Channel weights for speech recognition in cochlear implant users

Maureen A. Mehr, Christopher W. Turner, and Aaron Parkinson

Department of Speech Pathology and Audiology and Department of Otolaryngology, Head and Neck Surgery, University of Iowa, Iowa City, Iowa 52242

(Received 1 February 2000; revised 1 September 2000; accepted 5 September 2000)

The purpose of this study was to develop and validate a method of estimating the relative “weight” that a multichannel cochlear implant user places on individual channels, indicating its contribution to overall speech recognition. The correlational method as applied to speech recognition was used both with normal-hearing listeners and with cochlear implant users fitted with six-channel speech processors. Speech was divided into frequency bands corresponding to the bands of the processor and a randomly chosen level of corresponding filtered noise was added to each channel on each trial. Channels in which the signal-to-noise ratio was more highly correlated with performance have higher weights, and conversely, channels in which the correlations were smaller have lower weights. Normal-hearing listeners showed approximately equal weights across frequency bands. In contrast, cochlear implant users showed unequal weighting across bands, and varied from individual to individual with some channels apparently not contributing significantly to speech recognition. To validate these channel weights, individual channels were removed and speech recognition in quiet was tested. A strong correlation was found between the relative weight of the channel removed and the decrease in speech recognition, thus providing support for use of the correlational method for cochlear implant users. © 2001 Acoustical Society of America. [DOI: 10.1121/1.1322021]

PACS numbers: 43.66.Sr, 43.66.Ts, 43.66.Yw, 43.71.Gv [KRK]

I. INTRODUCTION

Perhaps the biggest advance in cochlear implants subsequent to their initial introduction has been the employment of multiple electrodes and/or channels for simulation of the auditory nerve. The multiple electrode array presumably takes advantage of the natural tonotopic organization of the cochlea and its innervation pattern to allow the speech processor to provide “place-frequency” information to the listener. Considering the range of speech recognition performance among individual cochlear implant users, it may be beneficial to help determine how individual cochlear implant users understand or recognize speech. At present, besides overall recognition scores, we have very few tools available to determine how the multichannel implant is functioning. A good deal more information is required regarding what behavioral measures such as threshold, comfort level, dynamic range, and/or various electrophysiological characteristics can tell us about how well a particular electrode, and its associated nerve fibers, can transmit speech information. Because most of today’s multichannel cochlear implant processors divide speech information into different frequency bands, with each frequency band corresponding to an electrode, or group of electrodes, it is desirable to investigate the extent to which individual bands or electrodes contribute to the overall understanding of speech.

Since some multichannel cochlear implant users have relatively good understanding of speech while others do not, it is important to determine the factors influencing their success. Previous studies have suggested that these differences may be attributed to signal processing (Gantz et al., 1988; Tyler and Tye-Murray, 1991; Waltzman et al., 1992; Wilson, 1993), individual device-programming strategies (Tyler et al., 1992) and rehabilitation (Lansing and Davis, 1988). Other studies have looked at electrophysiology (e.g., Black et al., 1987; Brown et al., 1990; Simmons et al., 1984), psycholinguistics (Knutson et al., 1991; McKenna, 1991), and psychoacoustic measures (Dorman et al., 1990; Hochmair-Desoyer et al., 1985; Shannon, 1993; Chatterjee, 1999). Few of these previous efforts have directly investigated how speech information is transmitted to the listener through the implant beyond just measuring an overall percent score.

