerance for spectral mismatches up to a few mms. Therefore an error of 1-2 mm from surgeon's estimation for the insertion depth (Skinner *et al.*, 2003), for example, would not change the mismatch performance patterns significantly. Ideally, in a clinical setting where there is only a limited time available to fit the patient with the best frequency-place map, a cochlear

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

distance range can be estimated for the electrode array to match the map. Then this map can be fine-tuned with a simple perceptual test containing only a few entries that are determined by observing the highest error patterns in the confusion matrix.

With normal hearing subjects we had simulated only 20 and 25 mm insertion depths. These are realistic values for reported full insertions of most implants but they do not represent partial or shallow insertions, which are usually defined as insertions less than 20 mm. Because Combi 40+ electrode array is deeply inserted and covers a long stimulation range in the cochlea it provides a powerful research tool to explicitly explore the effects of insertion depth. Therefore we added more experiments at the end where frequency-place map was compressed and expanded on to the electrode arrays at varying cochlear distances. Faulkner *et al.* (2003) showed with simulations that the information loss for insertion depths less than 19 mm would be detrimental for speech recognition. This is expected because those ranges include only frequencies higher than 1 kHz, thus excluding some crucial speech information at lower frequencies, as indicated by the Speech Intelligibility Index (SII) which predicts the speech recognition scores from the amount of acoustic information available (ANSI 1997). Blamey *et al.* (1992) and Skinner *et al.* (2003) found a significant correlation between open set speech recognition scores and the insertion depths of the electrode arrays of implant users. Yet Hodges *et al.* (1999) found no correlation for insertion depths from 17 mm to 25 mm even though SII predicts a significant drop in performance for this range as the insertion depth decreases. In both

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

studies performance from different patients with different insertion depths were compared. In such an experimental setting it is difficult to isolate the effect of varying insertion depth only from inter-subject variations. A more objective approach would be to change the insertion depth on the same implant user and observe the relative drop in the speech performance from the baseline performance at full insertion. We did two experiments with Combi 40+ users to explore the effects of the insertion depth: In the first one, the baseline condition was an array of 10 electrodes inserted deeply. From this baseline condition of deep insertion, the insertion depth was made shorter by turning off the most apical electrode for each following condition. In every condition speech performance was assessed with both a matched map and a compressed map where the total frequency range available was assigned onto the electrode array. As a result, the matched map assigns less and less acoustic input information to match the shallower insertion whereas the compressed map always provides the most information possible with an increased compression as the insertion gets shallower. For shallower insertions where the amount of information loss was detrimental an optimum map was found by compromising between matching and information loss. This experiment realistically simulates implants with partial insertions. Yet the number of electrodes (and hence the stimulation range) changes from condition to condition in addition to the insertion depth. To isolate the effect of the location of the stimulation only, the speech performance was evaluated with the same number of electrodes at three different insertion depths in the second experiment.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Note that the main purpose of this study is to understand the pattern recognition mechanism of the brain for speech recognition as well as how to possibly use that mechanism to our advantage to provide implant users with better maps for more efficient speech recognition. The results might be very different for other acoustic stimuli such as music or environmental sounds where the quality of the sound might be more preferable over identifying it accurately. Yet because all devices have at least three program settings for three different maps the patients can always have additional maps that they might favor for other acoustic stimuli.

This dissertation is organized as follows:

Chapter 2 introduces basic principles of cochlear implants, such as how they work and how they evolved with improving technology over time. Different hardware and programming properties of most commonly used implants with latest designs are summarized as well.

In the first experiments the conditions were simulated to test with normal hearing subjects because all parameters can easily be controlled and modified in such simulations. Also normal hearing subjects generally display more consistent results and clearer patterns compared to implant users. The noise-band vocoder technique used for the simulation experiments is introduced in Chapter 3.

In all experiments Greenwood mapping function, described in Chapter 4, was used to determine the corresponding frequency of a cochlear

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

distance from the round window. Compression and expansion of the frequency-place mapping are introduced in Chapter 4 using this Greenwood mapping function. The experimental method to evaluate the effects of compressed and expanded mapping on speech recognition is described, and the speech recognition results of the experiments from normal-hearing subjects are presented. The phoneme recognition scores are also analyzed for speaker gender. Speech Intelligibility Index model is used to isolate the effects of information loss only and it is slightly modify to account for spectral mismatch effects additionally.

Chapter 5 illustrates further simulation experiments with similar mismatch conditions such that compression and expansion applied on apical or basal end only while the other end was matched. In a following experiment compression and expansion were combined with spectral shifts to simulate the effects of electrodes that are shifted from the estimated cochlear location. In the last experiment speech-shaped noise is added to compression and expansion conditions.

This dissertation focuses on the instant effects of frequency-place compression and expansion on speech recognition where subjects are not given any time to practice. In Chapter 6, effects of adaptation are explored with two small experiments. In the first one, one normal hearing subject was retested to observe the changes in the scores over time. In the second, two subjects were trained on one condition and were retested over a time period of 10 days how such training would change the performance.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Chapter 7, the implants Combi 40+ and Clarion II, that are used in the rest of the experiments, are described in more details. The experiments in this chapter are similar to simulations and the results from implant users are compared to simulation results.

Chapter 8 describes insertion depth experiments with Combi 40+ users. In the first experiment the partial insertion is simulated by decreasing the number of active electrodes. At each insertion the results from matched and compressed maps are compared. In the second experiment, the number of electrodes is kept the same while the stimulation range is changed.

The experiments and the main findings are summarized in Chapter 9.

The importance of the results as an insight to the speech recognition mechanisms of the brain, as well as possible benefits for cochlear implant users are discussed in Chapter 10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2

COCHLEAR IMPLANTS

Sensorineural hearing loss is the most common type of hearing loss, affecting 23% of population older than 65 years of age[1], and around 17 million people in USA[2]. Main reasons for such hearing loss are presbycusis (progressive hearing loss from old age), Meniere's disease, ototoxicity, prolonged noise exposure, infections such as meningitis, and acoustic neuromas on rare occasions.

Sensorineural hearing loss can be due to damage in the hair cells inside the cochlea or the acoustic nerves. Hair cells convert the mechanical motion of the filling fluid inside the cochlea to electrical signals that are delivered to the auditory nerves. If the hair cells do not function anymore whereas the auditory nerves are still intact, this function of the hair cells can be restored by a cochlear implant, as shown in Figure 2.1. The implant picks up the sound from outside via a microphone that is located just on top of the ear in the headpiece and processes it to code into electrical signals that directly stimulate the auditory nerves. Therefore the whole mechanical system of the ear is bypassed, making the device very different than

[1] Hearing Loss, Timothy C. Hain, MD: http://www.tchain.com/otoneurology/disorders/
[2] American Speech, Language, and Hearing Association: http://professional.asha.org/resources/factsheets/hearing.cfm

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conventional hearing aids which only amplify the sound for people with difficulty in hearing.

In this chapter a brief overview of the cochlear implants will be presented.

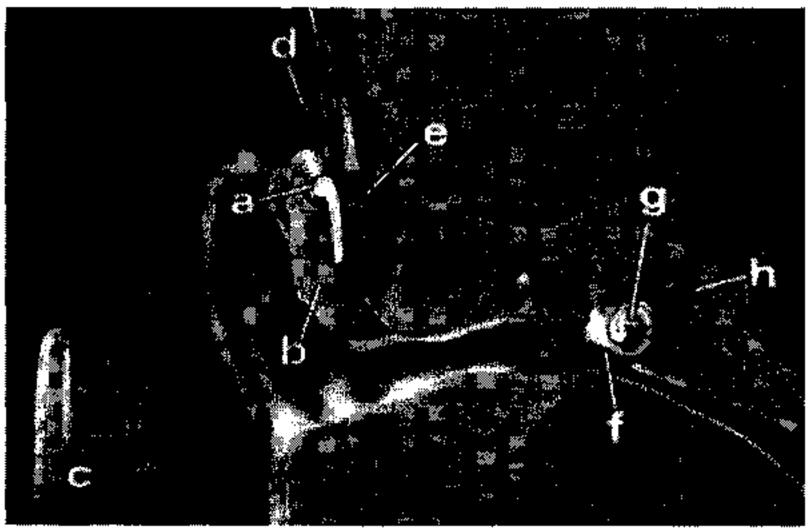


Figure 2.1: A typical cochlear implantation system[3]. Basic components of the system are: a. microphone, b. transducer, c. speech processor, d. transmitter, e. decoder, f. cochlea, g. electrodes, h. hearing nerve.

2.1. HISTORY

Even though the first attempts to stimulate the auditory nerves electrically were made as early as 18^{th} century, the idea of using electrical stimulation particularly for hearing came up in the 1950s. The advances in other medical devices in the 60s helped the development of cochlear implants tremendously, and the first implamantation with patients started.

[3] The Bionic Ear Institute, Melbourne, Australia: http://www.medoto.unimelb.edu.au/bei

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

William House first implanted patients in early 1960s and his work led to the first single-channel cochlear implant in early 1970s (Spelman, 1999). Two versions of single-channel processors have been widely used: House/3M and Vienna/3M (Loizou, 1998). They both stimulated the cochlea with a single electrode at one location in the cochlea. Therefore the device cannot deliver the spectral place information explicitly; it provides only some temporal information in the form of an amplitude-modulated pulse train. However, even though the resulting sensation was just like an amplitude-modulated pure-tone, it still enabled patients to understand some speech, especially when combined with lipreading.

The cochlea acts like a spectrum analyzer, responding to different spectral acoustic components at different tonotopic locations. The high frequencies are picked up at the base of the cochlea (which is closest to the round window), and the lower frequencies are picked up at the apex (deeper into the cochlea). Greenwood (1990) quantified this mechanism of encoding frequencies in the cochlea with a simple equation, as explained later in Chapter 4. The main issue with single-channel implants was that they did not make use of this frequency-place encoding property of the cochlea at all.

Multichannel implants were first introduced in 1980s. Because they stimulate the cochlea at multiple sites with an array of electrodes, speech is presented with spectral resolution by making use of the place theory. Increasing the number of electrodes increased the patients' speech perception performance significantly, although only up to a certain number of electrodes. The implants that are currently used (Clarion from Advanced

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bionics, Nucleus from Cochlear, and Combi from Med-El) are all using multichannel processing strategies (see Table 2.1).

Since the FDA approval of implants with adult users in 1984, approximately 59,000 patients around the world have benefited from the device, according to Food and Drug Administration 2002 data[4]. In USA about 13,000 adults and 10,000 children have received the device. They work best with adults who were deafened postlingually as well as young children who still have a high potential of plasticity to develop flawless language skills.

2.2. DIFFERENCES IN HARDWARE AND PROGRAMMING OF IMPLANTS

Even though the basic hardware components (such as microphone, speech processor, and electrodes) and the main idea for processing are similar in all devices, they still differ in several ways:

1. Transmission of signals through skin: Percutaneous vs. transcutaneous.

In percutaneous transmission the cable for data transmission goes through an opening in the skin whereas in transcutaneous transmission the signals are transmitted through the skin using RF. The transmitting coil is placed outside the skull and coupled to the receiver that is located under the skull via a magnet. Newer devices all employ transcutaneous transmission.

[4] National Institute on Deafness and Other Communicative Disorders: http://www.nidcd.nih.gov/health/hearing/coch.asp

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

<u>2. The stimulation mode of the electrodes:</u> Monopolar (MP) vs. bipolar (BP).

In monopolar mode all electrodes use one common ground electrode for return current path whereas in bipolar mode electrodes are activated in pairs. Even though BP stimulation is advantageous with smaller current spread and therefore less channel interaction, it requires more current than MP stimulation to achieve same loudness level.

<u>3. Type of stimulation:</u> Analog (AS) vs. pulsatile (PS) stimulation.

Depending on the signal processing employed by the device the stimulation is given either in an analog or pulse waveform that is amplitude-modulated to represent the signal intensity at that particular band. In pulsatile stimulation pulses are presented in a non-overlapping scheme to minimize the channel interaction.

<u>4. Signal processing strategy:</u> Waveform presentation vs. feature extraction.

Different signal processing techniques will be discussed in Section 2.3.

<u>5. Electrode design:</u>

In all implants used currently, electrodes are organized consistently with the tonotopic organization of the cochlea such that high frequency information is given to the electrodes that are located more basally and low frequency information is sent to the electrodes located more apically. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

array is inserted through the round window into the scala tympani, and usually up to 15-30 mm deep. Even though it is known that the more spectral channels available the better the speech recognition, there appears to be a limitation on the number of the electrodes to be used both due to the physiology of the neurons and also the excitation pattern of the stimulation. When the electrode array is fully inserted the number of the electrodes used changes from 6 to 22 depending on the model of the CI. Devices that use CIS processing system employ less number of channels (6-8) because all electrodes are activated in one cycle of stimulation. In strategies such as SPEAK where only the electrodes carrying most of the energy are stimulated in a cycle, a higher number of electrodes is used (22) but only 4-6 of them are actually stimulated.

Recent developments in electrode design resulted in a much improved contact between the electrodes and the nerves as well as deeper insertion. The latest progress can be summarized as follows:

1. Advanced Bionics: The Clarion electrode positioner system is designed to push electrodes closer to the inner wall of the cochlea for a better contact. Again from Advanced Bionics, the HiFocus electrodes allow focused stimulation toward the auditory nerve minimizing the channel interactions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Cochlear: Nucleus Contour electrode has a pre-curved shape which is flattened out during the insertion and curves once all electrodes are located inside the cochlea, producing a closer contact between the electrodes and the inner wall of the cochlea.

3. Med-El: Combi 40+ electrode array of 12 electrode pairs is designed such that the electrodes in a pair face opposite directions and they are all mounted on a soft base. Therefore they can bend during the surgery as they move along the scala tympani allowing an insertion as deep as 31 mm. Because this electrode array is specifically designed for deep insertion its electrodes are distributed with wider spacing on the electrode array covering 26.4 mm cochlear length (Hochmair *et. al*, 2000). The company also provides alternative shorter electrode designs in case of ossification or malformation.

The exact location of the electrode array inside the cochlea is usually not known precisely, but estimated from observations during the surgery. Yet during the fitting of the device the patient is given one of the predefined frequency-place mapping tables. Regardless of the array position all patients typically get similar frequency assignments. In this study we propose that with a little more customized frequency-place assignment patient's speech perception performance can be increased.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

These properties are summarized in Table 2.1 for most commonly used devices. Loeb (1990) and Loizou, (1998) provide more detailed reviews of implants including older versions that are not commercially available anymore.

2.3. SIGNAL PROCESSING STRATEGIES

As mentioned in Section 2.2, signal processing strategies can be categorized in two groups mainly: waveform presentation and feature extraction (Loizou, 1998). In the waveform presentation speech is filtered through bandpass filters and the envelopes from each bands are used to modulate an analog or pulsatile carrier. The stimulation current is adjusted accordingly. In the feature extraction strategies, however, formant information is extracted from speech and presented to the cochlea.

Main signal processing techniques are as follows:

1. Compressed Analog (CA):

One of the earliest signal processing techniques for multichannel implants, the speech is basically filtered through a bank of bandpass filters (usually 4-8 bandpass filters). After adjusting gain controls the processor sends the filtered analog signals to the electrodes which stimulate the electrodes simultaneously.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Continuous Interleaved Sampling (CIS):

One of the main concerns in CA processing was the channel interaction due to simultaneous stimulation. To minimize the channel interaction the CIS technique was developed. Instead of the envelopes from bandpass filters directly stimulating the neurons all at the same time, they modulate trains of biphasic pulses that are delivered to the electrodes sequentially.

3. F0/F2, F0/F1/F2, and MPEAK:

Formants of the speech are extracted by bandpass filtering at specific frequency regions and using zero crossing detectors at the filter outputs, then finally sent to the appropriate electrode. Formant information is very useful for vowel recognition, yet the strategy did not yield a very good consonant recognition and therefore was improved further to deliver better high frequency information which would be useful for consonant recognition (MPEAK).

4. Spectral Maxima Sound Processor (SMSP) and Spectral Peak (SPEAK):

In SMSP strategy 22 electrodes are used, yet not all of them are activated at the same time. Instead, after the speech is bandpass filtered, the 6 largest amplitude outputs are chosen with a spectral peak detector and only corresponding electrodes stimulate neurons (with those envelopes). The envelopes modulate biphasic pulses at a constant pulse rate. SPEAK

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

processor takes the stimulation in a more flexible direction where the number of electrodes activated change from 5 to 10 (out of 20) depending on the spectral content of speech, and the pulse rate of the simulation changes accordingly as well.

5. Advanced Combination Encoders (ACE):

ACE is a hybrid between CIS and SPEAK combining higher stimulation rate of CIS with SPEAK's selectivity of high energy bands.

The speech processing techniques used by various implants are summarized in Table 2.1.

2.4. COMMONLY USED IMPLANTS

Currently, the most commonly used implants are Clarion by Advanced Bionics, Nucleus by Cochlear, and Combi by Med-El. Some properties of these implants are summarized in Table 2.1.

All implants can now be used with a behind the ear processor as well in addition to the conventional body worn processor.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Implant Model	Electrodes			Stimulation		Processing Strategy
	Number	Spacing (mm)	Array (mm)	Mode	Type	
Clarion I	8	2	16	MP/ BP	AS/ PS	CA/CIS/ MPS
Clarion II	16	0.8/1.1	12/16.5	MP/ BP	AS/ PS	SAS/MPS/ CIS/ High-Res Broadband
Nucleus 22	22	0.75	16.5	BP	PS	SPEAK
Nucleus 24	22+2	0.75	16.5	MP/ BP	PS	SPEAK/CIS/ ACE
Med-El Combi	12	2.4	26.4	MP	PS	CIS/n of m
Med-El Combi 40+	12 pairs	2.4	26.4	MP	PS	CIS

Table 2.1: Device properties of the most commonly used implants.

2.5. LIMITATIONS

Even though cochlear implants have restored hearing to many people with profound hearing loss there is still a limit to the extent that a cochlear implant user can understand speech. Speech recognition by CI users is rarely as good as normal hearing (NH) people and there is a big variation in performances among CI users. For example, Friesen *et al.* (2001) measured speech recognition performances of CI users and normal hearing subjects listening to simulated implant processor (as explained in Chapter 3) in noise while changing the number of the spectral channels. They found that the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

average implant performance was poorer than average NH subjects' performance which was also confirmed by other studies. NH groups' performance increased up to 20 channels. The variation among CI users was observed in the study by comparing CI user performances to NH subjects' performances: CI users with better performance showed a similar pattern up to 7 channels whereas CI users with poorer performance had speech recognition performance improved only up to 4 channels. The study showed that most CI users are not able to make full use of the spectral information coming from devices fully.

Similar pattern of results was observed in this study, as shown in the following chapters. The average CI user speech recognition performance was generally lower compared to results with NH people from simulations even under the same processing conditions and there was considerably larger variation in CI user performances.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 3

SIGNAL PROCESSING FOR SIMULATIONS

For the simulations, the noise-band vocoder technique described by Shannon *et al.* (1995) was used. The noise-band vocoder is a powerful tool for speech experiments because it enables researchers to separate speech into its components, such as temporal (i.e., envelope information) or gross spectral properties, or fine spectral details, and manipulate them separately. Cochlear implant users get only crude spectral information from the place of the stimulation determined by the location of the electrodes along the cochlea. They also get envelope information from the amplitude-modulated current pulses that are delivered through each electrode. Yet, almost nothing else is known in electrode-neuron interaction, and it is believed that they do not get any fine spectral details of speech, which was also supported by many studies. The noise-band vocoder used in the simulations keeps most of the envelope (by envelope extraction at each channel) and gross spectral information (by having different channels at different center frequencies and as a result, delivering a distinct place information). Yet it deletes all fine spectrum by modulating noise bands which replace the fine spectral details.

In this simulation technique all bands nonselectively carry information at any given time instead of selecting the ones with the highest energy. Therefore it resembles Continuous Interleaved Sampling (CIS) strategy most

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

closely where all electrodes are activated in one stimulation cycle in an interleaved order. There are a few reasons for this choice:

1. It is easier to quantify the effects of spectral manipulations. When some acoustic information is missing or mismatched it is easier to calculate the proportion of that acoustic information to the whole acoustic content of the speech material and explain the results from this theoretical point of view.

2. There is not much difference in the speech recognition performances of normal-hearing people listening to CIS or Spectral Peak (SPEAK) simulations. Even though the sound quality of the processed tokens is different with different strategies the intelligibility does not vary significantly and therefore practically, the choice for strategy does not affect the resulting patterns of speech recognition performances.

3. CIS processing yields slightly more natural sounding stimulus in the simulations. A typical stimulation period for SPEAK strategy is 4 ms. In the simulations this means that every 4 ms the activated 6 bands are changed and a new set of 6 bands with highest energy are employed. In the simulation such abrupt changes contribute to the coarseness of the processed sounds.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Nevertheless, the frequency region of the speech is taken from 150 Hz to 10 kHz which is closest to SPEAK Table 9, a commonly used frequency-to-electrode mapping procedure. Hence, the simulation is based on CIS speech processing strategy with the input frequency range of SPEAK strategy. The resulting processor is a combination of several processing strategies, yet all values are chosen to simulate real implant parameters.

For this study, the vocoder was implemented in Matlab to generate the stimuli. First, speech materials were bandpass filtered into a number of contiguous frequency bands by 6^{th} order Butterworth filters. The -3 dB cut-off frequencies were determined according to Greenwood's (1990) frequency-to-place mapping equation, given in Equation 4.1 and described in Chapter 4 extensively. The exact frequency ranges and cut-off frequencies were determined depending on the specific experimental conditions. The speech envelope was extracted from each band by half-wave rectification and low-pass filtering using a 3^{rd} order Butterworth filter whose output was 3 dB down at a cut-off frequency of 160 Hz. The noise carrier bands representing the stimulation region in the cochlea were obtained from white noise by 6^{th} order Butterworth filters where the cut-off frequencies (-3 dB) were determined by the condition. The extracted speech envelopes were used to amplitude-modulate the noise carrier bands, and all modulated noise bands were combined to form the processed speech. The amplitude level of the processed speech was adjusted such that the original and processed tokens had the same overall RMS energy. The signal processing steps for a 4-band processor are summarized in Figure 3.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

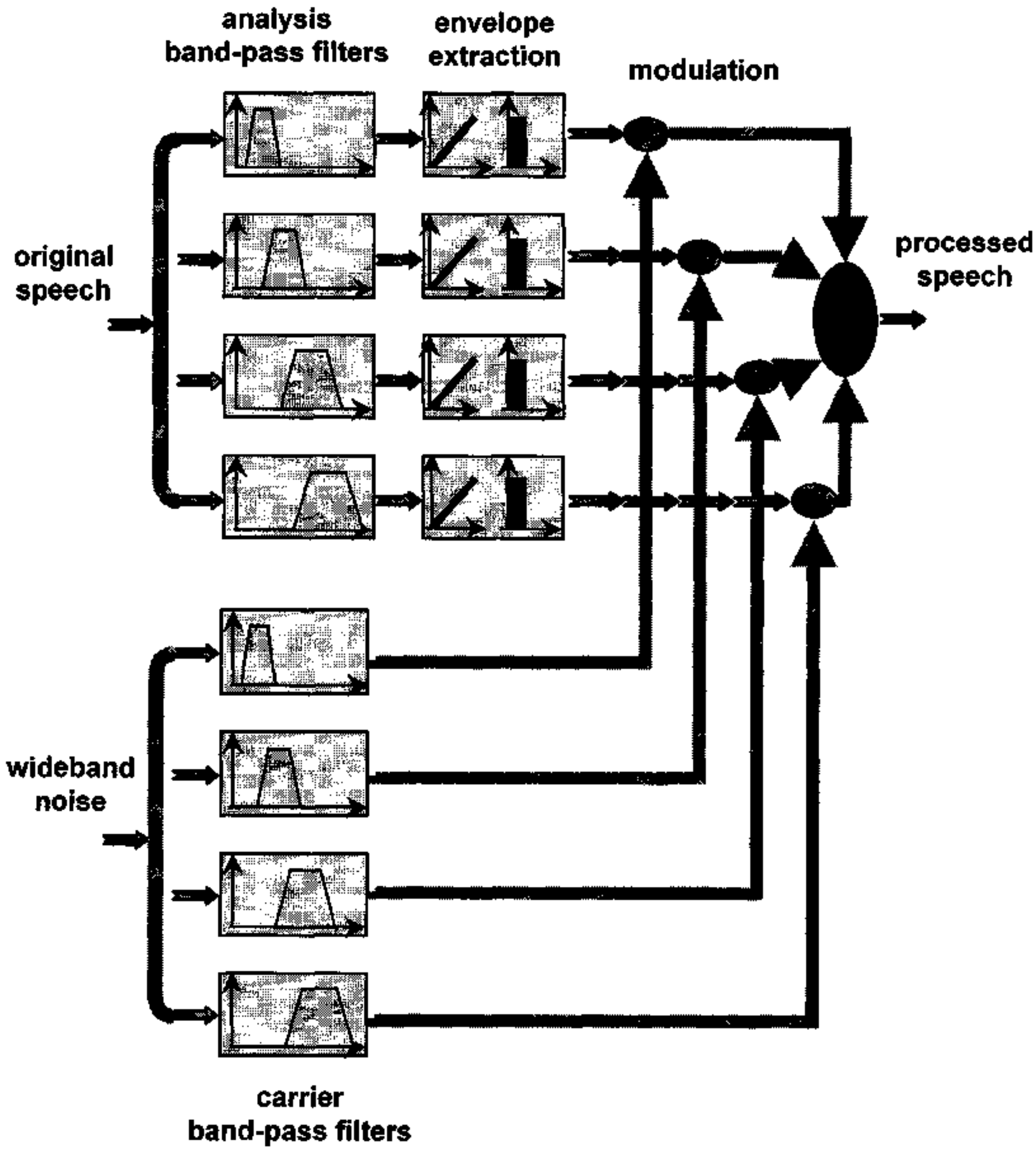


Figure 3.1: 4-band noise-band vocoder.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4

FREQUENCY-PLACE COMPRESSION AND EXPANSION

In most commercially available cochlear implants, the electrode array is typically inserted into the scala tympani, reaching a depth of 15-25 mm inside the round window. The active stimulation range is typically 12-17 mm in length for many commonly used implants (Table 2.1). As mentioned before, newer models have longer electrode arrays or deeper insertion depths but many patients who already have the devices have electrodes of those lengths shown in the table.

In this chapter, the resulting compression and expansion in the frequency-place mapping from such a configuration will be introduced. The effects of such spectral mismatch on speech recognition will be presented with percent correct scores from normal-hearing listeners tested with mismatch conditions simulated with the noiseband vocoder described in the previous chapter.

4.1. GREENWOOD FREQUENCY-PLACE MAPPING FUNCTION

Cochlea mainly behaves like a spectrum analyzer and each location on the cochlea best responds to a specific frequency. This frequency-to-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

place relation in the cochlea was expressed by Greenwood (1990) in an equation given by

$$f(x) = A(10^{a(L-x)} - k), \tag{4.1}$$

where *x* is the distance in mm from the round window in the cochlea and *f(x)* is the frequency in Hz that produces the most excitation at that distance in the cochlea. *L* is the cochlear length and assumed as *L=35 mm* for human in the study. 35 mm is widely accepted as the average human cochlea length but it can be a few mm shorter or longer as shown by Ulehlova *et al.*,(1987), Ketten *et al.*, (1998), and Skinner *et al.*, (2003). *a* is a species-specific constant inversely related to the size of the cochlea such that *a=2.1/L*, which comes out as *a=0.06* for *L=35 mm*. The fitting parameters *A* and *k* are determined by assuming that the whole cochlea can respond to frequencies from 20 Hz to 20 kHz. With this assumption the best fitting values are found as *A=165.4* and *k=0.88*. This equation implies a relatively linear sensitivity for low frequencies (at the apex of the cochlea) as a function of distance from the round window, and a logarithmic function for higher frequencies (at the base of the cochlea), as shown in Figure 4.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

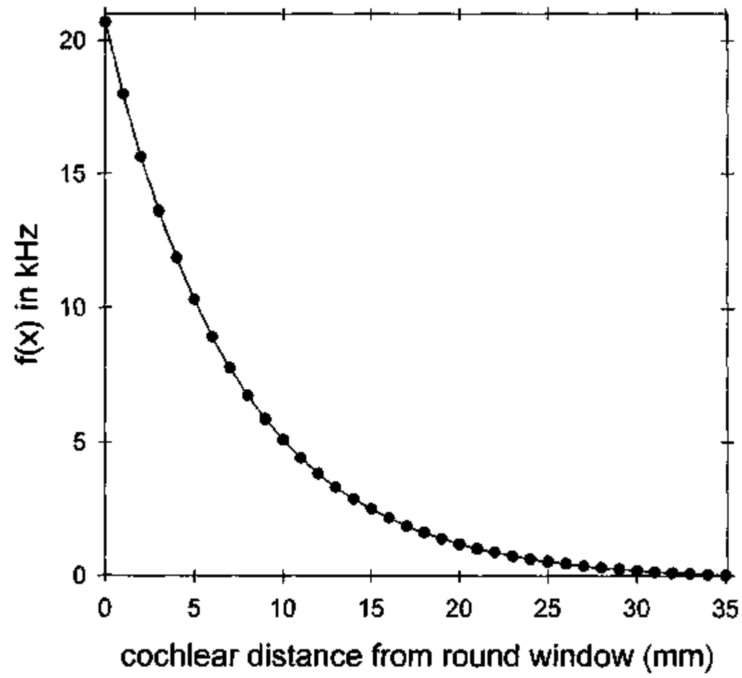


Figure 4.1: Greenwood's frequency-place mapping function.

Table 4.1 presents some values for the frequency-place mapping equation (Equation 4.1) with *L=35 mm, A=165.4, a=0.06, and k=0.88.* These values are used throughout the study to determine the matching tonotopic location of a specific frequency in the cochlea.

x	0 mm	5 mm	10 mm	15 mm	20 mm	25 mm	30 mm	35 mm
f (x)	20 kHz	10 kHz	5 kHz	2.5 kHz	1 kHz	500 Hz	200 Hz	~20 Hz

Table 4.1: Greenwood's frequency-place mapping values.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.2. FREQUENCY-PLACE COMPRESSION AND EXPANSION IN IMPLANTS

According to Equation 4.1, an electrode array of 16 mm with a 25 mm insertion stimulates a cochlear region that responds best to an acoustic frequency range of 500-6000 Hz in human. However, most commercial implant speech processors assign a much wider acoustic frequency range to this limited cochlear region. For example, the default frequency allocation of the Nucleus-22 implant (SPEAK strategy Table 9) assigns a frequency range of 150 Hz - 10 kHz to the electrodes (Cochlear Corporation, 1995). This acoustic range would normally cover a 25 mm range in the cochlea, rather than the 17 mm covered by the electrode array. Thus, mapping the larger acoustic frequency range onto the electrode array results in a linear compression in the frequency-to-place mapping.

Similarly, if the electrode array is as long as 26 mm and inserted to 31 mm, such as with Combi implant, it covers a cochlear region from 4 mm to 31 mm from the round window. This cochlear range corresponds to frequencies from 150 Hz to 12 kHz. Yet the body-worn processor for Combi can only deliver frequencies from 200 Hz to 5.5 kHz, which would normally cover a cochlear range of 20 mm. Thus a smaller acoustic range is mapped onto a wider cochlear range resulting in an expansion of the mapping.

As a result there are mainly three parameters affecting the speech perception in such a setting: the amount of acoustic information delivered, the accuracy of the cochlear location where this acoustic information is delivered, and the spectral resolution. By using noise-band vocoders with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

normal hearing subjects the effects of those parameters on speech recognition and the relationship to each other are explored. The results will allow us to understand which cues are more crucial for speech recognition so that more attention can be paid to those for better speech processing.

4.3. SIMULATION OF FREQUENCY-PLACE COMPRESSION AND EXPANSION

4.3.1. Experimental Method

4.3.1.1. Subjects

Six normal-hearing listeners, aged 26 to 34, participated in this part of the simulations. All subjects were native speakers of American English and had thresholds better than 20 dB HL at audiometric frequencies between 125 and 8000 Hz. One subject was excluded from the sentence recognition test because she was already familiar with the sentences in the database. Another 32-year-old subject was added to the sentence recognition test to maintain 6 subjects for each test.

4.3.1.2. Stimuli

The speech recognition tasks consisted of medial vowel and consonant discrimination, and sentence recognition. All stimuli were presented via a loudspeaker in a sound field at 70 dB on an A-weighted scale.

Consonant stimuli were taken from materials recorded by Turner *et al.* (1992, 1999) and Fu *et al.* (1998) at a 44.1 kHz sampling rate. Six

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

presentations (3 male and 3 female talkers) were made of 14 medial consonants /b d f g k m n p s ʃ t θ v z/, presented in an /a/-consonant-/a/ context. Tokens were presented in random order by custom software (Robert, 1998), and subjects were instructed to select the consonant they heard from the set of 14 possible consonants displayed on the screen. The resulting consonant confusion matrices were analyzed for information received on the production-based categories of voicing, manner, and place of articulation (Miller and Nicely, 1955). Chance performance level for this test was 7.14% correct, and the single-tailed 95% confidence level was 11.77% correct based on a binomial distribution.

Vowel stimuli were taken from the phoneme set recorded by Hillenbrand *et al.* (1995) at a 32 kHz sampling rate. The tokens were presented to the listeners in random order via custom software (Robert, 1998), and subjects were instructed to select the vowel they heard from the set of 12 possible vowels displayed on the screen. Ten presentations (5 male and 5 female talkers) were made of twelve medial vowels, including 10 monophthongs and 2 diphthongs presented in a /h/-vowel-/d/ context (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard, hayed, hoed). Chance level on this test was 8.33% correct, and the single-tailed 95% confidence level was 12.48% correct based on a binomial distribution.

Recognition of words in sentences was measured using the custom software (Tiger Speech Recognition System developed by Qian-Jie Fu) with Texas Instruments/Massachusetts Institute of Technology (DARPA/TIMIT) corpus of sentence materials (National Institute of Standards and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Technology, 1990). The sentences were of moderate-to-hard difficulty, such that individual words were difficult to predict from the context of the sentence, and sentences were spoken by multiple talkers. For each condition, the percent correct score was acquired for 20 sentences of varying length from each listener. The length of the sentences varied from 3 words to 12 words. The groups of 20 sentences were prepared such that the average word length per sentence was 6-8 words for each sentence. They were presented without any context information and no sentences were repeated to an individual listener. Sentences were not balanced for difficulty; so 20 sentences were used for each condition to obtain a sample that included varying levels of difficulty. In addition, the order of the presentation of sentences was completely randomized using a random number generator for each subject so that all subjects heard different sentences for different conditions. Consequently differences arising from varying difficulty of the sentences were randomly distributed across different conditions and different subjects. Subjects were asked to repeat what they had heard. The percent correct score was obtained by counting the percentage of words repeated correctly by the subject.

This study concentrates on effects of frequency-place compression and expansion on speech recognition without any learning effects. We typically observe a short-term adaptation to the test procedure by inexperienced subjects where all scores (regardless of the condition) increased slightly over the first three days of the testing. However, the scores remained more or less stable after this initial adaptation period. This was not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

observed with subjects who already had experience in similar experiments. To minimize any learning effects for the specific experimental conditions no practice was provided on any conditions prior to data collection, even for new subjects, and no feedback was given in any part of the testing. In addition, to reduce a possible adaptation to a particular condition, all conditions with all stimuli were presented to subjects in a completely random order. Therefore any effects of learning on scores would be distributed across different conditions with different subjects.

4.3.1.3. Signal Processing

The noise-band vocoder technique introduced in Chapter 3 was used to simulate the frequency-place compression and expansion conditions. The cut-off frequencies were determined according to Greenwood's (1990) frequency-to-place mapping equation given in Equation 4.1. The exact frequency ranges and cut-off frequencies were determined depending on the specific experimental conditions (as shown in Tables 4.2 and 4.3).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	15 – 9	2476	3080	3822	4736	5860	2476 - 5860
-3 mm (expansion)	17 – 7	1843	2663	3822	5459	7771	1843 - 7771
-1 mm (expansion)	19 – 5	1363	2301	3822	6289	10290	1363 - 10290
0 mm (matching)	20 – 4	1168	2138	3822	6749	11837	1168 - 11837
+1 mm (compression)	21 – 3	999	1985	3822	7243	13612	999 - 13612
+3 mm (compression)	23 – 1	722	1710	3822	8337	17990	722 - 17990
+5 mm (compression)	25 – 0	513	1471	3822	9594	23762	513 - 23762

Table 4.2: Frequency-place mismatch conditions for the 4-channel processor at the simulated 20 mm electrode insertion depth. The name of each condition represents the change in frequency range expressed in mm between the analysis and carrier bands. For each condition the table lists the following information for the analysis bands: cochlear location in mm from the round window, cut-off frequencies for a four-band processor, and total frequency range. Because the simulated electrode location was fixed, the noise carrier bands covered the frequency range from 1168 to 11837 Hz in all conditions, and the frequency partition of carrier bands was as shown in the center entry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	20 – 14	1168	1471	1843	2300	2864	1168 - 2864
-3 mm (expansion)	22 – 12	851	1262	1843	2663	3822	851 - 3822
-1 mm (expansion)	24 – 10	611	1081	1843	3080	5085	611 - 5085
0 mm (matching)	25 – 9	513	999	1843	3310	5860	513 - 5860
+1 mm (compression)	26 – 8	428	922	1843	3557	6750	428 - 6750
+3 mm (compression)	28 – 6	290	785	1843	4106	8944	290 - 8944
+5 mm (compression)	30 – 4	184	665	1843	4736	11837	184 - 11837

Table 4.3: Frequency-place mismatch conditions for a 4-channel processor at the simulated 25 mm electrode insertion depth. For each condition the table lists the cochlear locations of the analysis bands, cut-off frequencies of the band-pass filters, and the total analysis frequency range. The noise carrier bands were fixed between 513 and 5860 Hz with the partition shown as in the center of the table. The +5 mm condition is the one most similar to frequency-to-electrode assignment used in a cochlear implant with a full electrode insertion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.3.1.4. Mapping Conditions

Speech materials were processed using 4, 8, or 16 frequency bands. Two different electrode array locations were simulated, representing insertion depths of 20 mm and 25 mm from the round window. For all expansion and compression conditions the simulation region covered by the simulated electrode array was fixed at 16 mm (comparable to the typical length of the electrode array for many implant devices). The 20 mm insertion depth condition simulated an electrode array located between 4 and 20 mm from the round window, and the 25 mm insertion depth condition simulated a location between 9 and 25 mm from the round window. Because the simulation region was fixed at 16 mm, electrode locations were represented by noise bands that were 4 mm wide in terms of cochlear location for the 4 band condition, 2 mm wide for the 8 band condition, and 1 mm wide for the 16 band condition.

In the simulation, the noise carrier bands determine the cochlear location stimulated. The "acoustic analysis bands" are the filters used to process and extract the acoustic envelope information used to modulate the carrier bands (Figure 4.2). The distribution of carrier bands was kept fixed for each simulated insertion depth while the analysis bands were systematically altered to create the conditions of frequency-place expansion or compression.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

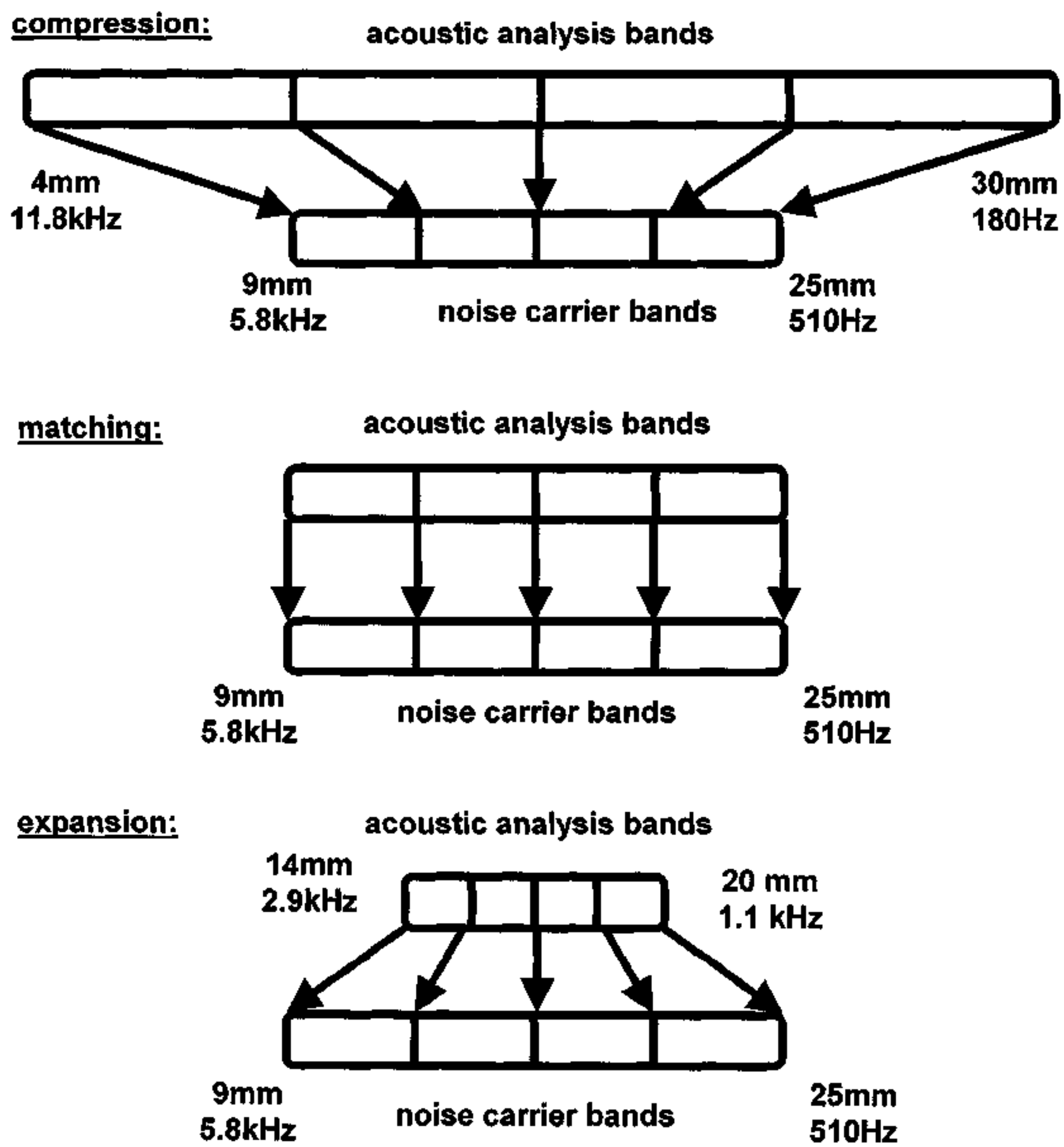


Figure 4.2: Frequency-place mapping conditions for 4-channel processor at the simulated 25 mm electrode insertion depth. The speech envelope was extracted from the analysis bands and used to modulate the noise carrier bands. The carrier bands were fixed for all conditions at the simulated 25 mm insertion depth (9-25 mm: 510-5800 Hz). The top panel shows the +5 mm condition schematically: the analysis bands are mapped onto wider carrier bands. The middle panel shows the 0 mm condition schematically, in which the analysis and carrier bands are matched. The lower panel shows the -5 mm expansion condition schematically, in which analysis bands are mapped onto wider carrier bands.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The cut-off frequencies of each analysis band and the related cochlear locations for the 4 band processors are summarized in Table 4.2 for the 20 mm simulated insertion depth and Table 4.3 for the 25 mm simulated insertion depth. Filters for the 8 and 16 band conditions were determined by dividing the four bands into two or four equal parts (in mm) and using Greenwood's (1990) formula to determine the acoustic frequencies for the band edges. The 0 mm condition refers to a perfect match between analysis and carrier bands. Thus, the cochlear location and frequency cut-offs listed for the 0 mm condition specify the fixed locations and frequencies of both analysis and carrier bands. Cochlear locations are all specified in terms of mm from the round window, using Greenwood's (1990) formula, assuming a 35 mm length for the human cochlea. In the +5 mm condition, the analysis band range was 5 mm wider than the carrier band range on both apical and basal ends, causing a frequency-place compression of approximately two octaves. Similarly, in the -5 mm condition, the analysis band range was 5 mm shorter on each end, causing a frequency-place expansion of about two octaves. The +5 mm compression condition in Table 4.1 most closely simulates the typical frequency-place compression observed in the standard clinical map of the Nucleus speech processor. Also, +5 mm compression and -5 mm expansion conditions for the simulated 25 mm insertion depth are also shown schematically in Figure 4.2, as well as the matching case. Here, one can see that the simulated electrode locations are from 9 mm to 16 mm from the round window and the frequency range used to modulate each noise carrier band is different for every condition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As the analysis filters were changed from -5 mm to +5 mm conditions, the amount of acoustic information delivered was changed. An important control condition was included to evaluate the effect of the varying amount of acoustic information. In these control conditions the analysis bands and noise carrier bands were always matched in frequency-place. These baseline conditions were not intended to simulate any electrode insertion depth or spacing, because in a cochlear implant the electrode position and length are fixed after the implant surgery. Rather, the baseline conditions only assess the effect of changing the overall acoustic bandwidth. Performance in the baseline condition indicates the effect of the gain or loss of acoustic information resulting from the expansion or truncation of the analysis frequency range. The difference in performance between the baseline condition and the compression-expansion condition is due to the frequency-place distortion only.

4.3.2. Speech Recognition Results

Percent correct scores for consonants, vowels, and sentences were obtained with 4, 8, and 16-channel processors at simulated insertion depths of 20 mm and 25 mm. In Figures 4.3-4.10, the number of channels increases from 4 to 8 to 16 in the left, middle and right panels, respectively, of each figure. The average percent correct scores of six subjects, corrected for chance [*p = 100*(score-chance)/(100-chance)*], are plotted for consonant and vowel recognition. The average score of six subjects is plotted for sentence recognition. Within each panel the filled symbols present results

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from the baseline conditions in which analysis and carrier bands were always matched, and the open symbols present results from the experimental conditions in which the frequency-place mapping was expanded or compressed.

4.3.2.1. Consonant Recognition

The consonant recognition scores are presented in Figure 4.3 for the 20 mm simulated insertion depth and Figure 4.4 for the 25 mm simulated insertion depth.

First consider the data from the baseline conditions (filled symbols), where the analysis and carrier bands were always matched. Consonant recognition increased only slightly as the analysis (and matched carrier) bands were widened (+5 mm baseline condition) from the simulated electrode range. However, as the analysis and carrier bands were narrowed (-5 mm baseline condition) there was a loss of acoustic information (due to the reduced acoustical bandwidth) that resulted in lower consonant recognition. This reduction in consonant recognition was more severe in the simulated 20 mm insertion depth condition (Figure 4.3) because more low-frequency information was eliminated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

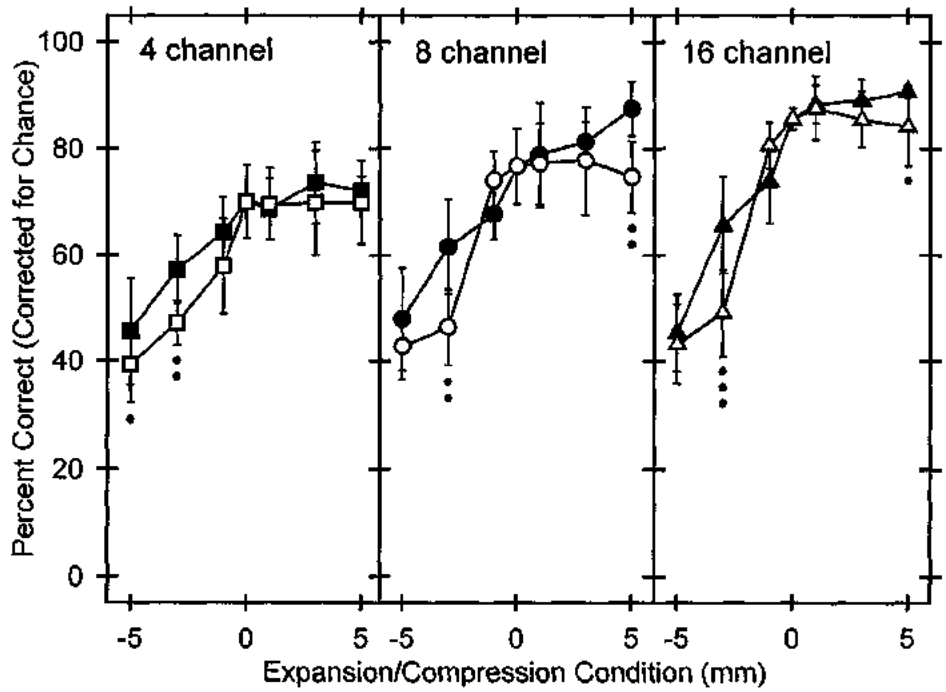


Figure 4.3: Consonant recognition percent scores for the simulated 20 mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping. The number of spectral bands increases from 4 to 8 to 16 in the left, middle and right panels, respectively. The percent correct scores represent the average performance of 6 normal-hearing subjects, corrected for chance, and the error bars represent the standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band frequency range. Open symbols denote the compression-expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied. The dots under the scores denote significant differences between the baseline (filled symbols) and mismatch (open symbols) conditions: one dot indicates p<0.05, two dots for p<0.01, and three dots for p<0.001.

Next consider the data from the experimental mismatch conditions (open symbols), where the analysis filter frequency range was smaller, equal to, or larger than the simulated electrode length, resulting in frequency-place expansion, matching, or compression, respectively. Note that performance in the mismatched conditions was always equal to or poorer than the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

baseline conditions. Thus, there are two contributing factors to the reduced performance: (1) the reduction in the amount of information delivered, and (2) the distortion in frequency-place mapping.

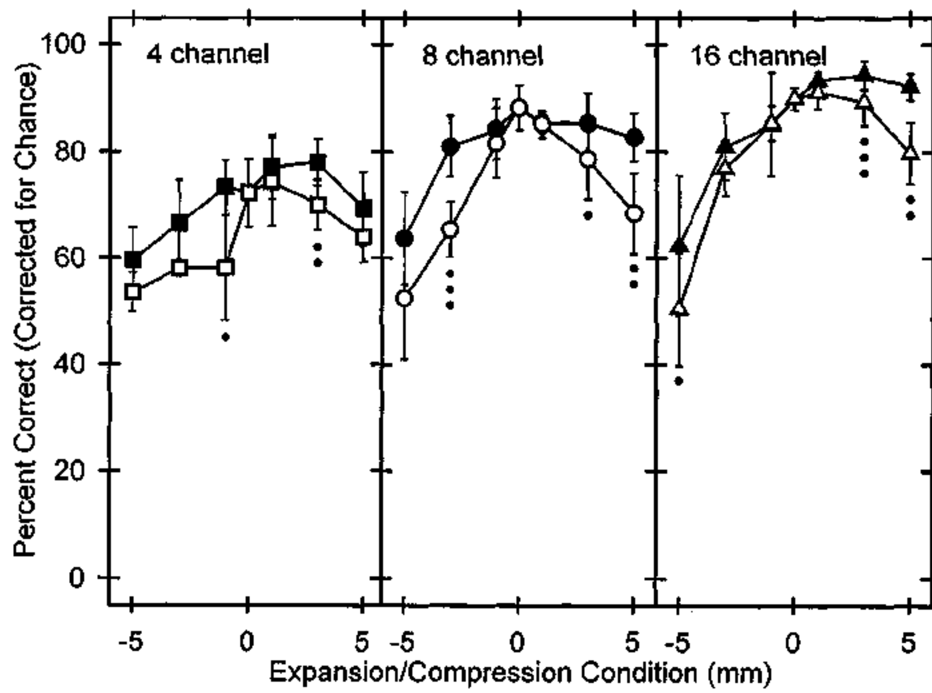


Figure 4.4: Consonant recognition, similar to Figure 4.3, for carrier bands simulating a 25 mm insertion depth.

A one-way repeated-measures ANOVA test was used to assess the significance of the drop in the performance with expansion/ compression mismatch conditions from the matched condition. Each run included only the expansion or compression percent scores in addition to the 0 mm matched percent score at a particular insertion depth with a specific number of channels. The baseline scores were not included in the ANOVA to isolate

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the effect of expansion or compression only on speech recognition. The baseline condition was compared to the corresponding mismatch condition with paired t-tests.

expansion at 20 mm insertion	F(3,15)	p	compression at 20 mm insertion	F(3,15)	p
4 channel	54.60	<0.001	4 channel	0.01	1
8 channel	89.05	<0.001	8 channel	0.23	0.87
16 channel	117.79	<0.001	16 channel	0.66	0.59
expansion at 25 mm insertion	**F(3,15)**	**p**	**compression at 25 mm insertion**	**F(3,15)**	**p**
4 channel	15.10	<0.001	4 channel	7.40	<0.01
8 channel	62.60	<0.001	8 channel	26.99	<0.001
16 channel	90.26	<0.001	16 channel	15.38	<0.001

Table 4.4: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for consonant recognition at 20 mm and 25 mm simulated insertion depths.

The analysis revealed that all frequency expansion conditions reduced performance significantly from the 0 mm matched condition. Corresponding F and p values are shown in Table 4.4. A substantial amount of this drop was due to the loss of acoustic information, as indicated by the filled symbols. An additional drop was observed for some conditions when the frequency-place mapping was expanded (open symbols).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A paired t-test analysis compared the baseline (filled symbols) and expansion (open symbols) performance. Conditions that are significantly different are indicated with stars on the bottom of the figure. One star denotes significant difference with a p value of $p<0.05$, two stars denote $p<0.01$, and three stars denote $p<0.001$. The analysis shows a significant difference for -3 mm expansion for most processors at both insertion depths. The difference was generally not significant for -5 mm expansion, possibly because the performance was limited by a floor effect (previous studies have found consonant recognition scores as high as 30-40% even for single-channel noise processors, indicating that this level of performance is possible using only temporal cues: Van Tasell *et al.*, 1987; Shannon *et al.*, 1995). The difference was also not significant for -1 mm expansion, which produced performance similar to the 0 mm matched condition.

Thus, even though the cochlear tonotopic representation of the spectral information was expanded, resulting in improved spectral resolution within the pattern, performance was poorer than the matched condition. This result suggests that improved resolution in the spectral domain does not necessarily improve speech recognition, probably because the information is not in the appropriate cochlear place. In the present experiment, which did not provide any practice or time to accommodate to the new mapping, expansion in frequency-place mapping always resulted in poorer consonant recognition. It is possible that additional practice with the experimental processors would have resulted in improved performance.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Frequency-place compression did not have a significant effect on consonant recognition for the 20 mm simulated insertion depth, yet there was significant decrease from the matched 0 mm condition (10%-20% drop at +5mm compression) in performance for the 25 mm simulated insertion depth (as shown in Table 4.4). For the extreme condition of frequency-place compression (+5 mm) there was a significant reduction in performance of 10-15% relative to the baseline condition for 8 and 16 channel processors.

There was no clear difference between the pattern of results for 4, 8, and 16 channels other than the overall improvement in performance with more channels.

