RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 24 (2000)
Indiana University

Auditory Learning and Adaptation after Cochlear Implantation: A Preliminary Study of Discrimination and Labeling of Vowel Sounds by Cochlear Implant Users[1]

Mario A. Svirsky,[2,3,4] Alicia Silveira,[3] Hamlet Suarez,[3] Heidi Neuburger,[2] Ted T. Lai,[2] and Peter M. Simmons[2]

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This work has been supported by grants from NIH-NIDCD (R01-DC03937, Principal Investigator: M. Svirsky; and training grant T32-DC00012, Principal Investigator: D. Pisoni); the Deafness Research Foundation; the National Organization of Hearing Research; the American Academy of Otolaryngology-Head and Neck Surgery and from BID/CONICYT (Uruguay; Principal Investigators: H. Suarez and M. Svirsky).
[2] Department of Otolaryngology-Head & Neck Surgery, Indiana University School of Medicine, Indianapolis, IN.
[3] Facultad de Medicina y Hospital Maciel, Uruguay.
[4] Department of Biomedical Engineering, Purdue University.

Auditory Learning and Adaptation after Cochlear Implantation: A Preliminary Study of Discrimination and Labeling of Vowel Sounds by Cochlear Implant Users

Abstract. This study examined two possible reasons underlying longitudinal increases in vowel identification by cochlear implant users: improved labeling of vowel sounds and improved electrode discrimination. The Multidimensional Phoneme Identification (MPI) model was used to obtain ceiling estimates of vowel identification for each subject, given his electrode discrimination skills. Vowel identification scores were initially lower than the ceiling estimates, but they gradually approached them over the first few months post-implant. Taken together, the present results suggest that improved labeling is the main mechanism explaining post-implant increases in vowel identification.

Introduction

Cochlear implants (CI's) have been shown to be a safe and effective treatment for profound sensorineural deafness in postlingually deafened adults and in prelingually deaf children. It is not surprising that speech perception scores of the latter continue to increase many years after implantation, because these children must develop an oral language system and develop speech perception and production skills after receiving a CI. However, the full benefit that CI's provide to postlingually deafened adults is not instantaneous either. Normally, asymptotic speech perception performance (i.e., identification of vowels, consonants, words, sentences or running speech) is not apparent for at least several months or weeks after implantation. What are the reasons for this delay, and what is the nature of the underlying process that is indexed by increases in speech perception scores?

There are at least two kinds of reasons for this process. First, CI users may improve their *discrimination* skills over time. To understand speech, CI users must be able to discriminate sounds along relevant acoustic dimensions. For example, vowels are reasonably well characterized by the first two or three formant frequencies (Peterson & Barney, 1952; Hillenbrand, Getty, Clark, & Wheeler, 1995). For CI users, different formant frequencies result in stimulation of different electrodes and thus different cochlear locations. Consequently, the ability to discriminate stimulation delivered to different electrodes may be an important prerequisite to identifying vowels. Unfortunately, the number of intracochlear electrodes is relatively small (22 at most) compared to the number of discriminable steps in a normal hearing cochlea, a factor that limits the spectral information that CI users can receive. Further limitations are due to the fact that even with this relatively small number of electrodes, the ability of CI users to discriminate stimulation delivered to different electrodes is less than perfect. Given that electrode discrimination ability is likely a limiting factor in CI users' reception of spectral information, longitudinal improvement in this ability may underlie the observed improvement in, for example, vowel perception.