For example, Brown et al. (1990) reported that more rapid recovery functions on individual electrodes using an electrophysiological forward-masking paradigm correlated with higher overall speech recognition in implant users. On the other hand, Chatterjee (1999) found an apparently opposite relation using a behaviorally measured forward-masking task, that is, slower recovery times were related to better overall speech recognition. The actual relationships between characteristics of the implants’ ability to transmit speech information and various electrophysiological or behavioral measures may have been obscured in these studies by across-subject comparisons, in which other variables such as duration of deafness, etc., were also contributing. In addition, the single dependent variable (overall speech recognition) is a function of the combination of the characteristics of all the implant users’ functioning channels, whereas the behavioral and/or electrophysiological measures can be obtained individually across the multiple electrodes and/or frequencies. While there have been a few preliminary experiments in various laboratories in which the entire broadband speech

---

*Current address: Cochlear Corporation, Englewood, CO 80112.*
signal is input into individual electrodes or channels, in the present paper, we present a method which measures the ability of individual channels of the speech processor to transmit speech information appropriate for that channel under the condition of speech being presented to the implant user in their normal mode of listening.

A number of investigators have used the methods of the Articulation Index (e.g., ANSI, 1969; French and Steinberg, 1947; Fletcher and Galt, 1950) to determine frequency-importance information for listeners with acoustic hearing (both with and without sensorineural hearing loss). This method uses a quantity between zero and one to represent the proportion of speech information available in that frequency region to the listener. This available information in a particular frequency region is then multiplied by a frequency-importance, or ‘‘weighing’’ factor. To obtain frequency-importance information, a rather time-consuming process of using high- and low-pass filtering is utilized. In these filtering experiments, listeners are presented speech information under conditions of restricted frequency ranges.

An alternative approach to determining the importance of various frequency bands is offered by the correlational method (Doherty and Turner, 1996; Turner et al., 1998). This is an adaptation (for speech) of the method originally described by Lutfi (1995) and Richards and Zhu (1994) for use in psychoacoustic experiments. In the correlational method, speech is presented to the listener in a broadband listening condition which is more typical in normal listening situations. Turner et al. (1998) showed that the relative importance of various frequency regions in understanding speech was often different, depending upon whether the listener was tested under conditions of broadband listening versus restricted frequency ranges. Another advantage to the correlational method is that all channels are able to be tested at the same time, which has the potential to be less time consuming than filtering experiments.

The knowledge of which channels within an individual’s implant are functioning well for speech recognition and which are not may also have some clinical applications in terms of choosing an appropriate programming strategy for the speech processor. A study of Hanekom and Shannon (1996) showed that listeners’ speech recognition performance was a function of the set of electrodes chosen from a seven-electrode speech processor. In another study by Zwolan et al. (1997), a reduction in the number of electrodes used resulted in some cochlear implant users having significant improvements in their speech recognition scores while others showed a decline in performance. These studies suggest that it may be possible to maximize benefits to the listener based on the number and choice of electrodes used. If we can determine which channels or electrodes are contributing most to the user’s understanding of speech, it is possible that this information could then be used to “tune” the processor strategy in individual users to take maximum advantage of the fully functioning electrodes and/or to eliminate or redirect speech information away from the poorly functioning electrodes.

In the present study, we were interested in evaluating the application, reliability, validity, and efficiency of the correlational method in determining how important each frequency region is to the speech recognition of both cochlear implant users and normal-hearing listeners. The frequency weighting functions determined for normal-hearing listeners in the Turner et al. (1998) study suggested similar listening strategies among listeners. In experiment 1 we ask, “How does the listening strategy or weighting functions of listeners with normal-hearing compare to the weighting functions determined for cochlear implant users?” Ultimately, we would like to determine if the weighting functions can be used to specify which channels or electrodes are working most efficiently and which are not. In experiment 2 we describe the results of a test of the validity of the correlational method for use with cochlear implant users.

II. EXPERIMENT 1

In experiment 1, we obtained weighting functions from both normal-hearing listeners and cochlear implant users in order to determine how individual implant users make use of particular frequency regions as compared to normal-hearing listeners.