4.3.2.2. Consonant Feature Analysis

Information transmitted on the consonant features of place, manner, and voicing is plotted in Figures 4.5 and 4.6 for 20 mm and 25 mm simulated insertion depths, respectively. Within each figure the top, middle and lower panels show the percent of information transmitted on place, manner, and voicing, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

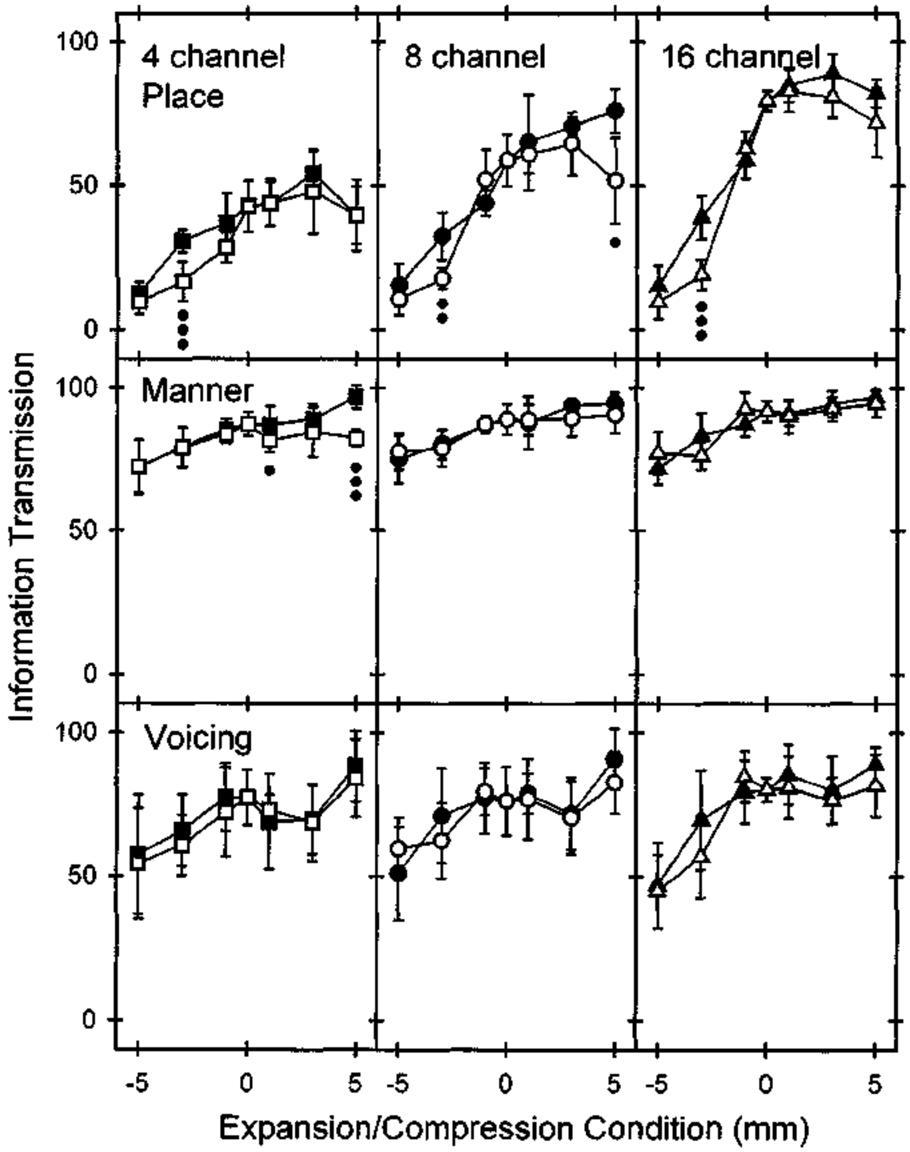


Figure 4.5: Information transmission percent scores for consonant features at 20 mm simulated insertion depth as a function of frequency-place mismatch conditions. The features are grouped into production-based categories of place, manner, and voicing, in the top, middle and bottom rows, respectively. The number of spectral bands increases from 4 to 8 to 16 in the left, middle and right panels, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

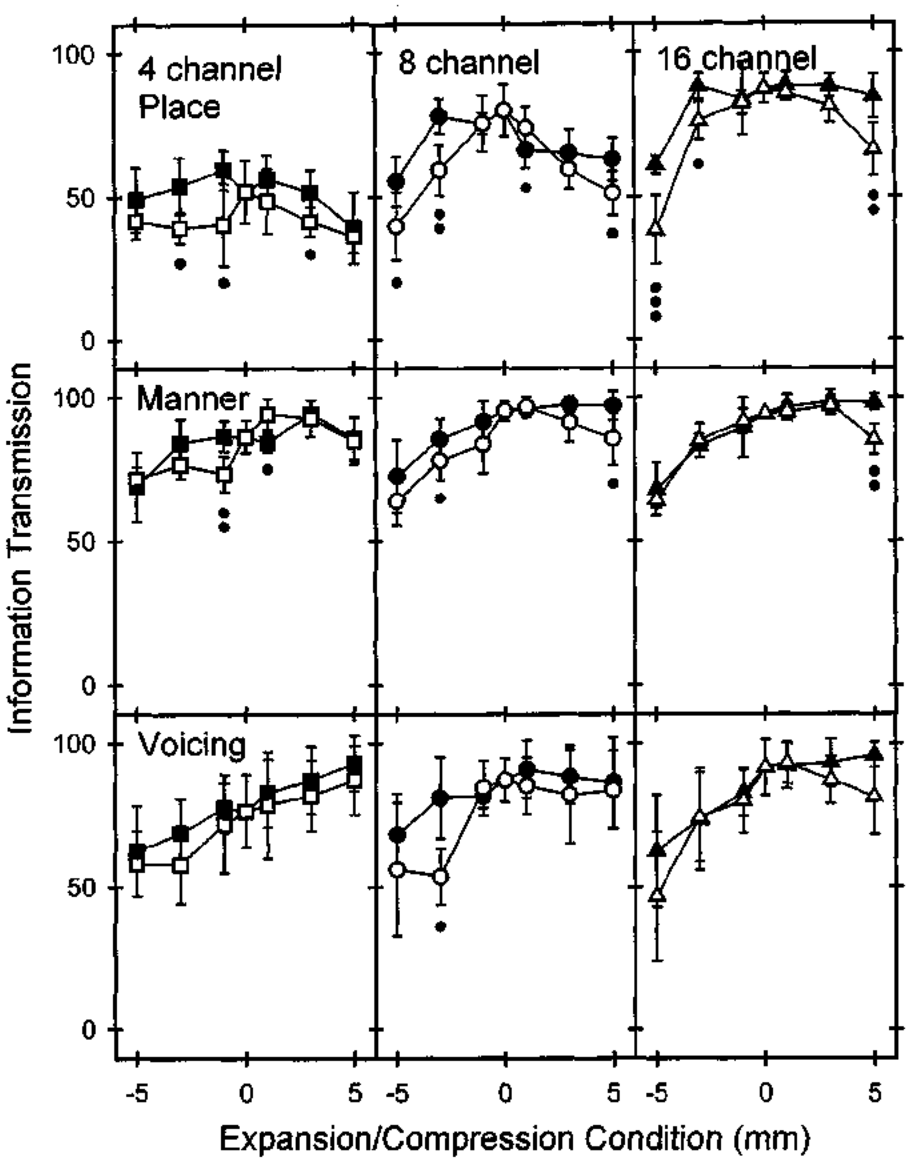


Figure 4.6: Consonant feature information transmission, similar to Figure 4.5, for carrier bands simulating a 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Information transmission percent scores are calculated from the confusion matrices where the diagonal entries are the numbers of correct answers and the off-diagonals are the confusions. A measure for the transmission of information is the covariance between the input and the output as given by

$$Cov(x,y) = -\sum_{i,j} p_{ij} \log \frac{p_i p_j}{p_{ij}}, \tag{4.2}$$

where p_i, p_j, p_{ij} are directly related to the frequencies of occurrences of stimulus *i*, response *j*, and joint occurrence of stimulus *i* and response *j*, respectively. Next, the covariance is converted to the information transmission percent score by normalization such that it yields 100% when the subject identifies all phonemes accurately (Miller and Nicely, 1955).

Note that information received on manner at 20 mm simulated insertion and on voicing at both insertion depths were not affected by frequency-place compression and manner information received at 25 mm simulated insertion depth was only slightly affected. Yet both manner and voicing information transmission scores dropped significantly with expansion (See Table 4.5 for corresponding F and p values.). Also the compression/expansion mismatch scores were similar to baseline scores for manner and voicing implying that these features are mostly affected by the bandwidth of acoustic information. The performance for both features was similar for different number of channels and for the two simulated insertion depths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expansion at 20 mm insertion	place		manner		voicing	
	F(3,15)	p	F(3,15)	p	F(3,15)	p
4 channel	63.42	<0.001	9.09	<0.01	6.90	<0.01
8 channel	63.01	<0.001	10.18	<0.001	12.97	<0.001
16 channel	347.90	<0.001	21.15	<0.001	25.77	<0.001
compression at 20 mm insertion	**place**		**manner**		**voicing**	
	F(3,15)	p	F(3,15)	p	F(3,15)	p
4 channel	1.55	0.24	1.51	0.25	4.47	<0.05
8 channel	1.48	0.26	0.16	0.92	3.52	<0.05
16 channel	2.49	0.10	2.43	0.11	0.62	0.62
expansion at 25 mm insertion	**place**		**manner**		**voicing**	
	F(3,15)	p	F(3,15)	p	F(3,15)	p
4 channel	2.19	0.13	8.98	<0.01	6.10	<0.01
8 channel	38.92	<0.001	24.90	<0.001	16.32	<0.001
16 channel	55.35	<0.001	72.88	<0.001	26.62	<0.001
compression at 25 mm insertion	**place**		**manner**		**voicing**	
	F(3,15)	p	F(3,15)	p	F(3,15)	p
4 channel	7.46	<0.01	3.77	<0.05	1.19	0.35
8 channel	25.38	<0.001	5.06	<0.05	0.75	0.54
16 channel	15.59	<0.001	15.43	<0.001	6.04	<0.01

Table 4.5: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for consonant feature recognition at 20 mm and 25 mm simulated insertion depths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In contrast, place information was strongly affected by expansion at both simulated insertion depths, and by compression at 25 mm simulated insertion depth (Table 4.5). The overall pattern of the performance changed from 20 mm insertion to 25 mm insertion and the information received increased with increasing number of channels. This pattern is very similar to the consonant recognition results of Figures 4.3 and 4.4 and therefore it appears that the overall shape of consonant recognition performance was primarily determined by the loss of place information. This observation agrees well with previous studies, which found that manner and voicing cues are more robust to spectral manipulations (Shannon *et al.*, 1998).

4.3.2.3. Vowel Recognition

Vowel recognition scores for simulated 20 mm and 25 mm insertion depths are presented in Figures 4.7 and 4.8, respectively. Vowel recognition was much more strongly affected by frequency-place mismatch than consonant recognition; performance decreased significantly from the matched 0 mm condition with both expansion and compression mismatch conditions (F and p values obtained by one-way repeated-measures ANOVA are given in Table 4.6). Overall performance improved as more spectral bands were used but the pattern of results was similar for 4, 8, and 16 bands. For 8 and 16 channels, vowel recognition decreased as the analysis frequency range was reduced (baseline condition going from +5 mm to - 5 mm), and a further drop in recognition was seen for both frequency-place compression and expansion as a result of the mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

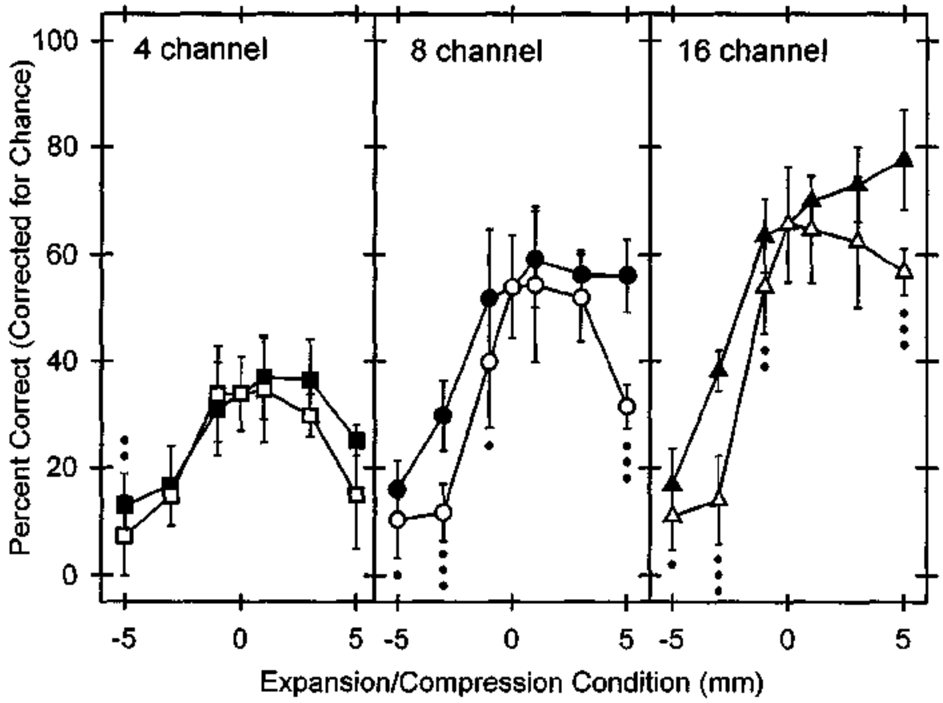


Figure 4.7: Vowel recognition percent scores for carrier bands simulating a 20 mm insertion depth, as a function of compression or expansion in the frequency-place mapping. The number of spectral bands increases from 4 to 8 to 16 in the left, middle and right panels, respectively. The percent correct scores represent the average performance of 6 normal-hearing subjects, corrected for chance, and the error bars represent one standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band range. Open symbols denote the compression-expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied. Dots under the scores denote significant differences between the baseline (filled symbols) and mismatch (open symbols) conditions: one dot indicates $p<0.05$, two dots for $p<0.01$, and three dots for $p<0.001$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

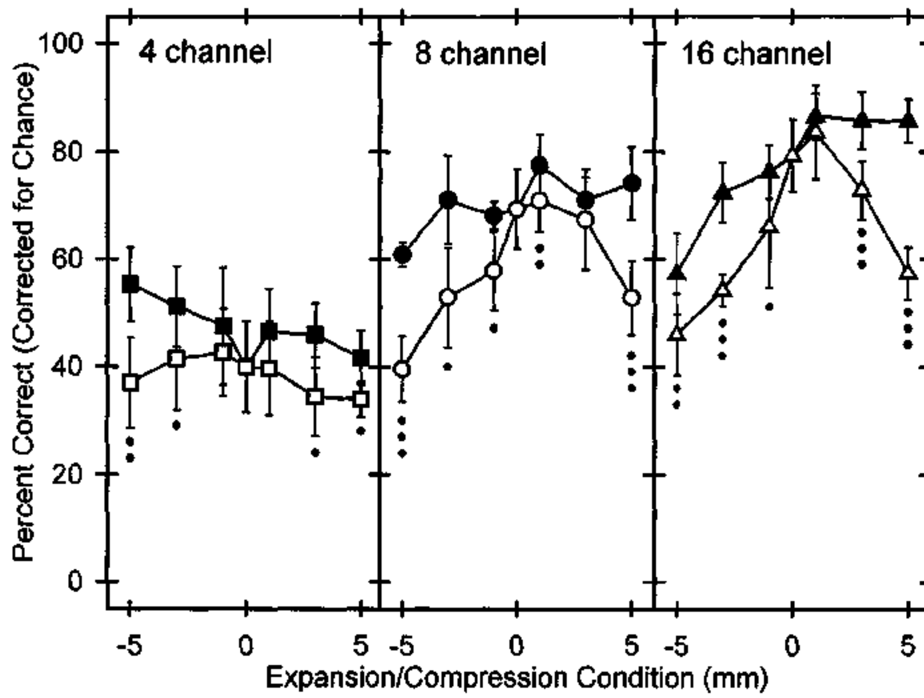


Figure 4.8: Vowel recognition percent scores, similar to Figure 4.7, for noise carrier bands simulating a 25 mm electrode insertion depth.

As in consonant recognition the dots under the percent correct scores in each panel indicate significant difference between baseline (filled symbols) and mismatch (open symbols) conditions determined by paired t-test.

For 8 and 16-channel processors both frequency-place expansion of -5 mm and compression of +5 mm resulted in 20-30% drop in recognition compared to the matched condition (the flat performance with the 4-channel processor at 25 mm insertion depth may have been limited by the overall poor level of performance). Note that the Nucleus cochlear implant

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

processor mentioned above typically uses a frequency-place assignment that is similar to the +5 mm compression condition, which produced a significant reduction in vowel recognition.

expansion at 20 mm insertion	**F(3,15)**	**p**	**compression at 20 mm insertion**	**F(3,15)**	**p**
4 channel	40.26	<0.001	**4 channel**	14.28	<0.001
8 channel	78.55	<0.001	**8 channel**	14.86	<0.001
16 channel	235.08	<0.001	**16 channel**	2.05	0.15
expansion at 25 mm insertion	**F(3,15)**	**p**	**compression at 25 mm insertion**	**F(3,15)**	**p**
4 channel	1.29	0.31	**4 channel**	1.95	0.17
8 channel	32.67	<0.001	**8 channel**	22.31	<0.001
16 channel	39.54	<0.001	**16 channel**	32.94	<0.001

Table 4.6: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for vowel recognition at 20 mm and 25 mm simulated insertion depths.

Although a reduction in the acoustic frequency range normally causes a drop in performance, Figure 4.8 shows an improvement in vowel recognition for the baseline expansion condition for 4-channel processor at 25 mm insertion ($F(6,30)=4.13$, $p<0.01$). This improvement could be due

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

either to an increase in resolution or to a better frequency partition of the analysis bands.

Note that there was probably a floor effect for the expansion conditions with the simulated 20 mm insertion, where increasing the number of channels did not increase the intelligibility (Figure 6, leftmost data point of each panel). Both -5 mm and -3 mm expansion results were close to chance level, which might be due to the loss of all low-frequency information below 1850 Hz (-3 mm condition) or below 2476 Hz (-5 mm condition).

4.3.2.4. Analysis of Phoneme Recognition Performance by Speaker Gender

All phoneme results were reanalyzed by gender of the speaker. The scores are replotted for phonemes spoken by female or male speaker separately in Figures 4.9-4.16. In all figures filled symbols are percent correct scores when the subjects listened to female talkers only, and the open symbols are scores when the subjects listened to male talkers only. As in the previous figures, the number of channels increases from left to right in each panel. Figures 4.9-4.12 show the percent correct scores as a function of frequency-place compression and expansion as reanalyzed for talker gender, and Figures 4.13-4.16 show the baseline scores, where only the range of acoustic information was changed and all frequencies were matched, as reanalyzed for talker gender.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

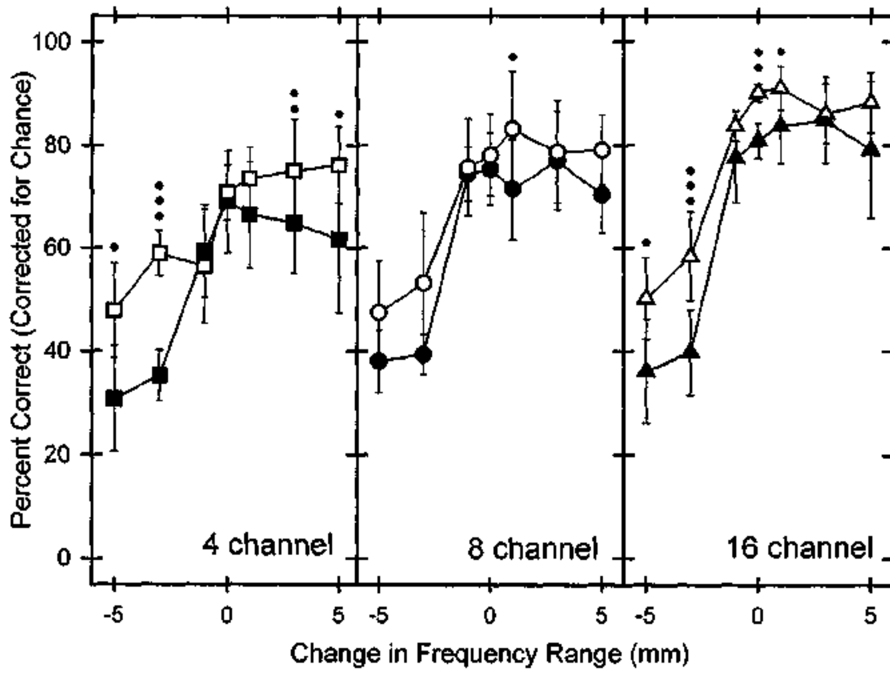


Figure 4.9: Consonant recognition percent scores for the simulated 20 mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping, reanalyzed for talker gender.

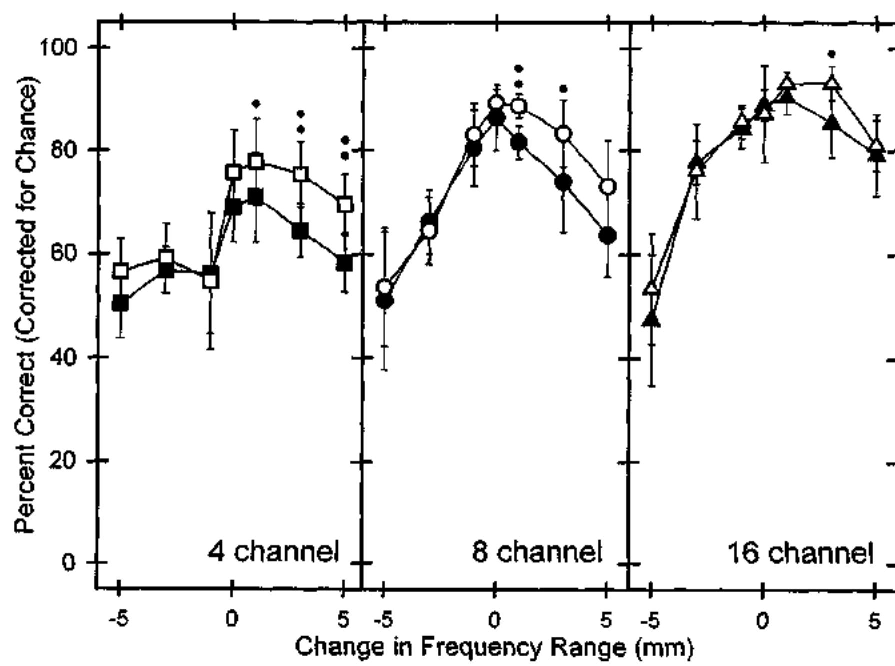


Figure 4.10: Consonant recognition scores, similar to Figure 4.9, for carrier bands simulating a 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

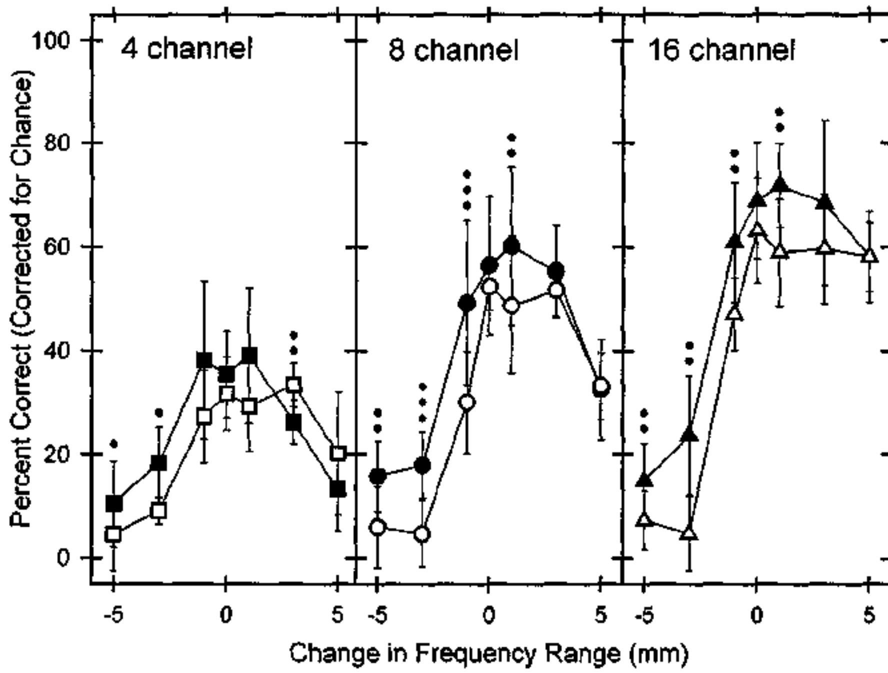


Figure 4.11: Vowel recognition scores similar to Figure 4.9, for carrier bands simulating a 20 mm insertion depth.

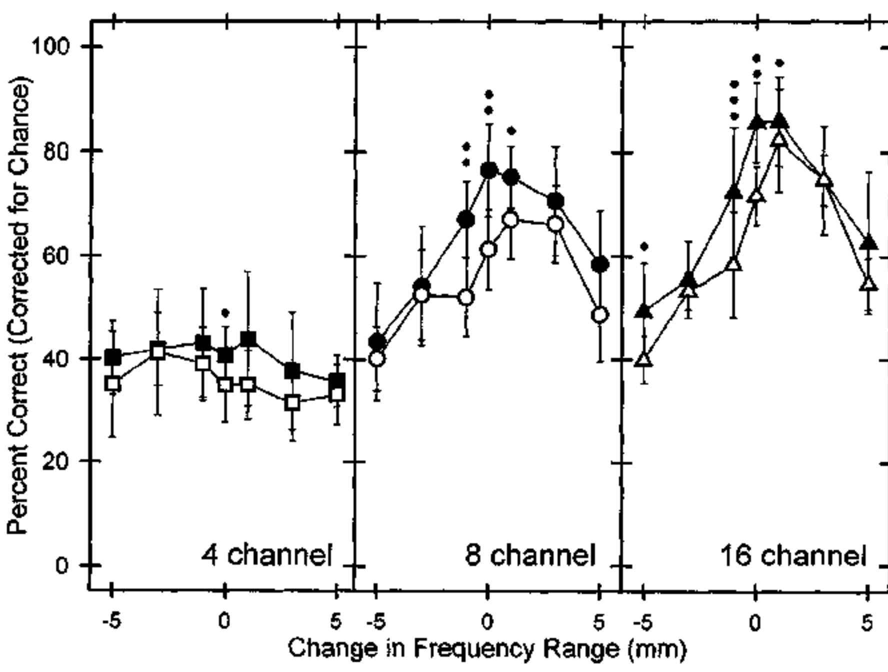


Figure 4.12: Vowel recognition scores, similar to Figure 4.9, for carrier bands simulating a 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

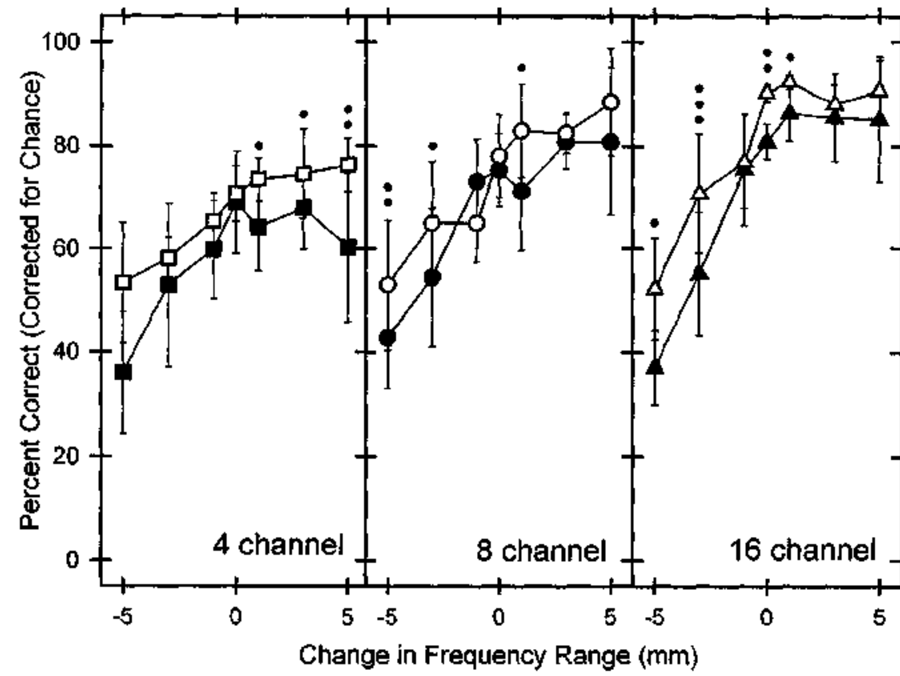


Figure 4.13: Consonant recognition baseline percent scores for the simulated 20 mm electrode insertion depth as a function of analysis band range that matches the noise carrier range, reanalyzed for talker gender.

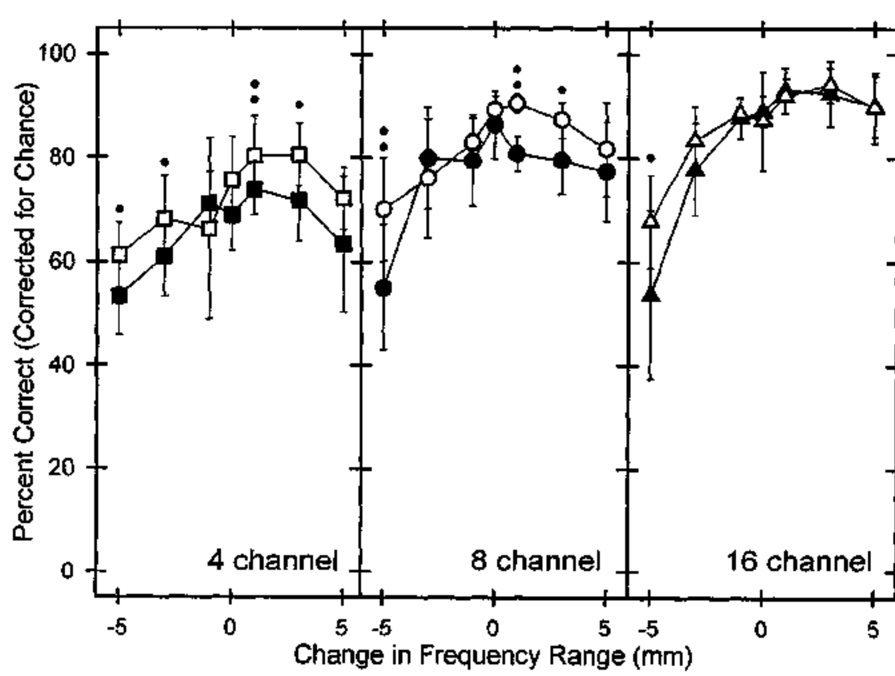


Figure 4.14: Consonant recognition baseline scores, similar to Figure 4.13, for carrier bands simulating a 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

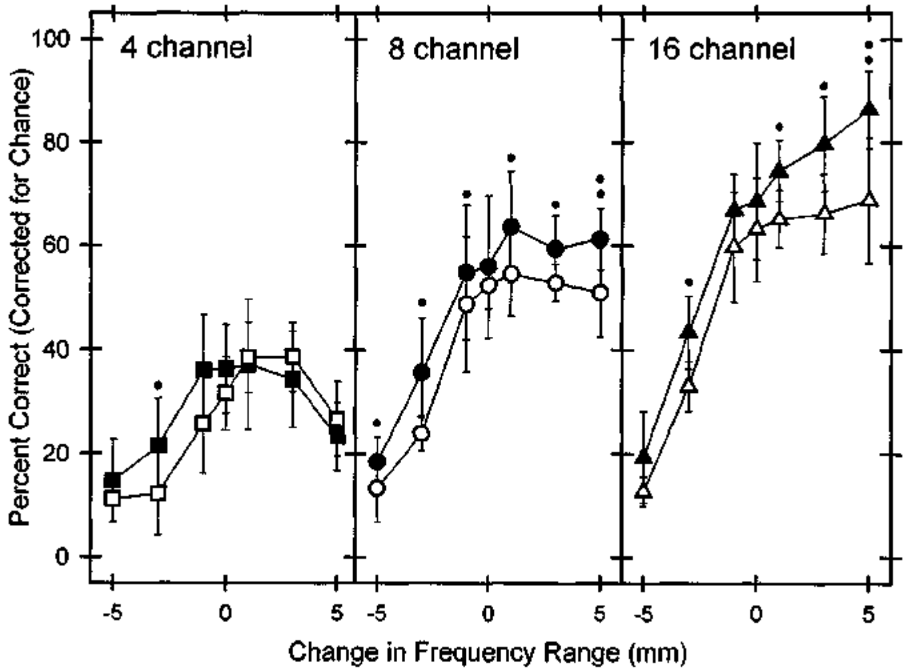


Figure 4.15: Vowel recognition baseline scores, similar to Figure 4.13, for carrier bands simulating a 20 mm insertion depth.

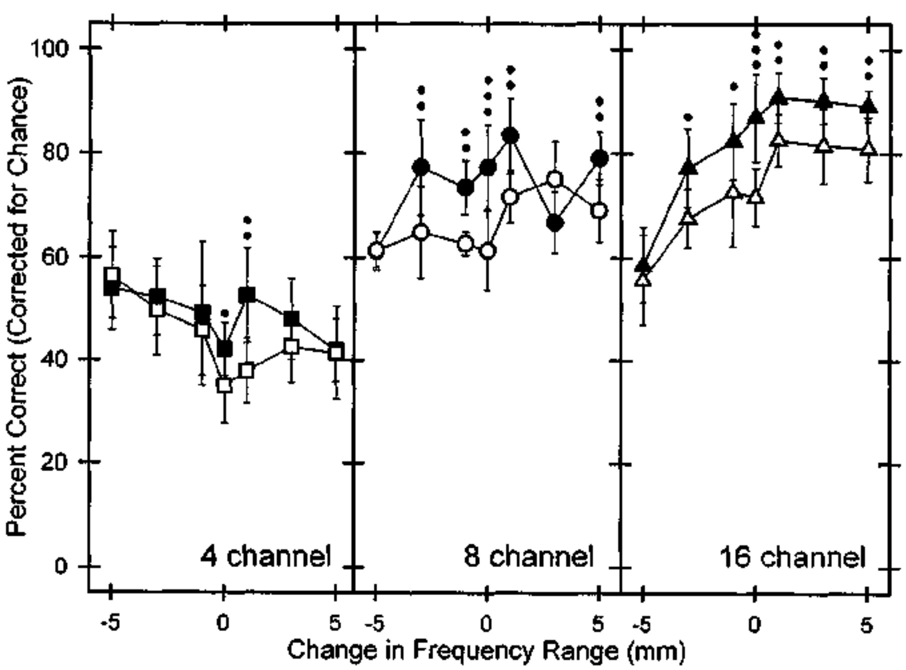


Figure 4.16: Vowel recognition baseline scores similar to Figure 4.13, for carrier bands simulating a 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In general, vowels spoken by females were recognized better than male talkers, and consonants spoken by males were better recognized than female talkers. This was actually expected because the spectral range of speech produced by female speakers is generally higher than male speakers. Presumably more useful spectral information is inside the stimulation region for female speakers, even when a lot of low frequency information was excluded. Similar pattern was observed with both mismatch and baseline conditions with both speakers of both genders.

4.3.3. Sentence Recognition

The percentage of words recognized in TIMIT sentences is presented in Figures 4.17 and 4.18 for 20 mm and 25 mm simulated insertion depths, respectively. Due to the limited number of sentence sets available, the matched baseline performance was measured only for extreme mismatch conditions (-5 mm expansion and +5 mm compression). For all numbers of channels and both simulated insertion depths the best performance was obtained when the analysis and carrier bands were matched.

Note that the drop in performance for frequency-place expansion (-5 mm) was dramatic compared to the matched condition (Table 4.7). Performance at the -5 mm condition drops to 5% correct for the 20 mm simulated insertion depth, and 15% for 25 mm insertion depth for all three spectral resolutions. For the 16-channel processor this drop was 75-80 percentage points. Although much of this loss was due to the information loss (filled symbols), there was an additional 20-40 point drop in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

performance due to the frequency-place expansion at the 25 mm simulated insertion depth.

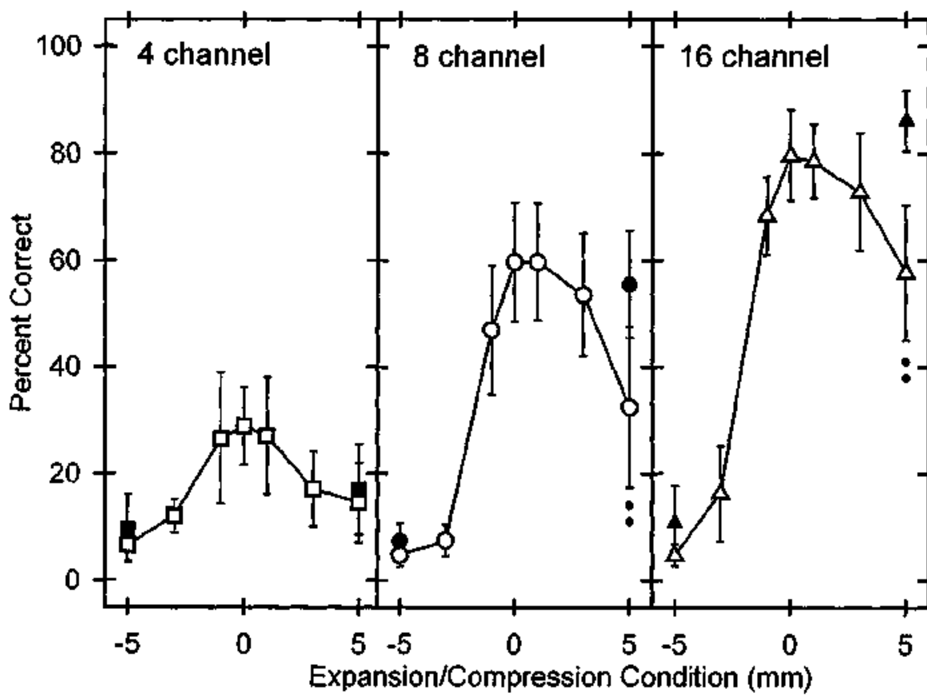


Figure 4.17: Average sentence recognition percent scores of 5 subjects for noise carrier bands simulating a 20 mm electrode insertion depth, as a function of compression or expansion in the frequency-place mapping. Open symbols denote the compression-expansion conditions where the carrier bands were fixed and the analysis bands were varied, and filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band frequency range. The dots under the scores denote significant differences between the female (filled symbols) and male (open symbols) talker conditions: one dot indicates $p<0.05$, two dots for $p<0.01$, and three dots for $p<0.001$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

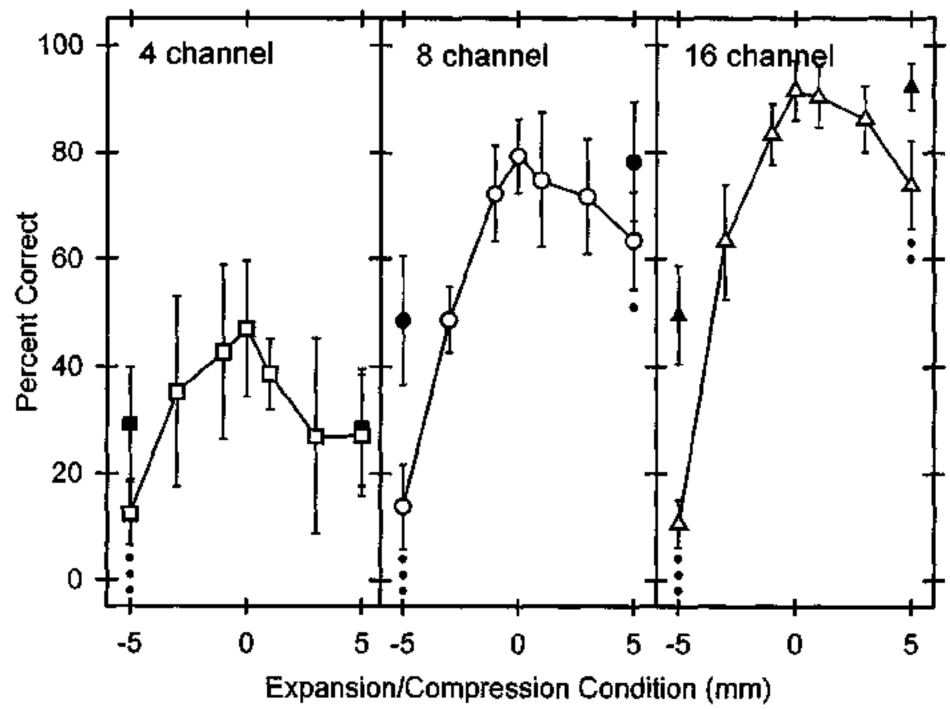


Figure 4.18: Sentence recognition percent scores, similar to Figure 4.17, for noise carrier bands simulating a 25 mm electrode insertion depth.

A smaller drop in performance was observed for frequency-place compression. There was a drop of 15-25 percentage points from the 0 mm baseline condition to the +5 mm baseline condition, even though the overall frequency range increased by about two octaves. There was an additional drop of 20-30 percentage points from the +5 mm baseline to the +5 mm compression condition. This +5 mm compression condition is similar to the typical frequency-place assignment used in Nucleus cochlear implants.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expansion at 20 mm insertion	F(3,15)	p	compression at 20 mm insertion	F(3,15)	p
4 channel	18.53	<0.001	4 channel	6.87	<0.01
8 channel	66.88	<0.001	8 channel	8.76	<0.01
16 channel	256.22	<0.001	16 channel	13.91	<0.001
expansion at 25 mm insertion	**F(3,15)**	**p**	**compression at 25 mm insertion**	**F(3,15)**	**p**
4 channel	11.64	<0.001	4 channel	4.29	<0.05
8 channel	147.89	<0.001	8 channel	3.59	<0.05
16 channel	233.12	<0.001	16 channel	22.17	<0.001

Table 4.7: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for sentence recognition at 20 mm and 25 mm simulated insertion depths.

4.4. SPEECH INTELLIGIBILITY INDEX MODEL MODIFIED FOR MISMATCHED FREQUENCY BANDS

A post-hoc Tukey test of overall results showed that most expansion conditions significantly reduced the speech recognition performance, with the exception of -1 mm expansion. On the other hand there was an optimum range from 0 mm matching to +3 mm compression in most cases, and the performance dropped significantly with more compression than 3 mm. Such a tolerance of a few mm to spectral mismatch was actually expected from the results of spectral shift studies mentioned before. Yet the asymmetry of this tolerance around the 0 mm matched condition implies that there is also a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

good compromise region (with compression of a mm or two) by having a little more information versus presenting that information to the correct tonotopic location. This pattern can actually be partly explained by a modified version of Speech Intelligibility Index (SII). SII was developed to determine the contribution of individual frequency bands to speech intelligibility and so to predict speech recognition performance under adverse listening conditions without actually running speech tests (ANSI 1997, S35).

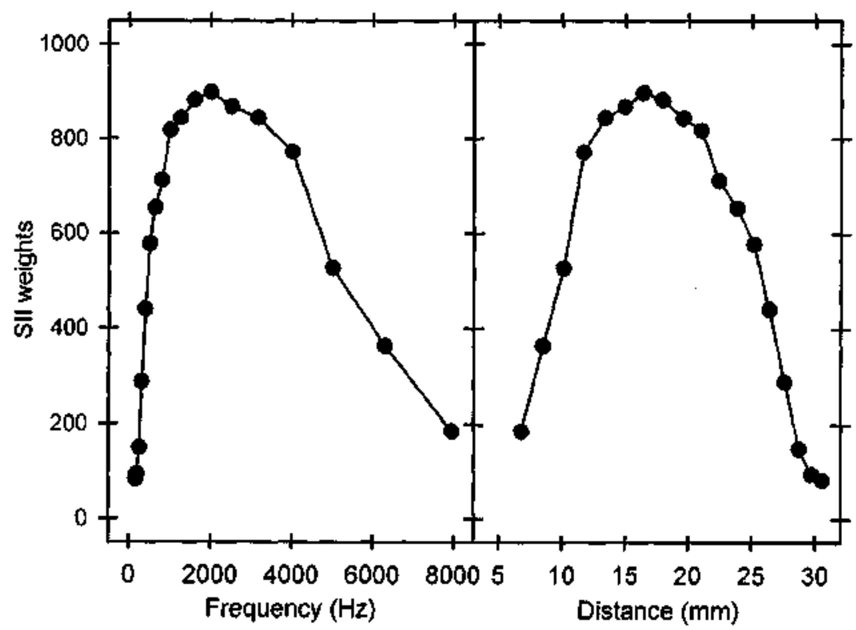


Figure 4.19: SII band importance function as a function of frequency in Hz, as shown in the left panel, and as a function of cochlear distance from the round window in mm, as shown in the right panel.

A sample band importance function is given in Figure 4.19 (one-third octave band importance function from Table 3 in ANSI 1997, S35). This function can be interpreted as the weights for different frequency bands for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

their relative contribution to speech intelligibility. The panel on left in Figure 4.19 shows these weights as a function of frequency in Hz, while the panel on right shows the same weights as a function of cochlear distance from round window in mm. The shape of this band importance function implies that frequencies around 1-3 kHz (15-20 mm from round window) contribute most to speech intelligibility. According to SII the wider the bandwidth of speech the more frequencies contribute and therefore more information is available to listener increasing the recognition of speech. The overall contribution from a number of spectral bands (n) is an additive function:

$$\mathrm{PC}(\%) = \sum_{i=1}^{n} w_i bw_i , \qquad (4.3)$$

where w_i is the weighting coefficient for the specific band i from the band importance function in Figure 4.19 and bw_i is the relative bandwidth of that frequency band. In the present study the frequency bands are first determined in mm and then converted to frequencies in Hz by Greenwood mapping function. Consequently all parameters are cochlear distances and this way it is easier to relate the results from simulations to the results in implants where the stimulation bands are real locations along the cochlea in mm. Similarly, the relative bandwidth bw_i is defined as the bandwidth of the band in mm divided by the total frequency range of the acoustic input in mm. Note that both weight and bandwidth in the equation are unitless because SII works on relative contributions of individual bands only, and does not predict

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

absolute percent correct scores. When SII is applied to experimental data a common procedure is to fit a sigmoid function to transfer SII predictions to real scores. In this study we preferred to skip this step both to keep the model as simple as possible, and also because we are only interested in relative drops from the best scores.

The vowel recognition scores with a simulated 8 channel electrode array at 20 mm insertion depth (Figure 4.7) is reproduced in Figure 4.20, left panel. Open symbols denote the frequency-place mismatch conditions, where the range of output bands was kept 16 mm in length (from 4 mm to 20 mm) and the acoustic information assigned onto these bands was made shorter or longer. 0 mm denotes perfect match where the frequency range perfectly matches tonotopic cochlear range represented by the noise carrier bands. The filled symbols are baseline conditions where the stimulation range was made longer or shorter to match the assigned acoustic range, so they do not represent any realistic cochlear implant setting. These baseline results show purely the effect of the amount of acoustic information available to the listener under the idealistic listening condition when all the acoustic information is always delivered to its matching tonotopic range in the cochlea. Equation 4.3 explains why this baseline condition is a monotonically increasing function of the acoustic input range (and the matched stimulation region). As the bandwidth gets wider more acoustic information becomes available to the listeners and they are expected to have better speech recognition scores. The prediction of the SII model by Equation 4.3 is given in right panel with filled symbols. SII predicts an increase as long as the total

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

bandwidth increases, yet the simulation data saturates for most compression conditions (or might even get worse for very wide channel bandwidths such as 4 channel processor, see Figure 4.7). This implies there might be an effect of resolution that comes from individual channel bandwidth; increasing individual bandwidth of spectral bands decreases resolution. Consider a 4 band processor. If the 4 bands span 1-3 kHz they can produce a good speech recognition performance because they cover important frequencies in that range. This performance might drop if they cover a range of 100-10k Hz due to the reduced spectral resolution in spite of the larger information available. Also in the case of low number of channels such as this 4 channel processor some frequency distinctions might be more important to have, such as a band transition near 1.5 kHz, which helps to differentiate between low and high formants. As the number of channels increases the bandwidth per channel decreases and the baseline performance looks more like the SII predicted baseline performance shown in the figure. These factors, such as the number of channels, individual channel bandwidth, and important frequency divisions are all missing in the conventional SII model.

For the same reasons SII cannot predict the absolute performance level either. It can only explain relative changes in speech intelligibility. Instead of fitting a sigmoid function to convert the relative scores to absolute scores, we simply adjusted the baseline in Figure 4.20 such that the percent correct score at 0 mm matched condition was the same level as the simulations, using Equation 4.3 only to calculate the relative drop.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

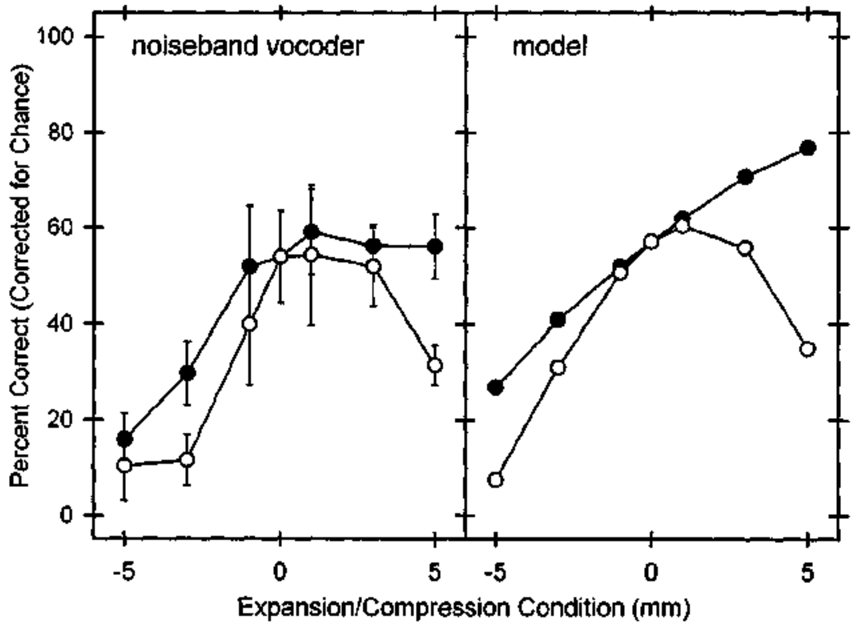


Figure 4.20: Frequency-place mismatch conditions (open symbols) and baseline conditions (filled symbols) for a 8 band noiseband vocoder at 20 mm simulated insertion depth, and the performance predicted by the SII mismatch model.

So far, the standard SII model given in Equation 4.3 explains only the baseline performance, based on the amount of acoustic information available. The SII model does not have a term to account for the drop from the baseline with frequency-place expansion and compression. We modified Equation 4.3 to include the deterioration that comes from the frequency-place mismatch. The shift of the frequency band from its matching tonotopic location decreases the contribution of that band to overall intelligibility and this drop is shown as a subtractive term in Equation 4.4.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$PC(\%) = \sum_{i=1}^{n}(w_i bw_i - \alpha w_i bw_i disp_i^2), \qquad (4.4)$$

$$PC(\%) = \sum_{i=1}^{n}(w_i bw_i - \alpha w_i bw_i disp_i^2),$$

where $disp_i$ is the displacement of the center frequency, i.e., the distance in mm of the center frequency of the band to the matching location in cochlea computed from the Greenwood equation. For a band around 1-3 kHz, which has the largest contribution to speech intelligibility, a shift in the center frequency causes a bigger drop from baseline condition. The weights w_i in Equation 4.3 are interpolated from the band importance function in Figure 4.19 by fitting a 6th order polynomial and evaluating the weights at the center frequencies of the bands in the simulated frequency-place mismatch conditions. We assumed an “inverse square law” for the deterioration effect of displaced bands, which is shown by $disp_i^2$ in the equation. The shape of $disp_i^2$ does not change much around the matching point.

α is the free (fitting) parameter of the model. The bigger the value for α the faster the drop in the performance is with increasing mismatch. The model does not predict absolute performance, but instead estimates the relative drop from the perfect match condition. Therefore the value of α was obtained by finding the α that gave the best fitting mismatch curve to simulation results. With the experimental data given in the left panel of Figure 4.20 the best fit was obtained when $\alpha=0.06$. The prediction of the modified model for the mismatch performance for the same processor with $\alpha=0.06$ is shown in the right panel of Figure 4.20, with open circles. The mismatch curve is generally consistent with simulation results except the missing floor effect that was observed in simulations starting at -3 mm

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expansion condition. Similar to experimental data, the mismatch curve closely follows the baseline performance around 0 mm matched condition, and drops more with increasing mismatch. For these specific testing conditions, the model actually predicts the best performance at +1mm compression, which might be due to a compromise between more acoustic information vs mismatch. Again similar to experimental data, the mismatch curve predicted by the model is not symmetrical around 0 mm matched condition; the drop from the baseline condition is larger with expansion conditions compared to compression.

Even though the SII prediction shown in Figure 4.20 displays the general pattern of the experimental data, the model actually does not predict all results from all conditions. For example it predicts higher scores for baseline conditions at +3 mm and +5mm compression. One possible reason is the lack of a term to represent the number of channels in Equation 4.3 because original SII model does not take spectral resolution or bandwidth of the bands into account. Another shortcome of the model is that it also ignores the temporal correlation between the bands. As a result the model predicts the performance curves only partially because its prediction is just based on the contribution from the acoustic information available.

The purpose of this simple model was to show how the present results might be explained by a combination of the SII importance function for speech and a metric relating to the degree of mismatch. Further refinements are needed for the model to cover all mismatch conditions and to predict absolute percent correct scores.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 5

VARIATIONS OF FREQUENCY-PLACE MAPPING CONDITIONS

In the previous chapter it was shown with noiseband vocoder simulations that speech recognition is best when the acoustic information is assigned to the tonotopically matching cochlear place. Usually there is an optimum speech perception range with a small compression of a few mm (<2 mm) where a slightly wider range acoustic information is mapped to tonotopic places relatively close to their correct locations.

In this chapter, the effect of frequency-place compression and expansion on speech recognition is further explored with:

1. Frequency-place compression and expansion on apical or basal end only:

In Chapter 4 the frequency-place mapping was symmetrically compressed or expanded at both apical and basal ends of the stimulation region represented by the noise carrier bands. However for understanding speech not all frequencies are equally important, as it was also shown by SII band importance function in Figure 4.19. For example, Shannon *et al.* (2001) created holes in different spectral regions of the stimulation range of the Nucleus cochlear implant by turning off several electrodes and observed the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

largest drop in the speech recognition when the hole was in the apical region around 500 - 2k Hz. To assess the relative importance of the cochlear location of the matched frequencies on speech perception the mapping was separated into two conditions. In apical mismatch, frequency-place map was compressed or expanded on the apical end of the noise carrier band range while the basal end was always matched. In basal mismatch the apical end was always matched while the map is compressed or expanded on the basal end.

2. <u>Compression and expansion at both ends combined with shifted electrodes:</u>

All the tests have been designed assuming the exact insertion depth of the electrode array is known so that the exact matching frequencies can be calculated using the Greenwood's frequency-place equation. In reality this is rarely the case; the surgeon can estimate the location of the electrode array only from observing how many electrodes were inserted during the surgery. If the array is only partially inserted it can be easier to have an estimation for the insertion depth. However if all of the electrodes were inserted we can only assume that insertion depth is deeper than the electrode array length, yet it can be off as much as 4 to 5 mm. To simulate this more realistic implant condition an uncertainty in the exact location of the electrode array is added to compression and expansion conditions such that electrodes are located further or shallower than the assumed insertion depth and then the mismatch conditions are applied. This uncertainty in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

location of the array is represented by a shift between the acoustic analysis bands and the noise carrier bands in addition to the mismatch introduced by compression and expansion.

3. <u>Holes redistributed:</u>

Shannon *et al.* (2002) created spectral holes in the noise carrier bands to represent damaged neurons which do not transmit any acoustic information to the brain. The effects of such spectral holes on overall hearing were explored by changing the size and the location of the hole. They observed that speech perception dropped with increasing hole size and holes in the apical region had larger effects. They created the holes with four different maps:

(a) All the information inside the hole was simply discarded, but the rest of the analysis bands were assigned to the matching carrier bands.

(b) All information inside the hole was reassigned onto the adjacent band next to the apical end of the hole.

(c) All information inside the hole was reassigned onto the adjacent band next to the basal end of the hole.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(d) All information inside the hole was split and reassigned onto both adjacent bands.

There was no significant difference between the performances of these four maps. In this experiment we tried a different map, where all the acoustic information was reassigned onto the whole range of the carrier bands, instead of reassigning them onto single bands similar to the compression conditions on each side of the hole. This map provides more acoustic information in the expense of introducing some mismatch between the analysis and carrier bands.

4. Compression and expansion at both ends in background noise:

This experiment is another attempt to create more realistic listening conditions. All simulation tests are run in soundproof booths with no other competing sounds but in real life such an ideal listening situation rarely exists. Also implant users always have much more difficulty understanding speech in noisy environments. To simulate noisy environments speech-shaped noise was added at varying SNR values. The changes in patterns of compression and expansion results with added noise were observed to find out the highest background noise level that the matched map can tolerate before losing its advantage over compressed and expanded maps.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.1. EXPERIMENTAL METHOD

5.1.1. Subjects

Normal-hearing listeners aged between 22 and 38 participated in the study. All subjects were native speakers of American English and had thresholds better than 20 dB HL at audiometric frequencies between 125 and 8000 Hz.

5.1.2. Stimuli and Signal Processing

The speech recognition tasks consisted of discrimination of the same medial vowels and consonants from Chapter 4. All stimuli were presented via a loudspeaker in a sound field at 70 dB on an A-weighted scale, without lip-reading. Due to the limited number of sentences they were not used in these experiments.

All phonemes were processed using the noise-band vocoder described in Chapter 3.