Second, CI users may improve their vowel identification over time due to improved *labeling* of speech sounds. The percepts elicited from a cochlear implant are different from normal acoustic hearing, and listeners may be initially unable to label the sounds they hear. In other words, different speech sounds may be distinct from each other (i.e., discrimination may not be the limiting factor in the identification of these sounds), but they may not sound the way listeners *expect* them to sound, leading to identification errors. The possibility of labeling errors is quite plausible, given the fact that cochlear implants do not stimulate the entire neural population of the cochlea but only the most basal 25 mm at best because the electrode array cannot be inserted completely into the cochlea. Therefore, cochlear implants may stimulate cochlear locations that are more basal and thus elicit higher pitched percepts than normal acoustic stimuli. For example, when the input speech signal has a low frequency peak (e.g. 300 Hz), the most apical electrode is stimulated. The neurons stimulated in response to this signal may have characteristic frequencies of 1000 Hz or even higher. This pronounced modification of the peripheral frequency map might lead to errors in identifying speech sounds, unless the auditory nervous system of CI users is adaptable enough to successfully "re-map" the place frequency code in the cochlea. This adaptation, however, may require weeks or months and may underlie the improvement in speech perception observed after cochlear implantation in postlingually deafened adults. It is important to remember that an individual listener may suffer both labeling and discrimination limitations to speech perception simultaneously.

The goal of this study was to assess the contribution of discrimination and labeling to vowel identification by Spanish-speaking cochlear implant users. To this end, vowel identification and electrode discrimination were assessed longitudinally over several months, starting the day of initial stimulation. The Multidimensional Phoneme Identification (MPI) model (Svirsky, Meyer, Kaiser, Basalo, Silveira, Suarez, Lai, & Simmons, 1999; Svirsky, 2000) was used to obtain "ceiling estimates" of vowel identification, representing the best performance a listener could achieve, given his electrode discrimination skills.

Materials and Methods

The subjects were seven cochlear implant users, all of them native Spanish speakers, implanted by Dr. Suárez in Montevideo, Uruguay. They all had postlingual, profound-to-total deafness. Six of them used the Nucleus-22 device while the remaining one used the Med-El Combi-40 device.

Vowel identification and electrode discrimination were measured up to 6.5 to 32 months after initial stimulation. The first testing session for 6 of the 7 subjects took place immediately after initial stimulation, before they heard any other speech sounds through the implant. Vowel identification testing was repeated at the end of the initial stimulation session for two of the seven subjects. The other subject (subject 4, who was the Med-El user) was tested for the first time 10 days after initial stimulation. Vowel identification testing was done by presenting each one of the five Spanish vowels ten times, in random order, in j-vowel-d context (where "j" indicates the Spanish velar fricative). Each one of the ten repetitions of each vowel was separately uttered, recorded and presented. The speaker was either a male or a female speaker of Uruguayan Spanish, whose utterances were recorded on a personal computer. Results were scored as the percentage of correct responses for the 50 stimuli. In one case (Subject 2, session done three months post-implant), both the male and the female vowels were presented, and the analyses described below were conducted separately for the male speaker and the female speaker datasets.

Electrode discrimination (or, equivalently, discrimination of place of stimulation in the cochlea) was assessed with a pitch-ranking task. Two adjacent electrodes were stimulated in sequence, for 500 ms each with a 500 ms pause in between, and the subject had to say which one of the two sounds was higher pitched. All sounds were presented at maximum comfortable level, and these levels were balanced for equal loudness prior to presentation. Due to the limited available testing time, only nine pairs of adjacent electrodes were tested in Nucleus-22 users: these were pairs 1-2, 2-3 and 3-4 (at the basal end of the electrode array); 9-10, 10-11 and 11-12 (in the middle of the array); and 17-18, 18-19 and 19-20 (at the apical end of the array). Each pair of electrodes was stimulated eight times in random order, with the more basal electrode being stimulated first about 50% of the time. The Med-El user had 11 active electrodes and in her case all pairs of adjacent electrodes were tested. Average d', an index of the ability to correctly pitch rank electrodes was calculated for each subject based on a procedure similar to that described by Levitt (1972). A d' of 0 indicates no discrimination, d'=1 indicates minimum discrimination ability, and a d' that is greater than 3 indicates near perfect discrimination.