A. Methods

1. Subjects

Data were obtained from five listeners (four females and 1 male) with normal hearing, and with a mean age of 32 years (range from 21–48 years), having scored at least 95% correct on the speech tests presented. The audiometric thresholds for each listener were 20 dB HL or better at octave test frequencies from 250 to 8000 Hz. Seven adults (four females and three males) using the Med-El CIS-Link 6 Channel Processor and six active electrodes of the Ineraid cochlear implant were recruited through the Department of Otolaryngology at the University of Iowa. They had a mean age of 62 years (range 44–79 years) at the time of testing. The cochlear implant users had a minimum of 12 months experience with the Med-El processor.

2. Speech materials

Our speech test consisted of a subset of the UCLA version of the Nonsense Syllable Test (NST), which uses six lists of consonant-vowel syllables. The vowels /a, i, u/ were paired with 22 different consonants and spoken by both a male and female talker for a total of 132 different speech tokens. These nonsense syllables were chosen to provide a rather large set of speech materials, which would include nearly all consonants in several vowel contexts and more than one talker, to reduce possible effects of token-specific cues for speech recognition. The various CV syllables were randomized when presented to the listener. Stimuli were digitized at 44.1 kHz (anti-alias filter of 20 kHz) and stored on a laboratory computer. The speech stimuli were digitally filtered into six frequency bands at frequencies corresponding to the configuration of the CIS speech processor programming strategy. FIR digital filters were implemented using MATLAB routines and provided less than 1 dB of passband ripple and 40 dB of stopband attenuation. This resulted in the six filtered speech bands corresponding to the
six electrodes of the cochlear implants used in this study. The frequency divisions were logarithmic and are listed in Table I.

### 3. Procedures

The correlational method was used for estimating speech weighting functions. On each trial, the root-mean-square (rms) level within each of the six frequency bands was calculated for the token to be presented. Six bands of noise, each filtered to match the frequency range of the corresponding speech band, were then mixed with the appropriate speech band at a randomly chosen signal-to-noise ratio. The signal-to-noise ratio in each speech band was chosen (in 2-dB steps) from a rectangular distribution that was 24-dB wide, the midpoint of which was determined from pilot testing (described below). On each trial, the signal-to-noise ratio chosen was independent of the level of noise in each of the other five bands. The six bands (speech plus noise) were then recombined and played to the listener. The weight placed upon a frequency band was estimated by a point biserial correlation between the signal-to-noise ratios in a band and the responses (whether correct or incorrect). That is, those frequency bands for which varying the signal-to-noise ratio had little or no effect on the subject’s performance received a low weight, and those bands in which the varied signal-to-noise ratio had a large effect on performance received a high weight. The six raw correlations for each subject were transformed to relative weights by summing their values and expressing each band’s weight as the raw correlation divided by this sum. Thus the relative weights of the six channels sum to 1.0. These six relative weights are then plotted as a function of the frequency band to yield a frequency-weighting function.

Each listener was allowed to listen to the nonsense syllables to become familiar with the consonant choices and response box. The response box consisted of 22 buttons, each labeled with one of the consonant alternatives. Listeners were given visual feedback in the form of a light located adjacent to the button corresponding to the correct answer after each trial. The listener was then presented with 6 CV lists of 50 trials each for a total of 300 trials to determine an overall percent-correct score in quiet. The listener was then acquainted with the speech stimuli that included the randomized noise added to the speech. Based on the overall percent-correct scores, a mean signal-to-noise ratio was then determined by additional preliminary testing to yield an approximate one-third decrease in recognition scores from those obtained in the quiet condition (i.e., from 60% correct to 40% correct). This is necessary because the correlational method requires that listeners make some recognition errors due to the added noise. For normal-hearing listeners, mean signal-to-noise ratio chosen was approximately 0 dB. For the majority of implant users, this level was approximately +12 dB.

Analyzed data consisted of results taken from the 132-item speech materials consisting of six lists at 200 trials each for a total of 1200 trials in the randomized noise condition. Data consisting of fewer trials were also analyzed to determine the number of trials necessary to yield a significant raw correlation in at least two or more frequency bands. Due to the occasional variability in day-to-day performance of cochlear implant users, we obtained 300-trial speech scores in quiet during each session for a baseline measure of the patient’s performance on the day of testing.