5.2. APICAL AND BASAL MISMATCH

5.2.1. Experimental Conditions

In apical mismatch the map is compressed or expanded on the apical (low-frequency) end of the stimulation region represented by the noise carrier bands while the basal (high-frequency) end is always matched. Figure 5.1 shows apical compression, matching, and expansion for a 4-band

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

processor at 25 mm simulated insertion depth. Basal end of the cochlear stimulation region is fixed at 9 mm from the round window, i.e. the high-frequency end of the acoustic information is always matched at 5.8 kHz. The exact frequency ranges and cut-off frequencies of the bands are given in Tables 5.1 and 5.2 for this processor at 20 mm and 25 mm insertion depths, respectively.

Similarly, in basal mismatch the map is compressed or expanded on the basal (high-frequency) end only while the apical (low-frequency) end is always matched. Figure 5.2 shows basal compression, matching, and expansion for a 4-band processor at 25 mm simulated insertion depth. The low end of the cochlear stimulation region is fixed at 25 mm from the round window, i.e. the low-frequency end of the acoustic information is always matched at 510 Hz. The exact frequency ranges and cut-off frequencies of the bands are given in Tables 5.3 and 5.4 for this processor at 20 mm and 25 mm insertion depths, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

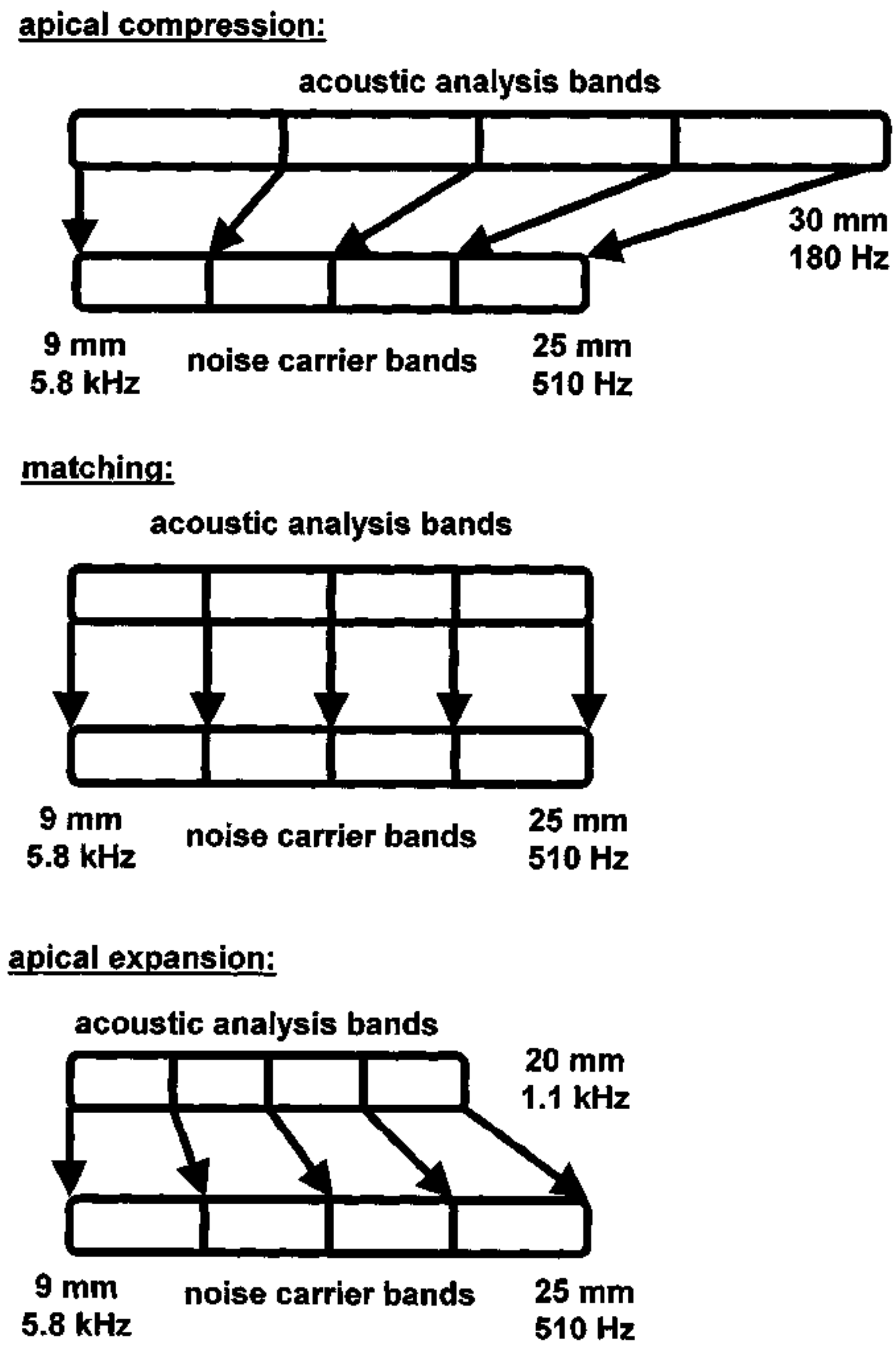


Figure 5.1: Frequency-place compression and expansion on apical end only for 4-channel processor at 25 mm simulated insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

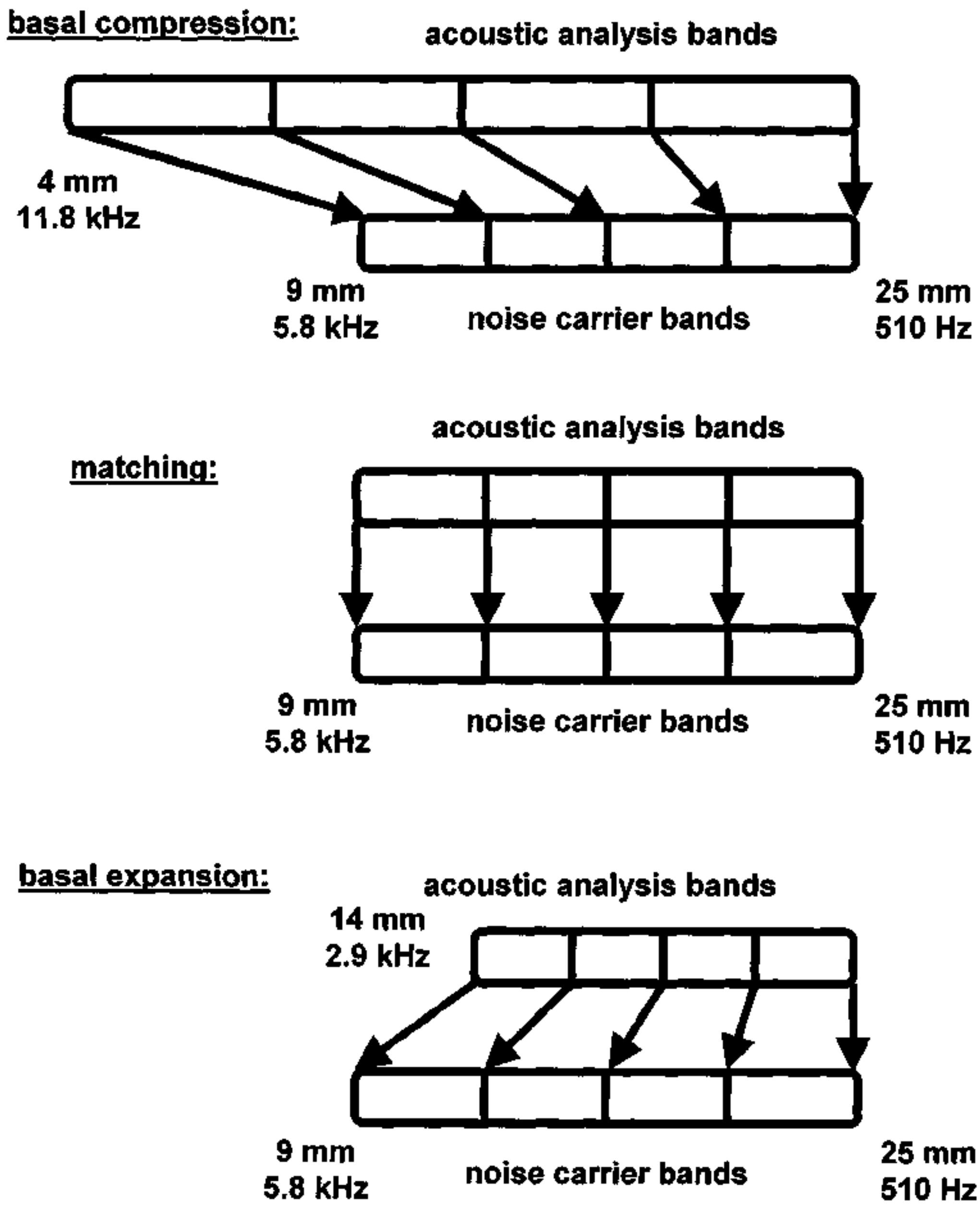


Figure 5.2: Frequency-place compression and expansion on basal end only for 4-channel processor at 25 mm simulated insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	15 – 4	2476	3687	5459	8049	11837	2476 – 11837
-3 mm (expansion)	17 – 4	1843	2970	4736	7502	11837	1843 – 11837
-1 mm (expansion)	19 – 4	1363	2387	4106	6992	11837	1363 – 11837
0 mm (matching)	20 – 4	1168	2138	3822	6749	11837	1168 – 11837
+1 mm (compression)	21 – 4	999	1913	3557	6515	11837	999 – 11837
+3 mm (compression)	23 – 4	722	1528	3079	6071	11837	722 – 11837
+5 mm (compression)	25 – 4	513	1214	2663	5656	11837	513 – 11837

Table 5.1: Frequency-place mismatch conditions on apical end only for the 4-channel processor at the simulated 20 mm electrode insertion depth. The name of each condition represents the change in frequency range expressed in mm between the analysis and carrier bands on apical end. For each condition the table lists the following information for the analysis bands: cochlear location in mm from the round window, cut-off frequencies for a four-band processor, and total frequency range. Because the simulated electrode location was fixed, the noise carrier bands covered the frequency range from 1168 to 11837 Hz in all conditions, and the frequency partition of carrier bands was as shown in the center entry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	20 – 9	1168	1776	2663	3962	5860	1168 – 5860
-3 mm (expansion)	22 – 9	851	1416	2301	3687	5860	851 – 5860
-1 mm (expansion)	24 – 9	611	1124	1986	3432	5860	611 – 5860
0 mm (matching)	25 – 9	513	999	1843	3310	5860	513 – 5860
+1 mm (compression)	26 – 9	428	886	1710	3193	5860	428 – 5860
+3 mm (compression)	28 – 9	290	693	1471	2970	5860	290 – 5860
+5 mm (compression)	30 – 9	184	536	1262	2762	5860	184 – 5860

Table 5.2: Frequency-place mismatch conditions on apical end only for a 4-channel processor at the simulated 25 mm electrode insertion depth. For each condition the table lists the cochlear locations of the analysis bands, cut-off frequencies of the band-pass filters, and the total analysis frequency range. The noise carrier bands were fixed between 513 and 5860 Hz with the partition shown as in the center row of the conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	20 – 9	1168	1776	2663	3962	5860	1168 - 5860
-3 mm (expansion)	20 – 7	1168	1913	3080	4907	7771	1168 - 7771
-1 mm (expansion)	20 – 5	1168	2060	3557	6071	10290	1168 - 10290
0 mm (matching)	20 – 4	1168	2138	3822	6749	11837	1168 - 11837
+1 mm (compression)	20 – 3	1168	2218	4106	7502	13612	1168 - 13612
+3 mm (compression)	20 – 1	1168	2387	4736	9263	17990	1168 - 17990
+5 mm (compression)	20 – 0	1168	2568	5459	11430	23762	1168 - 23762

Table 5.3: Frequency-place mismatch conditions on basal end only for the 4-channel processor at the simulated 20 mm electrode insertion depth. The name of each condition represents the change in frequency range expressed in mm between the analysis and carrier bands on basal end. For each condition the table lists the following information for the analysis bands: cochlear location in mm from the round window, cut-off frequencies for a four-band processor, and total frequency range. Because the simulated electrode location was fixed, the noise carrier bands covered the frequency range from 1168 to 11837 Hz in all conditions, and the frequency partition of carrier bands was as shown in the center entry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	cochlear location of analysis bands (mm)	band-pass filter cut-off frequencies for 4 channels (Hz)					frequency range of analysis bands (Hz)
-5 mm (expansion)	25 – 14	513	817	1262	1913	2864	513 - 2864
-3 mm (expansion)	25 – 12	513	886	1471	2387	3822	513 - 3822
-1 mm (expansion)	25 – 10	513	960	1710	2970	5085	513 - 5085
0 mm (matching)	25 – 9	513	999	1843	3310	5860	513 - 5860
+1 mm (compression)	25 – 8	513	1039	1985	3687	6750	513 - 6750
+3 mm (compression)	25 – 6	513	1124	2301	4570	8944	513 - 8944
+5 mm (compression)	25 – 4	513	1214	2663	5656	11837	513 - 11837

Table 5.4: Frequency-place mismatch conditions on basal end only for a 4-channel processor at the simulated 25 mm electrode insertion depth. For each condition the table lists the cochlear locations of the analysis bands, cut-off frequencies of the band-pass filters, and the total analysis frequency range. The noise carrier bands were fixed between 513 and 5860 Hz with the partition shown as in the center row of the conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.2. Results of Apical and Basal Mismatch

Seven normal-hearing listeners between ages 29 and 35 participated in this part of the study. Three participated in both experiments, two participated in apical mismatch experiment only and the remaining two participated in basal mismatch experiment only, resulting in five subjects for each experiment.

5.2.2.1. Apical Mismatch Results

Figures 5.3-5.4 show average percent correct scores of five subjects from vowel recognition and Figures 5.5-5.6 show scores from consonant recognition tests with 4, 8, and 16 band processors at 20mm and 25 mm simulated insertion depths. The number of channels increases from 4 to 8 to 16 in the left, middle, and right panels, respectively, in each figure. Within each panel the filled symbols present results from the baseline conditions in which analysis and carrier bands were always matched, and the open symbols present results from the apical mismatch conditions. All scores are corrected for chance level.

Similar to compression and expansion from both ends, the best performance was obtained at 0 mm matching point with a tolerance region of a few mm compression. Apical expansion and further compression both dropped the performance.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

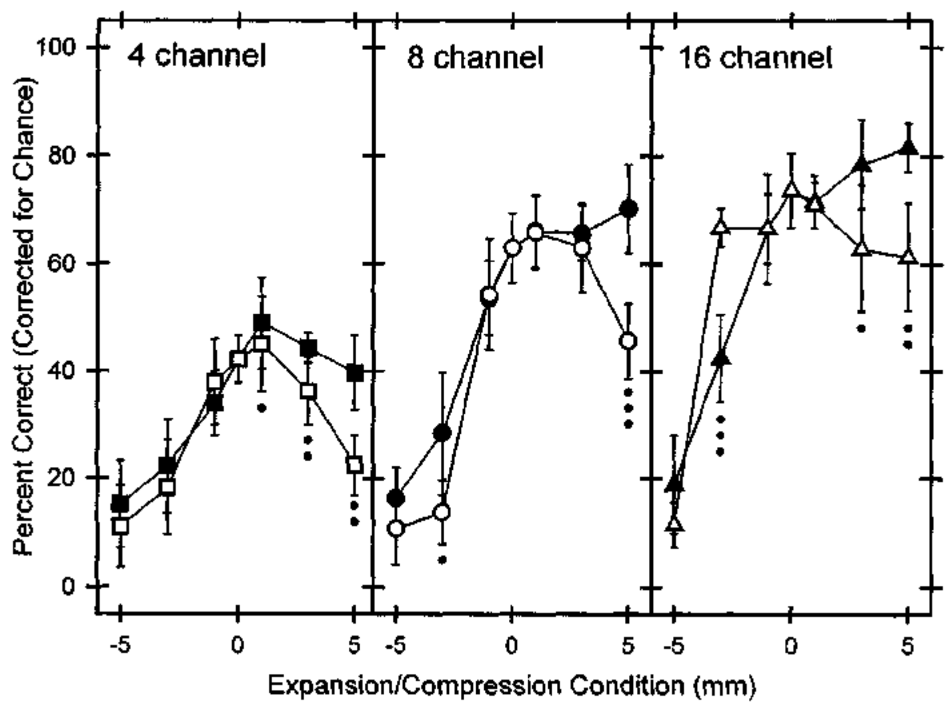


Figure 5.3: Vowel recognition percent scores for the carrier bands simulating 20 mm insertion depth, as a function of apical mismatch. The number of spectral bands increases from 4 to 8 to 16 in the left, middle and right panels, respectively. The percent correct scores represent the average performance of 5 normal-hearing subjects, corrected for chance, and the error bars represent one standard deviation. Filled symbols denote the baseline condition where the carrier bands were always matched to the analysis band range. Open symbols denote the compression-expansion conditions where the carrier bands were fixed and the analysis bandwidth was varied. Dots under the scores denote significant differences between the baseline (filled symbols) and mismatch (open symbols) conditions: one dot indicates $p<0.05$, two dots for $p<0.01$, and three dots for $p<0.001$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

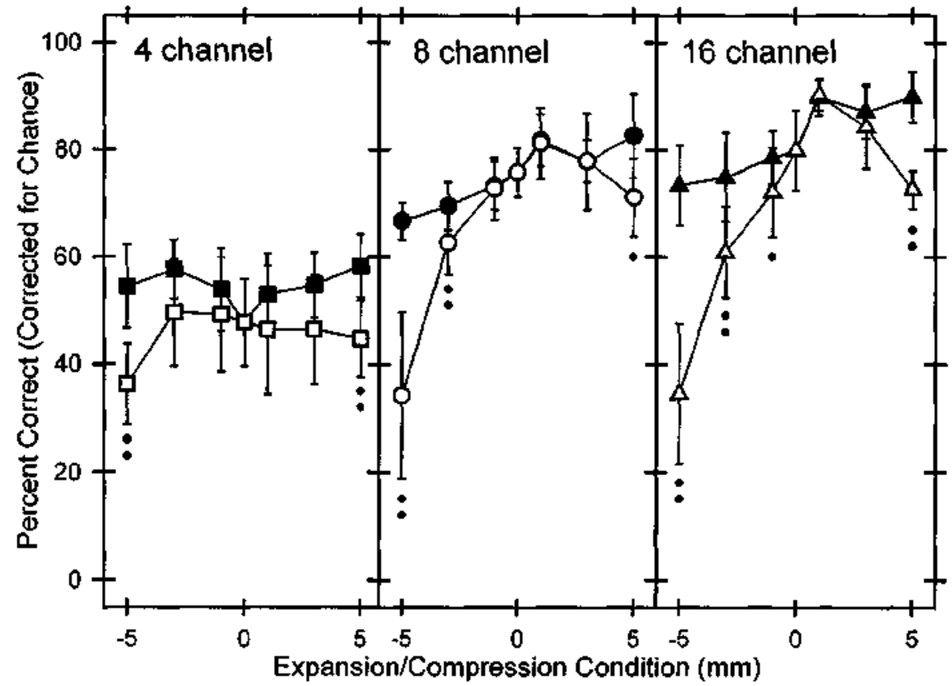


Figure 5.4: Vowel recognition percent scores for the carrier bands simulating 25 mm insertion depth, as a function of apical mismatch.

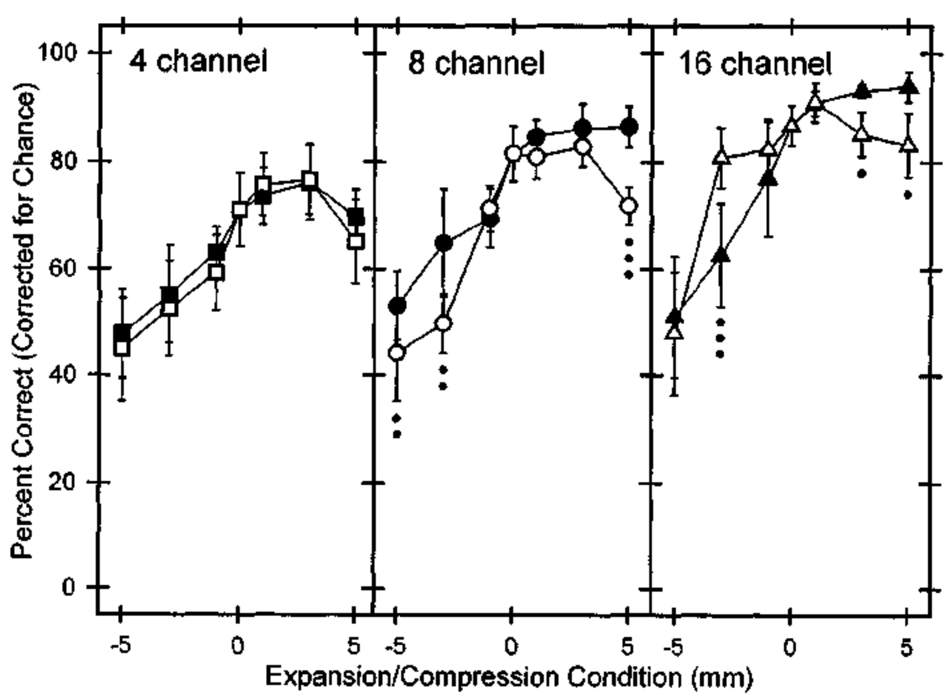


Figure 5.5: Consonant recognition percent scores for the carrier bands simulating 20 mm insertion depth, as a function of apical mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

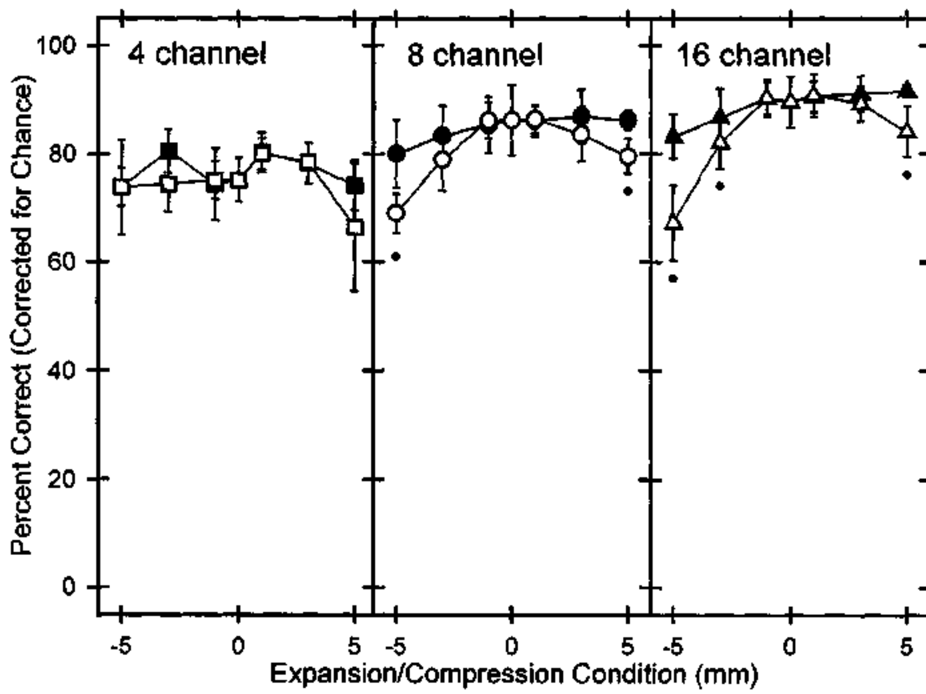


Figure 5.6: Consonant recognition percent scores for the carrier bands simulating 25 mm insertion depth, as a function of apical mismatch.

In matched baseline conditions shown with filled symbols the stimulation range was made changed to match the acoustic analysis band range. Both the noise carrier band and the analysis band ranges were fixed at the basal end. They were made wider from the apical end only. The resulting baseline performance was almost the same as when the stimulation region was made longer at both ends as shown in Figures 4.7 and 4.3 for vowels and consonants, respectively, at 20 mm simulated insertion depth. In the 0 mm matched condition the stimulation region covers the area from 4 mm to 20 mm from the round window, responding best to a frequency range of 1.2 kHz -12 kHz (as shown in Table 5.1). 12 kHz is too

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

high to carry much useful information, so the baseline condition is mostly dominated by including more (or less) acoustic information at the low end of the stimulation region. The same electrode array inserted deeper to 25 mm covers a tonotopic range from 9 mm to 25 mm from the round window, covering the frequencies 500 Hz -6 kHz (as shown in Table 5.2). Most vowel spectral energy is still much lower than 6 kHz, so baseline performance at 25 mm for vowels with apical mismatch looks similar to mismatch at both ends, yet consonant baseline performances are pretty much flat at this insertion depth. Consonants use much higher frequencies than vowels and changing frequencies around 500 Hz does not affect performance much as long as higher frequencies (up to 6 kHz in this case) are included.

Remember that the baseline conditions show the effect of information loss only with all spectral bands aligned perfectly, i.e. changing the length of the stimulation region. Theoretically any drop from the baseline condition is due to the mismatch between the carrier and analysis bands only.

A t-test was applied to find the significant drops with apical mismatch from the related baseline conditions. Resulting p values are shown by the filled circles under the scores in each figure: one circle for $p<0.05$, two circles for $p<0.01$, and three circles for $p<0.001$.

Overall the pattern of the results with apical mismatch were similar to mismatch applied at both ends, except a smaller drop from baseline in some cases, and a much flat mismatch curve with consonants at 25 mm. A comparison of the results is presented in Section 5.2.2.3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.2.2. Basal Mismatch Results

Figures 5.7-5.8 show average percent correct scores from vowel recognition and Figures 5.9-5.10 consonant recognition scores with 4, 8, and 16 band processors at 20mm and 25 mm simulated insertion depths. All scores are corrected for chance level.

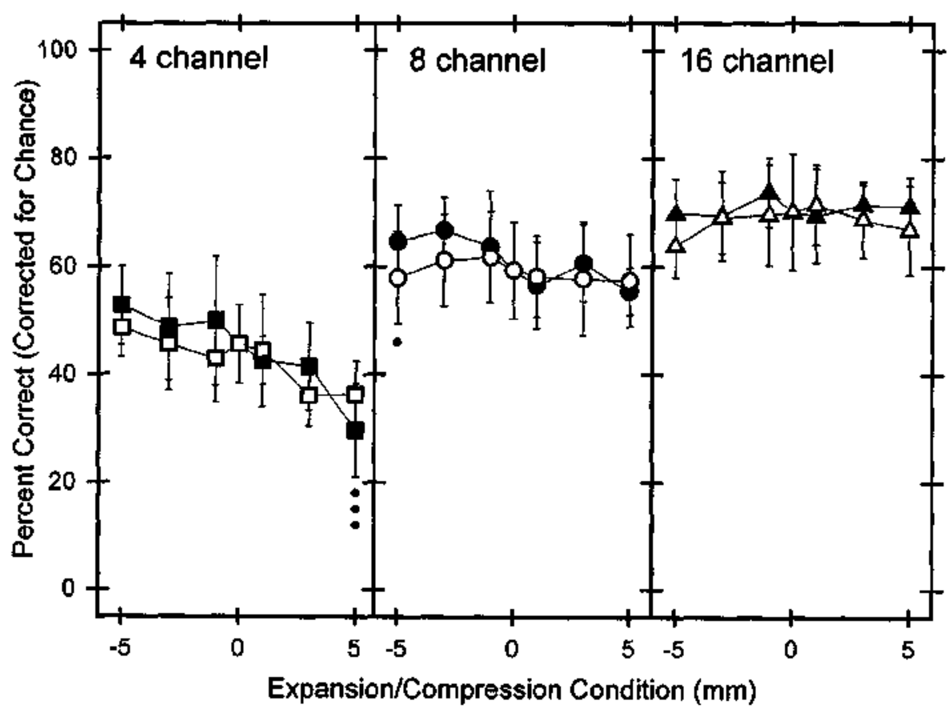


Figure 5.7: Vowel recognition percent scores for the carrier bands simulating 20 mm insertion depth, as a function of basal mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

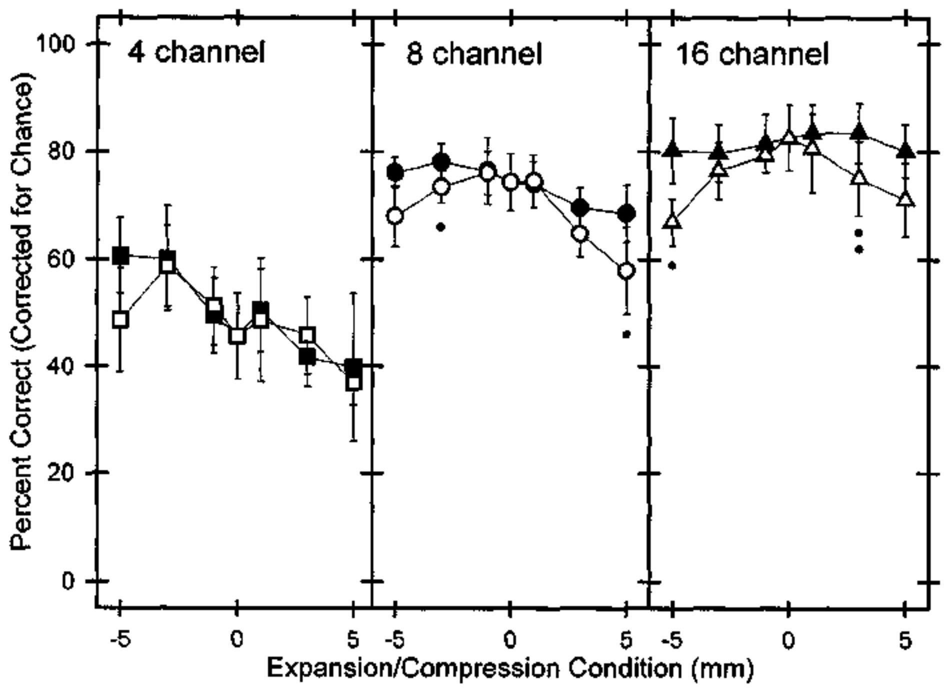


Figure 5.8: Vowel recognition percent scores for the carrier bands simulating 25 mm insertion depth, as a function of basal mismatch.

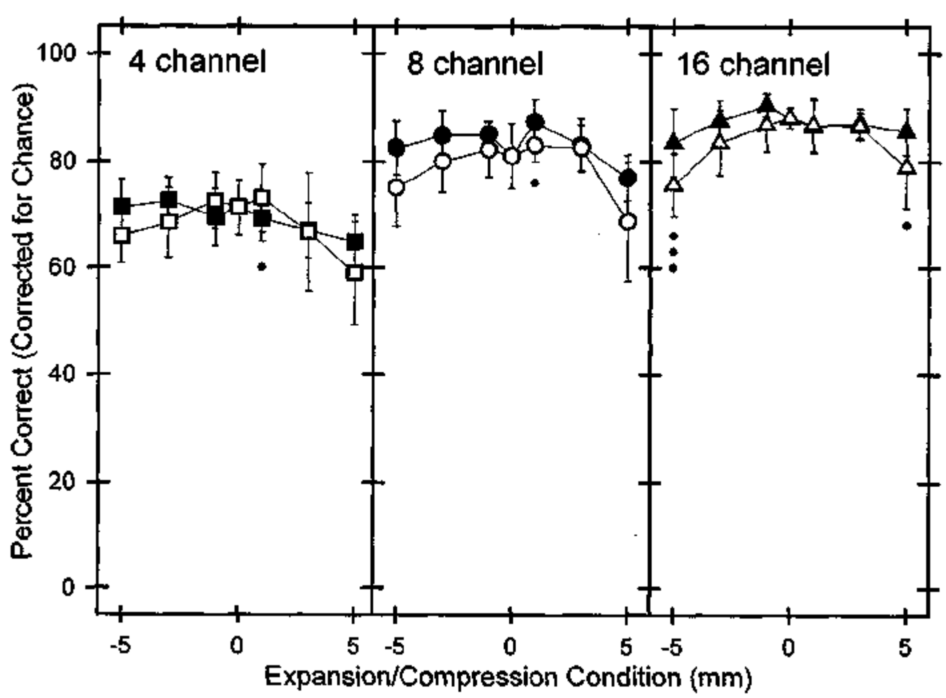


Figure 5.9: Consonant recognition percent scores for the carrier bands simulating 20 mm insertion depth, as a function of basal mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

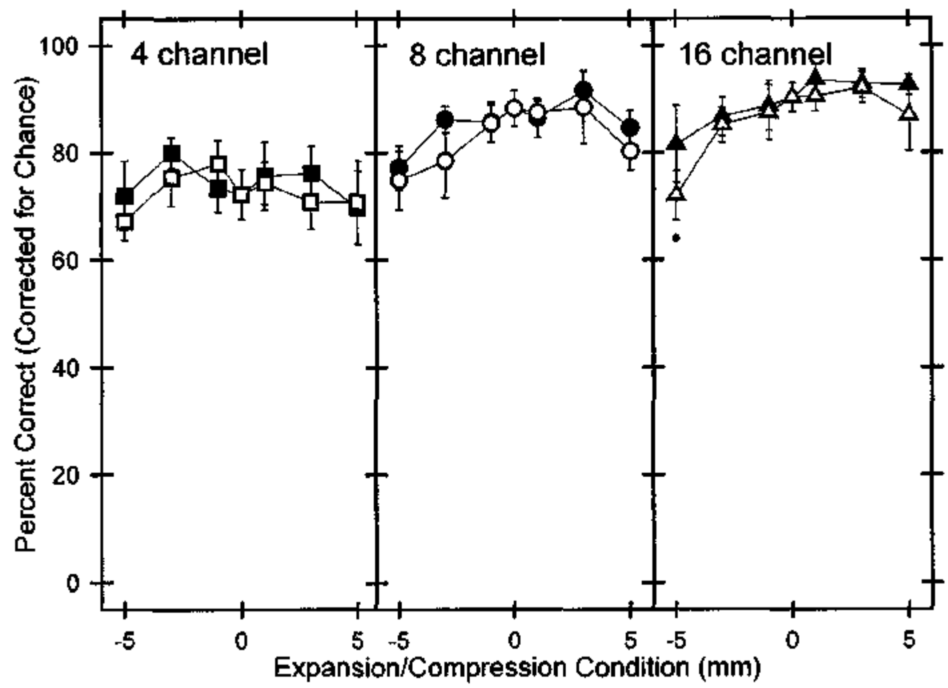


Figure 5.10: Consonant recognition percent scores for the carrier bands simulating 25 mm insertion depth, as a function of basal mismatch.

All consonant and vowel recognition curves were much more flat with basal mismatch, with almost no change with any expansion condition and only a small effect of compression. One interesting observation is that the vowel recognition baseline performance with 4 channels actually decreased when stimulation region was made wider at the high-frequency end ($F(3,12)=9.87$, $p<0.001$ for 20 mm insertion, and $F(3,12)=7.60$, $p<0.001$ for 25 mm insertion). A similar drop was observed when stimulation region was made wider at both ends, but only for vowels processed with 4 channels at simulated insertion depth of 25 mm. As mentioned in Chapter 4, this might be due to the change in the resolution in the important spectral region for vowels or a better frequency partition (for formant discrimination). Both

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

factors would have a more dominant effect with less number of channels, i.e. with 4 channels compared to 8 or 16 channels. But here we see a similar drop even with 8 channels ($F(3,12)=4.74$, $p<0.01$ for 20 mm insertion, and $F(3,12)=6.24$, $p<0.001$ for 25 mm insertion), and only at 16 channels the baseline performance becomes flat ($F(3,12)=0.64$, $p=0.70$ for 20 mm insertion, and $F(3,12)=1.09$, $p=0.40$ for 25 mm insertion).

A more detailed analysis of results is given in the following section.

5.2.2.3. Comparison of Results of Apical and Basal Mismatch to Results of Mismatch at Both Ends

A one-way repeated-measures ANOVA test was applied to all conditions. The resulting F and p values are summarized in Table 5.5 for vowel recognition, and Table 5.6 for consonant recognition tests. In every entry the first line is the ANOVA results for mismatch applied at both ends of the stimulation region, while second and third lines are ANOVA results for apical and basal mismatch, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expansion at 20 mm insertion	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**	**compression at 20 mm insertion**	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**
4 channel	40.26 52.60 1.64	<0.001 <0.001 0.23	**4 channel**	14.28 25.92 8.86	<0.001 <0.001 <0.01
8 channel	78.55 109.53 0.58	<0.001 <0.001 0.64	**8 channel**	14.86 15.08 0.05	<0.001 <0.001 0.99
16 channel	235.08 140.39 2.83	<0.001 <0.001 0.08	**16 channel**	2.05 6.17 2.41	0.15 <0.01 0.11
expansion at 25 mm insertion	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**	**compression at 25 mm insertion**	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**
4 channel	1.29 8.49 4.5	0.31 <0.01 <0.05	**4 channel**	1.95 0.22 4.81	0.17 0.88 <0.05
8 channel	32.67 44.18 3.49	<0.001 <0.001 0.05	**8 channel**	22.31 5.69 26.5	<0.001 <0.05 <0.001
16 channel	39.54 47.96 22.20	<0.001 <0.001 <0.001	**16 channel**	32.94 11.06 11.23	<0.001 <0.001 <0.001

Table 5.5: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for vowel recognition at 20 mm and 25 mm simulated insertion depths. In every entry the first line shows ANOVA results for expansion and compression at both ends of stimulation region, while second and third lines show ANOVA results for apical and basal mismatch, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expansion at 20 mm insertion	both (F(3,15)) apical (F(3,12)) basal (F(3,12))	p	compression at 20 mm insertion	both (F(3,15)) apical (F(3,12)) basal (F(3,12))	p
4 channel	54.60 41.64 5.46	<0.001 <0.001 <0.05	4 channel	14.28 25.92 8.86	<0.001 <0.001 <0.01
8 channel	89.05 134.33 2.25	<0.001 <0.001 0.14	8 channel	14.86 15.08 0.05	<0.001 <0.001 0.99
16 channel	117.79 56.47 7.58	<0.001 <0.001 <0.01	16 channel	2.05 6.17 2.41	0.15 <0.01 0.11
expansion at 25 mm insertion	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**	**compression at 25 mm insertion**	**both (F(3,15)) apical (F(3,12)) basal (F(3,12))**	**p**
4 channel	1.29 8.49 4.5	0.31 <0.01 <0.05	4 channel	1.95 0.22 4.81	0.17 0.88 <0.05
8 channel	32.67 44.18 3.49	<0.001 <0.001 0.05	8 channel	22.31 5.69 26.5	<0.001 <0.05 <0.001
16 channel	39.54 47.96 22.20	<0.001 <0.001 <0.001	16 channel	32.94 11.06 11.23	<0.001 <0.001 <0.001

Table 5.6: F and p values calculated with one-way repeated-measures ANOVA for expansion and compression mismatch conditions for consonant recognition at 20 mm and 25 mm simulated insertion depths. In every entry the first line shows ANOVA results for expansion and compression at both ends of stimulation region, while second and third lines show ANOVA results for apical and basal mismatch, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

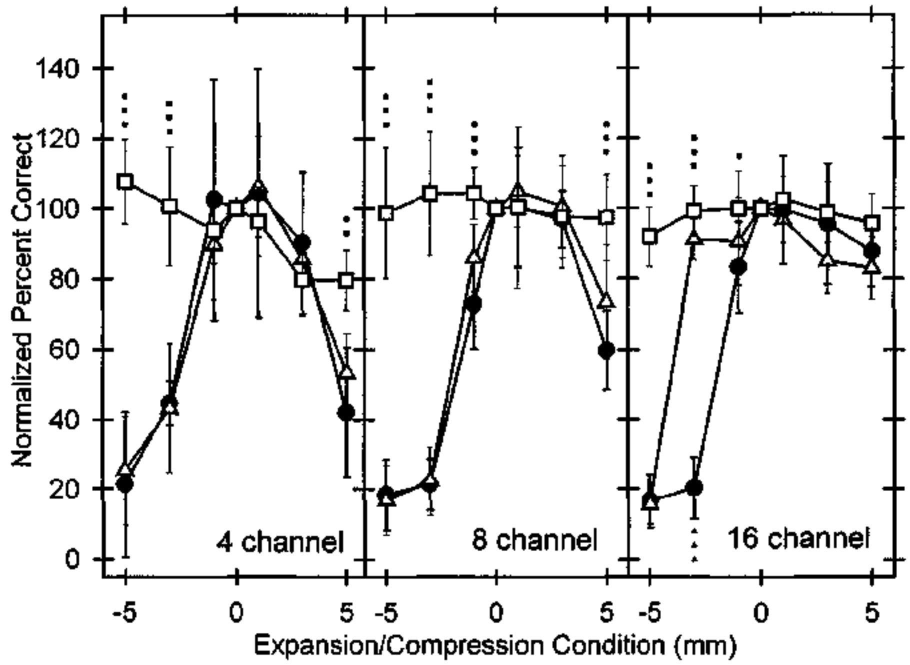


Figure 5.11: Normalized vowel recognition scores at 20 mm simulated insertion depth. The filled circles are normalized scores from Chapter 4 with mismatch at both ends. The open squares show the normalized scores with basal mismatch and the open triangles show the normalized scores with apical mismatch. The small triangles under the scores show significant difference between mismatch at both ends and apical mismatch, by an unpaired t-test: one triangle for $p<0.05$, two triangles for $p<0.01$, and three triangles for $p<0.001$. The small squares on top of scores show significant difference between mismatch at both ends and basal mismatch by an unpaired t-test: one square for $p<0.05$, two squares for $p<0.01$, and three squares for $p<0.001$.

Percent correct scores from all three mismatch maps (mismatch at both ends, apical and basal mismatch) were also replotted for an easier comparison as shown in Figures 5.11-5.14. This time the mismatch performance curves to be compared are scores from different subjects. To minimize this inter-subject variability the percent correct scores are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

normalized. The scores obtained with 0 mm perfect match condition were normalized to 100% and the other scores are plotted as relative drops from 100%. All scores are also corrected for chance level.

The filled circles in Figures 5.11 and 5.12 are the same average percent correct scores, normalized and replotted from vowel and consonant recognition tests with the standard mismatch conditions as shown in Figures 4.7-4.8 and Figures 4.3-4.4, respectively. The open triangles show the percent correct scores with apical mismatch conditions, normalized and replotted from Figures 5.3-5.6, and the open squares show the percent correct scores with basal mismatch conditions, normalized and replotted from Figures 5.7-5.10. As in the previous figures the number of channels increases from 4 to 8 to 16 in the left, middle, and right panels, respectively, in each figure. The filled triangles under the scores show significant difference between mismatch at both ends and mismatch at apical end only conditions, by an unpaired t-test: one triangle for $p<0.05$, two triangles for $p<0.01$, and three triangles for $p<0.001$. The filled squares on top of scores show significant difference between mapping at both ends and mapping at basal end only by an unpaired t-test: one square for $p<0.05$, two squares for $p<0.01$, and three squares for $p<0.001$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

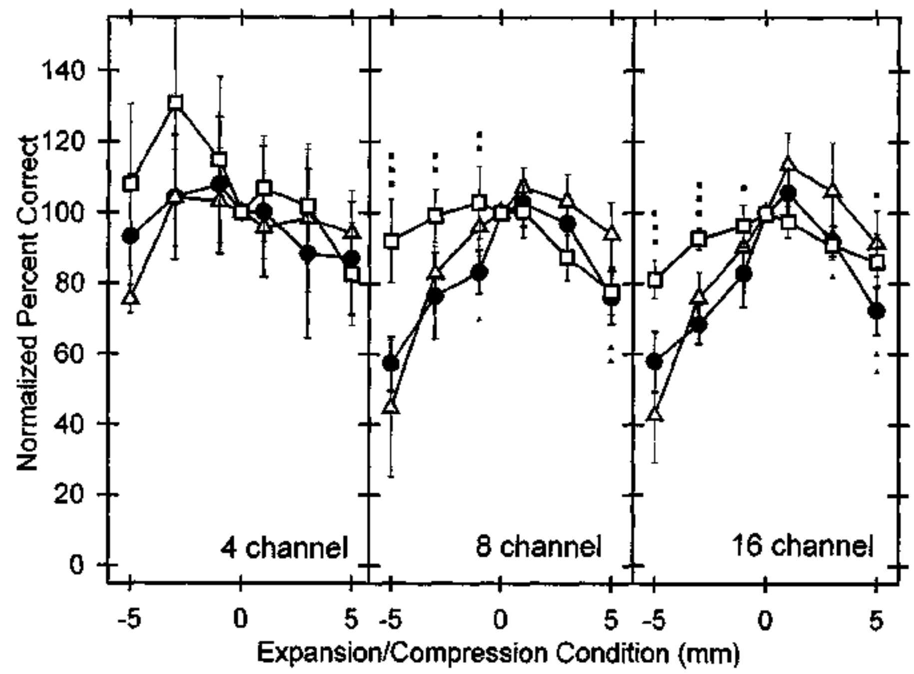


Figure 5.12: Similar to Figure 5.11, but the simulated insertion depth was 25 mm.

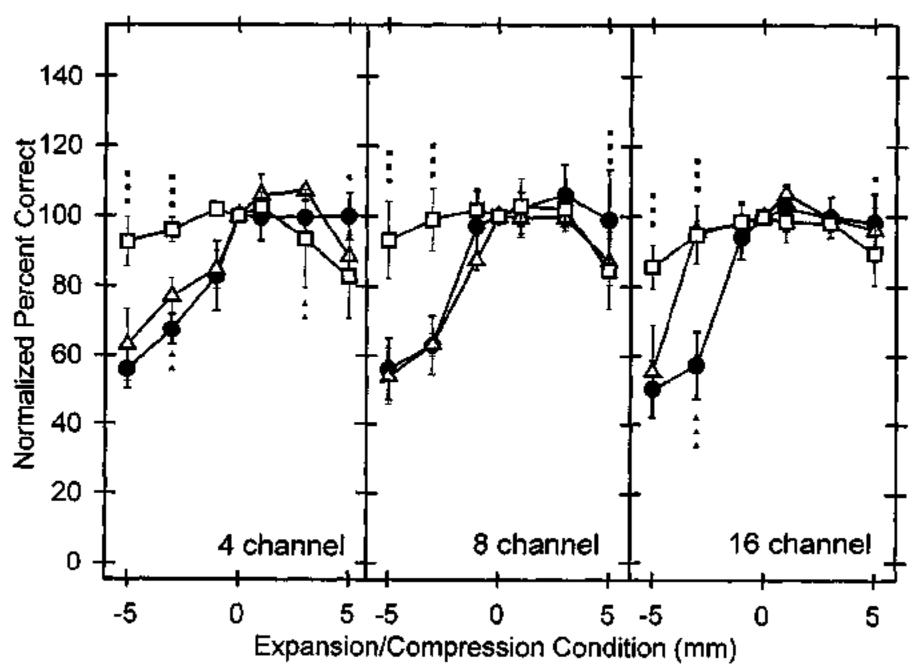


Figure 5.13: Similar to Figure 5.11, consonant recognition scores with simulated insertion depth of 20 mm.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

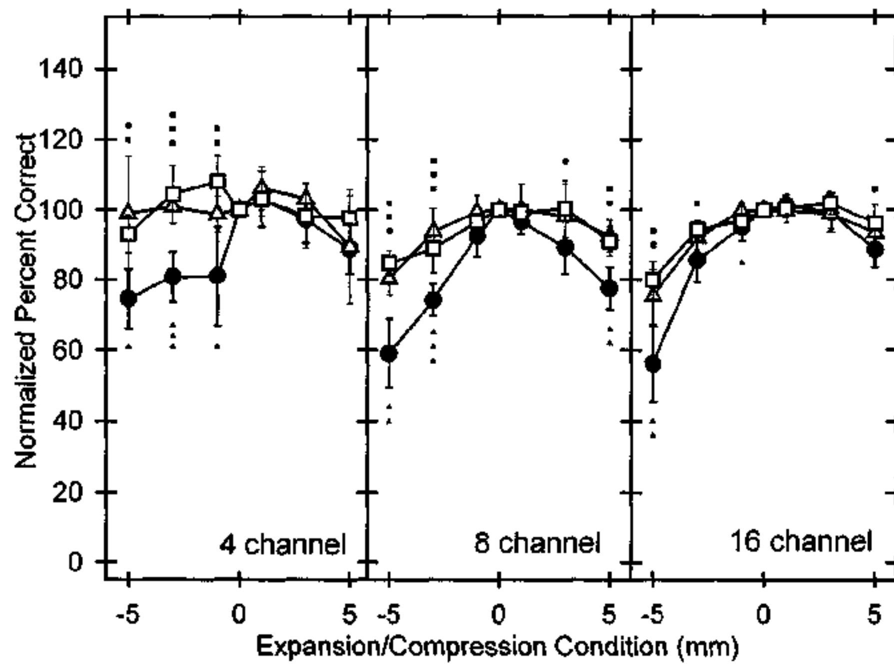


Figure 5.14: Similar to Figure 5.11, consonant recognition scores with simulated insertion depth of 25 mm.

In the figures the largest difference was observed with basal expansion compared to expansion at both ends of the stimulation region. For all stimuli at both insertion depths basal frequency-place expansion (shown with open triangles) resulted in significantly less drop from 0 mm matched condition. From Tables 5.5-5.6 we see that actually there is no effect of basal expansion on vowel recognition performance and only a minimal drop was observed for consonants at 20 mm simulated insertion depth. At 20 mm the noise carrier bands cover a range from 4 to 20 mm from the round window. 4 mm from the round window responds best to frequencies around 10 kHz by Greenwood mapping function, which is a much higher frequency than the primary speech frequency range. Even at -5 mm basal expansion

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

condition only frequencies higher than 6 kHz were lost. The most important spectral range with most spectral energy for vowels is much lower than 6 kHz, and therefore they are not affected by the loss of 6-10 kHz spectral information. Consonants need a little higher frequency range to be identified correctly. Therefore consonant recognition scores are affected by basal expansion little more compared to vowels. At 25 mm insertion depth there is more effect of basal expansion compared to 20 mm insertion, because at 25 mm the mismatched frequencies are more important both for vowels and consonants, yet the effect is still smaller compared to mismatch at both ends or apical mismatch.

Basal compression had a bigger effect on phoneme recognition scores compared to basal expansion, but again this effect was smaller compared to mismatch applied at both ends.

The results in this section also show that apical mismatch mostly resulted in similar performance to mismatch at both ends except for consonant recognition at 25 mm. At this insertion depth apical compression and expansion resulted in significantly higher scores, especially at extreme cases of expansion (-5 mm) and compression (+5 mm), than the mismatch conditions applied at both ends (see Table 5.6).

Similar to compression and expansion at both ends, we observed that the apical compression can be tolerated up to a few millimeters. +1 mm compression condition represents compromise between loss of information and compression. For some cases, as a result of this compromise, the best performance was obtained at +1 mm compression condition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.3. Implications for Implants:

The main conclusion from the comparison of the three mismatched maps is that matching the lower frequency end of the stimulation region is more crucial for good speech recognition performance. For insertions around 20 mm the low end of the stimulation region is right in the center of important frequency region for understanding speech. For deeper insertion it is still more advantageous to match the apical end for vowels, but the primary spectral region for consonants is already inside the stimulation region and so consonants can be perceived equally well up to a larger degree of expansion or compression. If electrodes are inserted deeper, such as 30 mm insertion of Med-El Combi, then the primary spectral region for vowel would also fit within the stimulation region and we would expect to see less effect of expansion or compression for vowels. Yet if the electrode array is short, high frequencies necessary for consonant recognition will be excluded, decreasing consonant recognition.

So far in the experiments only insertion depths of 20 mm and 25 mm were simulated. Deeper insertions than 25 mm and shallower insertions than 20 mm were tested later with implant patients (as will be shown later in Chapters 7 and 8). All results combined support the hypothesis that it is beneficial to speech recognition to match the most important acoustic information to its correct tonotopic location.

These results imply that for implants the most ideal setting would be a long array deeply inserted: that would cover a wide range of spectral information and this information could be presented to the matching place in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the cochlea. Yet because of the physical and physiological limitations the electrode array can not always be located deeply inside the cochlea. In addition, the cochlea is considerably narrower towards the apex, decreasing the physical distance between the auditory nerves on opposite sides of the spiral. Therefore, even if the electrodes can be inserted deeply it may still be difficult to stimulate the correct low-frequency auditory nerves independently. All these factors decrease the chances for the implant user to access correct low-frequency information.

5.2.4. Prediction by Modified SII Model

The modified SII model, described in Chapter 4, was used to predict the experimental data produced by mismatch at both ends, apical mismatch, and basal mismatch for a processor at 20 mm insertion. The same fitting parameter of $\alpha=0.06$ from Chapter 4 was used for all three maps. Figure 5.15 shows the predicted performance by the modified SII model for apical mismatch in the left panel, and for the basal mismatch in the middle panel. The scores are normalized such that the percent correct score with 0 mm matched condition is 100%.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

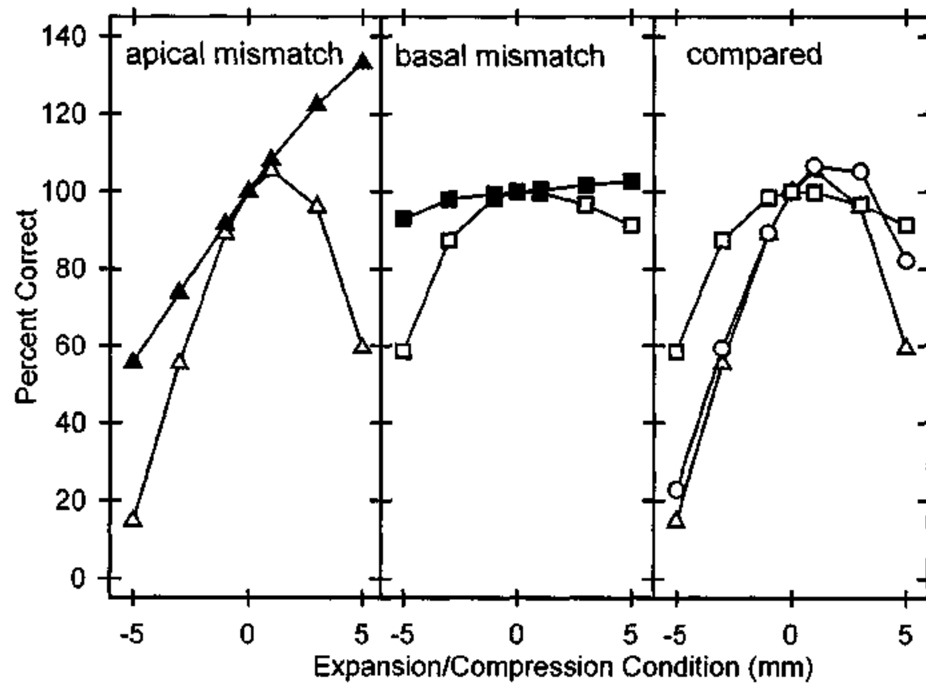


Figure 5.15: SII prediction for apical and basal mismatch as a function of frequency-place compression and expansion at 20 mm simulated insertion depth. The left panel shows the predicted performance for apical mismatch, and the middle panel shows the predicted performance for basal mismatch. The filled symbols in these panels show the baseline scores calculated using Equation 4.3 while the open symbols show the mismatch scores calculated with the additional drop from mismatched frequency bands, using Equation 4.4. Third panel shows all predicted scores from all three mismatch maps. The open circles show the scores with mismatch applied at both ends, the open triangles show the apical mismatch scores, and the open squares show the basal mismatch scores.

The filled symbols show the baseline scores where the analysis band was always assigned onto the same length as the carrier band range. The predicted baseline scores are calculated using Equation 4.3. Because both ranges are the same length they are always matched. Therefore the baseline performance is dominated by the amount of input acoustic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

information range. It increases steeply for apical mismatch conditions, similar to mismatch applied at both ends. Also compared to experimental data with apical mismatch we observe that the baseline performance is more similar to vowel baseline (shown in Figure 5.3) than the consonant baseline (shown in Figure 5.5). As mentioned before, SII predicts vowel recognition performance more closely because it is based on the amount of spectral information contributing to the overall spectrum and vowels are the phonemes that are mostly affected by spectral contents. In addition, the SII band importance function most probably fits the spectral contents of vowels better than consonants. The pattern of baseline performance in Figure 5.3 changes as the number of channels changes. Similar to the baseline performances with vowels when mismatch was applied at both ends as it was shown in Figure 4.7, the baseline drops with increasing range with 4 channels, stays the same from 0 mm matched condition to +5 mm compression with 8 channels, and constantly increases with 16 channels. SII can predict only the monotonic increase exhibited with 16 channels where the effect of spectral resolution (or bandwidth of individual bands) is minimal due to the large number of channels. The flat performance of the baseline condition for basal mismatch is predicted by the model successfully. Similarly, the experimental baseline performances dropped slightly with increasing acoustic input range with 4 and 8 channels, but the 16 channels processor resembles vowel recognition baseline performance very closely (Figures 5.7-5.10).