A previous study (Svirsky et al., 1999) provided evidence that the relevant perceptual dimensions for vowel perception in Spanish were A2 (i.e., the amount of energy delivered to electrodes encoding frequencies in the second formant region), F1 and F2 (i.e., the centers of gravity for stimulation pulses delivered to electrodes encoding the F1 and F2 formant frequency ranges, respectively, weighted by pulse amplitude). The MPI model (for a full description see Svirsky, 2000) was used to obtain ceiling estimates of vowel identification, assuming that F1, F2 and A2 were indeed the perceptual dimensions used by all subjects. The MPI model generates a predicted confusion matrix based on a listener's just noticeable differences (JND's) along the relevant perceptual dimensions. In this study, the JND's for the F1 and F2 dimensions were estimated as the inverse of the d' values that were derived from the pitch ranking experiment. Because JND for the A2 dimension was not measured in this study, predictions of vowel perception scores were obtained for each individual using a wide range of JND values for A2. Normally, the MPI model can be used to estimate a listener's maximum possible vowel (or consonant) identification performance, given his JND's for all the relevant perceptual dimensions. Instead of calculating a single ceiling estimate, in this study we obtained a range of values where the actual maximum was expected to fall. These ranges sometimes changed over time, as the listener's pitch ranking skills (and therefore the estimates for F1 and F2 JND's) increased or decreased. The ceiling ranges were compared to each listener's actual vowel identification scores. When the vowel identification scores fall within the ceiling range (which is partly determined by the listener's pitch ranking skills), this suggests that the listener may be labeling vowels in an optimal fashion, and vowel identification is only limited by his ability to discriminate these speech sounds. Conversely, when vowel identification scores are substantially below the ceiling range predicted by the MPI model, this represents strong evidence that the listener is not using all the acoustic information that is available to him, possibly due to limitations in vowel labeling.

Results

Table I shows electrode discrimination scores (d') for all subjects as a function of time. The first four subjects did not show any systematic improvement after the first testing session. In fact, Subject 1 showed a substantial decrease in d' between the day of initial stimulation and a second testing session seven months later. Subjects 5, 6, and 7, on the other hand, showed better discrimination in later testing sessions than they did on the day of initial stimulation.

Although a detailed description exceeds the scope of this manuscript, it should be noted that the MPI model provided good fits to the subjects' confusion matrices that were obtained at least a few months after initial stimulation. In other words, the model was able to predict which vowel pairs would be confused by the subjects as well as which vowel pairs would not be confused. In addition, the model predicted that, for given levels of frequency and amplitude discrimination, vowel identification scores would be higher when the female speaker was used than when the male speaker was used. This is precisely what happened during the three-month post-implant session for Subject 2, the only session where both the male and the female vowels were administered (see the top right panel of Fig. 1). These results provide some validation for the choice of dimensions employed in this study and for the MPI model itself, validating the use of ceiling estimates obtained with it.

Subject	Time after initial stimulation (months)	Average electrode discrimination (d')
1	0	2.3
	7	0.76
2	0	1.2
	3	0.94
	20	0.86
3	0	0.69
	0.1	1.02
	11	0.35
	25	0.64
4	0.3	3.21
	15	2.30
5	0	0.65
	3	2.17
	15	1.94
6	0	0.56
	4	1.75
7	0	0.56
	12	1.18

Table 1. Average electrode discrimination as a function of time. For subjects 1–4, discrimination did not improve after the day of initial stimulation (and it even decreased in some cases), but subjects 5, 6 and 7 did show better electrode discrimination when they were tested a few months after initial stimulation.

Figure 1 shows the vowel identification scores for subjects 1-4, as a function of time. Vertical bars indicate the ceiling ranges for subjects 1-3, as predicted by the MPI model. Predictions were not obtained for subject 4, who used a different device than the other subjects. The day of initial stimulation, subjects 1-3 obtained scores that were substantially lower than the ceiling ranges, but within 2-5 months they all reached vowel scores that were within, or quite close to the ceiling ranges.

Figure 2 shows similar data for subjects 5-7. Because in these cases d' increased with time, the ceiling ranges increased accordingly as a function of time. However, subjects 5 and 6 showed the same pattern as subjects 1-3, failing to reach their ceiling ranges at initial stimulation but reaching those ranges 3-6 months later. Subject 7 was the only one whose initial vowel scores were within his ceiling range at initial stimulation.