Stimulus presentations and experimental records were accomplished by a Macintosh Power PC 9500 computer and DigiDesign Audiomedia III, 16-bit digital-to-analog converters. For all subjects, sound field presentation of stimuli was used in a double-walled sound-attenuation chamber (IAC). A Crown (Model D-75A) amplifier was used to drive a single three-way loudspeaker (Pyle Mfg.) located 1.5 m directly in front of the subject. The cochlear implant users wore their speech processors for the experiment. Normal-hearing listeners heard the speech stimuli at a level of 70 dB SPL. For the implant users, the presentation level was set to the highest comfortable level for each subject, attempting to ensure that the majority of the speech range was above their thresholds. From the trial-by-trial experimental record stored on the computer, correlations were calculated for each band between the signal-to-noise ratio on each trial, and whether the subject scored correctly or not. When the number of trials equals 1200, a raw correlation of $r = 0.0564$ or higher is considered significantly different from zero at the $p < 0.05$ level of confidence (Lutfi, 1995). We then compared the weighting

<table>
<thead>
<tr>
<th>Band number</th>
<th>Frequency range (in Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>300–486</td>
</tr>
<tr>
<td>2</td>
<td>486–791</td>
</tr>
<tr>
<td>3</td>
<td>791–1283</td>
</tr>
<tr>
<td>4</td>
<td>1283–2062</td>
</tr>
<tr>
<td>5</td>
<td>2062–3384</td>
</tr>
<tr>
<td>6</td>
<td>3384–5500</td>
</tr>
</tbody>
</table>
functions determined for the cochlear implant users to the weighting functions determined for the listeners with normal hearing.

4. Results

Weighting functions determined for the normal-hearing listeners are plotted in Fig. 1. Similar listening strategies were demonstrated across listeners. Each channel or frequency region was approximately equal in its importance for understanding these speech tokens. Furthermore, all regions appeared to be making a contribution. As noted by the error bars, there was very little variation across listeners.

Weighting functions determined for six of the cochlear implant users are plotted in Fig. 2, with the solid line representing the mean normal weighting function and the dashed line representing the cochlear implant user’s weighting function. The seventh cochlear implant subject’s data were not used, as this subject did not yield significant raw correlations in any of the frequency bands [this was due to the subject’s very low speech recognition score in quiet (18%), which made it very difficult to find an appropriate level of noise to add to the speech which would degrade the recognition score]. Because the relative weights must add to one, a relative weight greater than the normal-hearing values does not necessarily indicate that the implant user does better in using this band than normal listeners. Rather, that particular frequency region is more important for that cochlear-implant listener relative to the other frequency bands.

In contrast to weighting functions of the normal-hearing listeners, weighting functions of the cochlear implant users showed unequal weighting across the various frequency regions. Of the 36 data points for the cochlear implant users (6 channels times 6 implant users), 20 of the relative weights lie outside the 95% confidence intervals of the estimate for average normal-hearing listeners’ relative weights. Channels were not contributing equally, and each implant user had at least one channel that was near zero. Four cases of negative correlations were obtained from the total of 36; however,
those negative correlations yielded negative relative weights very close to zero (the largest was \(-0.06\)). A zero weighting indicated that there was no significant relationship between the signal-to-noise ratio in that band and the subject’s performance. Speech recognition scores in quiet for each of the cochlear implant users are noted on each graph in Fig. 2, with no individuals scoring near 100%.

Combined individual weighting functions of the implant users are replotted in Fig. 3, along with the averaged results of the normal-hearing listeners. Although there are large individual differences, there appears to be a general trend among cochlear implant users for a lower weighting on bands 1 and 6 and a higher weighting on band 3. It can be easily seen from this plot that the cochlear implant users’ weighting functions are quite different from those for normal-hearing listeners, with each implant user having unequal weighting across the bands.