The additional drop from the baseline conditions comes from the mismatch of the analysis bands, as it was shown with a subtractive term in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Equation 4.4. The mismatch scores calculated using this equation are shown in the left panel for apical mismatch, and middle panel for basal mismatch in Figure 5.15. There was a bigger drop in performance from the baseline with apical mismatch compared to basal mismatch. Because apical mismatch is at the important frequency ranges for speech, mismatching these frequencies deteriorates the performance more than basal mismatch.

The third panel shows predicted scores with all three mismatch maps. The open circles show the scores with mismatch applied at both ends, the open triangles show the apical mismatch scores, and the open squares show the basal mismatch scores. Similar to experimental data with vowel recognition, apical mismatch resulted in a performance more similar to mismatch at both ends. Even though the basal mismatch performance was much more flat compared to the other maps, the model predicted a bigger drop with basal mismatch than was actually observed in the experiments.

5.3. COMPRESSION AND EXPANSION WITH SHIFTED ELECTRODES

In the simulations so far it has been assumed that the exact location of the electrode array inside the cochlea is known precisely so that the acoustical range can be matched to the stimulation range. In real life, this might not be the case. The insertion depth of the electrode array is estimated by the surgeon during the surgery and generally this is the only available information about the electrode array position. In extreme cases such as ossification or malformation of the cochlea the array might even bend during insertion increasing the ambiguity in the exact location of the array. Ketten *et*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

al. (1998) found a large variation in the insertion depths among Nucleus patients. Yet the insertion depth of the electrode array was only a few mms off from the estimated depth on average for each patient and we have shown that a few millimeter of mismatch can be tolerated as long as it is close enough to matching condition. Also the newer electrode designs provide deeper insertions as well as a consistency in the insertion depth from patient to patient. Nevertheless it is an interesting question to explore that how the expansion and compression would affect speech recognition if the electrodes were not at the estimated location and how much tolerance we have for such combined spectral mismatch.

In this experiment, the frequency range was expanded or compressed onto the noiseband carriers for an assumed insertion depth. In the case of implants this relates to the estimated value of the insertion depth by the surgeon. Then the carrier band range was shifted apically or basally to simulate an array whose actual location inside the cochlea differs from the estimated distance. The effect of such a shift between analysis and carrier bands on speech performance was explored when combined with compression and expansion.

5.3.1. Experimental Conditions

The compression and expansion mapping conditions were implemented with the noise-band vocoder for 8 and 16 channel processors at 20 and 25 mm simulated insertion depths. First, the analysis bands were matched on the noise carrier band range. Then the carrier bands

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(representing the electrodes) were shifted apically (further into the cochlea) by +1mm, +3 mm, and +5 mm. At each shifted position of carrier bands, the analysis band range was kept the same as it was matched to the carrier band range before apical shift. As a result as the carrier bands were shifted the mismatch between analysis and carrier bands increased. Then the carrier bands were shifted basally (towards the round window) by -1 mm, -3 mm, and -5 mm, while the analysis band range was again kept at the same range. These conditions show the effect of pure shift because there is no compression or expansion of the map yet. They simulate how the matched map results change if electrodes are shifted apically, i.e., inserted deeper, or shifted basally, i.e. inserted shallower than the surgeon's estimated depth.

After testing the matched map with shifted electrodes, this time the analysis band range was +5 mm compressed onto the carrier band range, when the carrier bands were at their actual location. While the acoustic range was kept at the same values determined by +5 mm compression with unshifted carrier band range, the carrier bands were shifted apically and basally. The same procedure was repeated with - 5 mm expansion.

Figure 5.16 summarizes matched (as shown in the middle portion), +5 mm compression (as shown in the top portion), and -5 mm expansion (a shown in the bottom portion) conditions with +5 mm apically shifted noise carrier bands. The electrodes are assumed to be inserted to 25 mm whereas the electrodes are actually inserted 30 mm. They are shifted +5 mm apically, i.e. further pushed into the cochlea from their normal location of 25 mm.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Note that even though the bands are set up to match the electrodes at 25 mm, they actually do not match anymore because the simulated electrodes are in 5 mm too far, and therefore it is not as ideal a condition as matching when the electrodes were not shifted. This condition now is a pure shift in the spectrum without any compression or expansion of the acoustical information. On the other hand compression does not look as bad as compression from both ends without the shift because an apical shift of +5 mm actually causes the apical frequencies to match and from apical mismatch experiment we can guess that this might actually improve the performance. Expansion also causes the frequencies to match at one end, but this is the high frequency end, and from basal mismatch experiment we know that matching the basal end does not change the speech performance significantly.

The conditions with matched (as shown in the middle portion), +5 mm compressed (as shown in the top portion), and -5 mm expanded (a shown in the bottom portion) maps with -5 mm basally shifted noise carrier bands are shown in Figure 5.17. Similarly, matching is not as ideal anymore, all acoustic information going to shifted locations on the stimulation region. In this case compression combined with basal shifting results in matched basal end of the stimulation whereas expansion with shifting results in matched apical end. Therefore we might see some improvement with expansion with the basal shift compared to expansion with electrodes that are not shifted.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

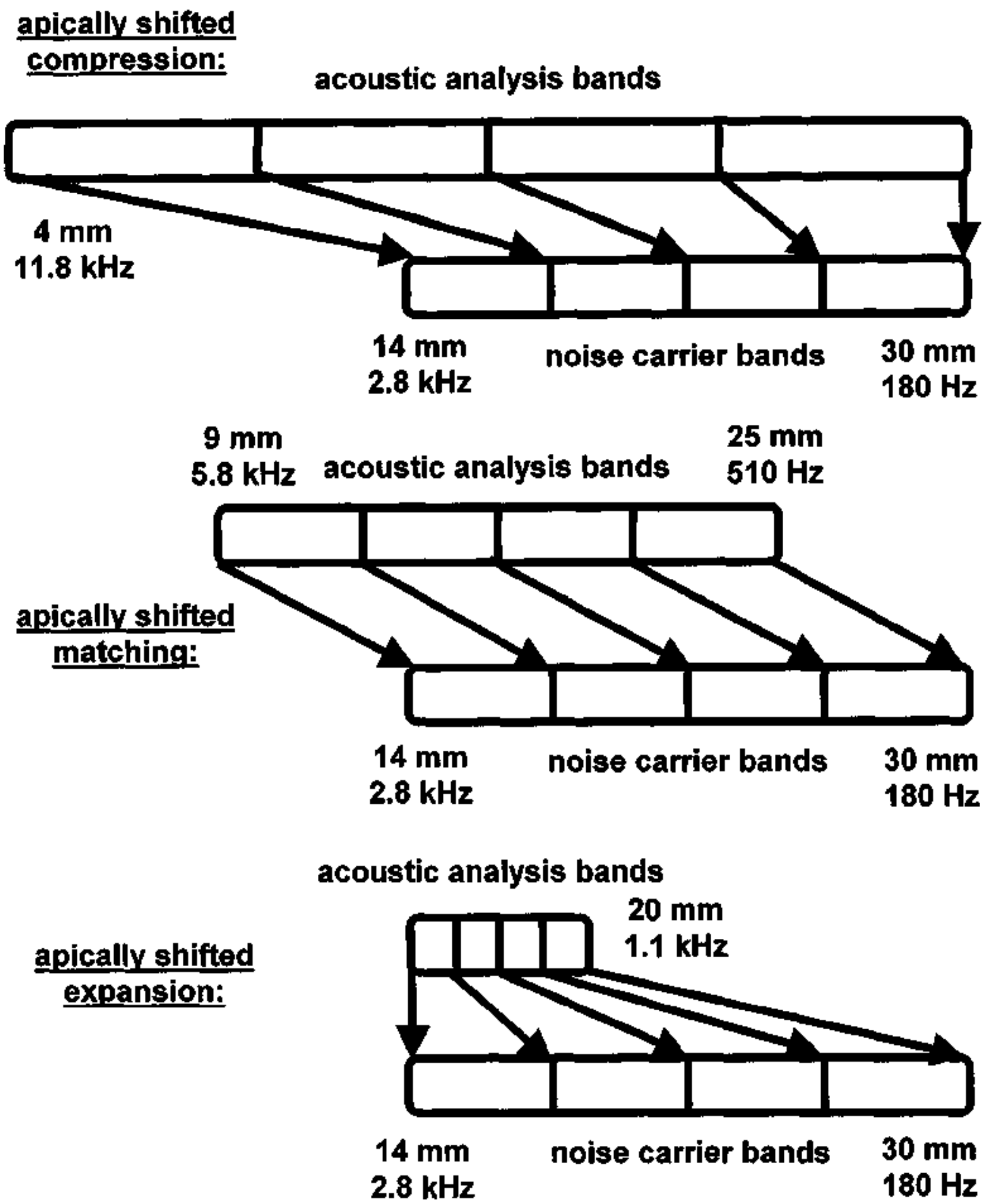


Figure 5.16: Frequency-place mapping conditions for 4-channel processor at the simulated 25 mm insertion depth, where the noise carrier bands are shifted +5 mm apically resulting in an actual insertion depth of 30 mm for simulated electrodes. The carrier bands are fixed simulating the shifted position of the electrode array (14-30 mm: 180-2800 Hz). The top row shows the +5 mm compression condition, the middle row shows the 0 mm matched condition, and the lower panel shows the -5 mm expansion condition when electrodes are shifted +5 mm apically.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

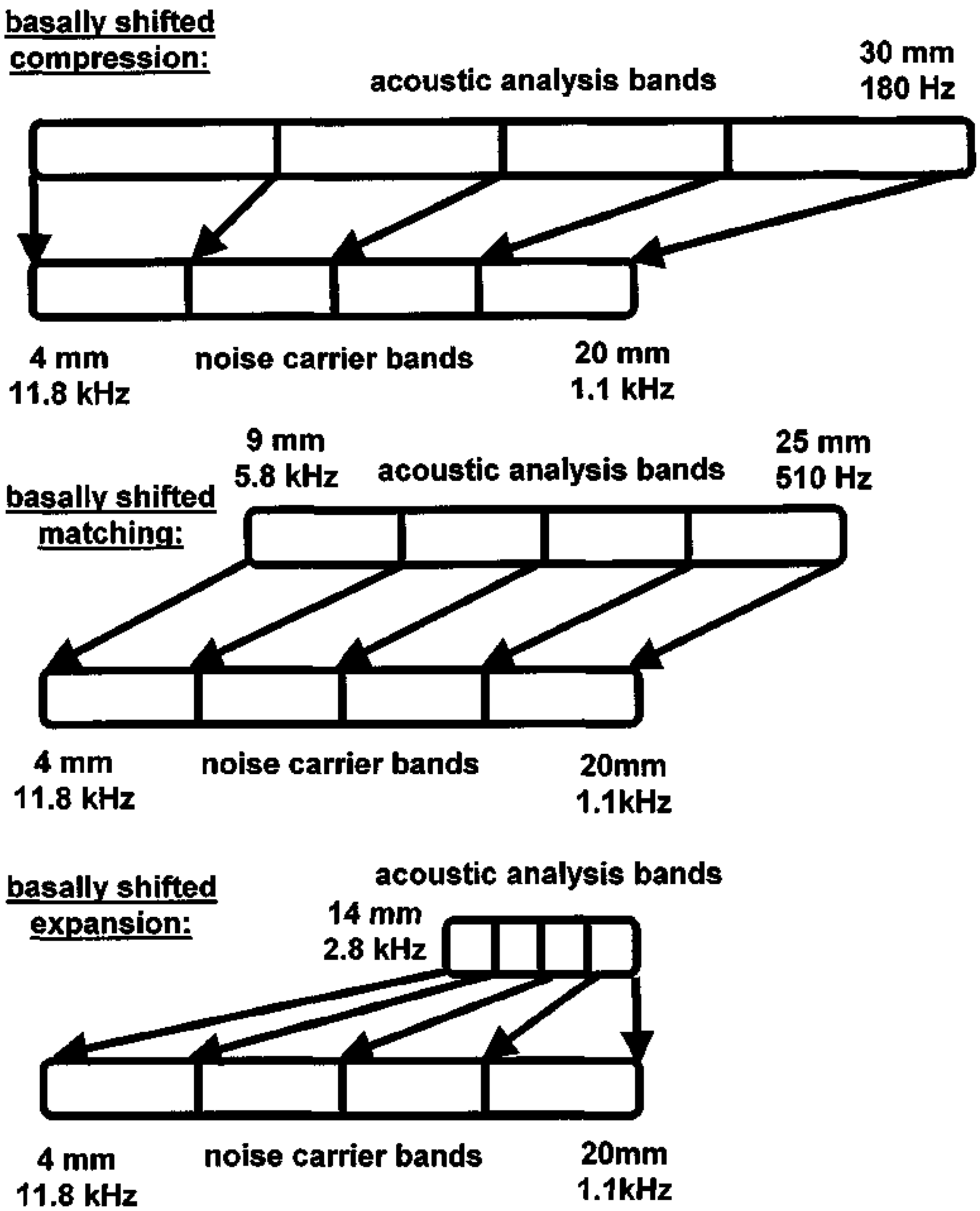


Figure 5.17: Frequency-place mapping conditions for 4-channel processor at the simulated 25 mm insertion depth, where the noise carrier bands are shifted -5 mm basally resulting in an actual insertion depth of 20 mm for simulated electrodes. The carrier bands are fixed simulating the shifted position of the electrode array (4-20 mm: 1.1-12 kHz). The top row shows the +5 mm compression condition, the middle row shows the 0 mm matched condition, and the lower panel shows the -5 mm expansion condition when electrodes are shifted -5 mm basally.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The simulated electrode locations and the analysis band ranges are given in mm for all compression and expansion conditions (-5, -3, -1 mm expansion, 0 mm matching, +1, +3, +5 mm compression) combined with electrode shifts (-5, -3, -1 mm basal shift, 0 mm no shift, +1, +3, +5 mm apical shift) in Table 5.7, for 25 mm insertion depth. Columns show the shift of the electrode array position while rows show the compression and expansion conditions. In every entry of the table the top line is the tonotopic range of the analysis bands and the bottom line is the tonotopic range of the carrier bands which imply the simulated electrode array range. Perfect match condition with no compression, expansion, or shift is the center entry in the table. The middle column shows the compression and expansion conditions when electrodes are not shifted, and the middle row shows pure shift of the bands without any compression or expansion. The apical shift of +5 mm shown in Figure 5.16 is the rightmost column of the table while the -5 mm basal shift shown in Figure 5.17 is the leftmost entry of the table.

Same table with corresponding frequency ranges (by Greenwood mapping function) is given in Table 5.8.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shift ► / mism ▼	- 5mm	-3 mm	-1 mm	0 mm	+1 mm	+3 mm	+5 mm
-5 mm	14-20mm 4-20mm	14-20mm 6-22mm	14-20mm 8-24mm	14-20mm 9-25mm	14-20mm 10-26mm	14-20mm 12-28mm	14-20mm 14-30mm
-3 mm	12-22mm 4-20mm	12-22mm 6-22mm	12-22mm 8-24mm	12-22mm 9-25mm	12-22mm 10-26mm	12-22mm 12-28mm	12-22mm 14-30mm
-1 mm	10-24mm 4-20mm	10-24mm 6-22mm	10-24mm 8-24mm	10-24mm 9-25mm	10-24mm 10-26mm	10-24mm 12-28mm	10-24mm 14-30mm
0 mm	9-25mm 4-20mm	9-25mm 6-22mm	9-25mm 8-24mm	9-25mm 9-25mm	9-25mm 10-26mm	9-25mm 12-28mm	9-25mm 14-30mm
+1 mm	8-26mm 4-20mm	8-26mm 6-22mm	8-26mm 8-24mm	8-26mm 9-25mm	8-26mm 10-26mm	8-26mm 12-28mm	8-26mm 14-30mm
+3 mm	6-28mm 4-20mm	6-28mm 6-22mm	6-28mm 8-24mm	6-28mm 9-25mm	6-28mm 10-26mm	6-28mm 12-28mm	6-28mm 14-30mm
+5 mm	4-30mm 4-20mm	4-30mm 6-22mm	4-30mm 8-24mm	4-30mm 9-25mm	4-30mm 10-26mm	4-30mm 12-28mm	4-30mm 14-30mm

Table 5.7: The electrode locations and the analysis band ranges (in mm) for all compression and expansion conditions (-5, -3, -1 mm expansion, 0 mm matching, +1, +3, +5 mm compression) combined with electrode shifts (-5, -3, -1 mm basal shift, 0 mm no shift, +1, +3, +5 mm apical shift), for 25 mm insertion depth. Columns show the shift of the electrode array whereas rows show the compression and expansion conditions. In every entry the top line is the range for the analysis bands and the bottom line is the range for the noise carrier bands. Perfect match condition with no compression, expansion, or shift is the center entry in the table. Corresponding column shows the compression and expansion conditions when electrodes are not shifted, and corresponding row shows pure shift of the bands without any compression or expansion. The apical shift of +5 mm shown in Figure 5.15 is the rightmost column of the table whereas the -5 mm basal shift shown in Figure 5.16 is the leftmost entry of the table.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shift ▶ / mism ▼	- 5mm	-3mm	-1mm	0mm	+1mm	+3mm	+5mm
-5mm	1.1kHz-2.8kHz	1.1kHz-2.8kHz	1.1kHz-2.8kHz	1.1kHz-2.8kHz	1.1kHz-2.8kHz	1.1kHz-2.8kHz	1.1kHz-2.8kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
-3mm	850Hz-8.9kHz	850Hz-8.9kHz	850Hz-8.9kHz	850Hz-8.9kHz	850Hz-8.9kHz	850Hz-8.9kHz	850Hz-8.9kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
-1mm	610Hz-6.8kHz	610Hz-6.8kHz	610Hz-6.8kHz	610Hz-6.8kHz	610Hz-6.8kHz	610Hz-6.8kHz	610Hz-6.8kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
0mm	510Hz-5.8kHz	510Hz-5.8kHz	510Hz-5.8kHz	510Hz-5.8kHz	510Hz-5.8kHz	510Hz-5.8kHz	510Hz-5.8kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
+1mm	430Hz-5kHz	430Hz-5kHz	430Hz-5kHz	430Hz-5kHz	430Hz-5kHz	430Hz-5kHz	430Hz-5kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
+3mm	290Hz-3.8kHz	290Hz-3.8kHz	290Hz-3.8kHz	290Hz-3.8kHz	290Hz-3.8kHz	290Hz-3.8kHz	290Hz-3.8kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz
+5mm	180Hz-2.8kHz	180Hz-2.8kHz	180Hz-2.8kHz	180Hz-2.8kHz	180Hz-2.8kHz	180Hz-2.8kHz	180Hz-2.8kHz
	1.1kHz-11.8kHz	850Hz-8.9kHz	610Hz-6.8kHz	510Hz-5.8kHz	430Hz-5kHz	290Hz-3.8kHz	180Hz-2.8kHz

Table 5.8: The same electrode location and the analysis band ranges as in Table 5.7, this time given in frequency ranges in Hz, for 25 mm insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.3.2. Results

In this experiment, 8 channel and 16 channel processors at 20 and 25 mm insertion depth were simulated. Figures 5.18-5.21 show the average percent correct scores for vowels and consonants from 5 normal hearing subjects with +5 mm compressed, matched, and -5 mm expanded maps, when electrodes were shifted apically or basally, as a function of the shift. 0 mm shift condition refers to unshifted noiseband carriers. Therefore the results for 0 mm shift condition are the same as compression, matching, and expansion results in Chapter 4, when the simulated electrodes were at their estimated location, as it was shown in Figures 4.3, 4.4, and 4.7, 4.8. The filled triangles show the scores when the analysis band range was matched to the estimated carrier band range, but the carrier band range was shifted apically or basally to simulate electrodes shifted from their estimated location (as shown in the middle row of Table 5.7). The extreme apical shift of +5 mm with matched map, which is the rightmost score of the filled triangles, is the condition that was shown in the middle row of Figure 5.16. Similarly, the extreme basal shift of -5 mm with matched map, which is the leftmost score of the filled triangles, is the condition that was shown in the middle row of Figure 5.17. The filled squares show the scores when the analysis band range was +5 mm compressed onto the estimated carrier band range and then the carrier band range was shifted apically or basally (as shown in the bottom row of Table 5.7). The filled circles show the scores when the analysis band range was -5 mm expanded in reference to the estimated carrier band range and when the carrier band range was shifted apically or

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

basally afterwards (as shown in the top row of Table 5.7). The small squares on the top of the scores denote the significance of the difference between the performances from compression with shift and match with shift: one square for $p<0.05$, two squares for $p<0.01$, and three squares for $p<0.001$. The small circles at the bottom of the scores similarly denote the significance of the difference between the performances from expansion with shift and match with shift. All scores were corrected for chance level.

Vowel recognition scores with 8 and 16 channel processors at 20 mm insertion depth are presented in Figure 5.18. The perfect match condition, where there was no compression, expansion, or shift, is shown by the middle score (at 0 mm shift) of the filled triangles. As expected, the perfect match resulted in the best performance compared to all compressed, expanded, and shifted conditions. A tolerance range of a few mm is observed for shifting with 16 channels; an apical or basal shift of ±1 mm results in almost the same performance as the perfect match condition. As the carrier bands shifted, the length of the analysis band range and the noise band range remained the same for the matched map, as it was shown in the middle rows of Figure 5.16 and 5.17. In this map there is no compression or expansion, and all changes in the speech recognition is solely due to the shifts between the bands. As a result, the performance drops from the perfect match score for shifts in either direction almost equally, because the bands are mismatched in equal amounts for apical or basal shift. The performance with the compressed map, as shown by the filled squares, did not drop when the carrier bands were shifted apically up to +3 mm from the no shift condition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This pattern was expected because, as it was shown in the top row of Figure 5.16, important frequencies for speech are actually matched better to the appropriate cochlear place as the carrier bands are shifted apically. The drop with the basal shift is much sharper because these frequencies are exposed to a larger degree of mismatch. For this reason the drop with the shift is not symmetrical around 0 mm shift condition for the compressed map. The performance with expanded map, as shown by the filled circles, generally resulted in a low performance at all shift conditions. This is most likely dominated by the large amount of information lost due to the limited range of the analysis bands, as it was shown in the bottom rows of Figure 5.16 and 5.17. The performance actually improved with increasing basal shift, again most likely due to the better matching of the important frequency ranges at the appropriate cochlear place at this particular condition, shown in the bottom row of Figure 5.17.

The patterns of performances were similar for 8 and 16 channels, except with 16 channels the top scores around perfect matching were higher than 8 channels, even though all scores dropped to similar values for ±5 mm apical and basal shifts for both processors.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

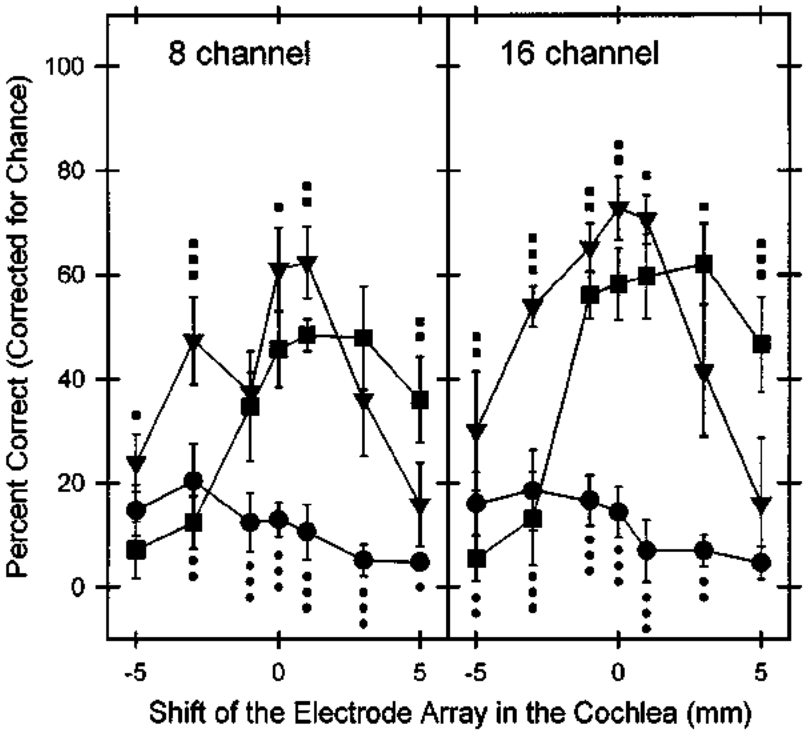


Figure 5.18: Vowel recognition percent scores at the simulated 20 mm electrode insertion depth for 8-channel and 16-channel processors, as a function of the shift in the simulated electrode array position represented by the noise carrier bands. The percent scores, corrected for chance level, are the average from 5 normal hearing subjects. The filled triangles show the scores with the matched map, with the perfect match condition in the center, where there was no compression, expansion, or shift. The analysis range remains the same as the carrier bands are shifted apically or basally. The filled squares show the scores with the compressed map as the carrier bands are shifted. The small filled squares on the top of the scores denote the significance of the difference between the compression and match conditions such that: one square for $p<0.05$, two squares for $p<0.01$, and three squares for $p<0.001$. The filled circles show the scores with the expanded map as the carrier bands are shifted. The small filled circles at the bottom of the scores denote the significance of the difference between the expansion and match conditions such that: one circle for $p<0.05$, two circles for $p<0.01$, and three circles for $p<0.001$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

All these changes with shifted carrier bands produce an interesting pattern between the three maps. For example, even though the matched map gives the best scores for unshifted carrier bands compared to compressed or expanded maps, the compressed map actually becomes more beneficial for conditions where carrier bands are shifted more than +3 mm apically. As explained above, with apical shifting important frequencies in the analysis band range of the compressed map move closer to their matching tonotopic locations. Similarly, at -5 mm basal shift, compressed map results in such a low score that expanded map actually becomes more advantageous in spite of the much shorter range of the analysis bands. Again the main reason is better matched frequencies that are important for speech understanding. Yet matched map still gives the best performance in most conditions. The matched map performance is significantly better than expanded map performance except the +5 mm apical shift condition, and it is significantly better than the compressed map except the +3 mm and +5 mm apical shift conditions.

Vowel recognition scores with 8 and 16 channel processors at 25 mm insertion depth are presented in Figure 5.19. Performance was similar to 20 mm insertion depth with matched and compressed maps, except the scores are higher around the unshifted carrier bands condition. Because the scores still drop to similar levels at ±5 mm apical and basal shift conditions, the drop with increasing shift is more visible at this insertion depth. Also expanded map resulted in much higher scores when insertion depth changed from 20 mm to 25 mm. Similar to results at 20 mm insertion, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

matched map gives better performance than the compressed maps for all basally shifted, unshifted, and +1 mm apically shifted carrier band conditions. Except for the +5 mm apical and -5 mm basal shifts it was also better than the expanded map.

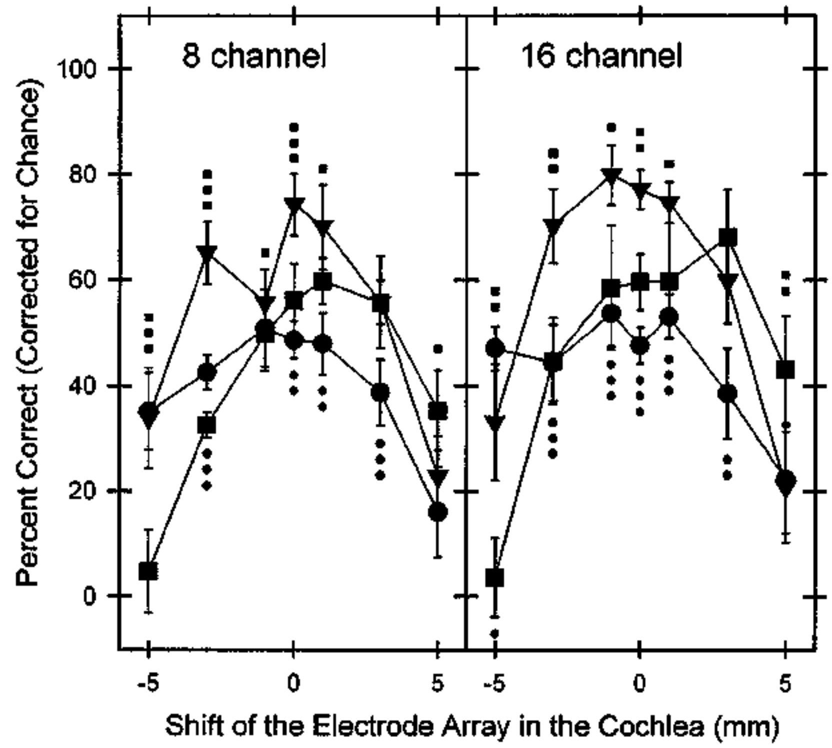


Figure 5.19: Similar to Figure 5.18, except the simulated insertion depth is 25 mm.

Average percent scores for consonant recognition with 8 and 16 channel processors at 20 mm and 25 mm insertion depths are presented in Figure 5.20 and 5.21, respectively. With consonants, scores with the matched and compressed map were similar, with the exception of +5 mm apical shift condition where compressed map produced a better performance

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

at 20 mm insertion and 0 mm unshifted condition where the matched map gave a better score at 25 mm insertion. Performance with expanded map was always worse except -5 mm basal shift condition. As mentioned before, consonant recognition is not as dependent on spectral content as vowels, therefore even at extreme spectral mismatch conditions subjects can get consonant recognition scores as high as 40%. For the same reason, spectral mismatch always has smaller effects on consonants. Van Tasell *et al.* (1987) observed consonant recognition of 30-40% even with no spectral cues, so such a low consonant recognition score may indicate that the listeners are unable to use spectral cues at all.

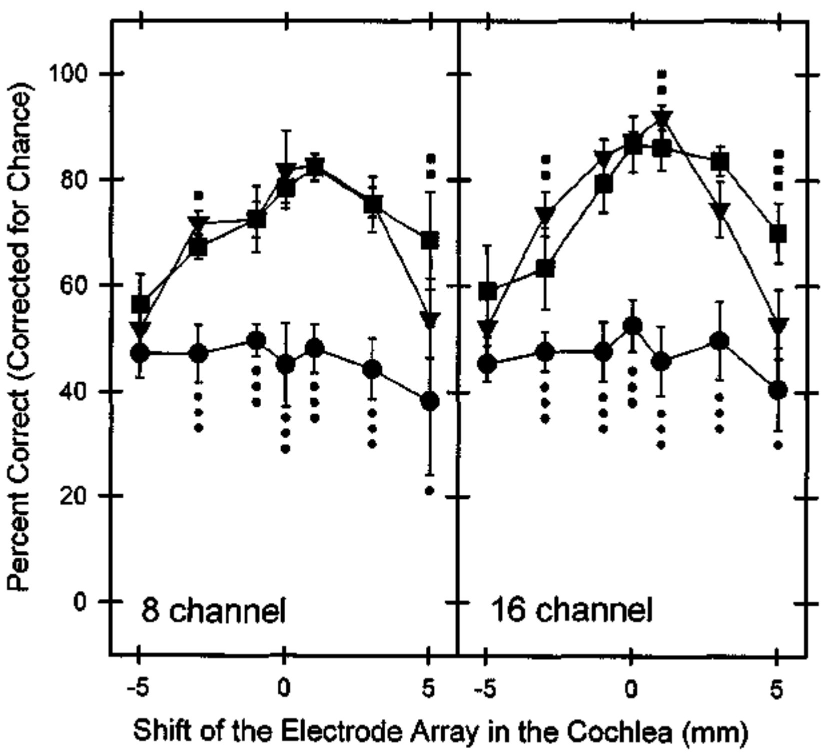


Figure 5.20: Similar to Figure 5.18, consonant recognition scores with the simulated insertion depth of 20 mm.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

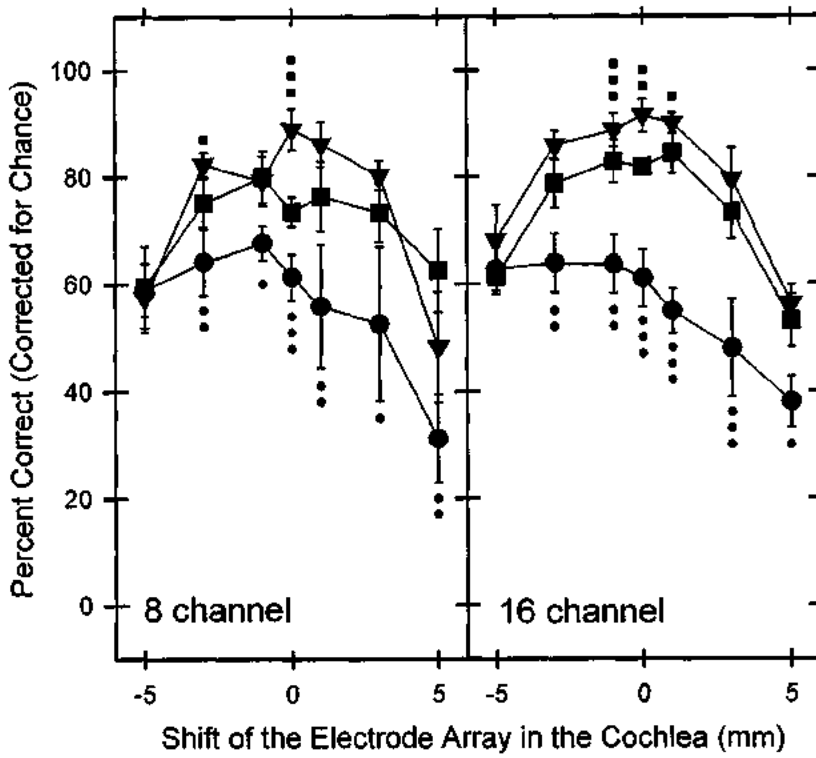


Figure 5.21: Similar to Figure 5.20, except the simulated insertion depth is 25 mm.

For application purposes, the overall results at both insertions mean that matching is better for the best vowel recognition if the electrodes of an implant user are off from their estimated insertion depth by a few mm; up to +2 mm for apically shifted electrodes, and up to -4 mm for basally shifted electrodes. This result is actually promising since these numbers are inside the error range between the actual location of the electrodes and the estimated values by the surgeons as reported by Ketten *et al.* (1998).

Only the noise bands that were apically shifted by +3 mm and +5 mm resulted in better speech recognition with compression. Apically shifted noise bands simulate electrodes that are inserted deeper than the insertion depth that was estimated by the surgeon. As it was shown by Marsh *et al.* (1993),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ketten *et al.* (1998), and Skinner *et al.* (2003), this is rarely the case; as a matter of fact, the electrodes are generally inserted shallower than the surgically estimated depth.

To summarize the overall results, a mismatch between the frequency information and the normal cochlear place for that information always produces a drop in performance, whether the mismatch is due to an apical/basal shift or due to compression/expansion.

5.3.3. Prediction by the Modified SII Model

The original SII model given in Equation 4.3 could not predict any of the results of this experiment because it does not take into account any mismatched between analysis and carrier bands. However when the modified equation (Equation 4.4) is used the basic pattern of the experimental data can be replicated nicely.

The predicted performances for the three maps by the modified SII model are given in Figure 5.22. Similar to experimental data the filled triangles show the predicted scores for the matched map, the filled squares show the scores for the compressed map, and the filled circles show the expanded map scores as a function of the shift of the carrier bands from the assumed insertion depth. Insertion depth was 20 mm. $\alpha=0.03$ was used to obtain a similar pattern to the experimental data. The only other free parameter was the score for the perfect matched condition because SII could not predict absolute scores. This score was again adjusted with the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experimental data for that specific condition and all the other scores were shifted up or down accordingly. -5 mm and -3 mm basal shift condition with compressed map and +5 mm apical shift condition with expanded map produced negative scores with the model, therefore they were manually set to 0.

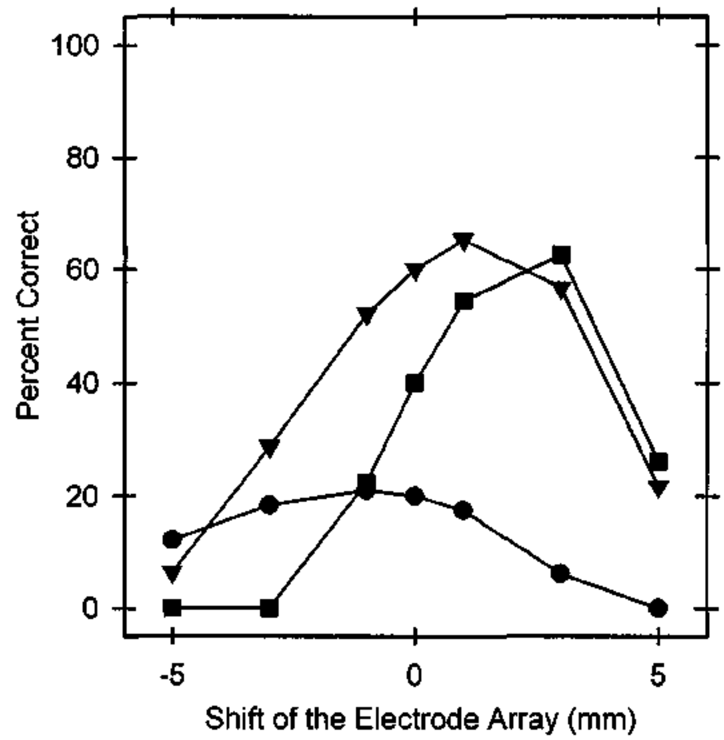


Figure 5.22: Percent scores for matched, compressed, and expanded maps as a function of the shift in the electrode array, as predicted by the modified SII model.

The pattern of the predicted data is similar to vowel recognition scores at 20 mm insertion presented in Figure 5.18. SII always gives better predictions for vowels than consonants as explained before. As in the experimental data, the SII scores also show that matched map gives the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

best scores for a range from -4 mm basal shift to +3 apical shift. Compressed map gives better scores with apical shifts larger than +3 mm compared to the matched map. Expanded map generally gives low scores, but the scores improve with some basal shift. At - 5 mm basal shift expanded map is even slightly better than both compressed and matched maps.

5.4. HOLES IN HEARING

Shannon *et al.*(2002) simulated regions of damaged neurons along the cochlear distance which create holes in the tonotopic representation of the spectral information. A noiseband vocoder was used to simulate a 20 band processor with the frequency-electrode mapping similar to SPEAK Table 9. Holes were created in the basal, middle, and apical areas by eliminating noise carrier bands. The frequencies related to the holes are given in Table 5.9. In the table, columns from left to right show the increasing size of the holes by the number of the carrier bands eliminated. In each condition, the size of the hole was increased by eliminating two more bands, and the relative drop from the baseline condition (the whole range of 20 bands with no hole in it) was observed. Rows show the different locations for the holes. The top line in every entry shows the number of the bands eliminated, numbering starting from the basal end, similar to Nucleus implant electrodes. The corresponding frequency ranges are presented in the bottom lines.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

	2 bands	4 bands	6 bands	8 bands
basal	5 - 6 4180-5740 Hz	4 - 7 3570-6730 Hz	3 - 8 3080-7880 Hz	2 - 9 2680-9240 Hz
middle	10 - 11 2030-2680 Hz	9 - 12 1770-3080 Hz	8 - 13 1550-3570 Hz	7 - 14 1350-4180 Hz
apical	15 - 16 950-1350 Hz	14 - 17 750-1550 Hz	13 - 18 550-1770 Hz	12 - 19 350-2030 Hz

Table 5.9: The frequency ranges of the spectral holes. Columns from left to right show the increasing size of the holes by the number of the carrier bands eliminated. Rows show the different locations for the holes. The top line in every entry shows the number of the bands eliminated, numbering starting from the basal end. The corresponding frequency ranges are presented in the bottom lines.

The results showed that holes in the apical area were more detrimental to speech than the holes in the basal or middle area. From the frequency ranges in Table 5.9 we can actually predict this result because the apical holes eliminate important frequencies for speech.

In following experiments instead of eliminating all information inside the hole, the information was reassigned to the adjacent carrier bands in three different ways: all on the apical neighboring carrier band, all on the basal neighboring carrier band, and split half and half to both bands. The results showed that there was no difference between the results obtained with the holes where the information inside was dropped or reassigned with any of these maps.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

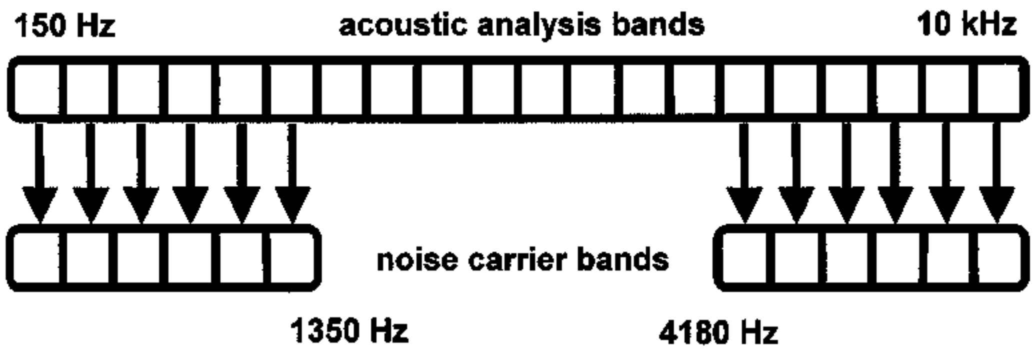
The 20 channel processor with a middle hole of 8 bands is shown in Figure 5.23. The top row shows when all the information inside the hole was dropped and the rest of the information was assigned on to the remaining noise band carrier. The middle shows when the information inside the hole was split onto the bands next to the hole. These two maps are the conditions from Shannon *et al.* study.

In this experiment, we reassigned the spectral information onto whole range of the remaining bands instead of the neighboring bands and compared these conditions to the dropped conditions. This mapping is shown in Figure 5.23, in the bottom row. The results of the study by Shannon *et al.*(2002) suggest that there should not be any difference between two conditions. Conventional SII model would predict that there should be no drop with increasing hole if the information was somehow presented onto the remaining region. Yet the results from this dissertation imply that there would be an additional effect of such mismatch compared to dropped conditions.

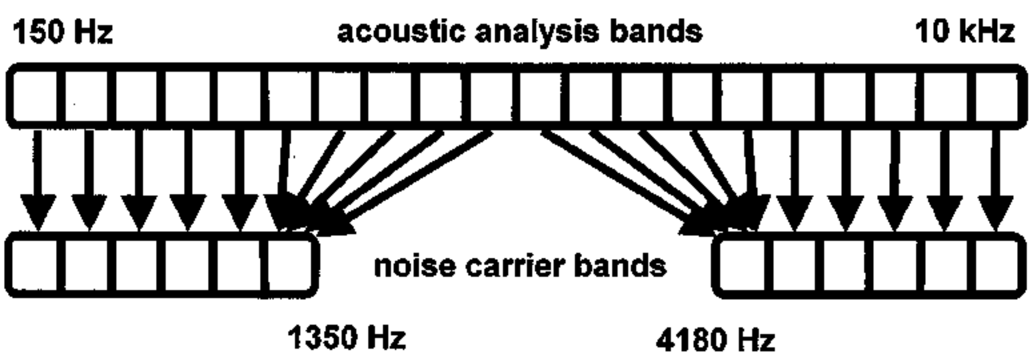
In the previous experiments the frequency ranges were always matched at the ends of the stimulation region. This time they will be matched inside the stimulation region at the ends of the holes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

20 electrodes with a hole of 8 bands dropped:



20 electrodes with a hole of 8 bands split on neighbor bands:



20 electrodes with a hole of 8 bands reassigned on carrier

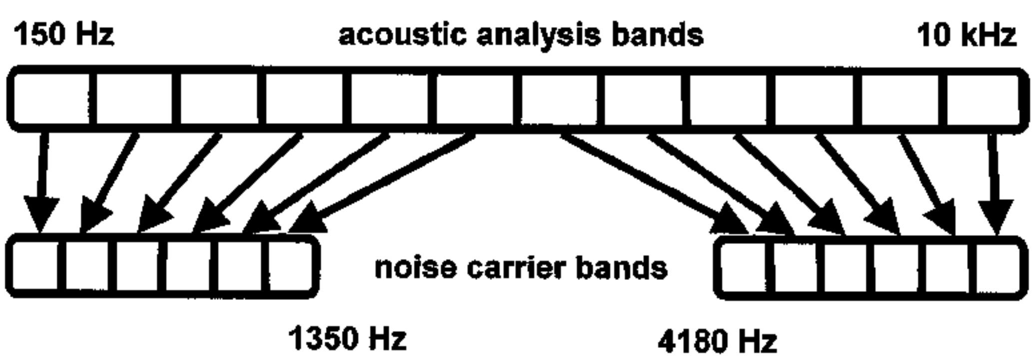


Figure 5.23: Three different maps to assign the analysis bands onto the remaining noise carrier bands with a hole of 8 bands in the middle range of the stimulation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.4.1. Experimental Method

To keep the consistency with Shannon study, the same 20 band processor was simulated with the noiseband vocoder. This time the frequency bands were not distributed in terms of cochlear distance, but with SPEAK Table 9 of the Nucleus system, for the same reason.

5.4.2. Results

The average vowel and consonant recognition scores of 7 normal hearing subjects are presented in Figures 5.24 and 5.25, respectively, as a function of the hole size. In each figure the left panel shows the effects of a hole in the basal area, the middle panel shows the effects of a hole in the middle area, and the right panel shows the effects of a hole in the apical area. The scores are both normalized such that the baseline performance is 100% and corrected for chance level. The filled symbols show the conditions when the analysis bands inside the hole were simply dropped, as shown in the top row of Figure 5.23. The open symbols show the conditions when the analysis bands inside the hole were reassigned onto the whole range covered by the remaining carrier bands around the hole, as shown in the bottom row of Figure 5.23. The error bars show one standard deviation. The filled circles under the scores show the significantly different scores by the two maps as calculated by a paired t test. Three circles show significance level of $p<0.001$, two circles show significance level of $p<0.01$, and one circle shows significance level of $p<0.05$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

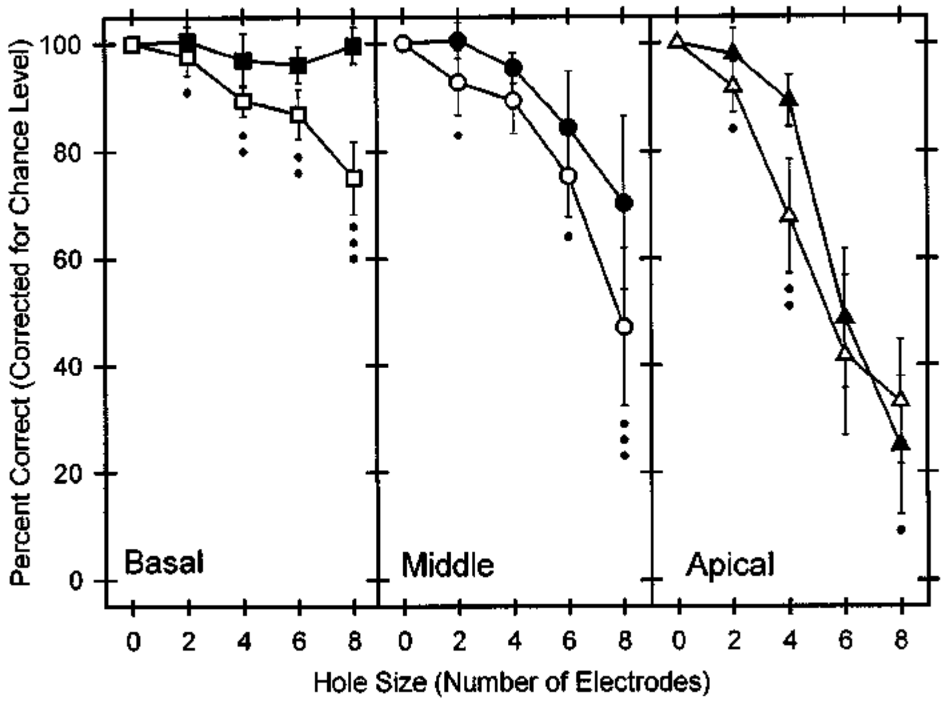


Figure 5.24: Vowel recognition scores as a function of the hole size in terms of the number of bands. The filled symbols show the scores from the dropped map where the information inside the hole was simply eliminated and the rest of the analysis bands were assigned on the matching carrier bands. The open symbols show the scores from the reassigned map where the whole range of the analysis bands was reassigned onto the remaining carrier bands. The error bars show one standard deviation. The filled circles under the scores show the significantly different scores by the two maps as calculated by a paired t test. Three circles show significance level of $p<0.001$, two circles show significance level of $p<0.01$, and one circle shows significance level of $p<0.05$.

The holes created with dropped bands, shown by the filled symbols, are actually the same conditions from the Shannon study. Similar to that study, the effect of the hole increased as the hole moved to middle and apical areas from the basal area. The effect was much more significant on vowels than consonants, as expected.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

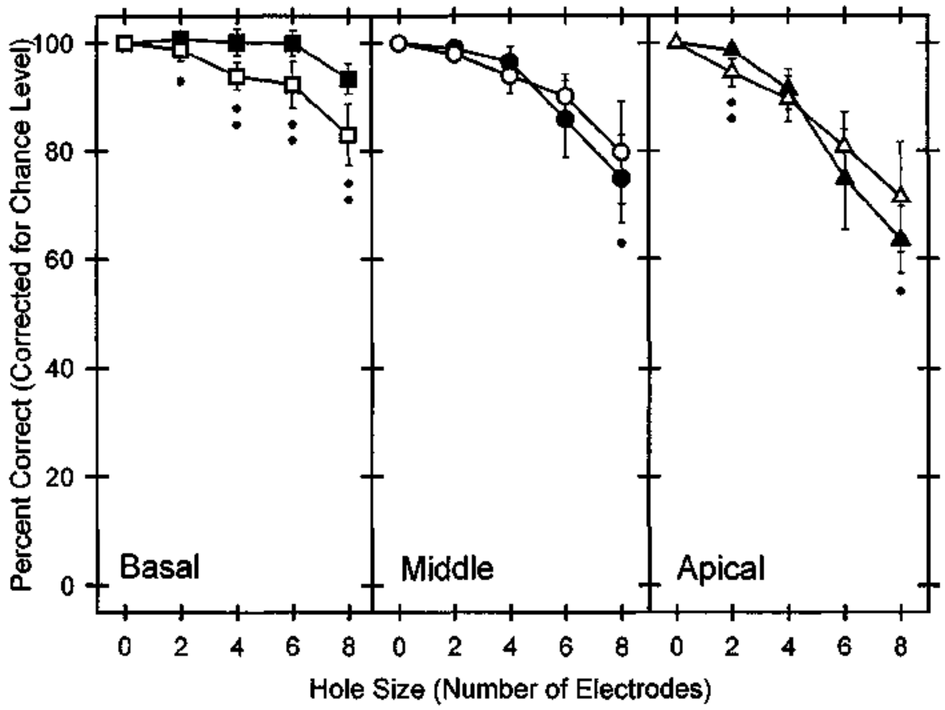


Figure 5.25: Similar to Figure 5.24, except the scores are the average consonant recognition scores.

The interesting result comes from the comparison of two maps. As expected from the previous experiments in this dissertation, the reassigned map resulted in different scores than the dropped map. In both maps the stimulation region represented by the carrier bands are the same. In the reassigned map more information was assigned onto the carrier bands but this also added a mismatch between the analysis bands and the carrier bands. As a result of this mismatch, there was a drop from the performance at some conditions compared with dropped map even though the reassigned map provided more acoustic information. The drop was significant for the holes in the basal area for both vowels and consonants. There was a smaller

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

drop for middle and apical holes, except the hole of 8 bands in the apical area, where there was a slight advantage of the reassigned map over the dropped map. The maps resulted in almost same performance for middle and apical holes with consonants.

These spectral holes can also be interpreted as the middle electrodes turned off in implants. This happens with some patients in cases of uncomfortable stimulations such as facial nerve stimulation. The fitting programs used in the clinics automatically redistribute the whole acoustic input range onto the remaining electrodes as shown in the bottom part of Figure 5.23. The results in this experiment imply that depending on the number and the location of electrodes to be turned off it might be more beneficial to the patient to deactivate those electrodes without changing the assignment of the other electrodes.

5.5. COMPRESSION AND EXPANSION IN NOISE

Many cochlear implant patients have much poorer speech recognition in background noise even though they face this problem in most real life listening conditions. In this experiment, speech-shaped noise is added to the frequency-place compression and expansion conditions to both simulate these more realistic listening conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.5.1. Experimental Conditions

The typical spectrum of speech-shaped noise is shown in Figure 5.26. This noise can simply be produced by filtering wideband noise with a 1st order Butterworth low-pass filter with a cut-off frequency of 1 kHz.

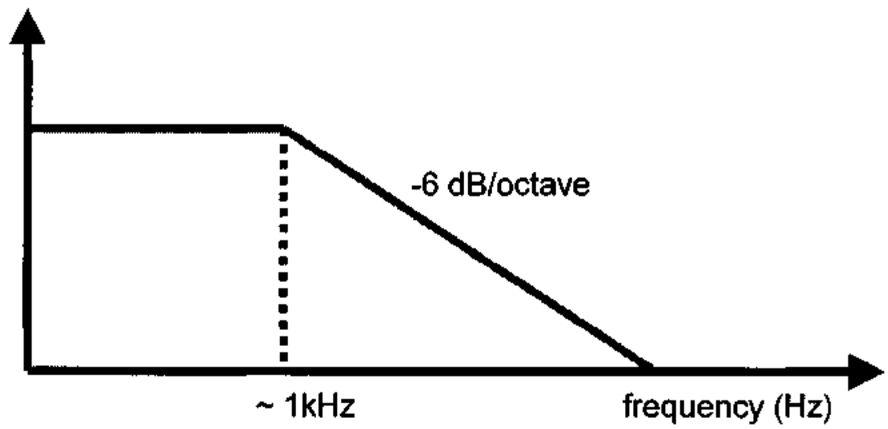


Figure 5.26: Spectrum of speech-shaped noise.

Same test materials were used as in the previous simulation experiments. The noise at SNR levels of -5 dB, -2.5 dB, 0 dB, 2.5 dB, 5 dB, 7.5 dB, 10 dB, and 12.5 dB was added before the speech materials were processed with the noiseband vocoder.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.5.2. Results

Figures 5.27-5.30 show the average vowel and consonant recognition percent scores from 5 normal hearing subjects with an 8 channel processor as a function of increasing noise level in dB with -5 mm expansion, 0 mm matching, and +5 mm compression conditions. The filled symbols show the scores with matched baseline conditions and the open symbols show the scores from the mismatch conditions. The error bars show one standard deviation. The circles on top of the scores show the conditions where there was a significant drop from the matched baseline condition score to the corresponding mismatch condition score.