Subjects 1 and 5, who were tested at the beginning and at the end of the initial stimulation session, showed a marked increase in vowel identification during the session. Both scored only 28% correct immediately after the implant was turned on, and they increased their scores to 58% (subject 1) and 48% (subject 5) by the end of the two-hour session.

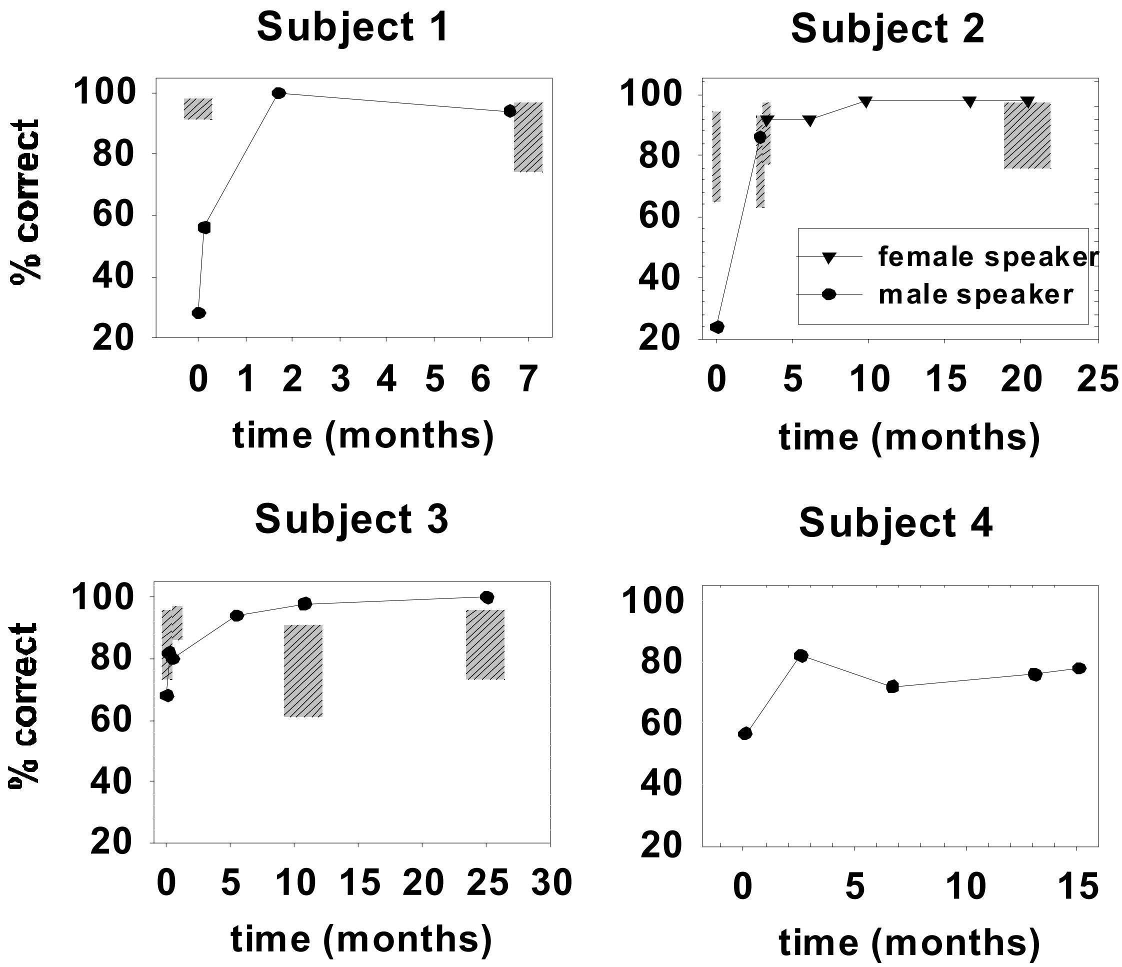

Figure 1. Vowel identification as a function of time for subjects 1-4, whose electrode discrimination skills did not increase post-implant. The vertical bars represent ceiling estimates of vowel identification performance for each individual, as estimated with the MPI model based on the listener's electrode discrimination. Estimates were not obtained for subject 4, who used a different device than the others. Vowel identification by subjects 1, 2 and 3 reaches ceiling estimates by 2-5 months post-implant.