We were also interested in the reliability and efficiency of the procedure in terms of the number trials required to obtain accurate weighting functions. Because the statistical significance of the raw correlations is dependent upon the number of trials, we compared the number of significant raw correlations obtained using 600 trials versus 1200 trials. The data indicated that 1200 trials were necessary in order to yield significant raw correlations in at least two or more frequency bands for these implant users. Therefore, weighting functions based upon fewer trials may not identify every band that makes a substantial contribution to speech recognition.

Relative weights for each subject in each band determined from the first 600 trials are plotted against relative weights determined from the full 1200 trials in Fig. 4. For a relative weight in the middle range \((\approx 0.16)\), the 95% confidence interval of a prediction based upon 600 trials relative to the value obtained after 1200 trials was calculated to be plus or minus 0.17. Thus the relative weights determined on the basis of 600 trials would be informative primarily in the sense of showing which bands were high contributors versus which bands were low contributors.

III. EXPERIMENT 2

In this experiment we addressed the issue of the validity of the channel weights for the implant users, as determined in experiment 1. If the relative weights obtained for an implant user are valid, then removing speech information in a band with a high relative weight should result in a substantial decrease in speech recognition performance. Furthermore, the removal of speech information in a band with a low relative weight should result in a small decrease. In experiment 2, we tested these relations.

A. Methods

1. Subjects

Three of the six adult cochlear implant users from experiment 1 provided usable data for this experiment. Experiment 2 was conducted several months following the data collection for experiment 1, and cochlear implant users were only available for a limited time block for our testing as they were visiting Iowa City for their regularly scheduled clinical checkups. Due to the elapsed time between experiments 1 and 2, each subject’s actual relative weighting functions may have changed from the values originally measured. Two of the original six implant users showed a large and statistically significant increase or decrease (13%–16%) in their speech recognition in quiet performance between the time of experiment 1 and experiment 2 (implant users LB and CL). Therefore, the validity of their previously measured weighting functions at the time of experiment 2 were suspect and they were not included in experiment 2. A third implant user from experiment 1 (patient KH) was having technical difficulties with her speech processor/implant on the day of experiment 2, so we were unable to collect reliable speech recognition scores from this patient. Thus only data from cochlear implant users HW, KB, and NS were obtained in experiment 2.
3. Results

As noted, there is a strong relationship \((r = 0.77)\) between the decrease in speech recognition measured and the relative weight of the channel removed. When the change in speech recognition with a channel removed was expressed in arcsin-weight form of speech recognition scores, the relation was essentially identical \((r = 0.73)\). Removal of a channel with a low weighting resulted in a nominal change in percentage points. Conversely, removal of a channel with a high weighting resulted in a larger (negative) change in percentage points. While the relationship in Fig. 5 between the weight of the channel removed and its effect upon speech recognition is striking, caution should be observed due to the limited number of individuals for which data could be collected. Additional research with more cochlear implant users is warranted.

IV. DISCUSSION

Results from these experiments indicated that the correlational method can provide a method for estimating the relative weight that listeners place on various frequency bands. This procedure yields estimates of how well each “channel” is contributing to the implant user’s understanding of speech in comparison to each of the other channels. Moreover, these measures can be obtained simultaneously for all channels in a broadband listening situation.

Results for normal-hearing listeners indicated that each listener used information in speech across frequency bands in a similar manner. This is in agreement with the similarity of band-weighting patterns across normal-hearing listeners reported in the Turner et al. (1998) study, which also used the correlational method. Additionally, it appears that the logarithmic frequency division of the channels produces approximately equal weights across frequency bands for the normal-hearing listeners for these speech materials. These two conclusions are further strengthened in that this same methodology yields very different weighting functions for individuals with cochlear implants. The equal-weighting across bands found for all normal-hearing listeners is not an artifact of the procedure itself. This “flat” weighting function and the similarity across listeners was most likely related to the large set of speech materials used in the present study. In Doherty and Turner (1996), the correlational method showed very different weights across channels and listeners. That study required the listeners to discriminate between only three speech tokens, and it was suspected that each listener developed a highly stimulus-specific strategy for the task, as opposed to a more general speech recognition strategy. We have recently begun testing a different set of speech materials (16-item, four-talker, /aCa/ consonant recognition) and have again observed approximately equal weighting across logarithmically divided bands for normal-hearing listeners.