As expected all performances dropped with increasing noise levels, except -5 mm expansion with vowels at 20 mm, which was already at the floor level and could not drop further with noise. Even though both compression and corresponding baseline performances both dropped, the difference that comes from spectral mismatch was significant for noise levels lower than 0 dB for vowels at 20 mm insertion and -2.5 dB for vowels at 25 mm insertion. Consonant recognition scores similarly dropped significantly with increasing noise levels. There was a significant drop in performance from baseline performance with compression conditions at 25 mm simulated insertion depth for SNR levels better than 0 dB.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

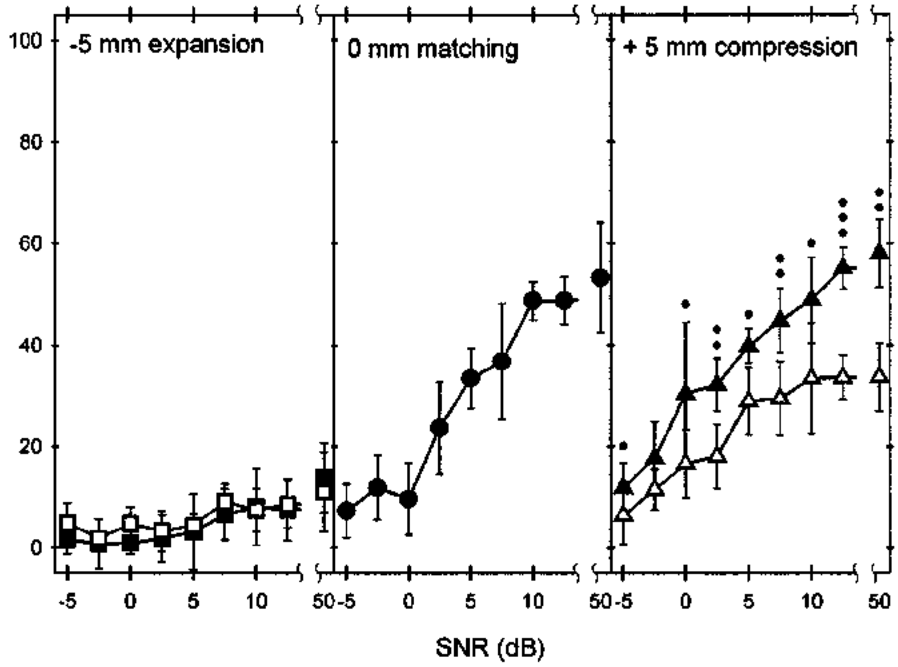


Figure 5.27: The average vowel percent scores from 5 normal hearing subjects with an 8 channel processor at 20 mm simulated insertion depth, as a function of increasing noise level in dB. The filled symbols show the scores with matched baseline conditions and the open symbols show the scores from the mismatch conditions. The error bars show one standard deviation. The circles on top of the scores show the conditions where there was a significant drop from the matched baseline condition score to the corresponding mismatch condition score.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

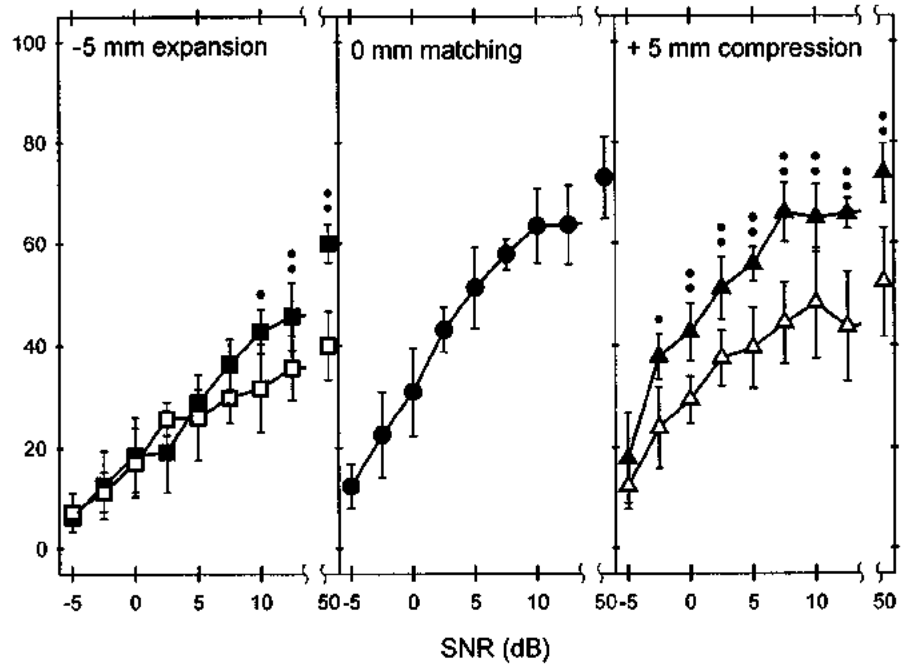


Figure 5.28: Similar to Figure 5.27, except the simulated insertion depth is 25 mm.

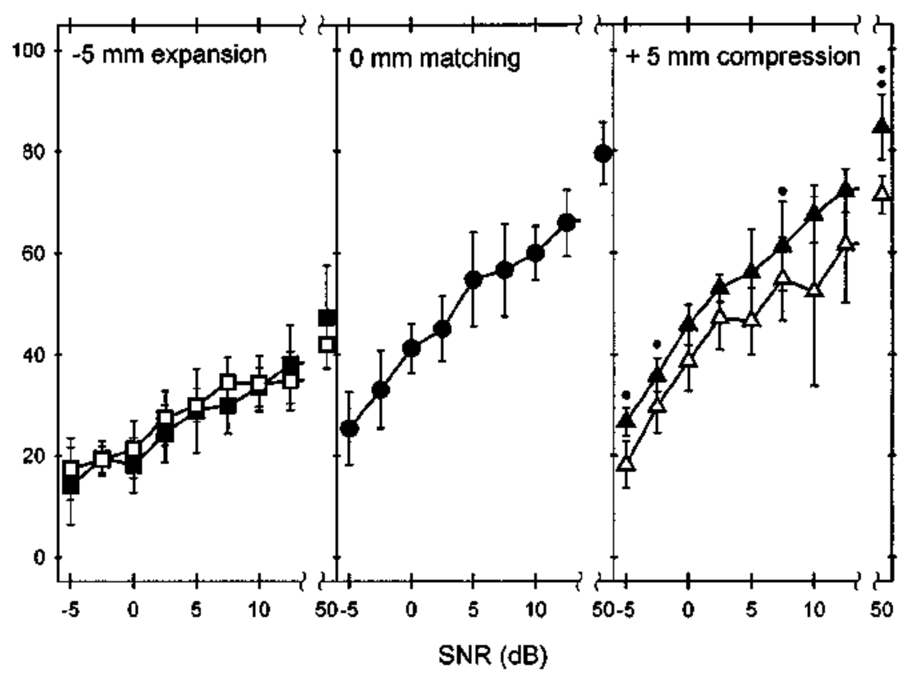


Figure 5.29: Similar to Figure 5.27, consonant recognition scores with the simulated insertion depth of 20 mm.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

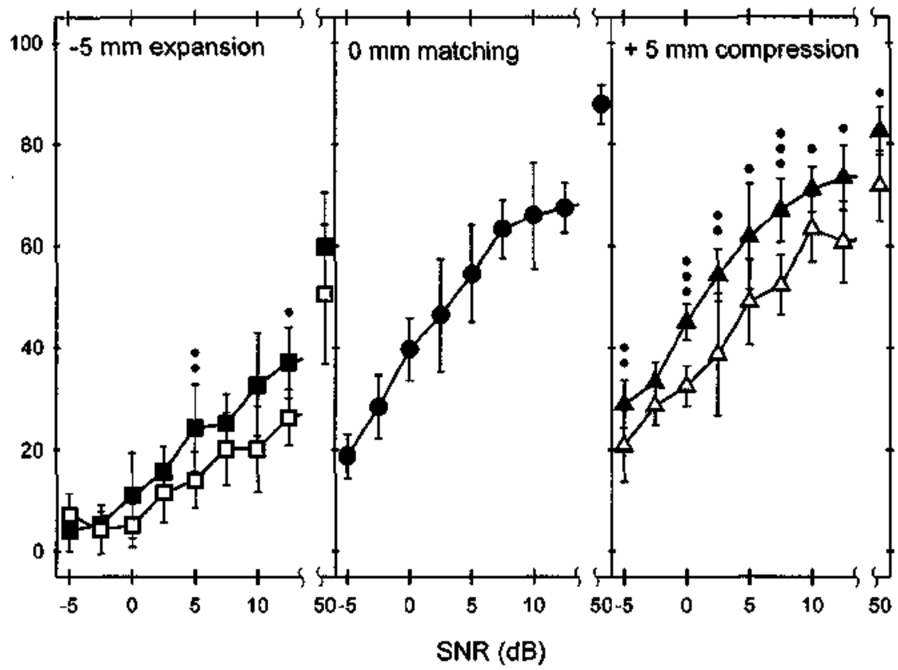
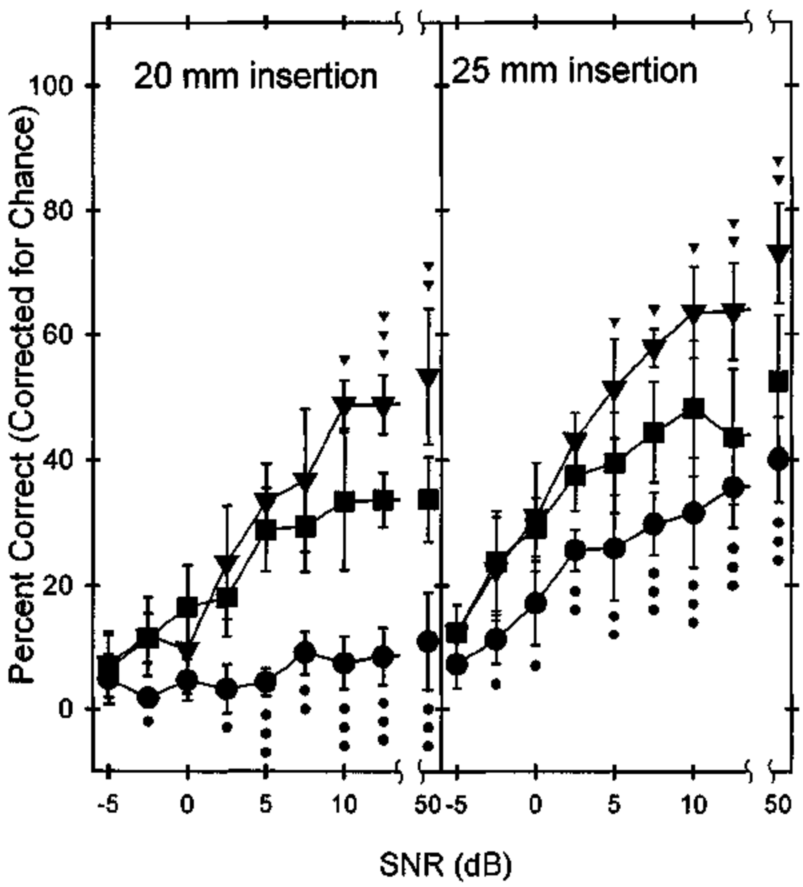


Figure 5.30: Similar to Figure 5.29, except the simulated insertion depth is 25 mm.

To determine how the overall mismatch pattern changes with noise, the mismatch conditions of +5 mm compression and -5 mm compression scores were replotted in Figures 5.31 and 5.32 and compared with the 0 mm matched condition scores. In each figure left panel shows the results from 20 mm simulated insertion depth, and the right panel shows the results from 25 mm simulated insertion depth. In each panel triangles show percent correct scores from 0 mm matched map, squares show the scores from +5 mm compressed map, and circles show the scores from -5 mm expanded map. The scores to the right at each panel are the scores from no noise conditions. The small triangles on top of the scores show the compression conditions that resulted in a performance that was significantly lower than the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

matched condition. Similarly, the small circles under the scores show the expansion conditions that resulted in a performance that was significantly lower than the matched condition.



Figures 5.31: Vowel recognition percent scores for the simulated 20 mm and 25 mm electrode insertion depths, as a function of SNR. The triangles show the scores with 0 mm matched map, the squares show the scores with +5 mm compressed map, and the circles show the scores with -5 mm expanded map. The small triangles on top of the scores show the statistically different scores between the matched and compressed maps, and the small circles in the bottom show the statistically different scores between the matched and expanded maps.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

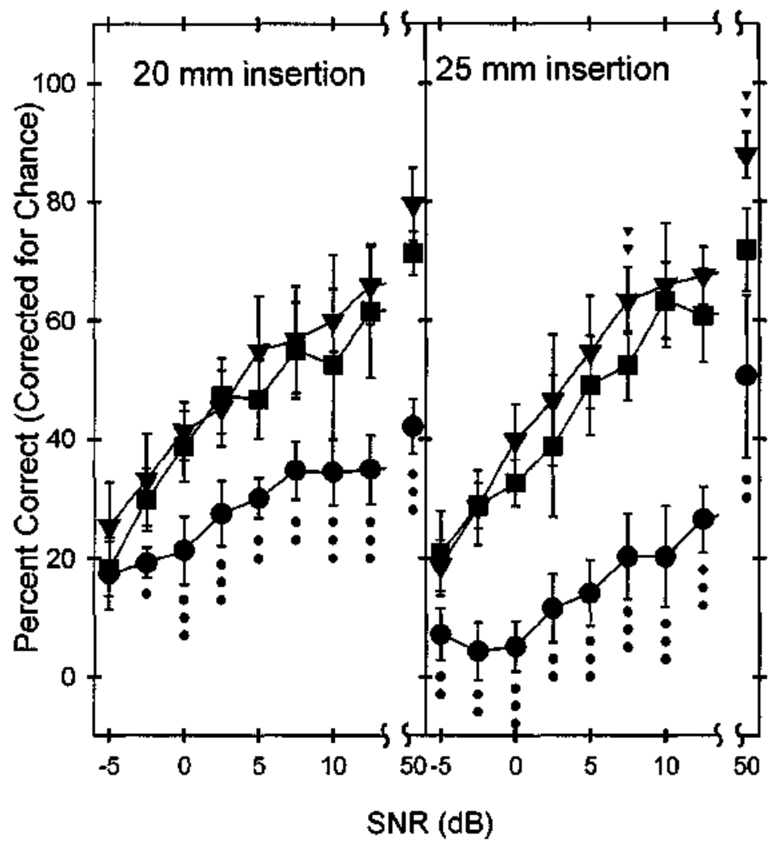


Figure 5.32: Similar to Figure 5.31, except the stimuli are the consonants.

The results show that -5 mm expansion always resulted in significantly lower scores compared to 0 mm matched map. Vowel recognition was better with matched map compared to the compressed map when there was no background noise, as it was shown in Chapter 4. This advantage of matching over compression disappeared for high levels of noise, i.e., for SNR levels lower than 7.5 dB at 20 mm simulated insertion depth and for SNR levels lower than 5 dB at 25 mm simulated insertion depth. Consonants had similar scores with matched and compressed maps at most noise levels.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this chapter, the hypothesis that it is important to match the acoustic input frequencies to the cochlear locations for good speech recognition performance was further confirmed with various spectral mismatch experiments. Diverse conditions were simulated such that electrodes that were shifted from their estimated location or background noise to find out the most optimum listening conditions. It was shown that some frequency regions are more important to match, mostly because they contribute more to speech information. The modified SII model was used to determine how much of the results came from information loss and how much came from spectral mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 6

EFFECTS OF ADAPTATION WITH SIMULATIONS

To the extent that electrical hearing is a completely different mode of hearing any new patient who gets the device needs considerable time to make use of the signals coming from the stimulated neurons and to associate them with actual sounds. Speech recognition improves for almost all patients after the surgery as they gain more experience with their devices. However it is still unknown if patients can adapt to any mapping that is assigned to their devices. As discussed in Chapter 5 several studies have shown significant improvement in speech recognition over a short period of time when subjects were trained with certain processors. However, these studies do not answer some questions: Is this only a short-term adaptation or if the subjects could have kept being trained would they improve further? Most studies show that people learn fast for the first few weeks but usually reach a limit and do not improve any further with more training. Another interesting point is that if they reach a limit in the improvement, and is that limit independent of the starting point?

For this study particularly the question is how performance changes with time if the patient uses compressed or expanded mapping extensively? Would the advantage of matching disappear over time or would the overall performance increase with more experience?

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

First, to see the effect of experience only without any specific training the test with compression-expansion at both ends is repeated with one subject. By the time the test was repeated the subject had gained around 100 hours of experience of listening to the simulated sounds, but did not receive any training on any particular condition. Consonant recognition percent scores are presented in Figures 6.1 and 6.2 for 20 mm and 25 mm insertion depths, respectively. Similarly, vowel recognition scores are presented in Figures 6.3 and 6.4. In each figure the filled symbols show the test results from the beginning of the experiment when the subject did not have any experience with noise-band vocoders. The open symbols are the scores from the same phoneme identification tests when repeated with the same subject after she had approximately 100 hours of noise-band vocoder processing exposure. With consonants there was not much improvement in the scores with experience. Most probably the subject was already making use of all cues available for consonant recognition even at the beginning of the experiment. When vowels are considered almost all scores improved reflecting the effect of the experience. All curves are elevated regardless of the condition keeping the advantage of matching over compression and expansion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

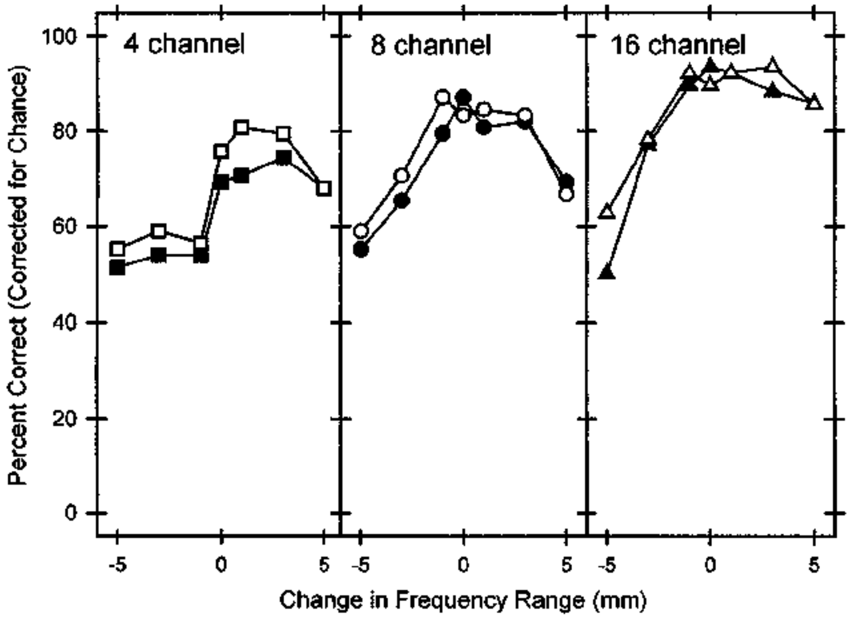


Figure 6.1: Consonant recognition percent scores for the simulated 20 mm electrode insertion depth, as a function of compression or expansion. The same test is repeated from Section 4.2.1 with one subject after she had 100 hours of experience with noise-band vocoders, as shown with open symbols. The consonant percent scores from the beginning of the test when the subject did not have any experience are shown with filled symbols.

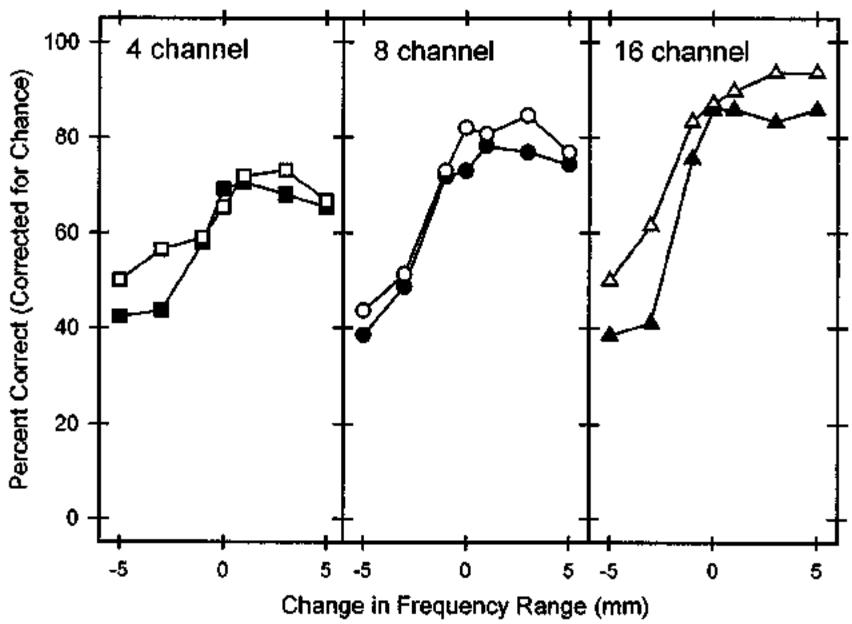


Figure 6.2: Similar to Figure 6.1, for 25 mm simulated insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

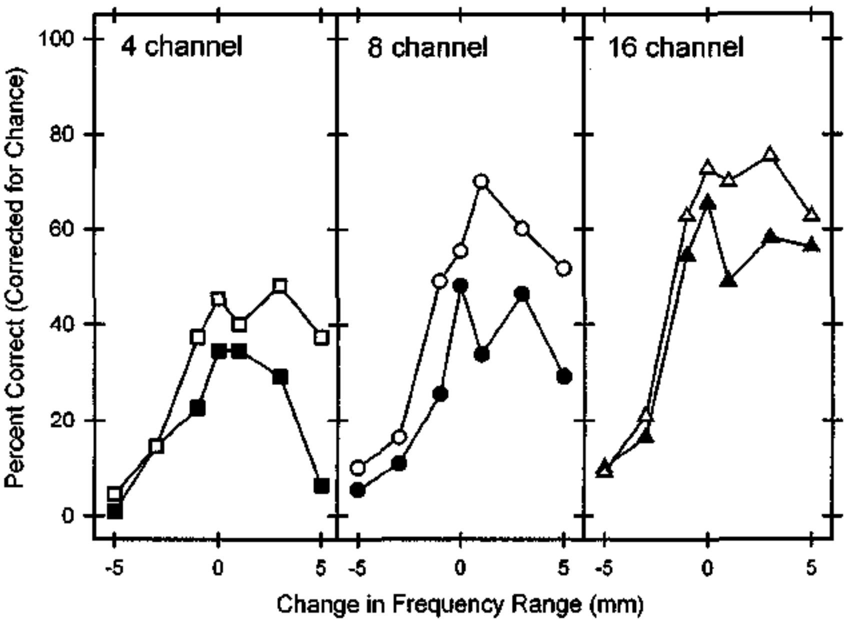


Figure 6.3: Similar to Figure 6.1, consonant recognition scores for 25 mm simulated insertion depth.

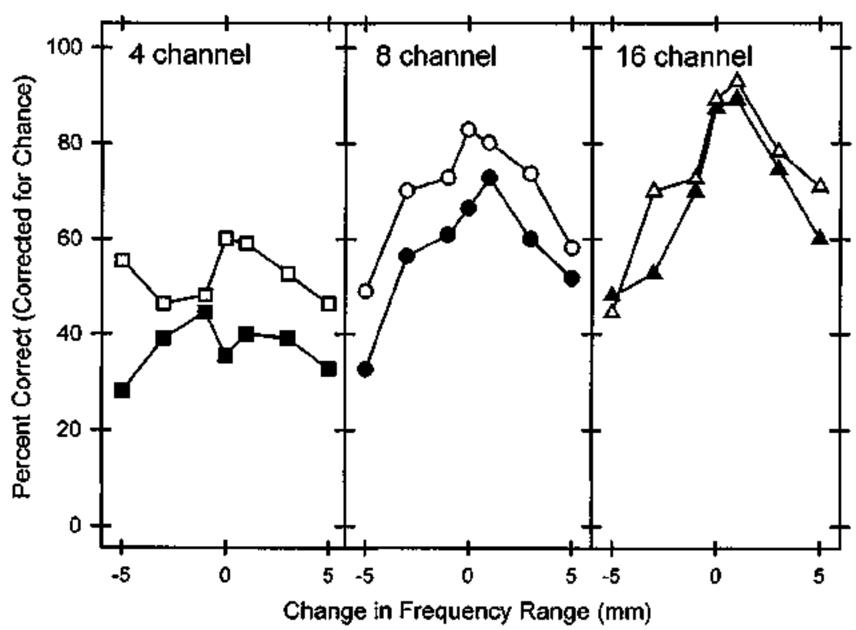


Figure 6.4: Similar to Figure 6.3, for 25 mm simulated insertion depth.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Next, a pilot study was designed as an attempt to train a subject more systematically with frequency-place mismatch conditions. In the studies mentioned before either normal-hearing subjects were trained with connected discourse in simulated conditions or CI patients wore a shifted map in their devices that they used for a certain time in their daily lives. Unlike those studies this subject was trained specifically on vowels because they were the ones which were significantly affected by compression-expansion conditions. The training condition was chosen as +5 mm compression with 8 channel processor at 20 mm simulated insertion depth. Figure 4.7 shows that this condition worsened speech recognition by %30 from matching condition. The subject did not have any experience prior to this experiment and she was given 30 minutes training on vowel recognition for 9 days starting from the first day. The training consisted of the same vowel set as used in the other vowel recognition tests and the subject was given feedback by highlighting the right answer right after she clicks her choice.

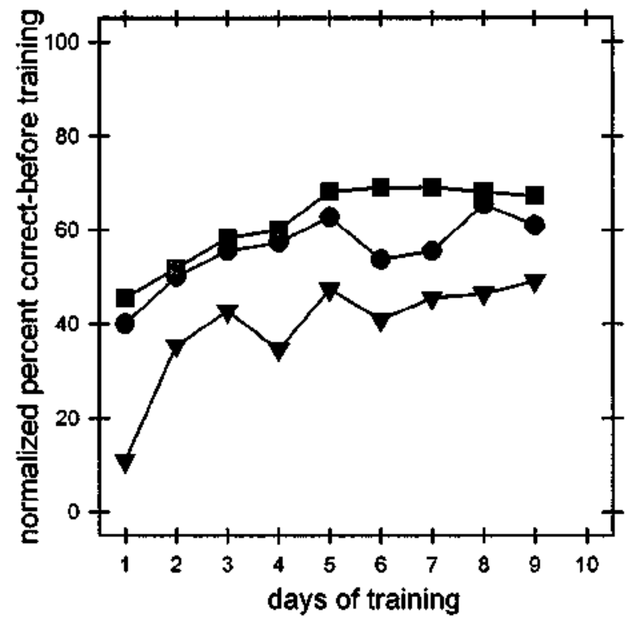
At each session the vowel recognition scores were obtained with matching, +5 mm compression, and +5 mm baseline compression condition both before and after training. In addition a set of 20 TIMIT sentences was tested with matched condition and another set of 20 sentences was tested with +5 mm compression after the training.

Figure 6.5 shows the vowel recognition percent scores before the training sessions. The triangles are the scores with +5 mm compression, whereas circles are scores from matching and squares from the baseline

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conditions. As expected all scores improve the first few days reaching an asymptotic value afterwards. Compression again gives poorer speech recognition compared to matching and matching curve is always between compression and baseline curves. However, when we consider the same test repeated right after the training session of 30 min., as shown in Figure 6.6, the compression scores improved immediately to the same values as matching. When the same subject was tested next day before the training session there was no visible effect of that training from previous day anymore. She was also tested with TIMIT sentences to see if this kind of training with phonemes would improve sentence recognition at all. The triangles in Figure 6.7 show the scores from compression and circles from matching. She consistently performed better with matching even right after training except for one day, which might have been caused by the uneven difficulty levels of the sentence sets.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Figures 6.5: Vowel recognition percent correct scores for a 8-channel processor at the simulated 20 mm insertion depth, as a function of the days. All scores are obtained from one subject before the training session and normalized for chance. The squares represent the +5 mm baseline compression condition, circles denote the matched condition, and triangles are scores from +5 mm compression condition.

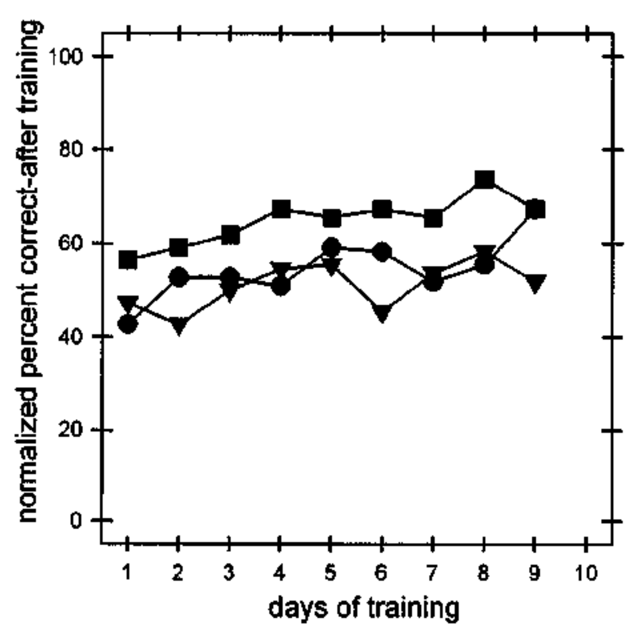


Figure 6.6: Vowel recognition percent correct scores for the same subject and the same processor when tested after the training session.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

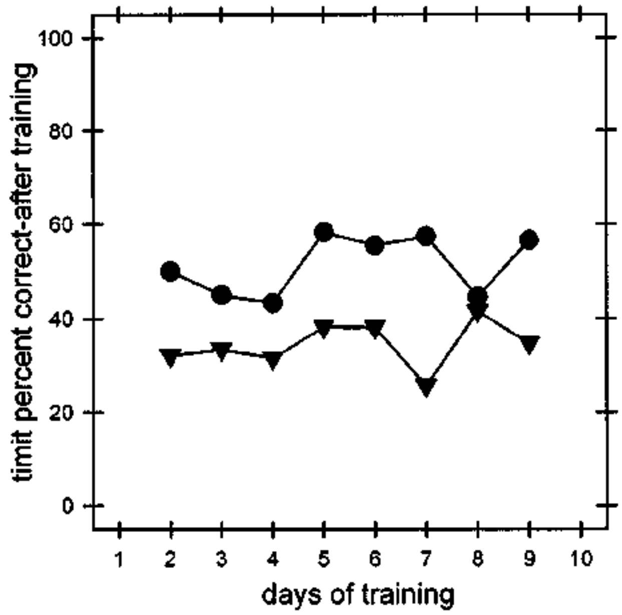


Figure 6.7: TIMIT sentence recognition percent correct scores for 8-band processors at 20 mm simulated insertion depth, as a function of the days of training provided, collected at the end of training session. The circles denote the performance with matched condition and the triangles show the scores for the +5 mm compression condition.

It is difficult to answer any of the questions asked above with these results. There is only a slight effect of training observed right after the training session which disappears next day. Training subject on vowels by providing a visual feedback only did not affect the subject recognition performance at all. Even though this pilot data implies that it is not easy to adapt to compression it still does not answer if a patient could adapt if she was using it extensively, i.e. during daily life at an extended period.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 7

FREQUENCY-PLACE COMPRESSION AND EXPANSION WITH IMPLANTS

In the previous chapters the effects of frequency-place compression and expansion on speech recognition were explored using acoustic simulations. Phoneme and sentence recognition were measured as a function of this mapping in terms of cochlear distance. These conditions were presented to normal-hearing listeners using a noise-band vocoder, simulating cochlear implant electrodes with different insertion depths and different number of electrode channels. The cochlear tonotopic range was held constant by employing the same noise carrier bands for each condition, while the analysis frequency range was either compressed or expanded relative to the carrier frequency range. Speech recognition in the matched condition was generally better than any frequency-place expansion and compression condition, even when the matched condition eliminated a considerable amount of acoustic information.

In this chapter we report results of similar conditions from implant patients. Six Med-El Combi 40+ users and one Clarion II user participated. Most conditions were run with Med-El users and a few conditions were repeated with the Clarion user. In this chapter these experiments will be described and results from implant users will be presented.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.1. IMPLANT SYSTEMS

7.1.1 Med-El Combi 40+ Implant System

The main reason for the choice of the Combi 40+ system was the flexibility of the standard fitting software of this device. At the time when the study was conducted other fitting software used in clinics did not allow the experimenter to vary the bandpass filter cut-off frequencies freely. Additionally the long electrode array (12 electrodes spaced 2.4 mm apart covering 26.4 mm in cochlea) and the deep insertion of the array (upto 31 mm) gave us the opportunity to test the effects of such parameters as insertion depth or different stimulation regions of the cochlea on speech perception.

Combi 40+ can be used with 2 types of processors:

1. CISPRO: CISPRO is a body-worn processor. The maximum range of the frequencies it can deliver to the implant is 200-5.5k Hz.

2. TEMPO+: TEMPO+ is the behind-the-ear (BTE) processor and comes with different wearing options which makes it more practical compared to body-worn processor. It can process frequencies from 200-8.5k Hz.

Even though the fitting software gives a wide range of choices for many fitting parameters as well as the frequency-to-electrode allocation, many audiologists do not make use of this flexibility due to time limitations as well as the uncertainty about how each parameter would affect performance of the patient.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.1.2 Clarion II Implant System

In this study we also used Clarion II system with High Focus II electrodes connected to a platinum processor. This processor gives the audiologist flexibility in many parameters, such as the stimulation mode, processing strategy, and a very high stimulation pulse rate. Yet the clinical software does not allow changes in the frequency range and for that purpose a special version of the clinical software was used which can read the filter coefficients from files generated by Matlab. Also, the whole electrode array in HF II is 12 mm long, a length that enables good speech recognition, but does not give as much flexibility to choose a variety of electrode configurations. The subject in this study had his electrodes implanted with a positioner which kept the electrodes close to the inner wall of the cochlea.

7.2. EXPERIMENTAL METHOD

7.2.1. Subjects

Six Combi 40+ users, aged 25-62, and one Clarion II user, aged 50, participated in experiments. All were reported to have full electrode insertion at surgery. All patients had been using hearing aids until their implant surgeries. Most of them did not use their hearing aids after getting an implant, except emergencies, and all described hearing with the hearing aid very low-quality compared to the implant. Detailed information about subjects is summarized in Table 7.1. Subject pseudonyms with M are for Med-El Combi users and with A is for Advanced Bionics Clarion user.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Only subjects M3, M4, and A1 are really postlingually deafened (becoming deaf after acquiring language). M1 became deaf as a child and all other subjects were prelingually deafened. All subjects have used oral communication as their main communication mode, while M2 and M5 have also used sign language frequently. All prelingually deaf and perilingually deaf subjects were born into hearing families As a result of this, they always preferred oral communication and also had been provided with speech correction therapies for long periods of time. All patients can converse over the telephone with their implants except M5 and M6 but, despite the difficulties, even they like using their phones extensively.

The baseline sentence scores given in Table 7.1 are for IEEE sentences which were too hard for subjects M5 and M6. So, they were retested with simpler sentences and were allowed to repeat as many times as they needed. Even with simpler materials their open-set scores were too low to allow observe any effects of spectral manipulations. They did not participate further after the first experiment.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subject	age	duration of profound deafness (years)-reason of deafness	experience with CI (years)	baseline vowel score (corrected for chance)	baseline consonant score (corrected for chance)	baseline sentence score (IEEE)
M1	39	30- high fever	2.5	60.00	55.26	38.22
M2	25	from birth- unknown	5	70.00	85.91	84.52
M3	62	12- noise exposure	1	68.18	70.18	92.81
M4	46	26 - unknown	2	82.50	86.67	93.94
M5	36	from birth- pregnancy rubella	3 total, 1 year with replacement	42.73	30.71	17.5* (HINT)
M6	40	from birth- unknown	4.5	44.55	52.63	12.8* (HINT)
A1	50	20 right ear, 9 left ear- unknown	5 total, 2.5 years with replacement	71.82	93.86	90.87

Table 7.1: Information about implant users. The sentence scores with asterisks for M4 and M5 were obtained with simple sentences because the patients could not identify any single word with standard sentences, and even with simple set they could obtain these scores only if they were allowed to repeat as much as they needed. Their baseline scores with 6 channels were also very low not leaving enough space to further deteriorate and observe the full effect of compression and expansion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.2.2. Hardware and Fitting System

The implants used in the experiments both had new generation electrodes and processors. The newer designs for electrodes allow deeper insertions and closer positioning of the electrode array to the spiral ganglia than before, which also gives a good opportunity for research because more experimental conditions can be created. They can also both go high rates of stimulation (18k pps for both Combi and Clarion). Yet in this study we were not interested in high rates, but instead focused on providing a good range of comfort level where the patient can have a good loudness growth for each electrode. The maximum rate for each electrode was usually limited to 1k-2k pps per electrode.

Med-El Combi 40+ Implant System:

A CISPRO and a TEMPO+ processor, fitting software and the fitting box were provided by Med-El for use in the experiments. BTE processor was used in the experiments for its wider range of frequencies. Because this research processor was programmed for every experimental condition the patient's maps given by the clinic were not changed at any given time. The only time patient's own processor was used was to asses her daily-life speech recognition skills with her own settings and also to train her for the speech tests when she first joined the study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABC Clarion II Implant System:

A Platinum processor was used in the experiments. A special version of the fitting program was provided by Advanced Bionics to create user defined bandpass filters.

The bandwidths for the bandpass filters were determined by Greenwood mapping function as in the simulations. In every condition the compressed or expanded cochlear distance in mm was divided by (number of electrodes-1) to create equally separated stimulation points. Then these stimulation distances in mm were converted to frequency in Hz with Greenwood equation. That same separation was also used as the bandwidth for each electrode. As a result the whole analysis frequency range was partitioned such that there were no spectral gaps or overlaps between adjacent frequency bands at -3 dB cutoff points.

7.2.3. Fitting Parameters

Even though this study focuses on the effects of the frequency-electrode mapping, there are many other parameters for CI users that have to be taken into account as well for comfortable settings. Every patient has different threshold and maximum loudness levels for each channel. To make the maximum use of the incoming acoustic information, the device settings should be set to these minimum and maximum levels, such that the patient can hear from the softest sounds to the loudest without any discomfort. In

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

addition patients have very different preferences. To keep the audibility at maximum as well as to ensure the patient's listening conditions are most comfortable, just as in the clinical setting, every patient is fitted with an optimal map before the test by carefully considering every parameter offered by the device and the fitting program. The relevant adjustable parameters are as follows:

IDR:

Input dynamic range (IDR) the intensity range of the acoustical signal coming to the microphone. Normal hearing people can hear sounds from 0 dB to 120 dB due to the appropriate compression and loudness growth mechanisms in the ear. When the nerves are stimulated with electrical current, the loudness growth is expansive. As a result the intensity of the sound can reach uncomfortable loudness levels much faster compared to normal hearing, limiting the useful dynamic range in electrical hearing to a much smaller range. The wide dynamic range of the acoustic input has to be compressed by the processor to match this much smaller electric dynamic range of the patient before it is sent to the electrodes for stimulation. This compression is achieved with an automatic gain controller (AGC) and IDR determines the range of the acoustic signal that will be fed into AGC. It was shown by Zeng *et al.* (2002) that important temporal fluctuations of multi-talker phonemes, which are similar to the stimuli used in this study, cover an amplitude range of up to 50 dB, and optimum settings for the IDR are found to be around 50-60 dB. The phoneme recognition

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

results with Clarion users confirmed this hypothesis. Cosendai *et al.* (2001) found 45 dB as the optimum setting for IDR.

IDR is fixed with Combi 40+ to a range of 60 dB but this range adaptively changes depending on the input trying to capture the most of this acoustic input. The fitting program for Clarion offers an option between 10 dB and 80 dB for IDR setting. For this study, taking the studies mentioned above as reference, 50 dB was used.

Sensitivity and Volume:

The wide dynamic range of the acoustic input is compressed by AGC inside the processor before it is sent to the electrodes for stimulation. AGC has one fast mechanism with a short time constant to compress sudden loud sounds in addition to the slow mechanism that works for slowly varying input such as speech. The time constants of these mechanisms are set by the companies and cannot be changed by the user. However, the compression ratio of AGC can be changed with the sensitivity knob of the processors. Changing the sensitivity is different than changing the volume. Volume equally turns up or down all incoming sounds. Sensitivity turns up or down the volumes of incoming sounds selectively. At low sensitivity level the compression is highest and as a result quieter sounds are suppressed more compared to louder sounds. Such sensitivity level can be useful when there is much background noise. On the other hand, with highest sensitivity the compression is at minimum, and all sounds are audible. In the experiments the subjects listen to the stimuli in the soundproof booth and therefore there

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is no danger of any background noise. Yet if IDR and AGC compression do not match, some sounds can be clipped. To prevent such distortions we made sure that sensitivity and volume settings do not distort the incoming sounds but still comfortable for the patient.

Stimulation Rate:

Stimulation rate was kept at a comfortable level for the user, usually around 1-2k pps per channel. This rate was high enough to have an appropriate pulse phase duration to obtain an efficient loudness growth at each channel. Both devices can go to very high rates but in this study we did not use any device at stimulation rates above 2 kpps/electrode.

Speech Processing Strategy:

Both TEMPO+ and Platinum processors were used with CIS only.

Stimulation Mode:

Combi 40+ is used with monopolar stimulation only. Clarion can have different electrode configurations such as pairwise or quadruple electrode stimulation for high-rate. In this study it was programmed for monopolar only as well.

The effects of other parameters on speech recognition are not relevant for present study and they were left at the standard values of the fitting software.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.2.4. Stimuli

In the study, Hillenbrand vowels were used for the vowel recognition test, similar to simulations.

For consonant recognition test we used consonants recorded by Shannon *et al.* (1999) at 44.1 kHz sampling rate. Six presentations (3 male and 3 female talkers) were made of 20 medial consonants /b tʃ d ð f g j k l m n p r s ʃ t v w y z/, presented in an /a/-consonant-/a/ context. This is actually a harder task than Fu Turner consonants because these recordings have more phonemes spoken by a larger number of speakers. Chance performance level for this test was 5%correct, and the single-tailed 95% confidence level was 8.27% correct based on a binomial distribution. Tokens were presented in random order by custom software (Robert, 1998), and subjects were instructed to select the phoneme they heard from a set of phonemes displayed on the screen. The map the subject was using was changed before every test.

For sentences two different sets were used: Harvard sentences (IEEE, 1969) spoken by a single male talker and HINT sentences (Nilsson et al, 1994) spoken by multiple talkers. IEEE sentences are phonetically balanced across lists and the predictability of the words is relatively low. To obtain the effects of the mismatch conditions fully patients who scored relatively low with IEEE sentences were retested with HINT sentences. HINT sentences are easier contextually and they are more similar to daily life natural speech. The also have fewer key words compared to Harvard sentences. HINT sentences are phonetically balanced across lists as well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Every list of sentences consists of ten sentences and implant users listened to two sets for each condition which were presented in random order. Attention was paid not to present the same sets to different subjects for the same condition.

7.3. RESULTS

7.3.1. EXPERIMENT 1: Shift in Frequency-Electrode Map to Confirm the Accuracy of the Assumed Insertion Depth

A significant difference between simulations and real implants is that in the simulations all parameters are controllable whereas there are many uncertainties in the implants. One major problem is the uncertainty in the exact location of the electrode array both in terms of insertion depth as well as the proximity to the spiral ganglia. Also we do not know how accurate the Greenwood mapping function (which characterizes the frequency sensitivity along the organ of corti) is because we do not know where the actual stimulation occurs. Variations in the lengths of individual cochleae contribute to the uncertainty as well (Ulehlova *et al.*, 1987). All these factors make it virtually impossible to accomplish a very precise frequency-place matching (unlike the simulations) and for that reason we have to make some assumptions. All we know is that all these patients have full insertion without any surgical complications as well as a range of possible insertion depths for the electrode array from company resources and some imaging studies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because Combi 40+ has 12 electrodes separated by 2.4 mm we can safely assume that the insertion is deeper than 26.4 mm, the total length of the whole electrode array. The widest range of frequencies that TEMPO+ processor can process is from 200 Hz to 8.5 kHz which corresponds to cochlear distances of 29.67 mm and 6.36 mm from the round window respectively (by Greenwood mapping function assuming a 35 mm length for the cochlea). To make use of this wide range of frequencies we assumed that the whole array was located between approximately 5 and 31 mm from the round window. This region covers a spectral range from 150 Hz to 10 kHz in the cochlea according to Greenwood equation. The stimulation region covered by the electrodes in the cochlea relates to the carrier bands in the simulations. We chose the middle 6 electrodes (electrodes 4-9) so that we can have enough space for both compression and expansion condition. The signal that will go through the bandpass filters of these electrodes comprises the analysis bands (such as in the simulations). With these assumptions these 6 middle electrodes are located from 12 mm and 24 mm in the cochlea with a total length of 12 mm. The center frequencies of the electrodes extent from 620 Hz and 3.88 kHz according to Greenwood mapping function. Because the electrodes are 2.4 mm apart, each electrode has a bandwidth that corresponds to this distance of 2.4 mm at that particular location. As a result, including the bandwidths of all electrodes, the entire analysis band range of the electrode array covers a cochlear distance from 10.8 mm to 25.2 mm with a total range of stimulation of 14.4 mm, and delivers the acoustic information of the frequency range of 495-4538 Hz.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

To confirm that our assumption for the insertion depth is actually reasonable we first took the analysis range of 620 Hz-3.88 kHz and passed this band-limited speech through 6 bandpass filters, such that the most apical filter's center frequency is 620 Hz and the most basal center frequency is 3.88 kHz. We assigned the band-pass filtered speech to a different set of electrodes for each condition as shown in Table 7.2. Hence for every condition the speech is processed the same way and is assigned to the same number of electrodes (6) that are located at different distances from the round window in the cochlea starting from apical end and moving to basal end. As a result, from condition 1 to 7 the cochlear location is shifted basally.

condition (most apical electrode)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
electrode array	1-6	2-7	3-8	4-9	5-10	6-11	7-12
center frequency range	620-3.88k	620-3.88k	620-3.88k	620-3.88k	620-3.88k	620-3.88k	620-3.88k

Table 7.2: Basal shift conditions for Combi 40+, as a function of the most apical electrode number.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

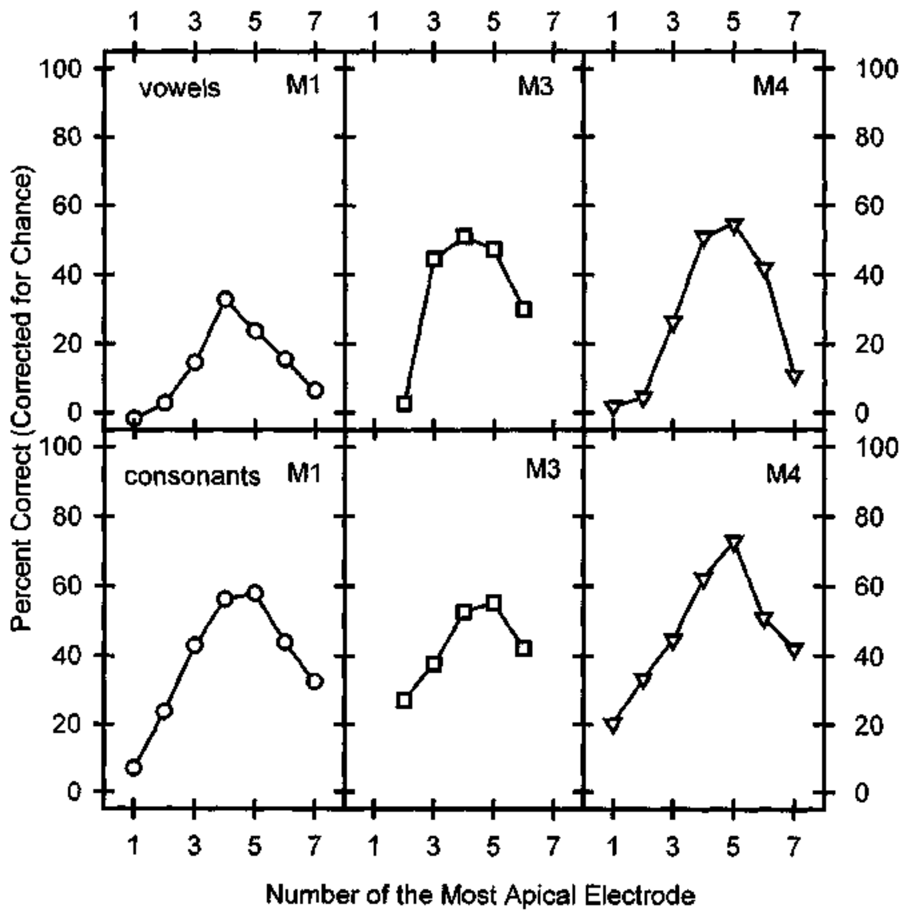


Figure 7.1: Basal shift scores from subjects M1, M3, and M4.

The percent correct scores from vowel and consonant recognition tests corrected for chance level [p = 100*(score-chance)/(100-chance)] with subjects M1, M3, and M4 are shown in Figure 7.1. We see peak vowel recognition performance around condition 4 for M1 and M3 (closest condition to our assumption of 31.2 mm insertion). For all consonants and vowels with M4, the highest scores were observed around condition 5. At this condition the electrode array is one electrode separation length shallower than the assumed 31.2 mm, i.e., 31-2.4=28.8 mm. These perceptual results imply that the insertion is somewhere between 29-31 mm. From simulations and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

other similar spectral shift experiments we know that CNS has a tolerance of a few mm to such spectral manipulations. To keep the uniformity among subjects as well as to keep the parameters similar to simulations we kept 31 mm as our insertion depth estimation even though in reality the insertion may be a few mms shorter. This should not affect the results too much due to the tolerance of the system.

7.3.2. EXPERIMENT 2: Effect of the Length of the Cochlea

When using the Greenwood frequency-place mapping function we always assumed that a cochlea is 35 mm long on average. The data collected by Ulehlova *et al.* (1987) showed an average length of 34.2 mm of the cochleae of 28 men. The extreme values were 28 mm and 40 mm. Ketten *et al.* (1998) estimated cochlear lengths of implant patients from CT scans and found an average length of 33 mm of 20 patients, with extreme values of 29 mm and 37.5 mm. In this experiment the assumed cochlear length was made shorter (-1mm, -2 mm, -3 mm, -4 mm shorter) or longer (+1 mm, +2 mm, +3 mm, +4 mm longer) than 35 mm and then the matching frequency range was calculated with Greenwood function. The change in the performance of the same subjects from the previous experiment when tested with a matched map determined by these changing cochlear lengths can be seen in Figure 7.2. The relatively sharp peaks in vowel recognition scores of M1 and M3 imply that these subjects might have cochleae that are 1-2 mm shorter than assumed 35 mm. M4 was not affected as much; her vowel test scores stayed the same for all lengths from 32 mm to 36 mm.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

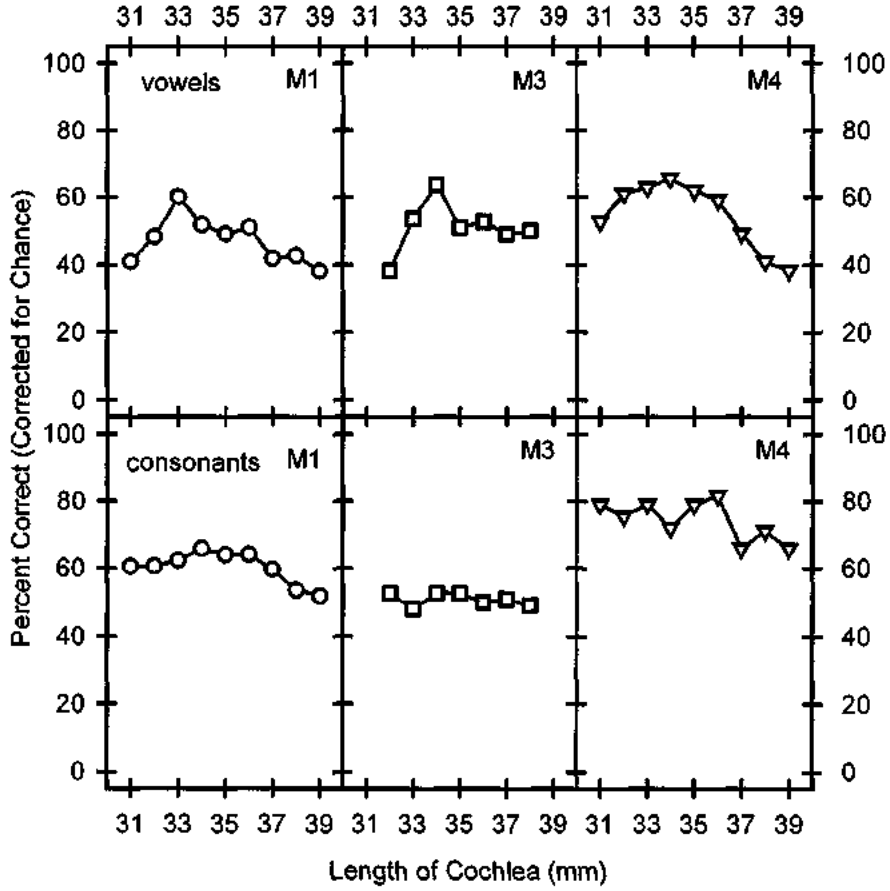


Figure 7.2: Effect of the length of the cochlea on frequency-place matching.

7.3.3. EXPERIMENT 3: Frequency-Place Compression and Expansion with 6 Middle Electrodes

7.3.3.1. Frequency-Place Compression and Expansion with Med-El Combi 40+

With the assumed insertion depth of 31 mm the array of 6 middle electrodes spans a cochlear range from 12 mm to 24 mm from the round window, and stimulates the region from 10.8 mm to 25.2 mm. This range was specifically chosen to imitate the conditions in the simulations so that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the data would be comparable. The speech spectrum range assigned on to these electrodes was made wider by +1 mm, +2 mm, +3 mm, +4 mm cochlear length at both ends for compression, and it was reduced by -1 mm, -2 mm, -3 mm, -4 mm cochlear length at both ends for expansion conditions. The corresponding frequencies for these conditions are given in Table 7.3.

frequency-place mismatch condition	**range of acoustic input (mm)**	**band-pass filter center frequencies for 6 channels (Hz)**						**frequency range of analysis bands (Hz)**
-4 mm (expansion)	20 – 16	1171	1323	1495	1687	1901	2141	1025 - 2367
-3 mm (expansion)	21 – 15	1003	1205	1450	1742	2078	2485	887 - 2762
-2 mm (expansion)	22 – 14	862	1106	1416	1801	2281	2882	752 - 3196
-1 mm (expansion)	23 – 13	728	1010	1373	1863	2506	3350	611 - 3847
0 mm (matching)	24 – 12	621	921	1342	1927	2740	3876	493 - 4522
+1 mm (compression)	25 – 11	523	839	1311	1996	3006	4419	394 - 5467
+2 mm (compression)	26 – 10	442	769	1278	2084	3297	5353	314 - 6305
+3 mm (compression)	27 – 9	363	700	1248	2145	3620	6242	247 - 7482
+4 mm (compression)	28 – 8	313	640	1220	2226	3973	6792	207 - 8082

Table 7.3: Frequency-place mismatch conditions with 6 middle electrodes (electrodes 4-9 located from 12-24 mm).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 7.3 shows the individual percent correct scores from all six Med-El patients tested on frequency-place compression and expansion conditions using the 6 middle electrodes as described above. Individual vowel recognition scores are plotted in the top row, consonant scores are in the middle row, and sentence scores are in the bottom row. The percent correct scores are corrected for chance for phonemes. M1 was tested twice with vowels and consonants (open and filled circles in the left top and left middle panels). M1 and M3 were tested with Harvard sentences (open symbols in the bottom row) as well as HINT sentences (filled symbols). As explained in Section 7.2.1, M5 and M6 were both prelingually deaf and could not acquire any meaningful score with open-set speech recognition tests. Their phoneme recognition scores are close to chance. They were not tested with sentences and did not participate in succeeding experiments.