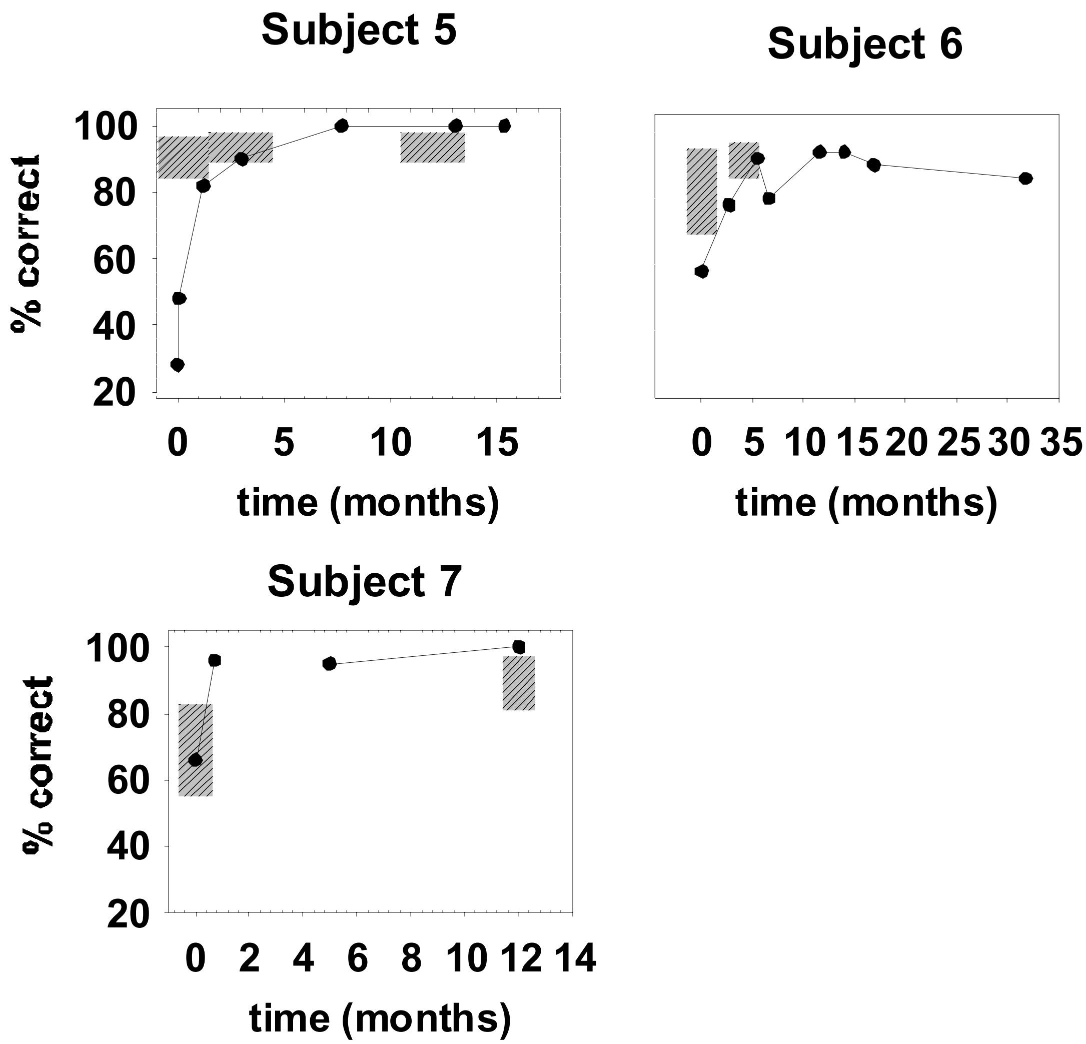

Figure 2. Vowel identification as a function of time for subjects 5-7, whose electrode discrimination skills did increase post-implant. Vowel identification by subjects 5 and 6 reaches ceiling estimates by 3-5 months post-implant, while subject 7 was the only one whose vowel identification scores were within the ceiling estimate range the day of initial stimulation.