Cochlear implant users in this study yielded very different weighting functions than did normal-hearing listeners. This suggests that the general characteristics of electrical simulation of the remaining auditory neurons in the auditory system of deaf individuals, along with the cochlear implant devices of the present study, do not completely capture the important aspects of normal speech transmission for implant users. This may be indicated by the common pattern across implant users noted in Fig. 3, in which channel 1 produced relatively low weights and channel 3 yielded higher weights. In support of the present pattern of results, Dorman et al. (1989) reported that the addition of a mid-to-high frequency channel of stimulation to the most apical (low-frequency) channel produced the largest improvements in speech recognition in their group of implant users. Channel 6, which stimulates the most basal location of the cochlea, received relatively low weighting by all the implant users. This result
might reflect a lack of viable neurons in the basal end of profoundly deafened individuals. It is probably the case that implant users have real differences between the effectiveness of individual electrodes in combination with the pattern of surviving neurons. However, other factors such as speech audibility for implant users as well as the dynamic range of the implant channel in relation to the speech signal may contribute to the ineffectiveness of various channels. It may also be the case that the age differences between the normal-hearing listeners and the implant users may be a factor in the pattern of weights obtained across the two groups. Further research may help us to understand these issues and the application of the correlational method as described in this paper may be useful in these investigations.

There are, however, some limitations of the correlational procedure, as described in the present study. The correlational method may be limited by the number of channels that can be tested at one time. If too many channels are tested at once, the obtained raw correlations may not reach significance unless the number of trials is very large. One could test fewer channels of a multichannel implant, but the weighting function determined would only relate to those channels chosen, as the estimated weights are relative to one another, and the desirable characteristic of a broadband listening situation might have to be sacrificed. In the present study, we found that using 600 trials did not always yield at least two channels with significant raw correlations, whereas 1200 trials did produce at least two significant correlations. Thus the procedure as described in this paper is probably too lengthy for standard clinical application at the present time. In addition, if the listeners’ speech recognition performance changes over time, weighting functions obtained at one session may not remain valid in subsequent experimental sessions. These concerns suggest that a more efficient testing procedure is desirable.

A related concern is the interaction between the difficulty of the speech materials for the range of implant users encountered, and the ability of the correlational procedure to produce significant correlations. Assigning the appropriate average noise level for each subject is very important. Too little noise will result in nonsignificant correlations. Thus the overall percent score in quiet for a patient should not be too low, otherwise the decrease in score from adding noise may result in the subject performing at chance level. This was seen in the failure of the correlational method with the seventh implant subject, for whom the speech materials were apparently too difficult.

In summary, the correlational procedure shows promise in its ability to assess the functional status and contribution to speech recognition of individual frequency bands (channels) in cochlear implant users. Additional research using this procedure may allow us to determine which factors involved in speech recognition through multichannel cochlear implants are critical in achieving optimal performance.

ACKNOWLEDGMENTS

This research was supported in part by NIDCD Grant No. 2 R01 DC00377-12. The project was the basis of an Honor’s thesis for the first author in the Department of Speech Pathology and Audiology at the University of Iowa. Valuable assistance was provided by Marla Ross, Ruth Sew-erson, Shelley Witt, and Abigail Wolaver of the Department of Otolaryngology at the University of Iowa. Helpful comments on an earlier version of this paper were provided by Paul Abbas and Bom Jun Kwon. In addition, many helpful criticisms and suggestions were provided by two anonymous reviewers as well as the associate editor.