The average scores from all Med-El patients (as shown with thick lines) are presented in Figure 7.4 on top of the individual scores (as shown with symbols and thin lines). The left panel shows the scores from vowel recognition, middle from consonant recognition, and the right panel from sentence recognition tests. The dashed line in the sentence recognition results shows the average score from patients M1 and M3 with HINT sentences (as shown with filled symbols in the bottom row of Figure 7.3). The solid line in the same panel shows the average score with IEEE sentences from patients M1, M2, M3, and M4 (as shown with open symbols in the bottom row of Figure 7.3).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

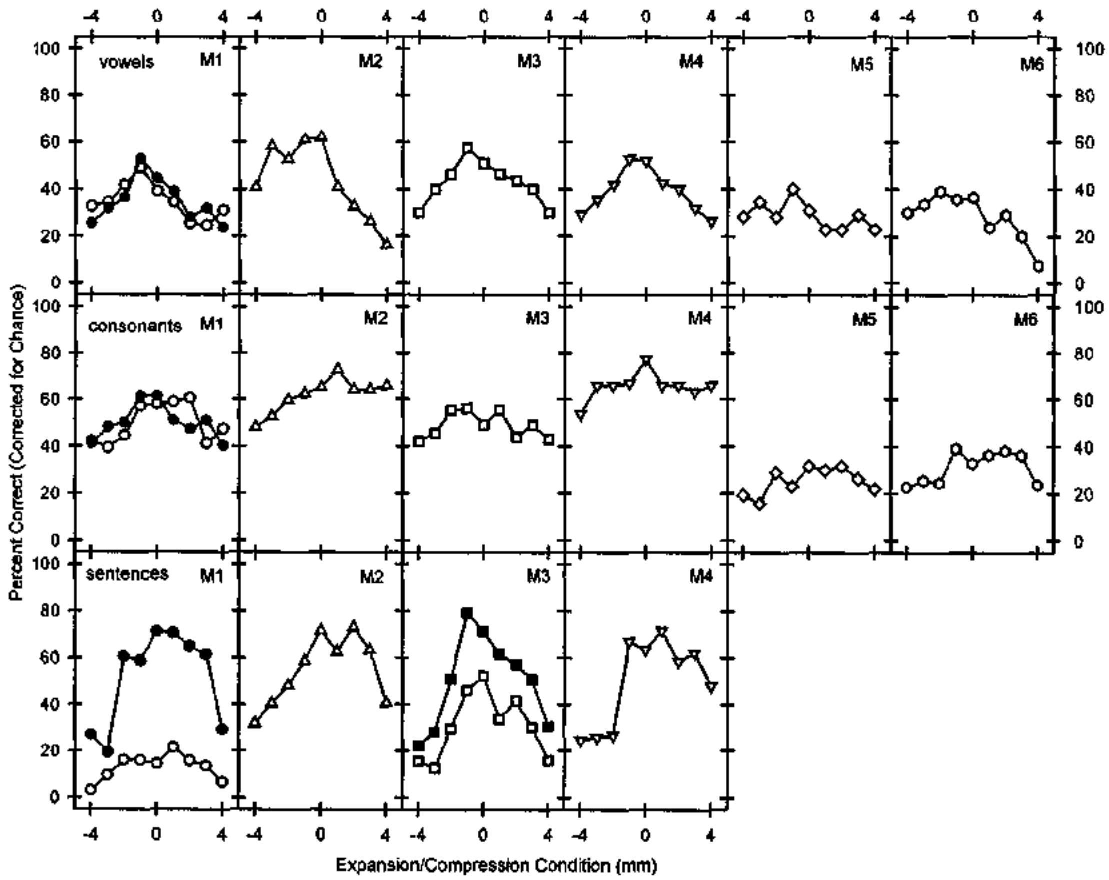


Figure 7.3: Individual percent correct scores of Med-El patients as a function of compression and expansion in frequency-place mapping. The top row shows the vowel recognition scores, the middle row shoes the consonant recognition scores, and the bottom row shows the sentence recognition scores. M1 was tested twice with vowels and consonants. M1 and M3 were tested with both IEEE and HINT sentences.

A one-way repeated-measures ANOVA was applied to results and corresponding F and p values are given in Table 7.4.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

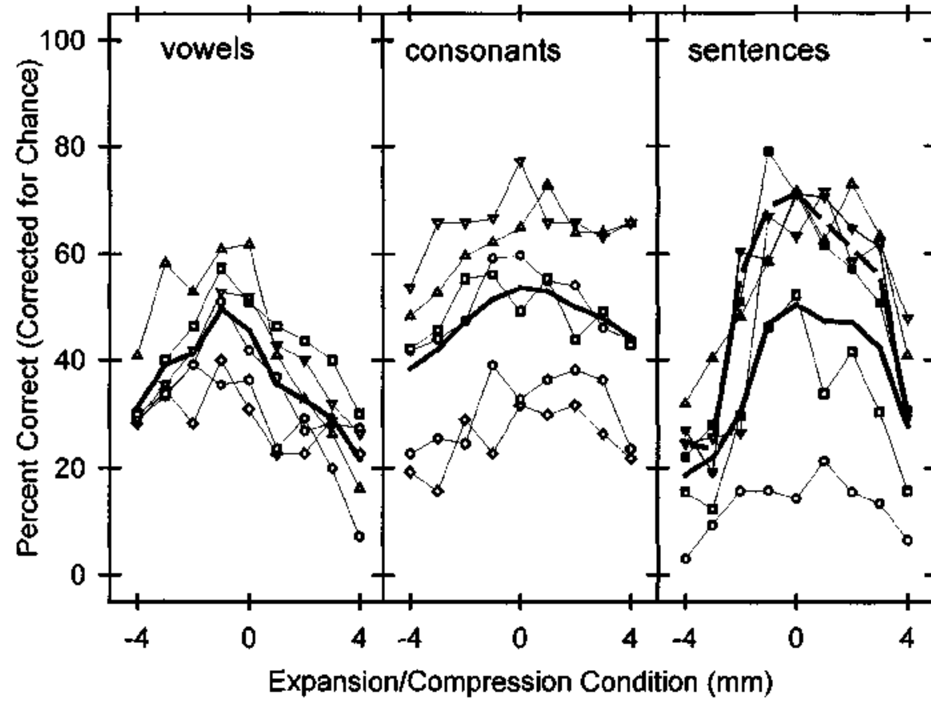


Figure 7.4: Average percent correct scores of six Med-El Combi 40+ users, shown with thick lines, superimposed on the individual scores from Figure 7.3, shown with symbols and thin lines.

expansion	F	p	compression	F	p
vowels **n=6** **F(4,20)**	13.65	<0.001	**vowels** **n=6** **F(4,20)**	13.41	<0.001
consonants **n=6** **F(4,20)**	15.89	<0.001	**consonants** **n=6** **F(4,20)**	5.11	<0.01
IEEE sentences **n=4** **F(4,12)**	10.63	<0.001	**IEEE sentences** **n=4** **F(4,12)**	7.61	<0.01
HINT sentences **n=2** **F(4,4)**	15.52	<0.01	**HINT sentences** **n=2** **F(4,4)**	34.14	<0.01

Table 7.4: One-way repeated-measures ANOVA analysis results for frequency-place expansion/compression percent correct scores of implant subjects.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For all stimuli there was a significant effect of compression and expansion as shown in Table 7.4. For vowels, a post-hoc Tukey test revealed that -3 mm, -2 mm, and -1 mm expansion scores were not significantly different than 0 mm matching. A tolerance region of a few mm was also observed in simulations, yet this tolerance was for a few mm of expansion in comparison to the tolerance to a few mm compression in simulations. In simulations, the effect of expansion on vowel recognition was worse than compression, whereas with Med-El patients a bigger drop in performance was observed with compression compared to expansion. As in the simulations, the effects of both expansion and compression were smaller on consonants than vowels. A Tukey test showed that there was no difference in results from -2 mm expansion to +3 mm compression, which is a much wider range of tolerance compared to vowels. Similar to simulations consonant recognition scores drop significantly with higher degrees of compression (+4 mm with Med-El patients) and expansion (-3 mm).

Vowel recognition scores peak around -1 mm expansion implying that a small amount of expansion might actually be beneficial for vowel recognition in contrast to consonants where a little compression seems to give the maximum performance. Sentence recognition is a good representation of daily life speech understanding and perception of vowels and consonants both contribute to sentence recognition. As a result, the best performance for sentences was obtained around 0 mm matching condition with a tolerance of ±2 mm, and the performance dropped significantly with further mismatch.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Consonant Feature Analysis:

Similar to simulations, information transmitted on the consonant features of place, manner, and voicing is plotted in Figure 7.5. The panels from left to right show the percent of information transmitted on place, manner, and voicing. Both place and manner recognition scores dropped significantly with expansion, as shown by one-way repeated-measures ANOVA in Table 7.5. The results for consonant feature analysis display a very large variation among subjects and some subjects perform better with increasing compression whereas some others had performed better around 0 mm matched condition. As a result the average performance with compression is relatively flat with a big standard deviation and it is difficult to observe a significant trend.

expansion	F(4,20)	p	compression	F(4,20)	p
place	6.15	<0.01	place	1.62	0.21
manner	10.59	<0.001	manner	0.96	0.45
voicing	2.24	0.10	voicing	0.92	0.47

Table 7.5: One-way repeated-measures ANOVA analysis results for consonant features of place, manner, and voicing.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

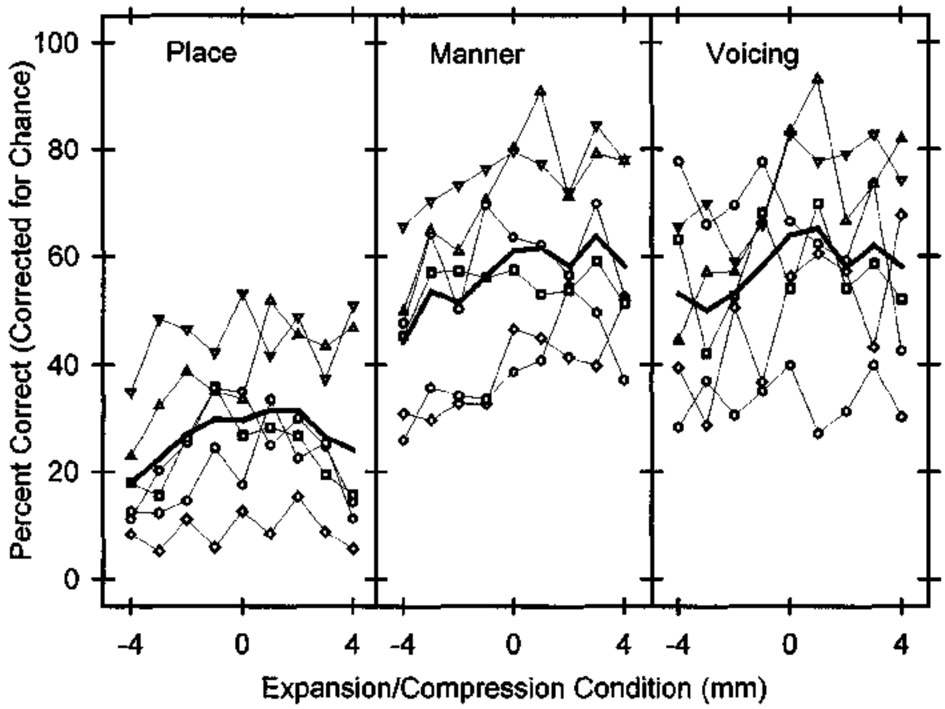


Figure 7.5: Information transmission percent scores for consonant features of six Med-El Combi 40+ users as a function of frequency-place mismatch conditions.

7.3.3.2. Frequency-Place Compression and Expansion with Clarion II

The same frequency-place compression and expansion conditions were repeated with subject A1 who is a Clarion user as shown in Table 7.1. The subject has HFII electrode array with a positioner. The HFII electrode array has 16 electrodes separated by 0.8 mm forming an array of 12 mm. He has an insertion of approximately 17 mm. Normally a 12 mm stimulation range from 5 to 17 mm from the round window would cover only 1.8-10 kHz by Greenwood mapping function. From simulations, experiments with implant users, and SII model predictions we would expect a much lower performance from A1 than he actually gets with his implant with these settings. Yet he does extremely well, probably because the positioner holds

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the electrode array right next to the inner wall of the cochlea. Greenwood mapping function holds for organ of corti, where the hair cells are located, and it is little further than this medial wall. The cochlear length is much different along the organ of corti (around 35 mm) than it is along the medial wall (around 25 mm). The differences in the radii of different trajectories inside the cochlea create this difference due to the spiral shape of the cochlea. As a result this shorter electrode array of 12 mm actually covers a wider range along the medial wall (12 mm out of 25 mm) than it would cover along the organ of corti (12 mm out of 35 mm). With this conversion we predict that a 12 mm long array at 17 mm positioned medially actually acts like a 20 mm long array inserted at 28 mm laterally in terms of the Greenwood function. With this correction this array would cover a spectral range of 300 Hz-6.8k Hz which is a more realistic range for the scores he gets with his implant.

6 electrodes out of 16 electrodes (i.e., electrodes 1, 4, 7, 10, 13, and 16) were chosen to apply frequency-place compression and expansion. Because the electrodes are already covering a wide range (20 mm) a wide range of expansion could be applied but for compression only +1 mm compression condition could be included. The percent correct scores from the single Clarion patient are shown in Figure 7.6. Again the top performance was observed around 0 mm matching with a tolerance of a few mm. This tolerance range was much wider (around 6 mm for phonemes, and around 4 mm for IEEE sentences) most probably because the stimulation range was much wider (20 mm compared to 12 mm) leaving a much bigger space for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mismatch at the ends and having more frequencies in the middle matched. It is not very clear if it is the ceiling effect or a real tolerance region around 0 mm because the subject's scores are close to 100%.

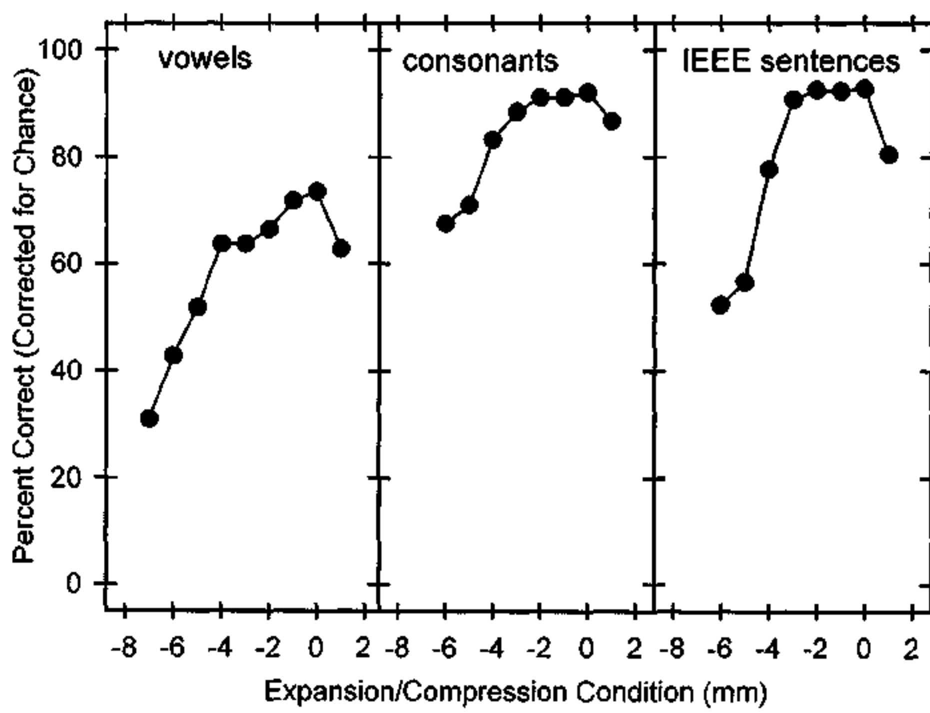


Figure 7.6: Percent correct scores from Clarion patient A1 as a function of frequency-place expansion and compression.

7.3.4. EXPERIMENT 4: Expansion with all 12 Channels of Med-El Combi 40+

As mentioned before the analysis band range of the 6 middle electrodes cover 14 mm of acoustical range which is actually slightly shorter than 16 mm used in the simulations. For this reason, as well as to make use

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the whole electrode array and see the effects with a more optimal setting, we used the widest stimulation range possible (all 12 electrodes). Because this range is already too wide (26.4 mm) covering more than the widest frequency range that the device can provide (200 Hz-8.5 kHz) we can apply only expansion. Yet it would still give an idea if the mismatch conditions deteriorate speech recognition in a better setting such as when more electrodes with a wider stimulation range and a deeper insertion are used.

As before an insertion depth of 31 mm was assumed, so the stimulation range is from 5 mm to 31 mm from the round window. The analysis range was made shorter than this cochlear range by -3 mm, -4 mm, -5 mm, -6 mm, -7 mm, and -8 mm from both ends. Figure 7.7 shows individual percent correct scores with these expansion conditions with 12 channels (open symbols) combined with results with 6 channels from Figure 7.3 (closed symbols).

Usually implant users performance increases up to 8 electrodes and then reach an asymptotic value. In the present data higher percent scores were obtained with 12 channels compared to 6 channels. In addition the 12 channel processor uses a much wider stimulation range and the electrodes are located deeper. Therefore we observe a replica of the pattern of expansion with 6 channels except with higher scores.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

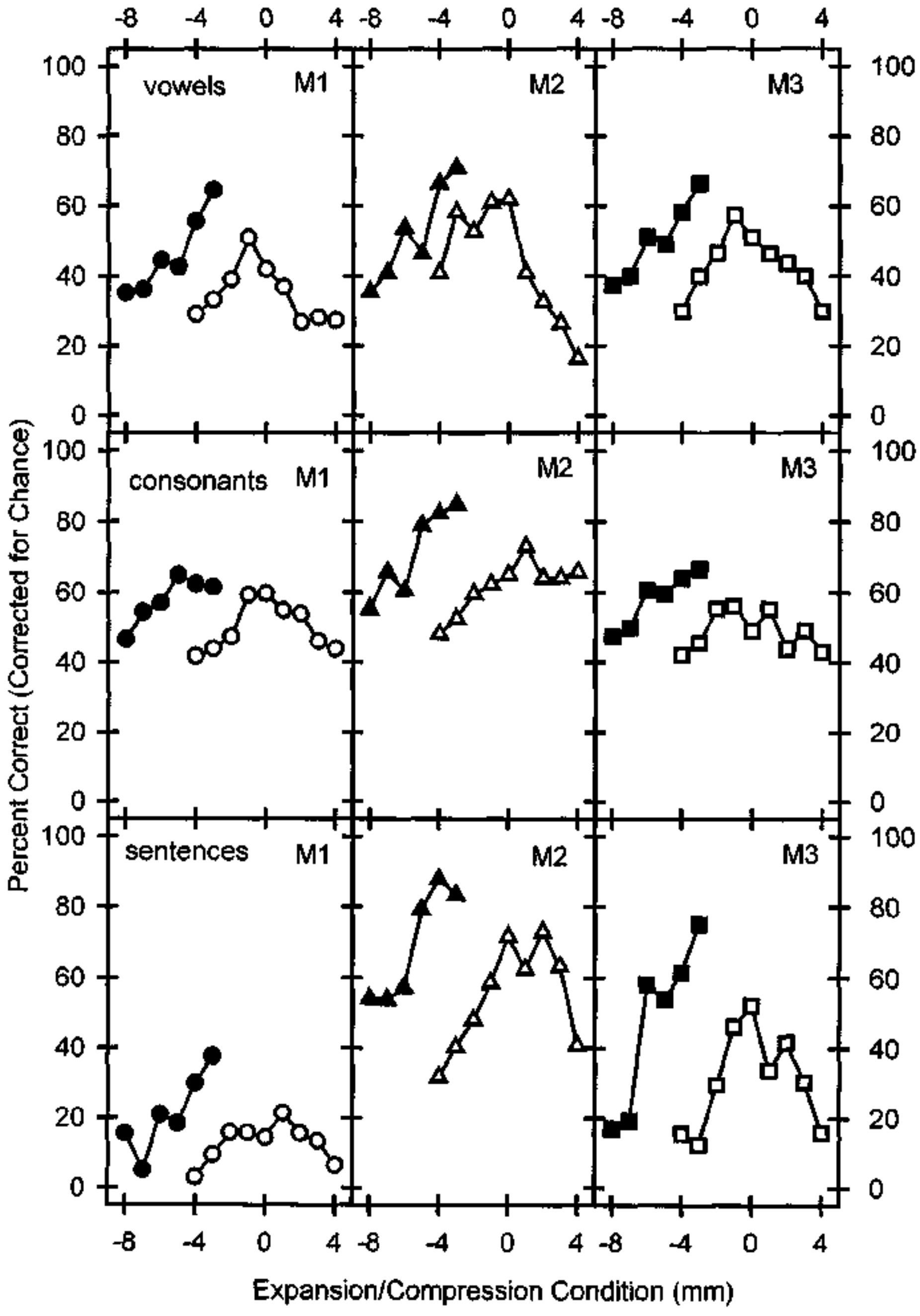


Figure 7.7: Individual percent correct scores for three implant subjects for frequency-place mismatch conditions with 6 channels (open symbols) and expansion condition with 12 channels (filled symbols).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 7.8 shows the average of the scores from Figure 7.7. The individual scores are replotted with small open symbols and the average is shown with thick lines.

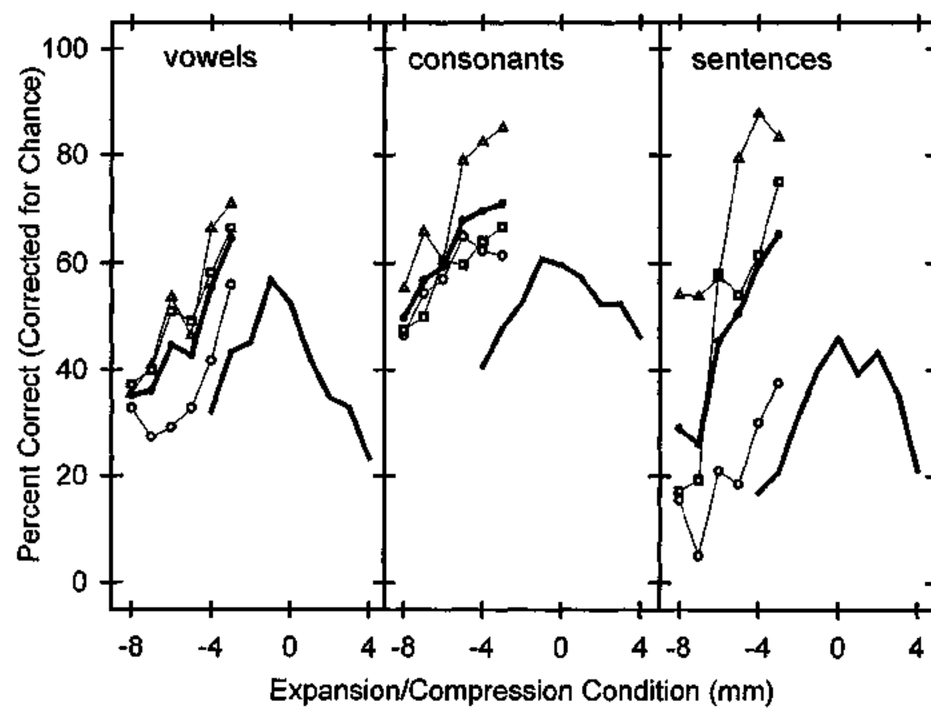


Figure 7.8: The average percent correct scores from subjects M1, M2, and M3. The small open symbols show the individual scores for 12 channel expansion conditions. The superimposing thick line is the average scores with 12 channels whereas the other line to the right of each panel is the average score of the same patients from expansion and compression with 6 middle electrodes.

A condition of special interest is -5 mm expansion with 12 channels. The body worn processor can process frequencies up to 5.5 kHz only. So even the widest range of frequencies that can be sent to electrodes is 200-5.5 kHz. Some patients had first been fit with their body worn processor. When they got the BTE processor they were simply given the same map

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from the body worn processor, which is similar to -5 mm expansion condition. As shown in Figure 7.8 they might be able to achieve vowel and sentence recognition scores that are 15-20% higher if provided with a wider range of frequency.

7.3.5. EXPERIMENT 5: Comparison with Preset Values of Med-El Tempo+

Tempo+ processor has a few preset values for low and high frequency end of the frequency range that the audiologist can easily choose with the standard fitting software. The lower frequency limit can be 200, 250, 300, or 350 Hz and the higher limit can be 3.5, 5.5, 7, or 8.5 kHz. Usually this range is divided into frequency bands by logarithmic scaling. Even though the newest version software (Studio+) is very flexible and allows the user set the bandpass filter cutoff frequencies freely audiologists typically do not alter this range because it was not shown before that assigning different frequency ranges was actually affecting the patient's speech recognition performance significantly.

In this experiment we retested subjects M1 and M2 with some of the preset values that the audiologists have been using to show that it actually makes a significant difference on phoneme recognition the way these frequencies are assigned.

Figure 7.9 shows the corrected percent correct scores of subjects M1, M2, and M4 from vowel recognition and consonant recognition tests with different preset maps of the fitting program. The same middle electrodes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(electrodes 4-9) were employed as in the previous experiment. The lines show the range of the speech recognition scores obtained with frequency-place expansion and compression in Experiment 2, i.e., the solid line shows the worst performance of the subjects and the dotted line shows the best. The x axis shows the higher end of the frequency range assigned onto these 6 middle electrodes. The open circles are the scores obtained when the lower end of the frequency was fixed at 350 Hz whereas with filled circles the low end of the frequency range was kept at 200 Hz. Especially for vowels even a small frequency difference at the apical end of the assigned frequency range makes a big difference for speech recognition performance. When the high frequency end of the assigned range was the same, both subjects performed significantly better when the low end was cut at 350 Hz (open circles) compared to 200 Hz (filled circles), i.e., adding more low frequencies (and so increasing the compression) clearly decreased the vowel recognition. It was already shown in simulations that manipulations in the low frequency end of the stimulation range can have a significant effect on speech understanding. Yet, as the open circles show, increasing the compression from higher frequency end can also decrease vowel recognition significantly, especially from 5.5 kHz to 7 kHz there is a sharp drop in performance. When the lower end is kept at 200 Hz the performance always stays at a floor level except when the high end is set to 5.5 kHz. This might be because of a better matching of some midrange frequencies. With consonants the pattern is very similar, except the changes in the performance are much smaller compared to vowels where the spectral manipulations have much bigger effect.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The dashed line is the average of performances of the same patients from Experiment 3 with 0 mm matched condition. In that condition the analysis band range matched to the electrode range was 500-4.5k Hz (from Table 7.3), which is the closest to the preset range of 350-3.5k Hz (shown by the leftmost open circle in both panels) and 350-5.5kHz (shown by the second leftmost open circle in both panels). With both vowels and consonants the best performance was obtained with these maps and they were comparable to 0 mm matching performance. The +4 mm compression condition from Experiment 2 is shown by the solid line in the figure. This condition had the analysis band range of 200-8k Hz (as shown in Table 7.3), which is the closest to the preset map range of 200-8.5k Hz (as shown with the rightmost filled circle in both panels). This preset map resulted in a much worse performance compared 350-3.5k Hz, even though it provided a much wider acoustical input frequency range.

These results show that choosing the right map for the patient is an important factor for maintaining a good speech understanding performance, especially if the patient has non-optimum settings such as smaller number of electrodes (6 electrodes) or a shorter array (12 mm). These settings were chosen on purpose to create relatively harder listening conditions to see the effects fully.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

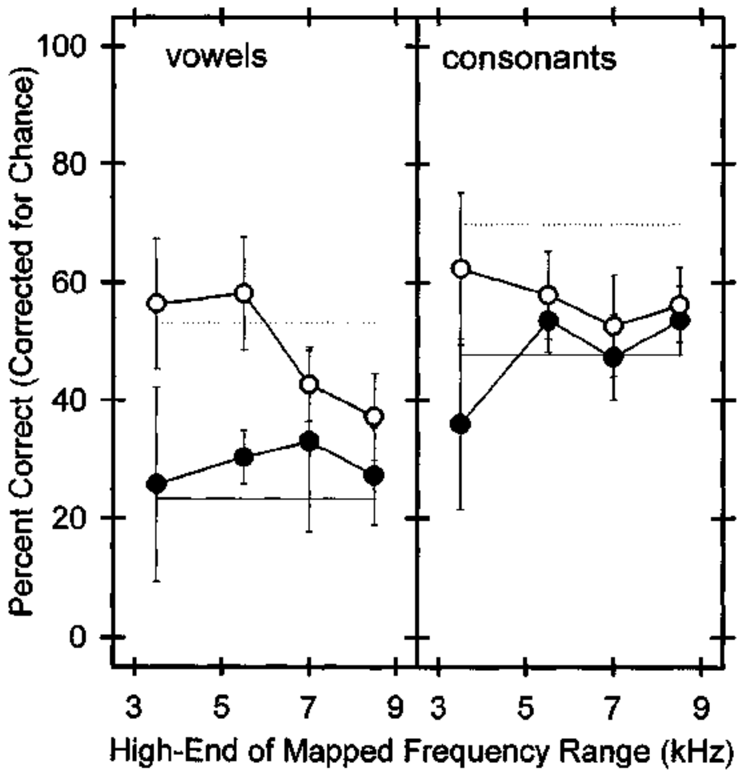


Figure 7.9: Average percent correct scores of subjects M1, M2, and M4 with Tempo+ preset maps. The high end of the frequency range assigned is given on the x axis. The lower end is 350 Hz for open circles, and 200 Hz for filled circles. The lines show the best and the worst performance of the same patients from Experiment 3.

7.3.6. EXPERIMENT 6: Mismatch on Apical or Basal End with 6 Middle Electrodes

7.3.6.1. Mismatch on Apical End with 6 Middle Electrodes of Med-El Combi 40+

As mentioned before, not all frequencies contribute equally to speech recognition. The contribution of each frequency band to speech intelligibility is determined by the band importance function of SII model (ANSI, 1997). This function peaks around 1-2k Hz meaning that the largest contribution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

comes from this range of frequencies. As a result, one would expect that it would be more crucial to match those important frequency ranges to get a good speech recognition performance. The frequency-place mismatch conditions were separated into two conditions, as in Chapter 5, to assess the relative contribution of compression/expansion from different frequency regions of the stimulation. First, the mismatch conditions were applied only on apical end while the basal end was matched (apical mismatch). Second, they were applied only at basal end while the frequencies were matched at the apical end (basal mismatch).

The percent correct scores with these distortions as well as the original scores with mismatch applied at both ends are shown in Figures 7.10 and 7.11. Implant users have a much bigger variation in the scores than normal hearing subjects compared to each other as well as in their performances from day to day. For that reason, one set of conditions for a specific experiment was always completed on the same day to minimize this daily performance variation. For example even though 0 mm matching condition is exactly the same in mismatch applied at both ends, at apical end only, and at basal end only, some subjects scored differently on different days.

The scores from Figure 7.3 are reproduced in Figure 7.10 for subjects M1, M2, and M3, in addition to the scores obtained with compression-expansion applied at apical end only. The scores when the mismatch was applied at both ends are shown with open symbols whereas scores from the apical mismatch conditions were shown with filled symbols.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

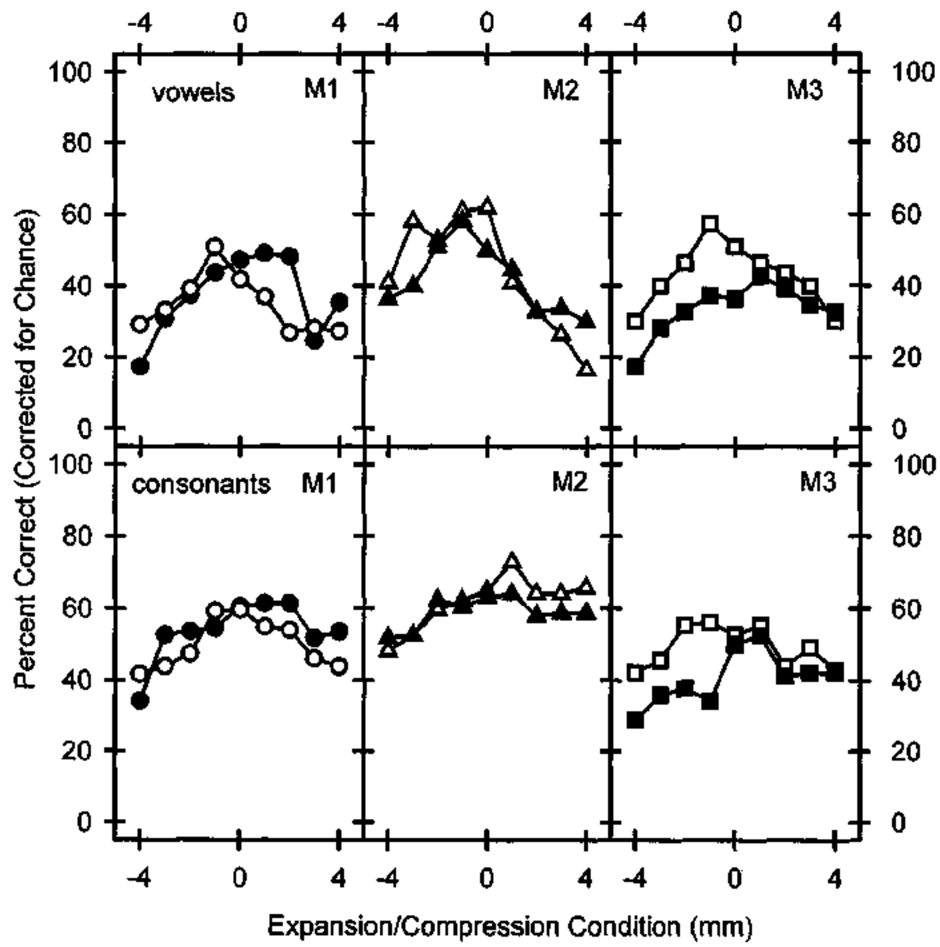


Figure 7.10: Individual percent correct scores of subjects M1, M2, and M3 with expansion/compression applied at both ends (as shown with open symbols) and applied at apical end only (filled symbols).

A large variation was observed across subjects unlike the simulations with normal hearing people. M1 had slightly better scores when the frequencies were compressed on apical end only while M3 did worse when frequencies were expanded on apical end only. M2 had slightly worse performance with apical compression with consonants. Therefore looking at individual scores it is difficult to observe a prevailing pattern in the performances.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

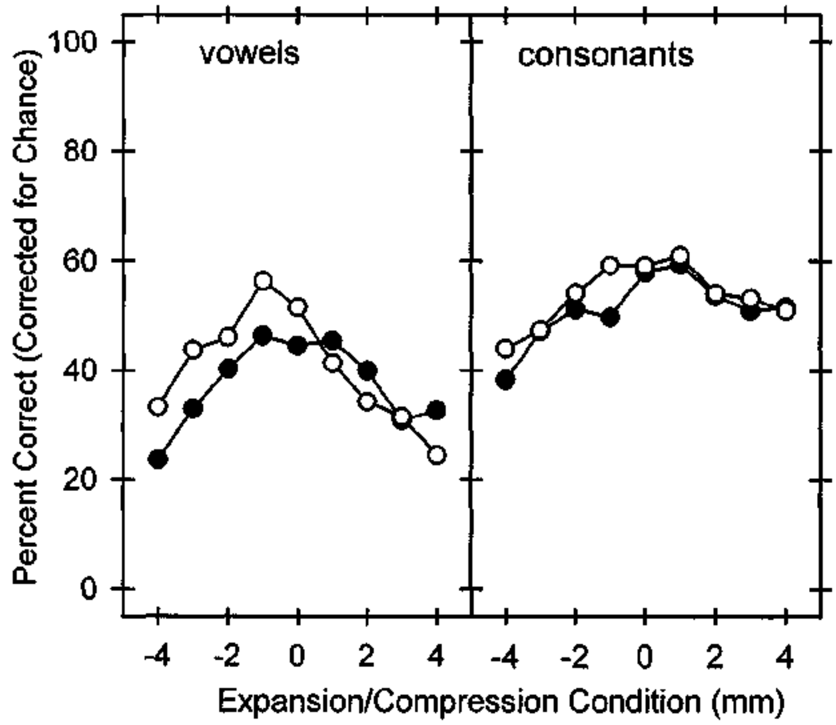


Figure 7.11: Average scores from patients M1, M2, M3. Open symbols show the same scores as Figure 7.3, where the frequency range was mismatched at both ends and filled circles show the average scores when the frequency range was mismatched on the apical end of the stimulation region.

The average percent correct scores are replotted in Figure 7.11 to explore if there was a pattern in common. The open circles denote the average scores when the mismatch was applied at both ends while the filled circles show when it was applied on apical end only. With vowels, expansion applied on apical end only seems to have resulted in lower performance compared to expansion applied at both ends. Yet it is actually difficult to tell if this is a real effect of different maps. This difference might as well be due to the inherent variations in the results of implant patients. They had scores varying by as large as 7% for the same 0 mm matched condition when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tested at different times (0 mm matched condition showed with filled and open circles in vowel recognition test in Figure 7.10). For consonants both maps resulted in almost identical performance. A paired t-test showed no significant difference between scores obtained with two maps for both vowels and consonants.

7.3.6.2. Mismatch on Basal End with 6 Middle Electrodes of Med-El Combi 40+

In simulations, we had seen that the basal mismatch conditions, where the frequency range was always expanded or compressed only at high-frequency end of stimulation region while the apical end was always matched, had much less effect on speech recognition. The high end of stimulation was already at 6 kHz for 25 mm insertion (even higher for 20 mm insertion), so a loss in such high-frequency information did not decrease the performance much.

With implants, only one patient (M1) had a performance comparable to simulations, as shown in Figure 7.12. In this figure, similar to Figure 7.10, the open circles are percent correct scores obtained when the frequencies were expanded or compressed at both ends, and the filled circles are the scores obtained when they were expanded or compressed at basal end only while the apical end was matched. As in the previous experiment, there is a large variation in results. M1 showed results similar to simulations where the scores were better with basal expansion in vowel recognition test compared to expansion applied at both ends. With consonants scores were better with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

both apical expansion and compression. Yet, especially with vowels, this pattern is not really clear. There was a difference of 10% in the scores for the same 0 mm matched condition when M1 was tested at different times, so some of the difference between two maps is probably due to this variation. M2 and M3 had almost the same scores with both maps.

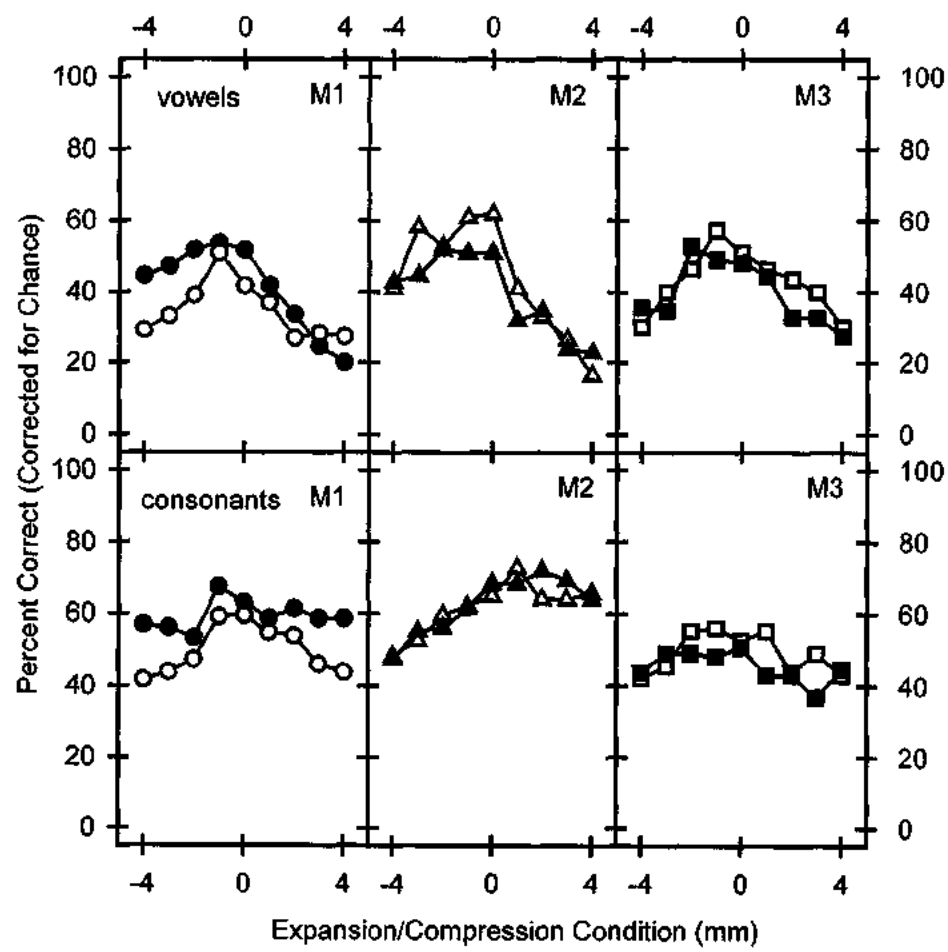


Figure 7.12: Individual percent correct scores of subjects M1, M2, and M3 with expansion/compression applied at both ends (as shown with open symbols) and applied at basal end only (filled symbols).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

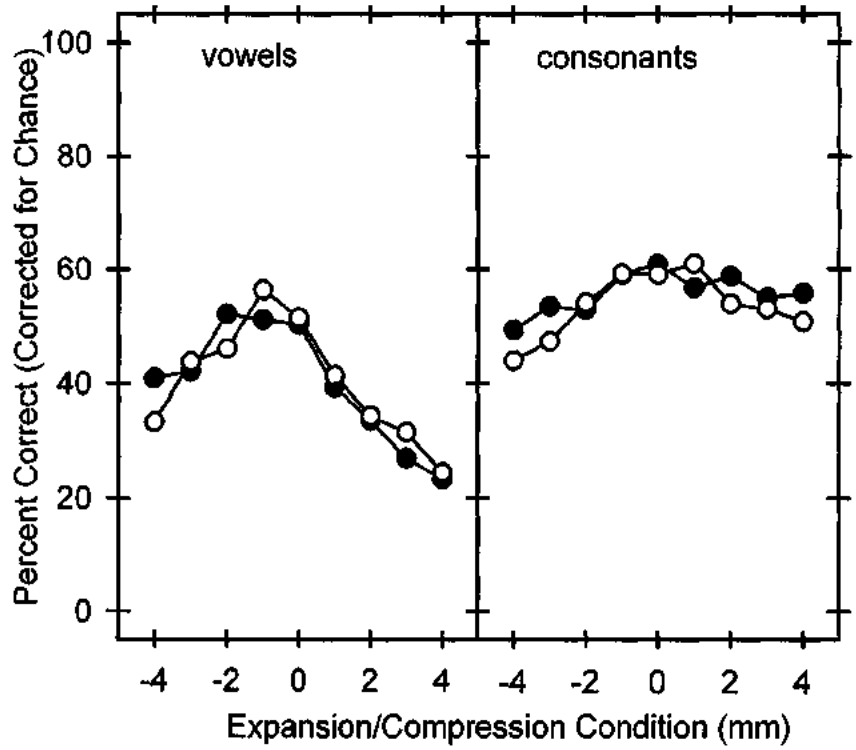


Figure 7.13: Average scores from patients M1, M2, M3. Open symbols show the same scores as Figure 7.3, where the frequency range was mismatched at both ends and filled circles show the average scores when the frequency range was mismatched on the basal end of the stimulation region.

The average data plotted in Figure 7.13 showed that there was no clear difference between the scores from two maps. A paired t test confirmed this observation. One reason that we have not seen the same effects as simulations might be that the stimulation region in this experiment is slightly different than the simulations. The high end of the stimulation range with 6 middle electrodes is 4.5 kHz which is lower than the high end of stimulation of simulations (6 kHz). As a result, the stimulation range in this experiment contains lower frequencies compared to simulations, and therefore expanding or compressing the map at the basal end only might have a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

bigger effect on speech recognition compared to the simulations. Another factor with implants is that subjects usually use a non-optimal map given in the clinic. As a result they get used to some spectral mismatch over time and some variation in the results might be due to this accommodation.

7.3.6.3. Mismatch on Apical or Basal End with 6 Electrodes of ABC Clarion II

Same setting from Experiment 1 was used for the Clarion user. Electrodes 1, 4, 7, 10, 13, 16 were activated. It was assumed that they covered a stimulation region of 20 mm from 8 mm to 28 mm from the round window. Such a wide range in the cochlea covers the frequencies from 300-6.8k Hz by Greenwood.

First the mismatch conditions were applied at apical end only while the basal end was matched. Again, because the stimulation range is already wide, mostly expansion conditions can be applied. The percent correct scores from apical mapping is shown in Figure 7.14 and the performance is very similar to when the mismatch was applied at both ends as shown in Figure 7.6. Second the mismatch conditions were applied at basal end only while the apical end was matched. The percent correct scores from this basal mapping are shown in Figure 7.15.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

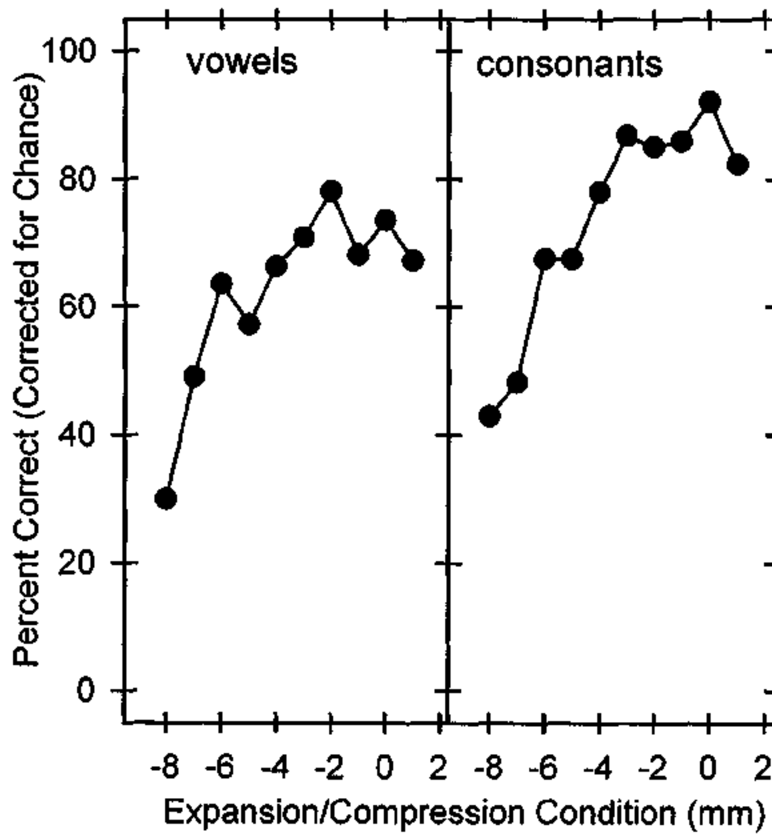


Figure 7.14: The percent correct scores normalized for chance level from Clarion patient when the frequency-place expansion and compression was applied on the apical end only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

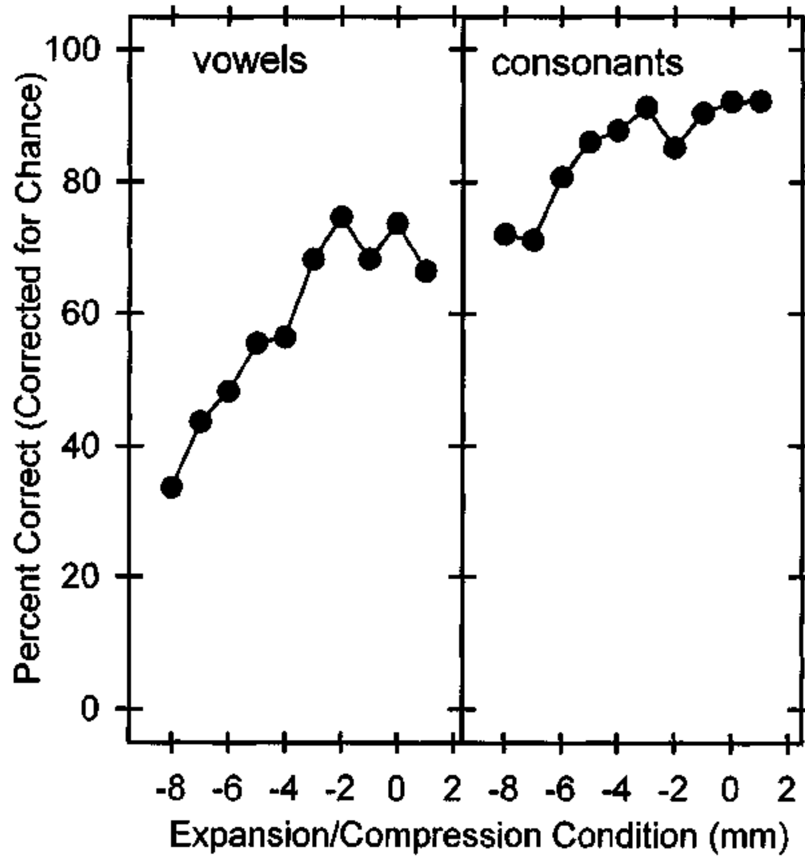


Figure 7.15: The percent correct scores normalized for chance level from Clarion patient when the frequency-place expansion and compression was applied on the basal end only.

All results from mismatch applied at both ends, at apical end only, and at basal end only are replotted in Figure 7.16, for comparison. The filled circles are from Experiment 3 when the mismatch was applied at both ends, open circles are from Figure 7.14, and open triangles are from Figure 7.15.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

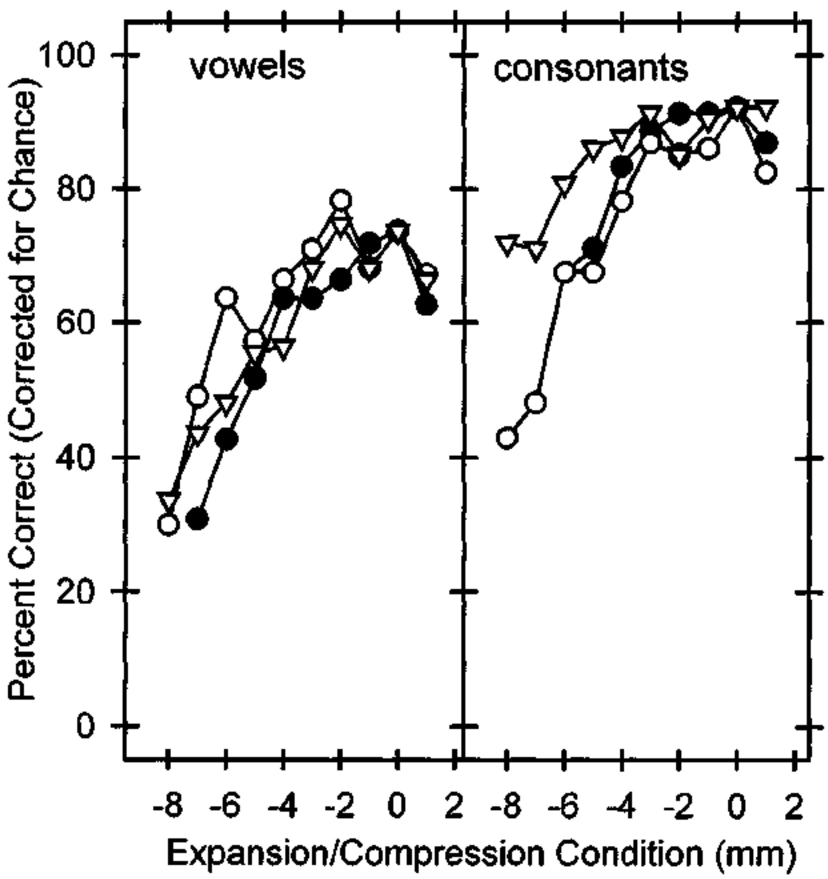


Figure 7.16: The percent correct scores normalized for chance level from Clarion patient when the frequency-place expansion and compression was applied on both ends (filled circles), on apical end only (open circles), and on basal end only (open triangles).

All performances with all maps produced similar performance except the basal mismatch with consonants. Less effect of expansion with basal map is actually in line with simulations, because the high end of the stimulation region is close to 7 kHz, which is too high compared to the most important acoustic information range for speech. SII model predicts the behavior from the vowel recognition that all three maps resulted in a similar

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pattern. A stimulation range of 8-28 mm from round window has most of the useful acoustic information as shown by the weights as a function of distance in Figure 4.11, and this function is symmetrical around the midpoint of 19 mm, which implies that both ends' contributions are equally important. Yet from Chapter 5 we also observed that SII model fits vowel recognition scores better than consonants where temporal properties also have an affect as well as spectral.

In this chapter frequency-place compression and expansion mapping was tested with implant users. In spite of the uncertainty in the exact location of the electrode array location, patients performed in similar patterns with frequency-place compression and expansion conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 8

FREQUENCY-PLACE COMPRESSION AND EXPANSION WITH ELECTRODES AT DIFFERENT INSERTION DEPTHS

In the previous chapter the effects of frequency-place compression and expansion on speech recognition were explored with Med-El Combi and ABC Clarion implant patients. Similar conditions to acoustic simulations were generated by changing the maps of the implants with fitting programs. Generally the results were similar to results from simulations.

In this chapter, the effect of insertion depth of the electrodes is explored by activating different number of electrodes located at different distances from the round window. Med-El Combi 40 + implant system was used in the experiments due to the long electrode array with deep insertion. In the first experiment, the insertion depth was changed from full insertion (where 10 electrodes were activated covering a large frequency range for stimulation) to shallower insertions by turning the most apical electrode off for each condition. This experiment realistically simulates partial electrode insertions where it was not possible to insert the whole array fully. Yet in addition to the insertion depth (and hence the stimulation range length) the number of electrodes changed as well. Shallow electrode insertions have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fewer stimulable electrodes and cover a smaller frequency range in cochlea. In a following experiment, the number of electrodes was kept the same while their location from the round window changed. These experiments will be described and results will be presented in the following sections.

8.1. EXPERIMENTAL METHOD

The same Combi 40+ users from previous study (M1, M2, M3, and M4, aged 25-62) participated in the experiments. They were mapped with Studio+ clinical fitting system. Similar to previous study, patients were wearing BTE TEMPO+ processor and all fitting parameters were kept at similar levels.

Hillenbrand vowels and Shannon consonants were used for the phoneme recognition test, and Harvard sentences were used for the sentence identification test, similar to previous studies.

8.2. RESULTS

8.2.1. EXPERIMENT 1: Frequency-Place Mismatch with Partial Insertion

In many implant patients the electrodes cannot be fully inserted into the cochlea due to physical abnormalities such as cochlear ossification or otosclerosis. As a result many patients have only a few electrodes inserted to a shallow depth covering mostly high frequencies. As mentioned before, in clinical settings the audiologist is given only a limited number of choices for the frequency range to be assigned onto this small region of stimulation, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

these choices do not offer many options for partial insertion. Because the effects of spectral mismatch had not been shown clearly before, the audiologist gives one standard map which assigns very low-frequency acoustic information to high-frequency locations of the cochlea. For insertion depths deeper than 20 mm (hence frequencies lower than 1 kHz) we have shown with simulations that matching gives the best or the same performance even in cases when there is a loss of acoustic information. Experiments described above with implant patients have supported this result as well. However we do not really know what would produce better performance for partial insertion (i.e., insertions shallower than 20 mm) because at that range (around 1 kHz) we start losing important speech information (as pointed out by Faulkner *et al.*, 2003). As a result, cutting the low frequency information completely might actually result in poorer performance than giving some low frequency information in a compressed scheme in these cases.

Several studies have addressed this question (Eyles *et al.*, 1996, Hodges *et al.*, 1999). Some were inconclusive because they were comparing different subjects' speech performance and trying to correlate it with CI user's array insertion depth. Yet this is not a meaningful comparison because as we know implant users' performances differ greatly, and most of the reasons for such variation are not known. The best way to make such a comparison would be to create partial insertion on the same patients and observe the changes from their baseline scores with full insertion. Combi 40+ is perfect for such a comparison due to the long electrode array covering

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ranges from 6-30 mm. Segments of this long electrode array can be used to simulate shallow insertion conditions within a single patient.

In this experiment, we simulated shallow insertions by turning off the most apical electrodes. All three patients who participated the study have full insertions. Therefore we could start from a deep insertion condition. In the deep insertion condition 10 middle electrodes (electrodes 2-11) were activated. Med-El electrode array is actually little longer than the widest frequency range that can be assigned. The best matching is obtained with these 10 electrodes whose center frequencies cover 244-7.5k Hz by Greenwood mapping function. With the assumed insertion depth of 31 mm this array covers a range of 7.2 mm to 28.8 mm from the round window. As a result, this condition simulates a deeply inserted (at 28.8 mm) array of 10 electrodes with a stimulation range of 21.6 mm. From this deep insertion condition the most apical electrode was turned off for each successive condition. Because the electrode separation is 2.4 mm, each condition created an insertion depth that was 2.4 mm shallower than the previous condition. The number of electrodes decreased by one as well.

In the first part of the experiment the same wide frequency range was assigned onto the electrodes for each condition. This wide analysis range of 184-8.9k Hz, which is similar to clinical fitting range of 200-8.5k Hz, is matching the stimulation range of 10 electrodes activated in the first condition. With each successive condition as the number of electrodes decreases the stimulation range decreases as well. Because the frequency range was kept constant the map was compressed further with each

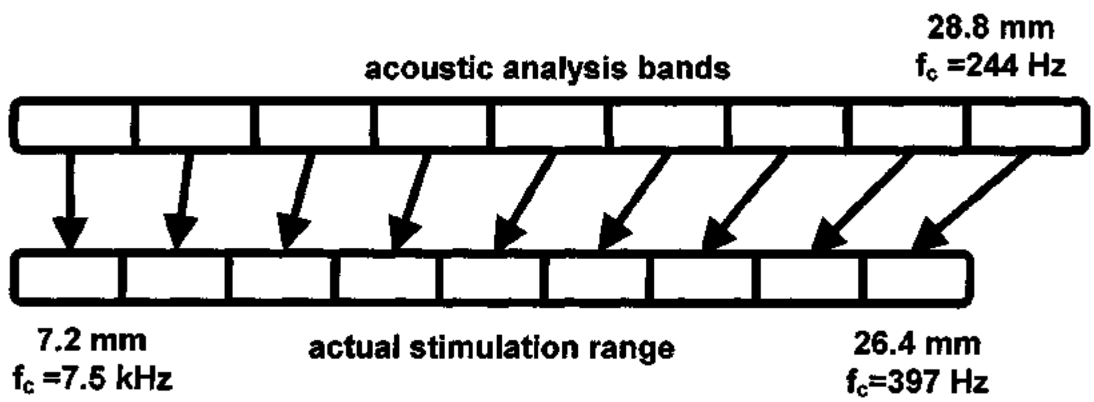
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

condition. This increasing compression is shown for partial insertion conditions of 9, 6, and 3 electrodes in the upper parts of the Figures 8.1-8.3, respectively. The frequency range assigned onto the electrodes is shown by the analysis bands on top. The stimulation region covered by the actual electrode array is shown by the bands under the analysis bands. These stimulation bands would refer to the noise carrier bands in the simulations. The center frequencies of both analysis and stimulation bands that are at the edges of the ranges, as well as the translated cochlear distance by Greenwood mapping function, are shown in the figures. These conditions simulate the clinical approach where the patients gets the same standard map from the clinic regardless of the actual location of the electrode array in her cochlea.