Discussion

All subjects showed the expected pattern of improvement in vowel scores over the first few months after initial stimulation. In four of the seven subjects, this change was not accompanied by an increase in electrode discrimination. Taken together, these data suggest that improvement in labeling of vowel sounds was the main mechanism underlying longitudinal increases in vowel identification for these four subjects. The other three subjects presented a more mixed picture, with simultaneous increases in electrode discrimination and in vowel identification. However, predictions obtained with the MPI model suggest that the improvement in electrode discrimination was insufficient to explain the pronounced increases in vowel identification observed in two of these three subjects: in their case, the improvement in vowel identification may have been due to parallel increases in discrimination and labeling skills. These results are consistent with those of Harnsberger et al. (in press), who asked cochlear implant users to select the regions of the F1-F2 plane that sounded like a given vowel. One goal of their study was to determine whether the basalward frequency shift imposed by cochlear implants results in systematic response biases in this task. No such bias was found, indicating that their subjects (who, unlike the subjects in the present study, had used their cochlear implants for at least one year) had learned to label vowels correctly, and their vowel perception was limited mostly by their ability to discriminate formant frequencies. An interesting direction for future research may be to use the method-of-adjustment task employed by Harnsberger et al. longitudinally, starting immediately after implantation, to directly measure changes in the CI users' ability to label vowels correctly. Additionally, it would be informative to image the cochleas of these subjects in order to obtain estimates of electrode location and cochlear length, which in turn would help refine estimates of the amount of basalward shift in these cochlear implant users.

The MPI model was used in this study to tease apart the effect of improved labeling and improved discrimination on speech perception by CI users. This kind of information may be clinically useful because it may suggest areas to be stressed during auditory rehabilitation following cochlear implantation. Subjects whose scores are well below their ceiling estimates may especially benefit from training designed to help them label speech sounds. Conversely, subjects who do reach their ceiling may benefit from training that helps them discriminate better along the acoustic and perceptual dimensions known to be important in speech perception.

References

Harnsberger, J.D., Svirsky, M.A., Kaiser, A.R., Pisoni, D.B., Wright, R., & Meyer, T.A. (in press). Perceptual "vowel spaces" of cochlear implant users: Implications for the study of auditory adaptation to spectral shift. *Journal of the Acoustical Society of America.*

Hillenbrand, J., Getty, L.A., Clark, M.J., & Wheeler, K. (1995). Acoustic characteristics of American English Vowels. *Journal of the Acoustical Society of America,* 97, 3099-3111.

Levitt, H. (1972). Decision Theory, Signal-Detection Theory, and Psychophysics. In E.E. David & P. B. Denes (Eds.), *Human Communication: A Unified View.* New York: McGraw-Hill. Pp. 114-174.

Peterson, G.E. & Barney, H.L. (1952). Control Methods Used in the Study of Vowels. *Journal of the Acoustical Society of America,* 24, 175-184.

Svirsky, M.A. (2000). Mathematical Modeling of Vowel Perception by Users of Analog Multichannel Cochlear Implants: Temporal and Channel-Amplitude Cues. *Journal of the Acoustical Society of America,* 107, 1521-1529.

Svirsky, M.A., Meyer, T.A., Kaiser, A.R., Basalo, S., Silveira, A., Suarez, H., Lai, T.T., Simmons, P.M. (1999). Learning How to Perceive Vowels with a Cochlear Implant: The Role of Discrimination and Labeling. Presented at The Association for Research in Otolaryngology, Twenty-Second Midwinter Research Meeting, Feb. 1999.