In the second part of the experiment, the acoustic input frequency range was always matching the stimulation range of the electrode array. Because the electrode array was getting shorter with each condition, the patient would get less information with each successive condition. These conditions are shown in lower portions of Figures 8.1-8.3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9 electrodes compression:



9 electrodes matching:

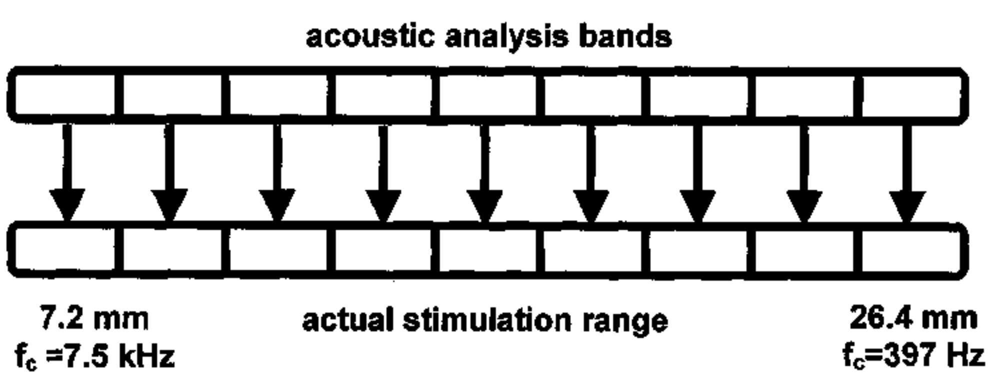


Figure 8.1: Partial insertion condition with 9 electrodes. In each map, the bands on top show the acoustical input range while the bands in bottom show the stimulation range in the cochlea determined by the position of the electrodes. The upper part of the figure shows the compressed map, where the wide acoustic range is assigned onto the stimulation range compressively. The lower part of the figure shows the matched map where the acoustic input range is truncated to match the actual stimulation range.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

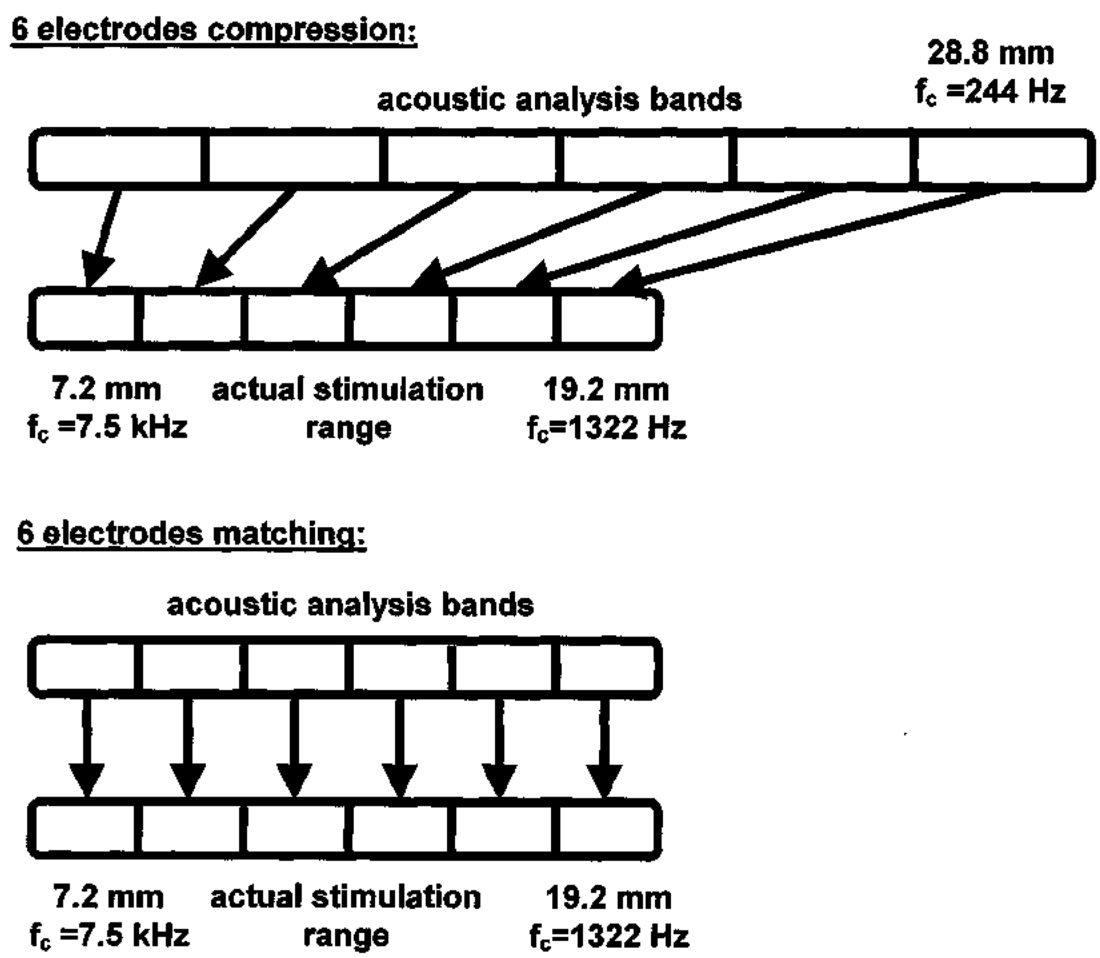


Figure 8.2: Similar to Figure 8.1, both compressed map and matched map for a partial insertion of 6 electrodes are shown.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

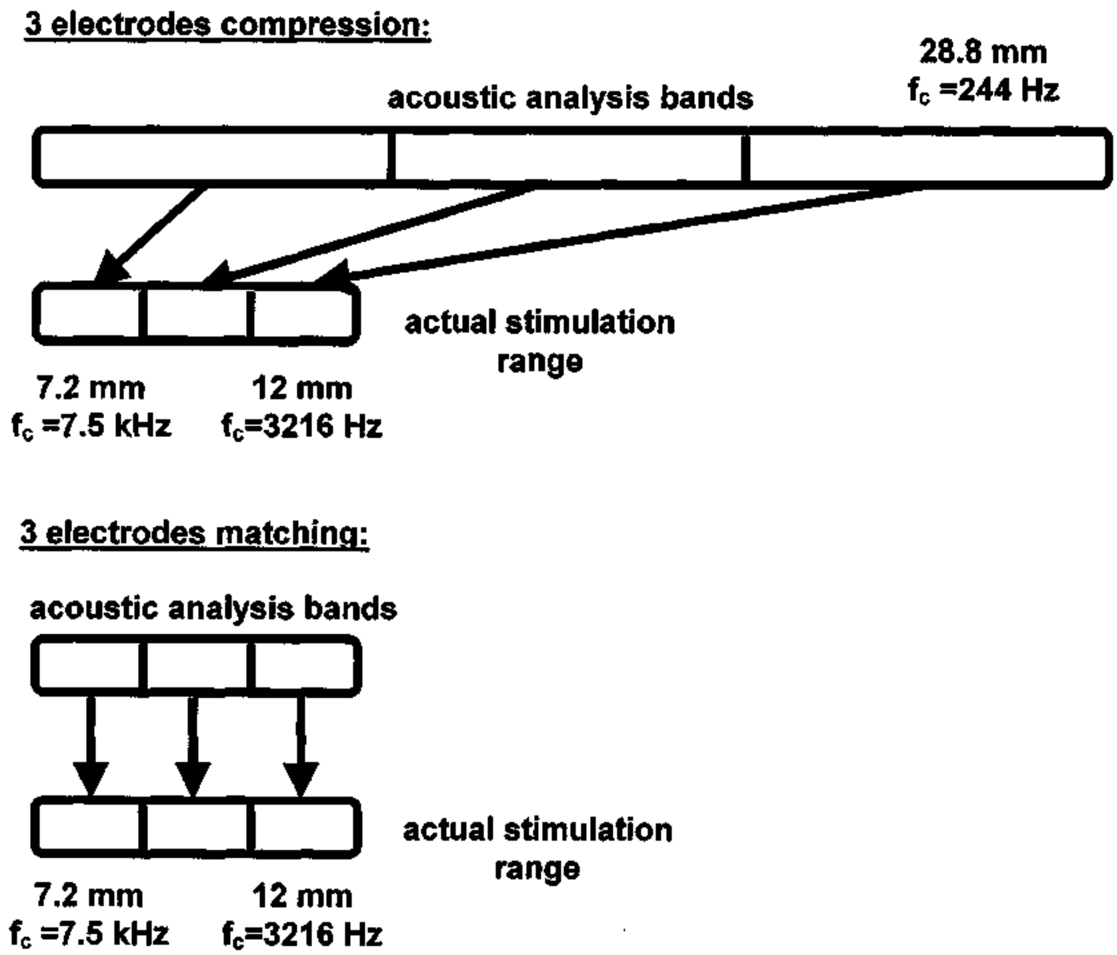


Figure 8.3: Similar to Figure 8.1, both compressed map and matched map for a partial insertion of 3 electrodes are shown.

All compressed and matched conditions for all number of electrodes are summarized in Table 8.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

condition	number of electrodes and electrodes employed	length of the electrode array	insertion depth	frequency range: center frequency range (Hz)/ total analysis range(Hz)	
10	10 (2-11)	21.6 mm	28.8 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	244 -7.5 k 184 - 8.9 k
9	9 (3-11)	19.2 mm	26.4 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	397 - 7.5 k 314 - 8.9 k
8	8 (4-11)	16.8 mm	24 mm	compressed	244 -7.5 k 184 - 8.9 k
				matched	610 - 7.5 k 495 - 8.9 k
7	7 (5-11)	14.4 mm	21.6 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	908 - 7.5 k 747 - 8.9 k
6	6 (6-11)	12 mm	19.2 mm	compressed	244 -7.5 k 184 - 8.9 k
				matched	1322 -7.5 k 1098 - 8.9 k
5	5 (7-11)	9.6 mm	16.8 mm	compressed	244 -7.5 k 184 - 8.9 k
				matched	1900 - 7.5 k 1586 - 8.9 k
4	4 (8-11)	7.2 mm	14.4 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	2703 - 7.5 k 2267 - 8.9 k
3	3 (9-11)	4.8 mm	12 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	3822 - 7.5 k 3216 - 8.9 k
2	2 (10-11)	2.4 mm	9.6 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	5382 -7.5 k 4538 - 8.9 k
1	1 (11)	single electrode	7.2 mm	compressed	244 - 7.5 k 184 - 8.9 k
				matched	-

Table 8.1: Frequency-place mapping conditions for varying insertion depths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Performance with both compressed and matched conditions are expected to decrease with decreasing number of electrodes. At each condition one more apical electrode is turned off. With compressed conditions this increases the mismatch between the acoustic input frequencies and the appropriate stimulation locations for those frequencies. For matched conditions, less information is provided to the patient. Yet the question is which condition would give the better performance when the patient has fewer number of electrodes that are inserted to a shallower depth, and whether there might be an optimum range between the most compressed map and the perfectly matched map.

The individual percent correct scores are shown in Figure 8.4. The top row presents vowel recognition scores, the middle row presents consonant recognition and the bottom row presents sentence recognition scores. The open circles show the percent correct scores with the compressed map shown in the upper portions of Figures 8.1-8.3 where the same center frequency range of 244-7.5k Hz was assigned on the partially inserted electrode array, resulting in increasing compression as the insertion depth gets smaller. The filled symbols show scores with the matched maps, as shown in the bottom portions of the same figures.

Figure 8.4 shows the trade-off between the amount of acoustic information delivered to the electrodes and the accuracy of the stimulation location. Even though there is a considerable variation in results across implant users, this trade-off was clear for every patient. The crossover point was 14.4 mm of insertion depth for M1, 17-19 mm for M2 and M4, and 15-18

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mm for M3. As a result, for moderate to deep insertions, there is a clear advantage of matching the acoustic information to the stimulation region over compressing the wide acoustic range (the default option in the clinical fitting system). More specifically, insertions from as shallow as 15-16 mm for vowels, and 19 mm for consonants and sentences to 26 mm resulted in better scores with matched map. For shallower insertions, patients M2 and M3 performed better with phonemes when mapped compressively with the wide acoustic range compared to the matched map, yet this was not very clear with M1 and M4. M2 exhibited similar pattern for sentences, where he had better scores with matched map for deeper insertions than 19.2 mm and had better scores with the compressed map for shallower insertions. On the other hand M3 had the best scores with the matched map with sentences. He was already performing at floor level with 6 channels and a wider acoustic range did not increase the scores for six or fewer electrodes.

Results from M2 and M3 generally support the idea laid out by Faulkner *et al.* (2003) that matching the acoustic range to the stimulation range can actually be detrimental to speech recognition for insertions less than 19 mm, yet results from M1 and M4 suggests that some patients might favor matched map even for shallower insertions. M1 had better phoneme recognition scores with matched map compared to compressed map even at 16.8 mm of insertion. M4 had better vowel recognition scores with matched map at all insertions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

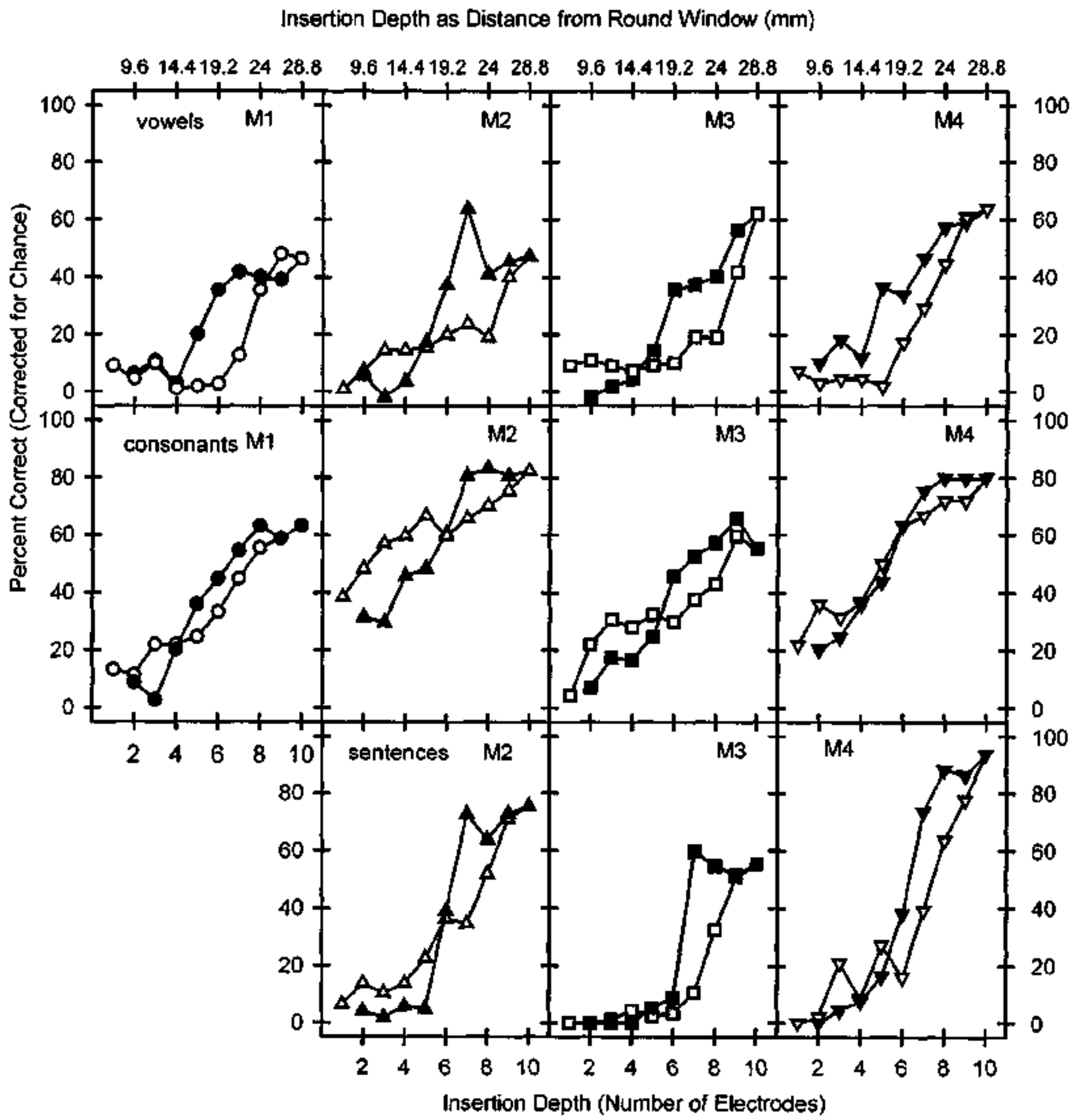


Figure 8.4: Individual percent correct scores from patients M1, M2, M3, and M4. Filled symbols are when the acoustic information range was matched to the stimulation region, and open symbols are when the entire acoustic bandwidth was compressed into the shorter stimulation region. Vowel and consonant recognition scores are corrected for chance level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A drop in the performance with the matched map is actually expected from SII model because the range of the acoustic information contributing to speech recognition decreases, as shown by the diamonds in Figure 8.5. An additional factor contributing to the decrease in the performance is the decrease in the number of electrodes, which most probably has the biggest effect when the number of electrodes employed is less than six. Yet SII model predicts the same performance for the compressed map for all number of electrodes regardless of the stimulation range because the same wide range of frequencies was available for each condition. Therefore the decrease in the performance is due to other reasons, such as decreasing number of channels and the mismatch between frequency bands and their corresponding stimulation locations in the cochlea. The prediction of SII model was modified to incorporate effects of mismatched frequency bands as given by the modified SII equation in Chapter 4 (Equation 4.4). The squares in Figure 8.5 show the predicted performance with the compressed map. The value of α was chosen such that crossover would be in a reasonable area, for this case the selection was $\alpha=0.025$. Similar to experimental data, better performance is estimated with the matched map for relatively deeper insertions, i.e. with number of electrodes that are higher than the crossover. The performance is estimated to be better with the compressed map for a smaller number of electrodes. The model predicts a steep increase in performance with the compressed map for very small number of electrodes. With implants it is known that if less than 6 electrodes are activated performance drops significantly. SII model cannot predict this

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

drop because it does not incorporate effects of spectral resolution. A simple modification of this model to account for spectral resolution should yield a better match to the experimental data.

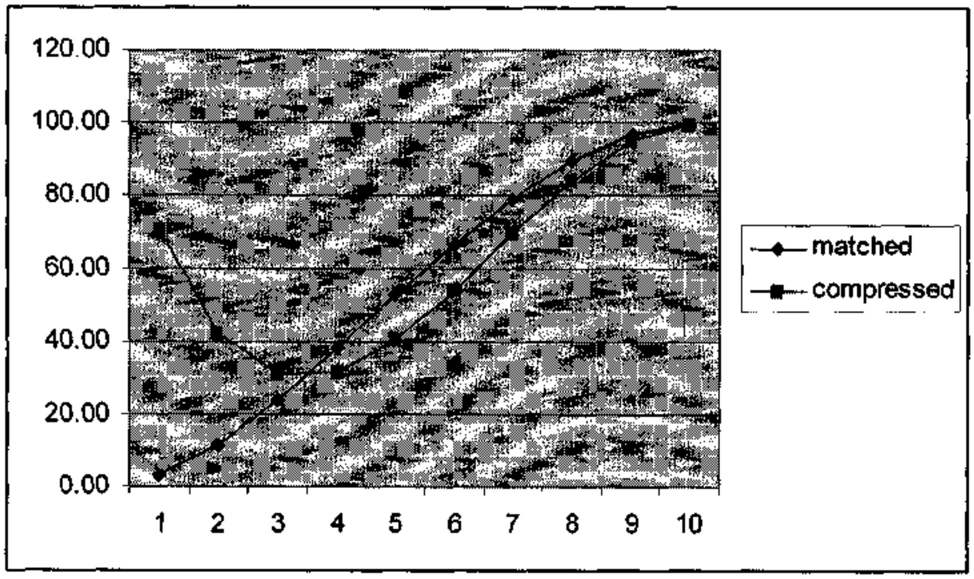


Figure 8.5: SII model prediction for matched and compressed maps as a function of number of electrodes.

In most cases, there was no difference in performance for both matched and compressed maps between 10 electrodes (28.8 mm insertion condition) and 9 electrodes (26.4 mm insertion condition). This might be due to a tolerance of few mm by the pattern recognition system. From acoustic simulations we would expect to see a bigger drop from matched to compressed map score. So this might be an indication for the decreased frequency selectivity of the spiral ganglia in the upmost turn of cochlea when stimulated electrically. The spiral ganglia are in the middle of the modiolus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and going up in the apical direction in the cochlea. The modiolus gets narrower at the apex so the spiral ganglia get closer together in a tight bundle, making the distinctive stimulation of each channel less likely. Even though we know that the speech recognition increases with increasing insertion depth, as shown by many studies for insertions up to 25 mm, the anatomy of the cochlea might put a limit on this improvement for deeper insertions (Cohen *et al.*, 1996b).

An interesting observation is that M2 and M4 always scored much higher with consonants in all implant experiments. For example, even at conditions where these subjects had similar vowel scores to both M1 and M3, when it came to consonants, they always had much higher consonant scores than both M1 and M3. This difference also contributed to their higher sentence recognition scores. This observation implies that subjects M2 and M4 make better use of temporal cues. Also their consonant scores were usually more flat compared to M1 and M3 with most mismatch conditions, which means that they are more robust to the spectral manipulations with consonants. Note that M2 still identified 40% of consonants and M4 identified 20% correctly even when they had only one electrode active (in Figure 8.4). At this condition there is no spectral resolution whatsoever even though a wide spectral range is provided. All information is carried in the envelope which amplitude-modulates the spectral content.

To extract a general pattern, the average of the scores were taken and replotted in Figure 8.6. Similar to Figure 8.4, filled circles show the average percent correct scores of all four patients for phonemes, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

average scores of M2, M3, and M4 for sentences when the patients had the matched map. Open circles show the average scores of all patients with phonemes, and patients M2, M3, and M4 with sentences, when they were tested with compressed map. The average scores show that the performance is almost the same for conditions where 9 or 10 electrodes were active with both matched and compressed maps. With decreasing insertion depth (i.e., decreasing number of electrodes) the performance started dropping with the compressed map right away (as shown by open circles). The performance did not drop as fast with matched map (shown by filled circles). Matched map gave better performance than the compressed map until 6 or fewer electrodes were active.

A ceiling effect at higher number of electrodes and a floor effect at very low number of electrodes minimize the difference between two maps.

For the comparison between two maps a paired t-test was used for phoneme recognition tests. The t-test shows significant advantage of matched map for phoneme tests at moderate insertions (6-8 electrodes for vowels, and 7-8 electrodes with consonants and sentences). Compressed map only resulted in significantly better consonant recognition scores at shallower insertions (2-3 electrodes). There was no significant difference in vowel and sentence recognition scores obtained with both maps.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

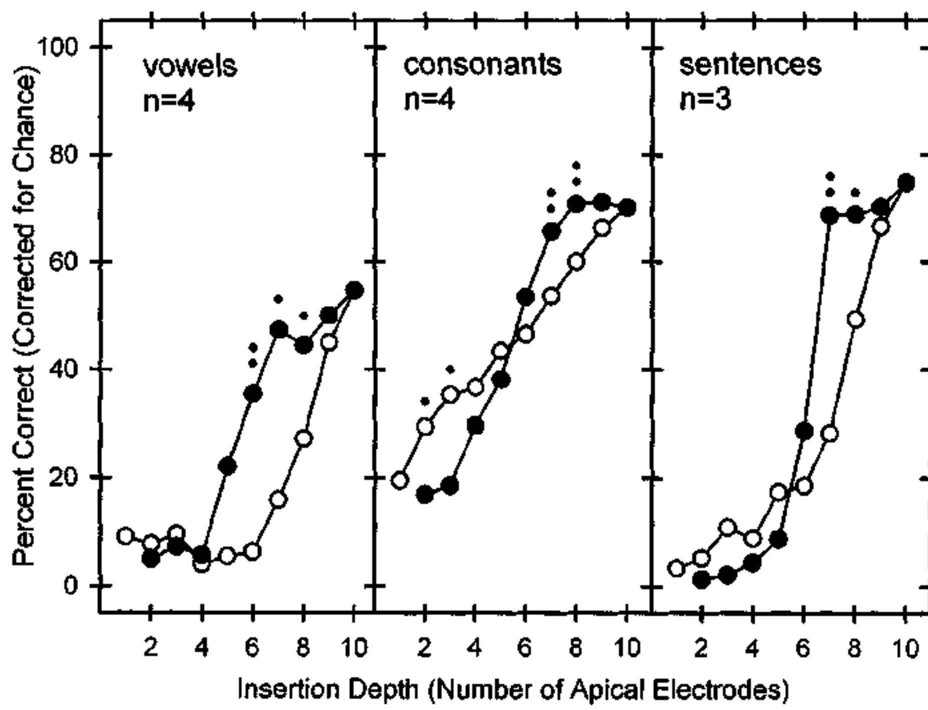


Figure 8.6: Average percent correct scores of M1, M2, M3, and M4 for phonemes, and M2, M3, and M4 for sentences. The open circles show the scores when a wide acoustic range was mapped to electrodes whereas the filled circles show the scores when the acoustic range was cut to match the stimulation range. Small dots on top of the scores show the significance level of the difference between the scores from two maps with a paired t-test: one circle for $p<0.05$, two circles for $p<0.01$.

Figure 8.4 shows the trade-off between the amount of acoustic information available versus the accuracy of the location where the information is mapped. The data in the figure shows two extremes, where either the widest range available is assigned, or it is cut exactly at the matching point at the expense of losing a lot of information. Yet there might be an optimum range in-between where some information is included by cutting off the acoustic range more in the middle but still keeping some accuracy in the location that the information is mapped.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The following experiments are designed to explore the question if there is such an optimum region at the insertion depths of 19.2 mm and 16.8 mm.

8.2.1.1. EXPERIMENT 1.1: 6 Electrodes at 19.2 mm Insertion Depth

As mentioned in the previous section, Figure 8.4 shows scores with two extreme maps, one with the widest possible acoustic input range mapped onto the electrode array (compressed map) and one with perfectly matching acoustic range where possibly a lot of useful speech information was cut off (matched map). The matched map at simulated insertion depth of 19.2 mm (with 6 electrodes active) resulted in significantly better vowel recognition, and only a very small improvement in consonant and sentence recognition, compared to the compressed map. Yet there might be an even better optimum point between these two maps by taking advantage of both maps, where a relatively wider acoustic range is mapped to a relatively accurate cochlear location. To find this optimal operating region as well as to explore how the scores change between the two maps, we kept the number of electrodes the same (electrodes 6-11 located from 7.2 mm to 19.2 mm covering 12 mm) and varied the frequency-electrode allocation from perfect matching to clinical settings in small cochlear distance steps. Hence we increased the compression on the apical end by providing more and more acoustic information. The conditions are summarized in Table 8.2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	range of acoustic input (mm)	band-pass filter center frequencies (Hz)						frequency range of analysis bands (Hz)
0 mm (matching)	7.2 – 19.2	1322	1899	2702	3822	5382	7555	1098 - 8.9k
+1 mm (compression)	7.2 – 20.2	1133	1685	2476	3609	5231	7555	922 - 8.9k
+2 mm (compression)	7.2 – 21.2	968	1493	2267	3407	5085	7555	772 - 8.9 k
+3 mm (compression)	7.2 – 22.2	824	1322	2075	3216	4942	7555	643 - 8.9 k
+4 mm (compression)	7.2 – 23.2	699	1168	1899	3035	4804	7555	531 - 8.9 k
+5 mm (compression)	7.2 – 24.2	590	1031	1736	2864	4669	7555	436 - 8.9 k
+6 mm (compression) (clinical setting)	7.2 – 25.2	495	907	1586	272	4538	7555	354 - 8.9 k
+7 mm (compression) (clinical setting)	7.2 – 26.2	412	798	1449	2549	4410	7555	284 - 8.9 k
+8 mm (compression) (clinical setting)	7.2 – 27.2	340	699	1322	2404	4286	7555	223 - 8.9 k

Table 8.2: Compression conditions for 6 electrodes inserted 19 mm deep and covering 12 mm in the cochlea.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 8.7 shows the individual percent correct scores from phoneme tests. 0 mm is the matched condition (which was shown with filled symbols in Figure 8.4) and +8 mm compression is the map with wide acoustic input (as shown with open symbols in Figure 8.4). In consonants performance did not change with increasing compression except the extreme case of +8 mm compression. The difference in vowel recognition scores between the two maps are clear with M1 and M3; the performance decreases with increasing compression after an optimum range of a few mm. M2 and M4 exhibit sharp optimums at +3 mm and +2 mm compression conditions, relatively. This pattern supports the hypothesis that at shallow insertions a compromise between the amount of information provided and the accuracy of the location that the information is presented can actually be beneficial. The optimum condition of +2 mm or +3 mm compression means that assigning acoustic information as low as 650 Hz (from Table 8.2) onto the electrodes improved the performance, but adding lower frequencies started to become harmful. If a patient had a shallow insertion of 19.2 mm the closest lower limit among the choices offered by the clinical program to this value would be 350 Hz, which is the +6 mm compression (from Table 8.2). At this condition subjects M1 and M3 already perform worse than the 0 mm matched condition where a lot of useful acoustic information (frequencies lower than 1 kHz) was missing. +6 mm, +7 mm, and +8 mm compressed maps are the only choices offered by the program and they are clearly not the optimal maps for such a shallow insertion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

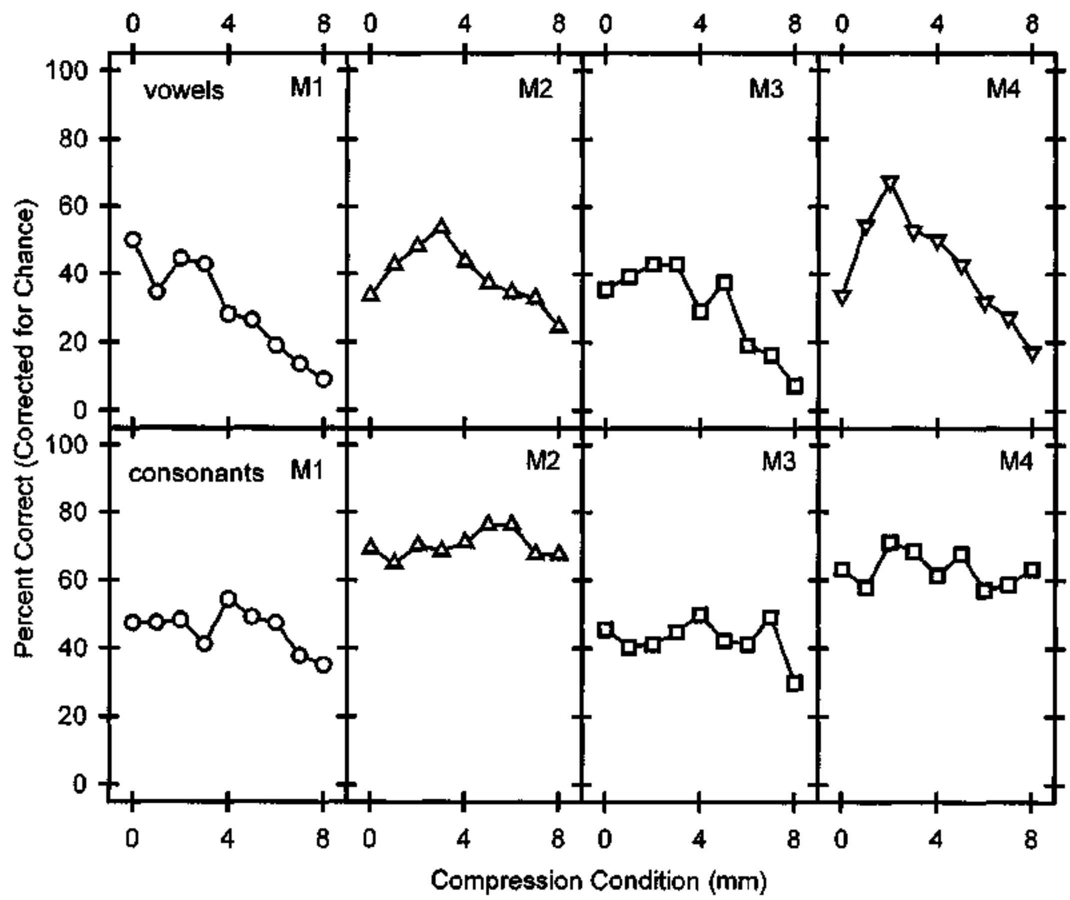


Figure 8.7: The frequency-range assigned to the electrodes is changed from perfect match to clinical settings in steps of 1 mm for 19.2 mm insertion depth.

The average performance of all four subjects is presented in Figure 8.8. A repeated-measures one way ANOVA shows that the vowel recognition performance decreased significantly with increasing compression ($F(8,24)= 14.06$, $p<0.001$). The peak performance was obtained with a compression of 2-3 mm. A post-hoc Tukey test showed that there was no significant difference between scores obtained from 0 mm matched condition up to +5 mm compression, but starting with +6 mm compression the scores were

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

significantly lower. +6 mm and higher compressions are the ones that can be achieved with the options of the current fitting program. Consonant scores do not change significantly but show a small peak around +4 mm compression.

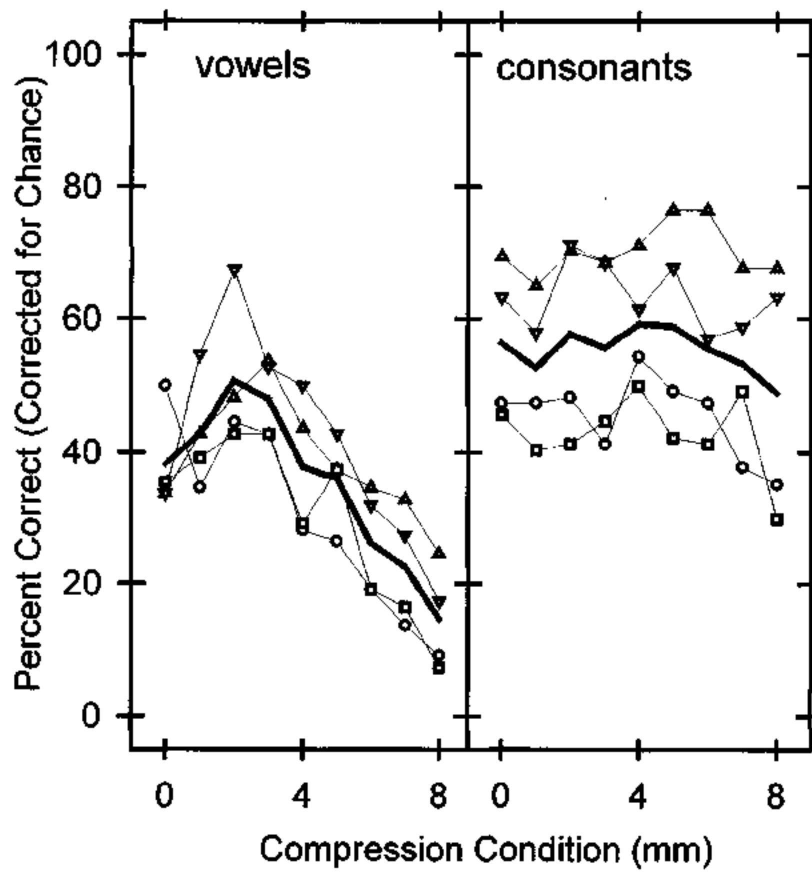


Figure 8.8: Average percent correct scores of subjects M1, M2, M3, and M4 with 6 electrodes at 19.2 mm insertion depth as the compression increases from 0 mm matched condition (as shown by thick lines). Thin lines with symbols show individual scores from previous figure.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

8.2.1.2. EXPERIMENT 1.2: 5 Electrodes at 16.8 mm Insertion Depth

A similar test was run with 5 electrodes at 16.8 mm covering 9.6 mm, representing a shallower insertion depth than the previous experiment. Recalling the scores from Figure 8.4, there was no clear difference between performances with 0 mm matched condition and the compressed map. The question is if there is an optimal region in-between where some compression helps by adding more useful acoustic information. To answer this question the acoustic range that was assigned onto those 5 electrodes was increased from 0 mm matched condition to the clinical map (+12 mm compression for this setting) in 1.5 mm steps. The test conditions with corresponding frequency ranges are summarized in Table 8.3.

The individual percent correct scores for this experiment are shown in Figure 8.9. As before the columns are scores from different patients. Top row shows the vowel recognition scores and the bottom row shows the consonant recognition scores. The drop in vowel recognition of M1 and M4 from Figure 8.4 can be seen here from 0 mm matched condition to +12 mm compression. Other subjects had the same scores for vowels and all subjects had the same scores for consonants with those two maps, as expected from Figure 8.4. Yet, there is a clear trade-off of the amount of acoustic information provided versus the accuracy of the location, and with the advantage of the trade-off the performance can actually be increased at a middle point between these two maps. This optimal range is sharper with vowels which are more sensitive to spectral manipulations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place mismatch condition	range of acoustic input (mm)	band-pass filter center frequencies (Hz)					frequency range of analysis bands (Hz)
0 mm (matching)	7.2 – 16.8	1899	2702	3822	5382	7555	1586 - 8.9k
+1.5 mm (compression)	7.2 – 18.3	1516	5103	3432	2293	7555	1226 - 8.9k
+3 mm (compression)	7.2 – 19.8	1205	1942	3080	4838	7555	941 - 8.9k
+4.5 mm (compression)	7.2 – 21.3	952	1641	2762	4586	7555	715 - 8.9 k
+6 mm (compression)	7.2 – 22.8	747	1384	2476	4347	7555	536 - 8.9 k
+7.5 mm (compression)	7.2 – 24.3	580	1164	2218	4121	7555	394 - 8.9 k
+9 mm (compression)	7.2 – 25.8	444	975	1985	3905	7555	282 - 8.9 k
+10.5 mm (compression) (clinical setting)	7.2 – 27.3	334	814	1776	3701	7555	193 - 8.9 k
+12 mm (compression) (clinical setting)	7.2 – 28.8	244	676	1586	3507	7555	123 - 8.9 k

Table 8.3: Compression conditions for 5 electrodes inserted 16.8 mm deep and covering 9.6 mm in the cochlea.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

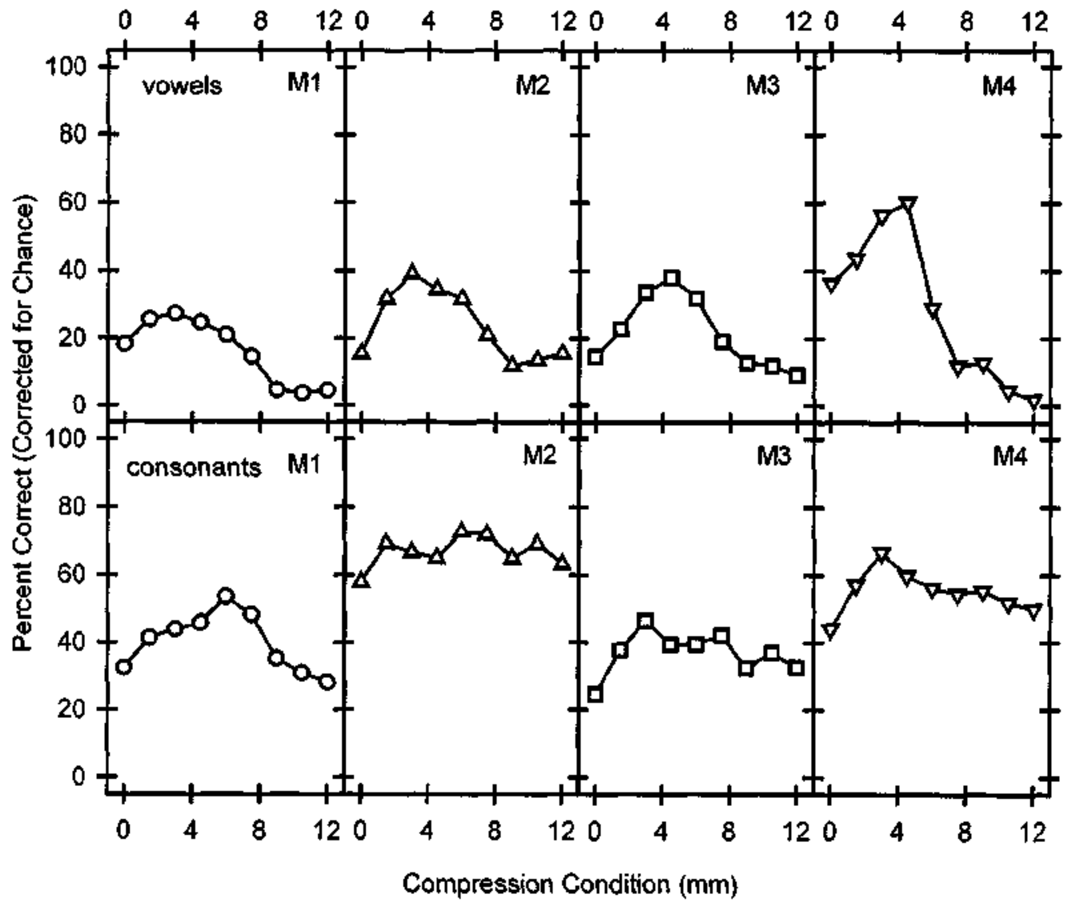


Figure 8.9: The frequency-range assigned to the electrodes are changed from 0 mm matched condition to clinical settings in steps of 1.5 mm for 5 electrodes at 16.8 mm insertion depth.

The average scores from all patients are shown with thick lines in Figure 8.10. The thin lines with symbols repeat the individual scores from Figure 8.9.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

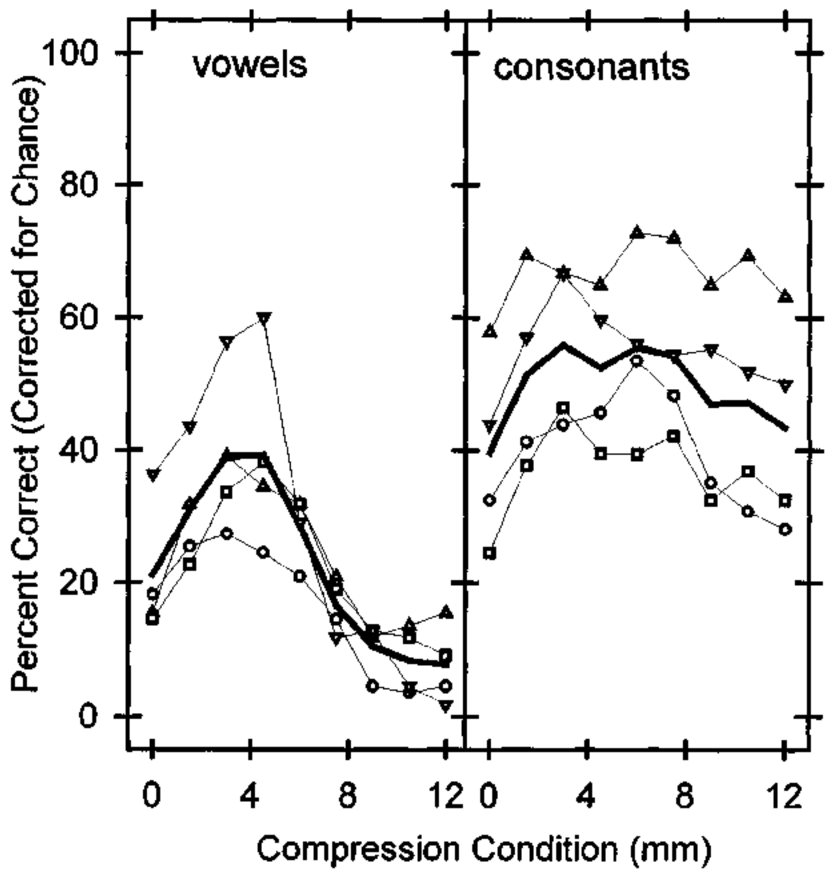


Figure 8.10: Average percent correct scores of subjects M1, M2, M3, and M4 with 5 electrodes at 16.8 mm insertion depth as the compression increases from 0 mm matched condition (as shown by thick lines). Thin lines with symbols show individual scores from previous figure.

A one-way repeated measures ANOVA shows that there is a significant effect of compression on both vowels ($F(8,24)=11.27$, $p<0.001$) and consonants ($F(8,24)=7.54$, $p<0.001$). A post-hoc Tukey multiple comparison test shows that there is no significant difference between +1.5 and +6 mm compression on vowel recognition, giving the best performance in this range. This range corresponds to frequencies with lower limits from 500 Hz to 1.2 kHz (as shown in Table 8.3). Increasing or decreasing compression results in significant drop in performance. For consonants the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

optimal range is much wider, any compression between 0 mm matched condition and +12 mm clinical setting map actually increases the consonant recognition.

This figure shows that by choosing a good map the performance can instantly be increased by 20-30%. At a low performance level (due to the shallow insertion) such an increase (from 20% to 45% for vowels, and from 40% to 55% for consonants for example) might make a big difference for the overall speech perception of the patient.

8.2.2. EXPERIMENT 2: 6 Electrodes Located at Different Insertion Depths

In the previous experiment we simulated partial insertions by changing the number of electrodes in each condition. As a result the total electrode array length and the stimulation region in the cochlea changed as well. In this experiment we measured the effect of the region of the cochlea stimulated by holding the number of electrodes the same and changing the insertion depth of this array of 6 electrodes. We activated electrodes 6-11 to create a shallow insertion around 19 mm, 4-9 for a typical insertion around 24 mm (such as in Experiment 3 in Chapter 7), and 2-7 for a deep insertion of 30 mm. The total range of active electrodes was always 12 mm (14 mm stimulation region). Therefore this experiment imitates an electrode array of 6 electrodes covering 12 mm in cochlear length at different stimulation regions. The stimulation region in the cochlea and the matching acoustic input ranges at these insertion depths are shown in Figure 8.11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Normally if we have a wide region for stimulation available we expect to see better performance with increasing insertion depth (due to abundance of acoustic information mapped to correct location). The exact nature of stimulation is still not very clear at the very end of the most apical turn, and as it was shown in Experiment 1 that the performance generally increased up to 24-27 mm of insertion, but usually reached a ceiling effect with further insertion. However in the present experiment the number of electrodes and hence the stimulation region is limited, and they cover only 12 mm. This means cochlear coverage from 7 mm to 19 mm for shallow insertion and because this range does not contain as much of the critical frequency range for speech we expect to see some drop in performance compared to other locations. Yet for the deep insertion of 29 mm a 12 mm length covers 17-29 mm from round window which will exclude part of the useful speech spectrum (frequencies>1.5 kHz). Because of that we might not see much improvement with deep insertion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

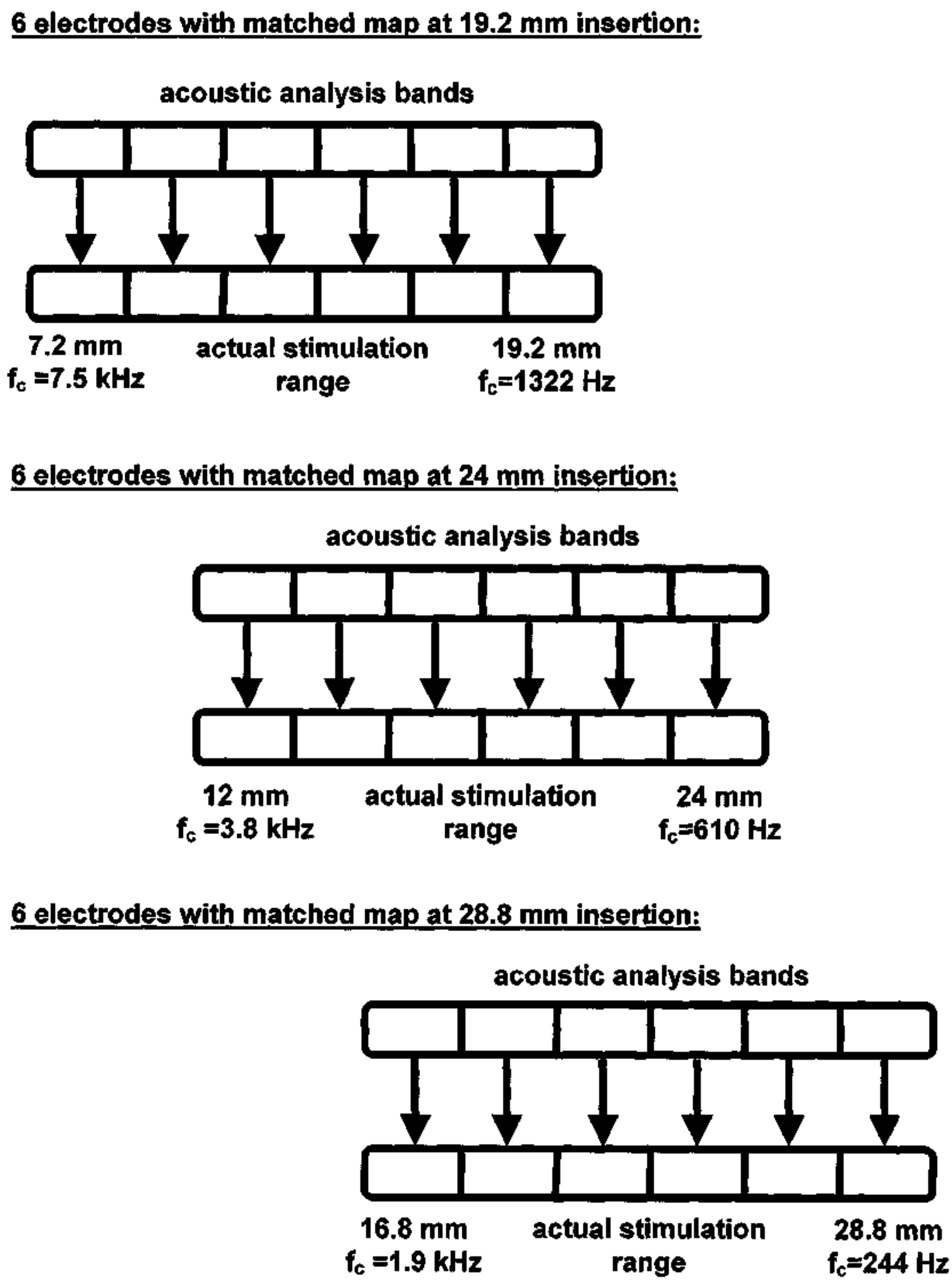
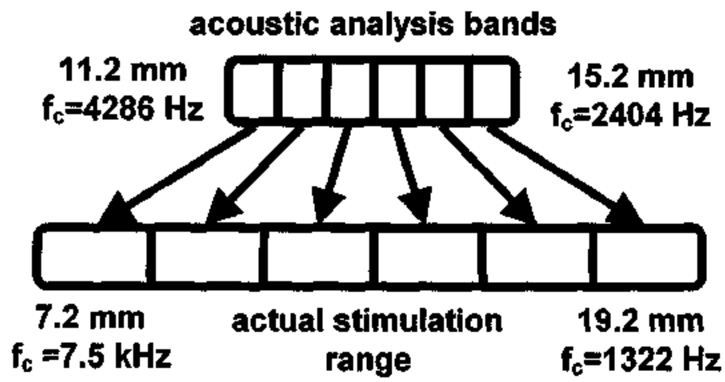


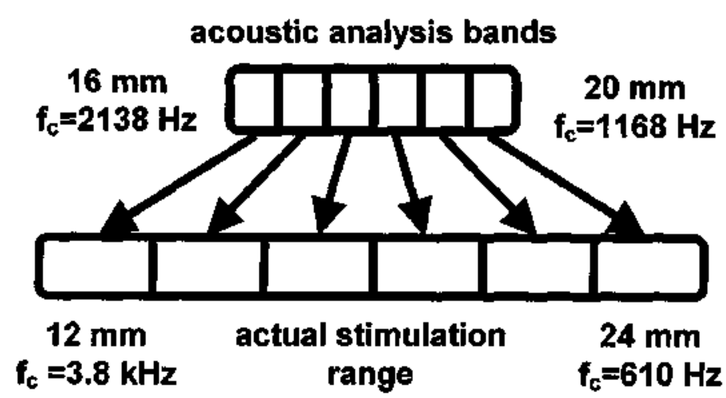
Figure 8.11: 6 electrodes activated with matched acoustic input at different cochlear regions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6 electrodes with -4 mm expansion at 19.2 mm insertion:



6 electrodes with -4 mm expansion at 24 mm insertion:



6 electrodes with -4 mm expansion at 28.8 mm insertion:

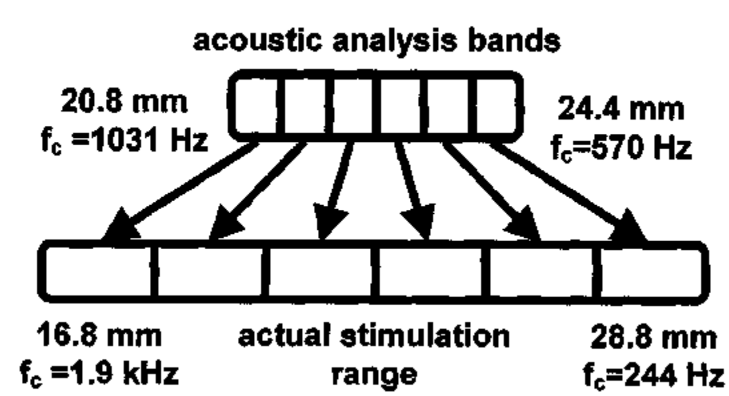


Figure 8.12: 6 electrodes activated with expanded acoustic input at different cochlear regions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

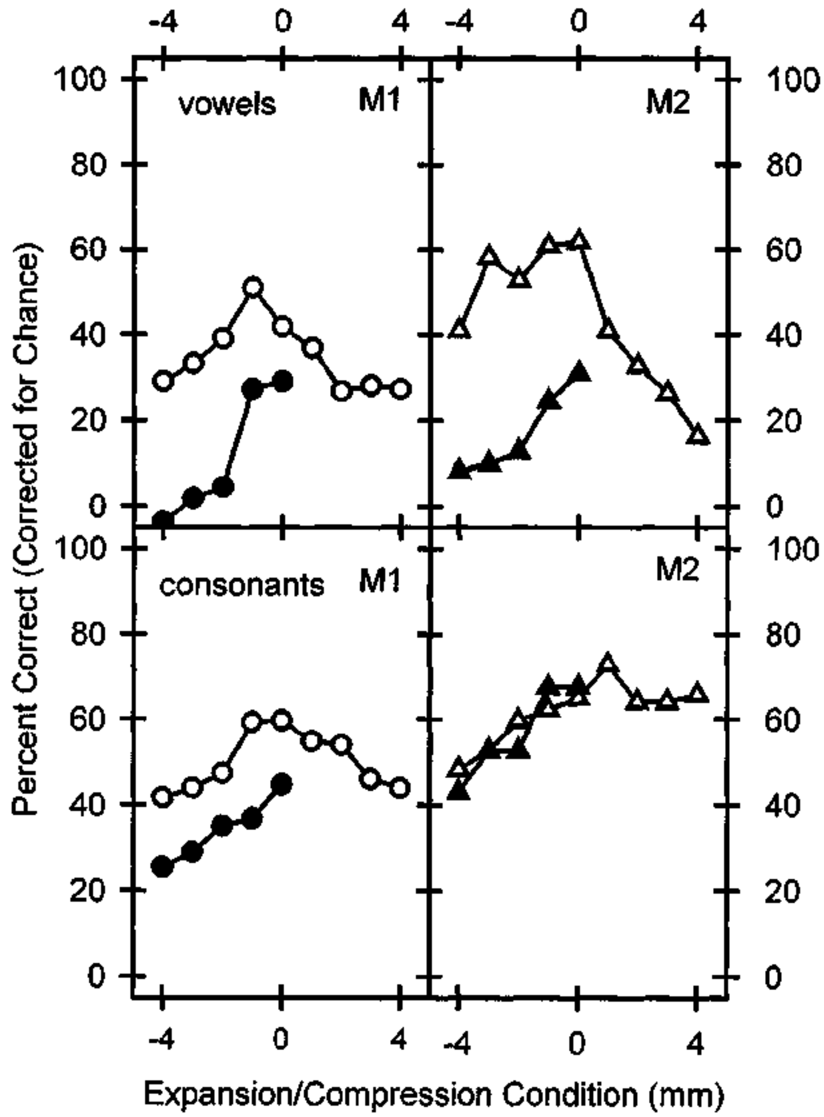


Figure 8.13: Individual frequency-place mismatch percent scores for 6 electrodes located shallow (19 mm, shown by filled symbols) and midway (24 mm, shown by open symbols) in the cochlea of subjects M1 and M2.

Figure 8.13 shows the individual percent correct scores with mismatch conditions when electrodes 4-9 were activated. These are the same scores replotted from Experiment 1, repeated here with open symbols. Filled symbols show percent correct scores with similar conditions when electrodes 6-11 were activated. The stimulation range in the cochlea is the same length

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of 12 mm because the number of electrodes used in both tests is the same (6 electrodes). As a result the only difference is how far this array of six electrodes is inserted, which is 24 mm for electrodes 4-9, and 19.2 mm for electrodes 6-11. Because the latter setting is already at the end of the whole stimulation region it leaves no spectral room to widen to create compression conditions. Therefore only expansion was applied with electrodes 6-11. The -4 mm expansion condition was shown in Figure 8.12 for all three insertion depths.

Similar to 24 mm insertion, expansion caused a big drop in speech recognition from 0 mm matched condition when the same array was at 19.2 mm. There is a large difference in vowel recognition scores obtained at two insertion depths, they drop significantly at shallower insertion. This is actually expected from the weights in Figure 4.11. The range covered by electrodes 4-9 is 12-24 mm from round window, and the analysis band range is 500-4.5 kHz. The range covered by electrodes 6-11 is 7.2-19.2 mm from the round window, with a corresponding analysis band range of 1-9 kHz. Even at 0 mm matched condition a lot of useful acoustic information is not available to the listener. From 0 mm matched to –2 mm expansion a sharp drop was observed, because at –2 mm all frequencies smaller than 1.5 kHz was missing, which is a very important range for vowel recognition.

The scores decreased in a similar pattern with M1 with consonants, yet M2 had exactly same performance with same number of electrodes placed 5 mm shorter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

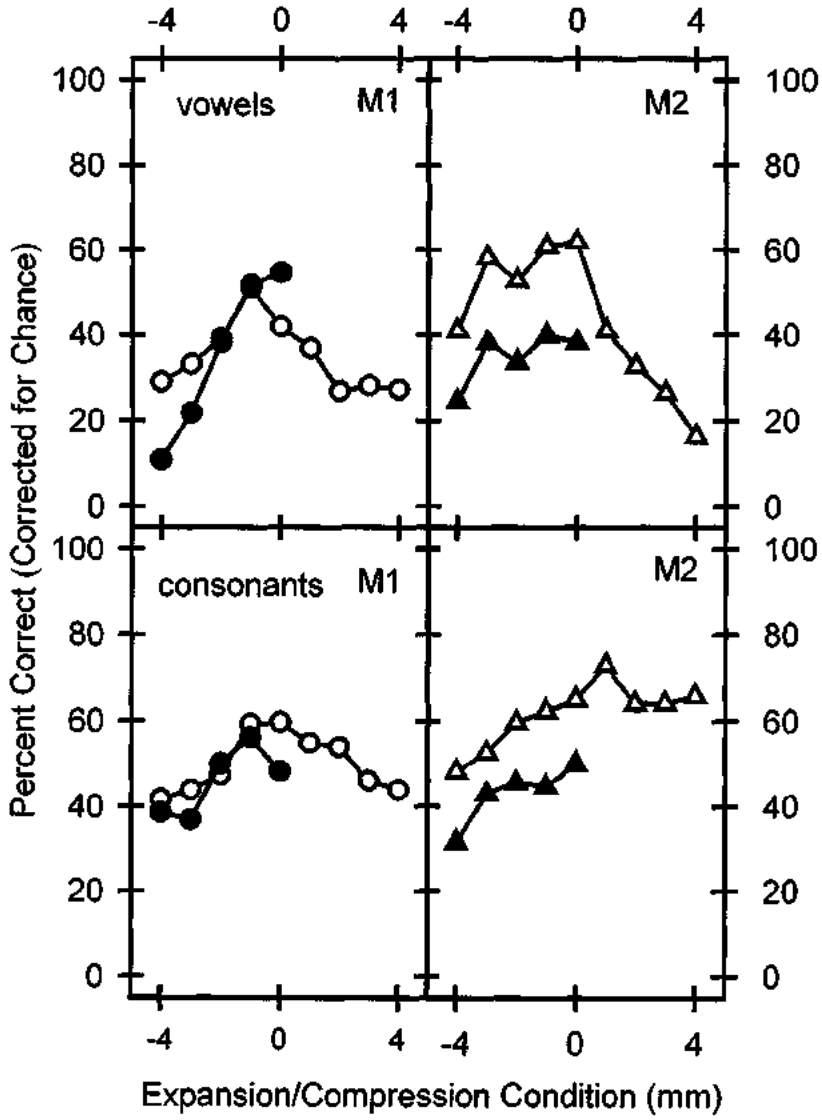


Figure 8.14: Individual frequency-place mismatch percent scores for 6 electrodes located deep (29 mm, shown by filled symbols) and midway (24 mm, shown by open symbols) in the cochlea of subjects M1 and M2.

Figure 8.14 similarly shows the individual percent correct scores, when mismatch conditions were applied at both ends of electrodes 4-9 (the same as in Figure 8.13, shown with open symbols), and electrodes 2-7 (shown by filled symbols). The range covered by electrodes 2-7 is 16.8-28.8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mm from the round window, with a corresponding analysis band range of 200-2.3 kHz. SII predicts a small drop in performance from midrange insertion, yet only M2 shows that drop. Also at this deeper insertion, scores of both subjects drop with increasing expansion, similar to midrange and shallow insertions shown in Figure 8.12.

To explore the general pattern, the average scores of M1 and M2 with all three settings are plotted in Figure 8.15. The open triangles show the performance when the frequency-place map was expanded and compressed with electrodes 4-11. The matched and -4 mm expanded maps were shown in the middle rows of Figures 8.11-12 for this array, respectively. The filled circles show the performance when the map was expanded with electrodes 6-11 activated simulating an array of 6 electrodes inserted up to 19.2 mm. The matched and -4 mm expanded maps were shown in the top rows of Figures 8.11-12 for this array, respectively. The filled squares show the performance when the map was expanded with electrodes 2-7 activated simulating an array of 6 electrodes inserted deeply, up to 28.8 mm. The matched and -4 mm expanded maps were shown in the bottom rows of Figures 8.11-12 for this array, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

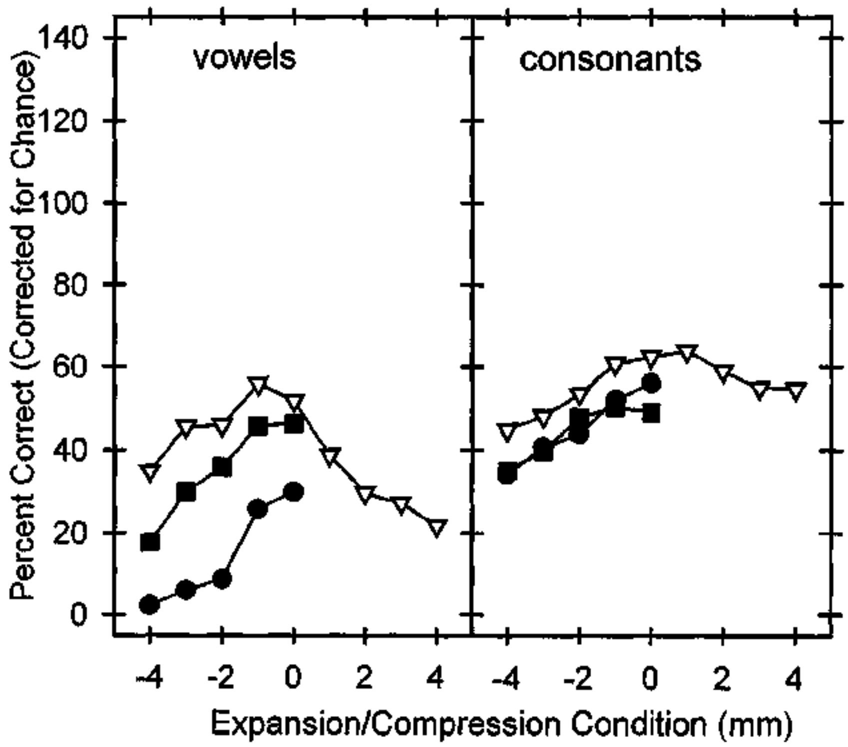


Figure 8.15: Average percent correct scores of subjects M1 and M2. Open triangles show the scores from Experiment 1 when the frequency-place expansion/compression was applied at both ends of the stimulation region of a 6 electrode array inserted to 24 mm. Filled circles are the scores from the same array with a shallow insertion of 19.2 mm. Filled squares show the scores from the same array with a deep insertion of 28.8 mm.

SII weights peak around 1.5-2 kHz as shown in Figure 4.11, and by Greenwood mapping function these frequencies correspond to a distance of 17-18 mm from the round window. The shape of this weight function is pretty much symmetrical around that peak point in cochlear distance. As a result an array covering 12-24 mm from the round window is expected to have the most acoustic information compared to arrays of the same length located

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shallower or deeper. On average the best performance was observed when the array was located in the middle of the weight function, i.e. from 12 to 24 mm from round window, as shown by open triangles in the figure. From the shape of the weight function, SII predicts a bigger drop for deeper insertion, shown by filled squares, than the shallower insertion, shown by filled circles. Yet the results show that the performance was much better with deeper insertion than the shallower insertion for vowels, and the same for consonants. SII takes into account only the amount of spectral information available, yet discrimination of formants is another strong cue for vowels and most formants are delivered in the apical region with deep insertion. This might be why the array inserted deeper results in better vowel recognition scores than expected. The drop is much smaller with consonants, and this might be the part that comes from spectral content, yet temporal cues available prevent a further drop.

In Experiment 1 the insertion depth was decreased by turning off the most apical electrode at each condition. Therefore the number of channels and the range of stimulation decreased as well as the insertion depth, which were all decreasing the performance collectively. Yet this experiment showed that even when the number of channels and the stimulation range length are kept the same, the performance changes just by placing the array at different depths from the round window, and at each insertion the performance drops with increasing expansion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

8.2.2.1. EXPERIMENT 2.1: Apical Mismatch with Electrodes 6-11

In Experiment 6 in Chapter 7 the basal end of the map was always matched and compression and expansion was applied at the apical end only of the stimulation region of electrodes 4-9. Similarly, we had an array of 6 electrodes in this experiment, yet the cochlear location of the array was changed by activating electrodes 6-11 instead of 4-9. The same apical mismatch conditions were tested with electrodes 6-11, which simulate an array of 6 electrodes inserted up to 19.2 mm. The results from both electrode settings were compared to observe the effect of the cochlear location of stimulation only.

The array at this cochlear location covers the cochlear range from 7.2 mm to 19.2 mm from the round window, which translates to an acoustic input range of 1-9k Hz, as it was shown in Table 8.2. The high frequency end of the stimulation region, i.e. 9 kHz, is too high to contain much useful speech spectrum, and most useful frequencies are more towards the apical end. As a result of this mismatch on the apical end such apical mismatch is expected to give same performance as the mismatch applied at both ends.

The experimental conditions are shown for this electrode array in Figure 8.16.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

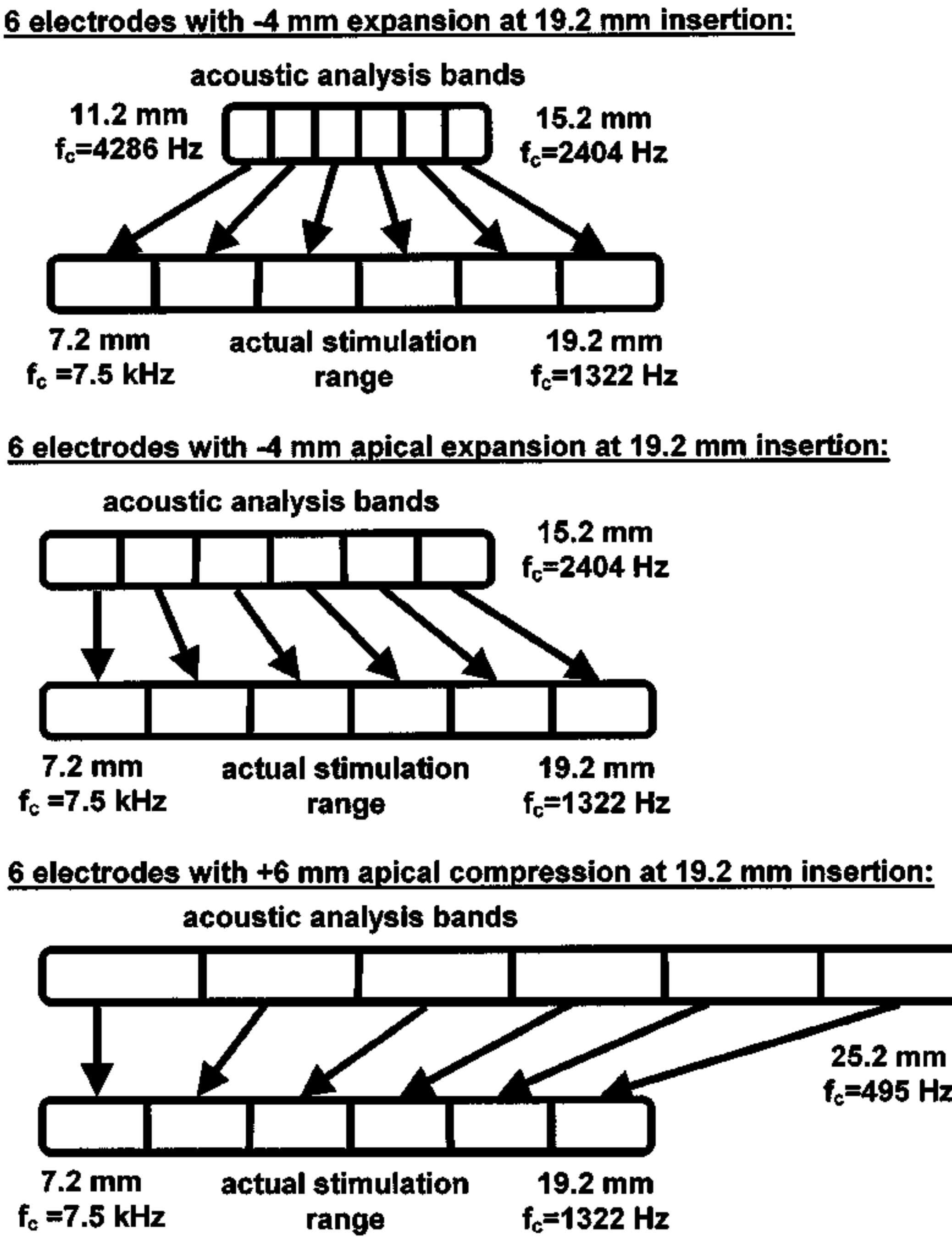


Figure 8.16: 6 electrodes at 19.2 mm insertion depth are simulated by activating electrodes 6-11. In this figure, the array is mapped with -4 mm expansion at both ends, -4 mm expansion on apical end only, and +6 mm compression on apical end only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

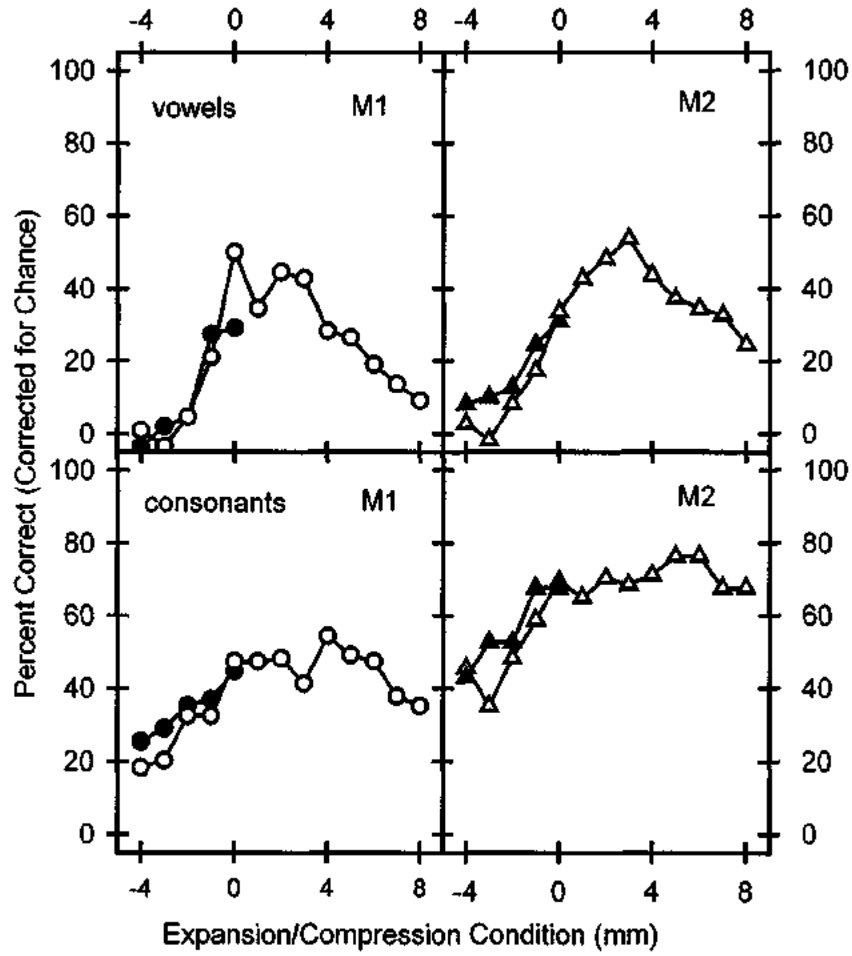


Figure 8.17: Individual percent correct scores of patients M1 and M2 with frequency-place mismatch applied at both ends of stimulation (shown by filled symbols) and on apical end only (shown by open symbols), with electrodes 6-11 activated.

The individual percent correct scores from patients M1 and M2 are shown in Figure 8.17. The filled symbols are the scores from previous experiment where frequencies were expanded at both ends of stimulation of the array at 19.2 mm insertion, as shown on the top row of Figure 8.16. The open symbols are the scores obtained when the frequency-place map was expanded or compressed at apical end only while the basal end was always

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

matched, as shown in the middle and bottom rows of Figure 8.16. Because electrodes 6-11 are at a very basal location in the whole array of 12 electrodes there is a big spectral space for compression, therefore a higher degree of compression could be applied. As expected expansion on apical end only and expansion on both ends both resulted in the same performance. The best performance was not obtained with 0 mm matched condition, but instead with a compression of a few mm. As it was also shown in Experiment 1.1, there is an optimum region for best speech perception at this relatively shallow insertion depth.

8.2.2.2. EXPERIMENT 2.2: Basal Mapping with Electrodes 2-7

In this experiment, electrodes 2-7 were activated to simulate the same array of 6 electrodes at a deeper insertion. Since such an array is located at low-frequency end of the widest frequency range that the device can handle, the map could be compressed only at basal end of stimulation. Electrodes 2-7 cover the cochlear range of 16.8-28.8 mm from the round window with an analysis band range of 200-2.3k Hz. Even the high end of the stimulation region is still low enough to carry important acoustical information, therefore this time we expect to see an effect of expansion or compression from basal end only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

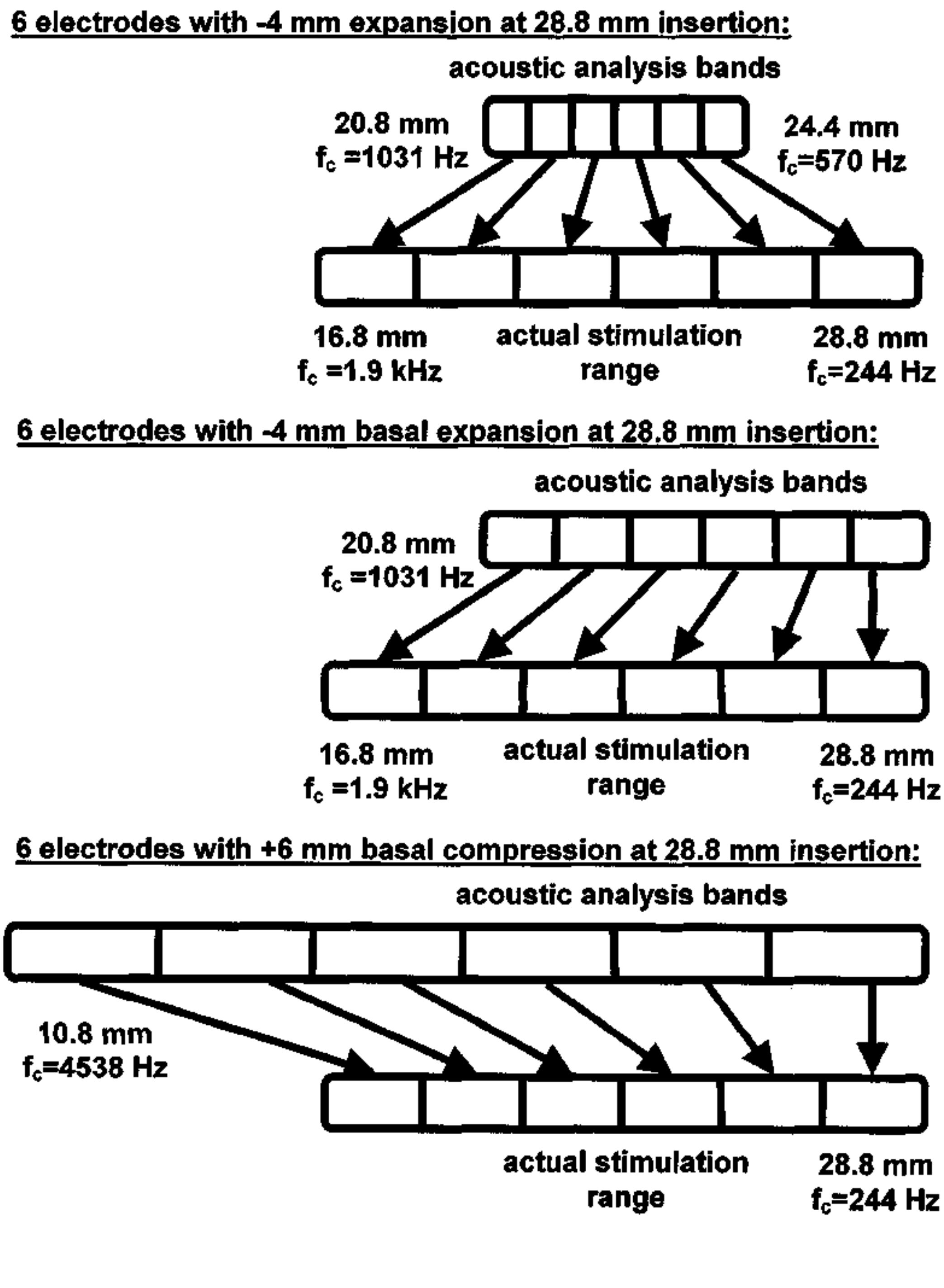


Figure 8.18: 6 electrodes at 28.8 mm insertion depth are simulated by activating electrodes 2-7. In this figure, the array is mapped with -4 mm expansion at both ends, -4 mm expansion on basal end only, and +6 mm compression on basal end only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

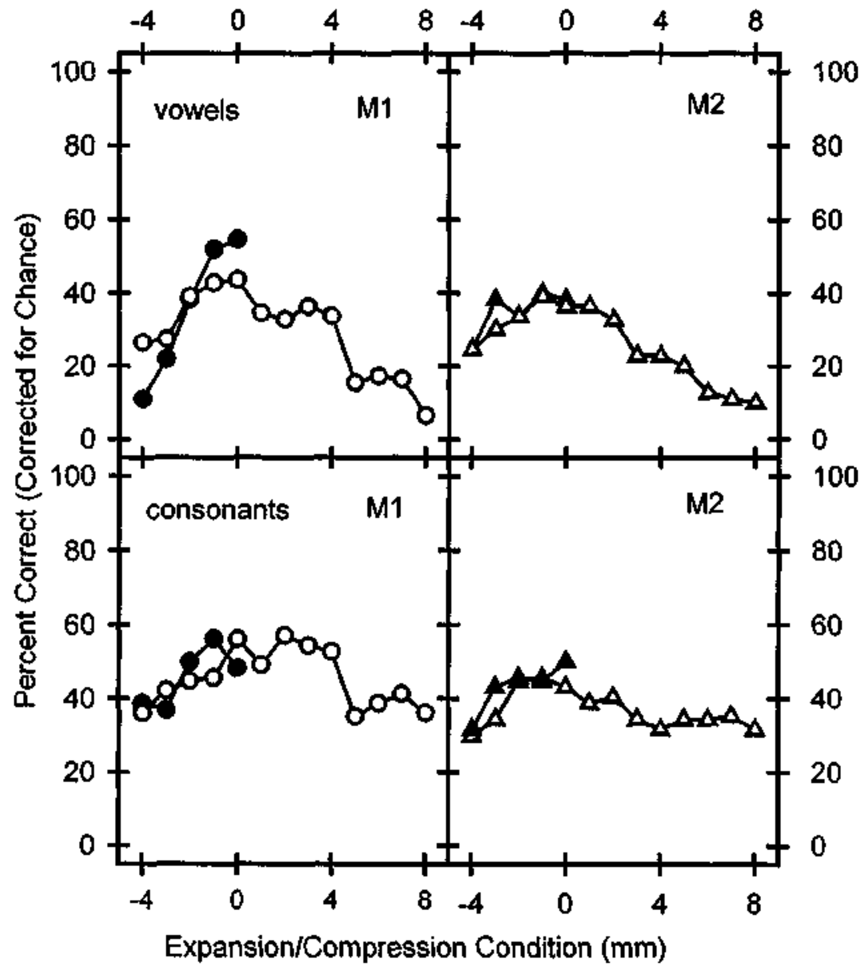


Figure 8.19: Individual percent correct scores of patients M1 and M2 with frequency-place mismatch applied at both ends of stimulation (shown by filled symbols) and on basal end only (shown by open symbols), with electrodes 2-7 activated.

The results from patients M1 and M2 are shown in Figure 8.19. The filled symbols are the scores from Experiment 2 where frequencies were expanded at both ends of stimulation of the array at 28.8 mm insertion, as shown on the top row of Figure 8.18. The open symbols are the scores obtained when the frequency-place map was expanded or compressed at

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

basal end only while the apical end was always matched, as shown in the middle and bottom rows of Figure 8.18. Because electrodes 2-7 are at an apical location a higher degree of compression could be applied. As expected expansion at basal end only produced similar performance to expansion at both ends. The drop was smaller with basal compression compared to expansion.

The overall results from this chapter show that it becomes more important to assign the appropriate acoustic input frequency range onto the electrodes when the settings are not optimal, such as implants with partially or shallowly inserted electrode arrays. The experiments showed that performance can change significantly with appropriate frequency-place mapping.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 9

SUMMARY OF RESULTS

1. Frequency-Place Compression and Expansion at Both Ends of Stimulation Range:

   Both with normal-hearing and implant subjects, the best vowel and sentence recognition performance was obtained when the analysis range was matched to the stimulation range, with a tolerance range of ±2 mm. Consonants were less affected than vowels and sentences.

   With implants it was further shown that the choice for the frequency range of the acoustic input with the clinical fitting program has a significant effect on speech perception.

2. Frequency-Place Compression and Expansion at Apical End of Stimulation Range:

   When the mismatch conditions were applied on the apical end only while the basal end was matched the percent correct scores were similar to mismatch conditions applied at both ends in simulations. The results were not as clear with implants. The overall results imply that it is more beneficial to match the frequencies contributing most to speech intelligibility (500 Hz - 3 kHz).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Frequency-Place Compression and Expansion at Basal End of Stimulation Range:

When the mismatch conditions were applied on the basal end only while the apical end was matched the performance did not change significantly in simulations. This observation supports the idea that it is more beneficial to match the frequencies contributing most to speech recognition and the high frequency basal end contributes little.

4. Frequency-Place Compression and Expansion in Noise:

Additive noise decreased all speech recognition scores, but did not change the overall pattern of results.

5. Holes in Hearing:

When tonotopic holes were simulated it was observed that the performance was generally better when the spectral information was completely deleted from the hole region and the remaining frequencies were mapped to matching locations, compared to the conditions where the spectral information was redistributed around the hole.

6. Frequency-Place Compression and Expansion with Shifted Electrodes:

Frequency-place matching is advantageous for both compression and expansion but one has to know the actual electrode location in the cochlea to achieve the best matched map. Simulations showed that if

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the electrodes are located deeper than the assumed location, compression gives the best performance and if the electrodes are located shallower than the assumed location, matched map keeps its superiority. This difference is explained by the importance of matching the most important frequencies for speech.

With implants, the acoustic range was held constant at mid-frequency range while the simulated array was shifted, simulating an error in estimated electrode location. A sharp drop was observed in both vowel and consonant recognition scores as the array was moved from the matched location.

In a similar experiment, the same number of electrodes was activated at three different insertion depths, all deeper than 19 mm. At all locations peak performance was around the map that best matched frequencies to the correct cochlear place.

7. The Effect of Insertion Depth:

When shallower insertions were simulated by turning off apical electrodes of implants, speech performance dropped sharply. For relatively deeper insertions the map that matched the acoustic input range to the stimulation region in the cochlea resulted in better vowel and sentence recognition over the map that compressed a wide acoustic input range onto the stimulation region. For shallower insertions some subjects had slightly better consonant recognition with a compressed map.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

At 19.2 mm and 16.8 mm insertion depths, it was shown that a compromise map between the matched and compressed maps can yield optimum performance. Such a map has little more acoustic information compared to matched map, but the distortion inherent in such a mild mismatch is not too detrimental.

8. Effect of Cochlear Length:

Results from three implant subjects implied that they might have cochleae that are 1-2 mm shorter than the assumed length of 35 mm. Yet due to the tolerance of the speech pattern recognition system in brain, still a matched map that produced good speech performance could be obtained.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 10

DISCUSSION

The best speech recognition performance was achieved in both acoustic simulations and implant experiments when the acoustic input frequency range was matched to its normal cochlear place. In most conditions, altering the frequency-place mapping more than a few mm by either compression or expansion resulted in poorer speech recognition. It was also observed that the place of stimulation in the cochlea affected the speech recognition. Better performance was obtained when the stimulation range was near the frequencies contributing to speech information most. With the implants it was shown that finding an optimum map (by compromising information loss and matching) is more important for diverse implant settings such as partial (or shallow) insertions.

These results provide an insight to how the pattern recognition mechanism in the central nervous system works, as well as to how we possibly can use that mechanism to our advantage to provide implant users with better maps for more efficient speech recognition. The significance of the results will be discussed in following sections.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

10.1. IMPLICATIONS FOR SPEECH PATTERN RECOGNITION

The present results illustrate the limitations of central pattern recognition mechanisms for speech, which may provide insights into the critical parameters of the pattern storage and retrieval process of central nervous system. For example, even though it is known that implant patients adapt over time and improve their understanding of speech, it is still unclear how much plasticity exists in these central nervous system mechanisms and whether the ability to accommodate to some types of alterations (e.g. frequency-place shift) might be easier than other types of alterations (e.g., nonlinear frequency-place distortion).

The pattern of results observed in the present experiment, when combined with previous results on frequency-place shifting (Fu and Shannon, 1999; Dorman *et al.*, 1997), warping (Shannon *et al.*, 1998), and frequency lowering (Braida *et al.*, 1979; Reed *et al.*, 1983) suggest that the central pattern recognition of speech is not stored in terms of an abstract pattern, but in terms of an absolute pattern. Speech recognition in healthy acoustic hearing can tolerate a small degree of distortion in this frequency-place pattern, probably to accommodate the natural range of variation in real-world listening conditions, e.g., differences in the gender of the talkers, talking speeds and styles, and different amounts of masking and interference in the listening environment. The results of the present study, combined with the results of previous studies on frequency-place distortions, suggest that speech patterns can tolerate only a relatively small amount of distortion (2-3 mm) in tonotopic space. If the peripheral representation of the pattern of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

speech information is shifted, warped, expanded or compressed beyond this tolerated cochlear distance of 2-3 mm, speech recognition will be significantly reduced. (It should be noted that none of these studies gave the subjects the opportunity to adapt to the distorted mappings.)

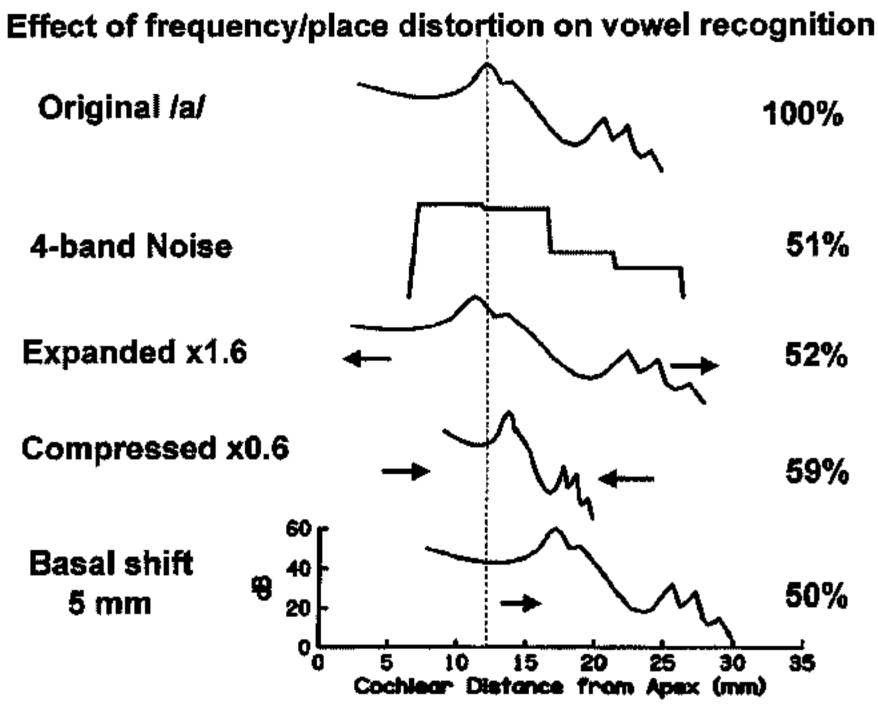


Figure 10.1: A review of various frequency-place distortions on vowel recognition. The top curve is the original spectral representation of the vowel /i/ plotted in terms of cochlear distance. The second curve shows the same vowel with a 5 mm apical hole. Next is the vowel represented by a four-band noise vocoder. The fourth curve is the frequency-place expansion with an expansion factor of 1.6 (10 mm extent expanded to 16 mm), followed by the frequency-place compression with a compression factor of 0.6 (26 mm extent compressed to 16 mm). Finally, the bottom curve shows the spectrum of an /i/ that has been shifted by 5 mm basally in the cochlea.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 10.1 presents a schematic representation of a vowel spectrum and the various types of distortion that result in a reduction in multi-talker vowel recognition to approximately 50% correct. These representations will be altered as they are processed by the cochlea and the central nervous system, but for simplicity they are shown here by their physical spectral representation in terms of distance along the cochlea. The top curve shows the original spectrum of the vowel /i/, presented in terms of mm along the cochlea. In this undistorted representation listeners will generally be able to identify 12 vowels at nearly 100% correct, even with multiple talkers and with the spectral resolution reduced to 16 channels. The second curve shows the same vowel in which spectral information has been removed to create a 5-mm hole in the apical spectral region, resulting in a drop to 50% correct recognition (Shannon *et al.*, 2001, also shown in Figure 5.24). The third curve shows the vowel represented by a three-band noise vocoder, which allows 46% correct on multi-talker vowel recognition (Fu *et al.*, 1998). The fourth curve shows the effect of a frequency-place expansion by a factor of 1.6, which results in 54% correct vowel recognition with 16 bands (-3 mm condition from Figure 4.8). The fifth curve shows the effect of frequency-place compression by a factor of 0.6, which results in 57% correct recognition (+5 mm condition from Figure 4.8). And the bottom curve shows the spectrum of an /i/ that has been shifted 5 mm basally in the cochlea, resulting in 44% correct vowel recognition (from Fu and Shannon, 1999). This comparison suggests that the central pattern recognition mechanisms are sensitive to the absolute tonotopic location of the cochlear pattern. If the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

frequency-place information is in the correct location, the central pattern recognition can tolerate the loss of a full octave of spectral information in the critical low-frequency region, or an extreme loss of spectral resolution – down to three bands. However, if the pattern is distorted by a frequency-place shift or compression or expansion, then speech recognition is impaired even with good spectral resolution. In terms of cochlear implants, even if an implant patient is able to use many electrodes effectively, their performance might be limited by distortion in the frequency-place mapping.

10.2. IMPLICATIONS FOR COCHLEAR IMPLANTS

The present study observed a reduction in vowel and sentence recognition when a frequency range was compressively mapped onto a smaller cochlear range, with both normal hearing and implant subjects. Even though a broader frequency range of acoustic information is presented in this condition, performance was reduced due to the distortion in the frequency-place assignment. This compressive frequency-to-place mapping is similar to the mapping used in Nucleus cochlear implant systems, in which the acoustic frequency range of 150 Hz to 10 kHz is typically mapped onto electrodes that occupy the cochlear locations that normally respond to an acoustic range of only 500-6000 Hz. This result implies that speech recognition in cochlear implants might be improved by as much as 20% if the frequency range for each electrode could be mapped according to the normal acoustic characteristic frequency of that cochlear location.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

How can a cochlear implant speech processor be adjusted to achieve the best mapping of frequency information onto the most appropriate cochlear place, given the variability in cochlear length and electrode insertion depth across patients? In implant listeners there is uncertainty in the exact location of the electrodes and further uncertainty as to the location of the stimulated neurons. Recent advances in imaging technology allow sufficient resolution to evaluate the depth of electrode insertion and to detect the presence of any kinks or abnormalities in the electrode carrier (Ketten *et al.*, 1998). However, these imaging procedures are costly, time consuming, deliver large doses of radiation, and may still not provide all of the necessary information. For example, even knowledge of the exact cochlear location of an electrode is no guarantee that the stimulation of neurons is actually occurring at that location. The actual stimulation location can be affected by the pattern of local nerve survival or by unusual current pathways due to bone growth and fibrous blockage. In addition, the actual site of stimulation may be in the spiral ganglion, whereas Greenwood's formula holds for stimulation at the basilar membrane. These factors produce additional uncertainties regarding the appropriate frequency-place mapping in implant patients.

Another uncertainty comes from anatomical and geometrical issues regarding the stimulation of deeper turns of the cochlea. Electrodes in cochlear implants do not always reside between 9 and 25 mm or between 4 and 20 mm inside the round window, the two conditions simulated in this study. The latest generation of CI electrodes, such as Clarion HiFocus,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Nucleus Contour, or Med-El Combi40+ offers deeper insertion, possibly up to 30 mm. Even though these specially designed electrodes make it possible to reach more apical locations inside the cochlea it is unknown if it is possible to stimulate the spiral ganglia corresponding to low frequencies. Cell bodies of the spiral ganglia from the apical turn of the cochlea are located in the modiolus of the cochlear middle turn, and so are physically (and presumably electrically) closer to electrodes in the middle turn than to the medial wall of the cochlea in the apical turn. Studies of pitch have shown little change in pitch with electrode location for electrodes that were deeply inserted into the apical turn, suggesting that there may be a point of diminishing returns in terms of electrode insertion depth (Cohen *et al.*, 1996). And even with the new electrode designs, the array cannot always be fully inserted due to cochlear ossification or otosclerosis. Thus, the actual location of the implanted electrode is difficult to determine accurately, and the location of the neurons actually stimulated by each electrode adds a further layer of uncertainty.

In simulations, such factors as actual stimulation locations can fully be controlled, at least within the constraints of the normal acoustic spread of excitation. In the experiments with implants this is not possible. Usually, the only information available is the estimated locations for the electrode arrays. We used these estimations to match the frequencies. Almost all patients performed best with matched maps but there was a large variation in the patterns of performances. This might partly be caused by such variations in the electrode array locations, nerve survival patterns, and cochlear lengths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

However the fact that we could see a peak performance with so many unknowns shows that this might be a good starting point for the fitting process. From this step, the optimal frequency-place alignment might best be determined functionally. A simple optimizing algorithm, paired with a sensitive phonetic contrast test could provide an efficient method for converging on an optimal frequency-place mapping for an individual patient, without the costs and risks of x-rays and CT scans. The present results may help to define the inherent trade-offs between electrode array insertion depth, number of electrodes, and frequency range.

The findings of this study are particularly important for diverse implant conditions such as partial or shallow insertions. When patients have long electrode arrays with deep insertions or arrays positioned closer to the inner wall which virtually behave like long arrays they can access most important frequencies for speech which are also mapped on matching stimulation locations. However, partially inserted arrays limit the stimulation range in the cochlea. This study showed that for relatively deeper insertions matched map was still advantageous. On the other hand, for shallower insertions a map with a compromise between information loss and compression resulted in the best speech recognition.

Another interesting implant case is the combined electric-acoustic hearing by implant patients who still have substantial residual hearing. With these patients, electrode array is usually inserted only shallowly in an attempt to preserve any residual acoustic hearing. Similar to partial insertions, in these special cases too it may be particularly important to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

assign the appropriate frequency-place mapping, because the electrically stimulated hearing must combine with residual acoustic hearing. Indeed, preliminary results (Turner and Gantz, 2001; Brill *et al.*, 2001) suggest that combined electric and acoustic hearing is best when the electrodes in the basal turn receive high frequency information that is matched to their tonotopic location.

Overall, it is possible that distortion in frequency-place mapping is responsible for at least part of the variability in performance across implant patients. If this is the case, then adjustments to the speech processor to produce a better match in frequency-place mapping may produce improvements in speech recognition.

10.3. TRADE-OFF BETWEEN SPECTRAL RESOLUTION AND OVERALL BANDWIDTH

Some of the present results indicate a trade-off between spectral resolution (number of bands) and overall bandwidth. For a given number of bands there may be an optimal bandwidth – too small a bandwidth would discard too much important speech information, and too wide a bandwidth would increase the frequency range of each band, reducing the relative resolution. Consider the baseline conditions from Chapter 4. In these conditions the analysis bands were always matched in frequency to the carrier bands, while the number of bands was held constant. For the 16-band conditions, a larger bandwidth generally produced better performance whereas for 8 bands the performance was generally unchanged as the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

overall bandwidth was increased relative to the standard matched condition (0 mm). However, when there were only 4 bands of spectral resolution available, there was a complex interaction between the bandwidth and spectral resolution. In many cases, performance dropped both when the bandwidth was increased and decreased relative to the standard matched condition. As bandwidth decreased the relative spectral resolution increased, but this was not sufficient to offset the loss of information. As bandwidth increased the additional spectral information was offset by the loss of relative spectral resolution (e.g., Figures 4.4, 4.7, and 4.8: compare 4 channel 0 mm and +5mm baseline conditions), resulting in poorer performance in spite of the larger bandwidth.

An expansion in the frequency-place mapping could theoretically improve speech recognition by spreading out the critical speech spectral region to a larger range in the cochlea. Echolocating animals have evolved such a strategy to provide better tonotopic resolution in the small frequency region of their echo signal. However, in the present study such expansion conditions mostly resulted in poorer speech recognition. The only exception was for the 4-channel processor with a 25 mm simulated insertion depth. This was likely due to an artifact of band edge placement: the 0 mm condition contained no band division in between 999 and 1843 Hz (see Table 4.3), while the -5 mm expansion condition contained a band division at 1472 Hz, which is an important frequency for distinguishing high from low second formant frequencies. In this particular condition, the contribution from this better frequency partition might have compensated for the loss of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

bandwidth. With a limited number of bands, the placement of the frequency divisions appears to be more important than the overall frequency range. Alternatively, the slight improvement in performance in the expansion condition could be due to the improved resolution in this condition. The small frequency range of 1168-2864 Hz was represented across a larger cochlear region that would have normally responded to a range of 513-5860 Hz. This expansive mapping may have helped recognition by stimulating a larger neural population with information from the smaller frequency range. Whichever explanation is correct (better band partition or expanded representation) the same effects were not observed with more than 4 bands.

10.4. POTENTIAL EFFECTS OF LEARNING

One aspect of speech pattern recognition not addressed by the present study is the potential effect of learning. Recent work has demonstrated that NH subjects listening to simulations of cochlear implants can improve their scores on speech recognition with only a modest amount of practice (Rosen *et al.*, 1999). Rosen *et al.* used noise-band vocoders in which the frequency-place mapping was shifted basally by as much as 6.5 mm. Listeners improved significantly in their ability to recognize phonemes and words with these shifted representations after only a few hours of training. However, their performance after this limited amount of training was still far poorer than their recognition with the unshifted speech. It is not clear if further training would allow complete recovery of performance to the unshifted levels. Fu *et al.* (2002) measured speech recognition in three

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cochlear implant listeners after a 3 mm apical shift in the frequency-to-electrode assignments. Initially, speech recognition was reduced dramatically. After 10 days of everyday experience with the shifted map there was a significant improvement in recognition, but then only little further improvement was observed over the next three months. This result suggests that there may be a limit to the amount of possible relearning. It is not clear if listeners would be able to adapt to a frequency-place compression or expansion over time. In the present experiments the emphasis was solely on speech pattern recognition with no practice or time for accommodation.

This dissertation parametrically explored the effects of mismatch in frequency-place mapping on speech recognition. The results have theoretical value since they help us understand the pattern recognition mechanisms in the brain, as well as practical value since they will also help with how to use these mechanisms effectively for implant users. By a series of experiments many different implant settings were simulated with both normal hearing and implant subjects, to fully understand the effects of many parameters such as the range of input frequencies, number and placement of the electrodes, background noise, spectral holes, length of cochleae. As a result, these findings can be used as guidelines achieving an optimum frequency-place distribution for implant patients.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REFERENCES

Advanced Bionics Corporation (2001). CII Bionic Ear Programming System.

American National Standards Institute (1997). Methods for Calculation of the Speech Intelligibility Index, S35.

Baskent, D. and Shannon, R. V. (2003). "Speech recognition under conditions of frequency-place compression and expansion," J. Acoust. Soc. Amer., 113(4), 2064-2076.

Blamey, P.J., Pyman, B.C., Clark, G.M., Dowell, R.C., Gordon, M., Brown, A.M., and Hollow, R.D. (1992). "Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients," Ann. of Otol., Rhinol., & Laryngol., 101(4), 342-348.

Braida, L.D., Durlach, N.I., Lippmann, R.P., Hicks, B.L., Rabinowitz, W.M., and Reed, C.M. (1979). "Hearing aids: a review of past research on linear amplification, amplitude compression, and frequency lowering," ASHA Mono., No. 19, 87-108.

Brill, S., Lawson, D., Wolford, R., Wilson, B., and Schatzer, R. (2001). Speech Processors for Auditory Prostheses, 11th Quarterly Progress Report, NIH Contract N01-DC-8-2105.

Cochlear Corporation (1995). Technical Reference Manual. Englewood, Colorado.

Cohen, L.T., Xu, J., Xu, S.A., and Clark, G.M. (1996a). "Improved and simplified methods for specifying positions of the electrode bands of a cochlear implant array," Am. J. of Otol., 17, 859-865.

Cohen, L.T., Busby, P.A., Whitford, L.A. and Clark, G.M. (1996b). "Cochlear implant place psychophysics: 1. Pitch estimation with deeply inserted electrodes", Audiol. Neuro-Otol., 1, 265-277.

Cosendai, G. and Pelizzone, M., (2001). "Effects of the acoustical dynamic range on speech recognition with cochlear implants," Audiology, 40, 272-281.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Daniloff, R.G., Shiner, T.H., and Zemlin, W.R. (1968). "Intelligibility of vowels altered in duration and frequency," J. Acoust. Soc. Amer., 44, 700-707.

Dorman, M.F., Loizou, P.C. and Rainey, D. (1997). "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," J. Acoust. Soc. Amer., 102(5), 2993-2996.

Echteler, S.M., Fay, R.R., and Popper, A.N. (1994). "Structure of the mammalian cochlea," Comparative Hearing: Mammals, ed. Fay, R.R. and Popper, A.N., Springer-Verlag, NY, 134-171.

Eyles, J.A., Boyle, P.J., and Burton, M.J. (1996). "Characteristic frequency mapping in subjects using the nucleus 22-channel cochlear implant system with partial and full insertions," Ann. of Otol., Rhinol., & Laryngol., Suppl. 116, 356-358.

Faulkner, A., Rosen, S. and Stanton, D. (2003). "Simulations of tonotopically-mapped speech processors for cochlear implant electrodes varying in insertion depth," J. Acoust. Soc. Amer., 113(2), 1073-1080.

Friesen, L. M., Shannon, R.V. and Baskent, D. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Amer., 110(2), 1150-1163.

Fu, Q.-J, Shannon R.V., and Wang, X. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," J. Acoust. Soc. Amer., 104(6), 3586-3596.

Fu, Q.-J. and Shannon, R.V. (1999). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," J. Acoust. Soc. Amer., 105(3), 1889-1900.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fu, Q.-J., Shannon, R.V., and Galvin, J. (2002). "Performance over time of Nucleus-22 cochlear implant listeners wearing a speech processor with a shifted frequency-to-electrode assignment," J. Acoust. Soc. Amer., 112(4), 1664-1674.

Greenwood, D.D. (1990). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Amer., 87(6), 2592-2605.

Gstoettner, W., Franz, P., Hamzavi, J., Plenk, H., Baumgartner, W., and Czerny, C. (1999). "Intracochlear position of cochlear implant electrodes," Acta Otolaryngol. (Norway), 119, 229-33.

Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Amer., 97(5), 3099-3111.

Hochmair, I., Kerber M., and Jolly, C. (2000). "Technological Challenges with the MED-EL Cochlear Implants," 2nd MEDICS Workshop Abstracts, Sulzbach, Germany.

Hodges AV, Villasuso E, Balkany T, Bird PA, Butts S, Lee D, and Gomez O (1999). "Hearing results with deep insertion of cochlear implant electrodes"",Am J of Otol, 20, 53-55.

Institute of Electrical and Electronics Engineers (1969). IEEE recommended practice for speech quality measurements.

Ketten, D.R., Skinner, M.W., Wang, G., Vannier, M.W., Gates, G.A., and Neely, J.G. (1998). "In vivo measures of cochlear length and insertion depth of Nucleus cochlear implant electrode arrays," Ann. Otol. Rhinol. Laryngol., Suppl. 175, 1-16.

Loeb, G.E. (1990). "Cochlear prosthetics," Ann. Rev. in Neurosc., 13, 357-371.

Loizou, P.C. (1998). "Mimicking the human ear: An overview of signal-processing strategies for converting sound into electrical signals in cochlear implants," IEEE Signal Proc. Mag., 15(5), 101-130.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Loizou, P.C. (1999). "Signal-processing techniques for cochlear implants: A review of progress in deriving electrical stimuli from the speech signal," IEEE Eng. in Med. and Biol. Mag., 18(3), 34-46.

Marsh, M.A., Xu, J., Blamey, P.J., Whitford, L.A., Xu, S.A., Silverman, J.M., and Clark, G.M. (1993). "Radiologic evaluation of multichannel intracochlear implant insertion depth," Amer. J. of Otol., 14(4), 386-391.

Miller, G.A. and Nicely, P.E. (1955). "An analysis of perceptual confusions among some english consonants," J. Acoust. Soc. Amer., 27(2), 338-352.

Nagafuchi, M. (1976). Intelligibility of distorted speech sounds shifted in frequency and time in normal children, Audiology, 15, 326-337.

National Institute of Standards and Technology (1990). DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus CD-ROM. Gaithersburg, MD.

Nilsson, M., Soli, S., and Sullivan, J. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise," J. Acoust. Soc. Amer., 95(2), 1085-1099.

Reed, C.M., Hicks, B.L., Braida, L.D., and Durlach, N.I. (1983). "Discrimination of speech processed by low-pass filtering and pitch-invariant frequency lowering," J. Acoust. Soc. Amer., 74(2), 409-419.

Robert, ME. (1998). CONDOR: Documentation for Identification Test Program. Los Angeles; House Ear Institute.

Rosen, S., Faulkner, A, and Wilkinson, L. (1999). "Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants," J. Acoust. Soc. Amer., 106(6), 3629-3636.

Skinner, M.W., Holden, L.K., Holden, T.A., (1997), "Parameter selection to optimize speech recognition with the Nucleus implant," Otolaryngol. Head Neck Surg., 117(3 Pt 1), 188-195.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Skinner, M.W., Ketten, D.R., Holden, L.K., Harding, G.W., Smith, P.G., Gates, G.A., Neely, J.G., Kletzker, G.R., Brunsden, B., and Blocker, B. (2003), "CT-Derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus 22 recipients," J. Assoc. Res. Otolaryngol., 3(3), 332-350.

Shannon, R.V., Zeng, F-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," Science, 270, 303-304.

Shannon, R.V., Zeng, F.-G., and Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues," J. Acoust. Soc. Amer., 104(4), 2467-2476.

Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., and Wang, X. (1999). "Consonant recordings for speech testing," J. Acoust. Soc. Amer., L71-L74.

Shannon, R.V., Galvin, J.G. III, and Baskent, D. (2001). "Holes in hearing," J. Assoc. Res. Otolaryngol., 3(2), 185-199.

Spelman, F.A. (1999). "The past, present, and future of cochlear prostheses- Accomplishments and challenges in treating sensorineural deafness through electrical stimulation," IEEE Eng. in Med. and Biol. Mag., 18(3), 27-33.

Stobich B., Zierhofer C.M., and Hochmair, E.S., (1999). "Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy," Ear Hear, 20(2), 104-116.

Tiffany, W.R. and Bennett, D.A. (1961). "Intelligibility of slow-played speech," J. Speech Hearing Res., 4, 248-258.

Turner, C.W., Fabry, D.A., Barrett, S., and Horowitz, A.R. (1992). "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," J. Speech Hearing Res., 35, 942-949.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Turner, C.W., Chi, S.-L., and Flock, S. (1999). "Limiting spectral resolution in speech for listeners with sensorineural hearing loss," J. Speech Hearing Res., 42(4), 773-784.

Turner, C. and Gantz, B. (2001). "Combining acoustic and electric hearing for patients with high frequency hearing loss," Abstracts of the 2001 Conference on Implantable Auditory Prostheses, Asilomar, CA, p 33.

Ulehlova, L., Voldrich, L., and Janisch, R (1987). "Correlative study if sensory cell density and cohlear length in humans," Hear. Res., 28, 149-151.

Van Tasell, D.J., Soli, S.D., Kirby, V.M., and Widin, G.P. (1987). "Speech waveform envelope cues for consonant recognition", J. Acoust. Soc. Amer., 82, 1152-1161.

Whitford, L.A., Seligman, P.M., Blamey, P.J., McDermott, H.J., and Patrick, J.F. (1993). "Comparison of current speech coding strategies," Adv. Otorhinolaryngol., 48, 85-90.

Wilson, B., Finley, C., Laeson, D., Wolford, R., Eddington, D., and Rabinowitz, W. (1991). "Better speech recognition with cochlear implants," Nature, vol. 352, 236-238.

Zeng, F.-G., Grant, G., Niparko, J., Galvin, J., Shannon, R.V., Opie, J., and Segel, P. (2002). "Speech dynamic range and its effect on cochlear implant performance," J. Acoust. Soc. Am., 111(1), 377-386.

Zierhofer, C.M., Hochmair I.J., and Hochmair, E.S. (1997). "The advanced Combi 40+ cochlear implant," Amer. J. of Otol., 18, S37-S38.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.