Effects of Stimulation Mode on Speech Recognition by Cochlear Implant Users

Qian-Jie Fu and Robert V. Shannon

Dept. of Auditory Implants and Perception, House Ear Institute, 2100 West Third St., Los Angeles, CA 90057

Abstract: This study investigated the interactive effect of both electrode configurations and frequency regions assigned to electrodes on speech recognition in three subjects implanted with Nucleus-22 multichannel cochlear implant. A 4-channel processor with continuous interleaved sampling speech processing strategy was implemented through a custom interface. The temporal envelopes from four broad frequency bands were used to modulate 500pps, 100 μs/phase pulse trains and then delivered to four electrode pairs. Ten different frequency allocations and three sets of 4-electrode configurations were used. Each frequency allocation represented different insertion depths of the four electrodes based on Greenwood’s place-to-frequency function. Preliminary results showed that the vowel score was highly dependent on the frequency allocations in all electrode configurations. Subjects with different insertion depths had similar recognition patterns as a function of the frequency allocations, indicating a possible accommodation to the pattern of speech information presented through their normal 20 electrode SPEAK processor. A similar recognition pattern as a function of the frequency allocations was observed within subjects for the electrode configurations which had the same return electrode, indicating that pattern recognition of speech sounds in electric hearing depends primarily on the electrically evoked discharge patterns at the return electrode.

INTRODUCTION

Cochlear implants are complex devices that allow control over many aspects of the pattern of stimulation delivered to the cochlea. Each of these parameters will change the electrically-evoked peripheral temporal-spectral pattern of neural information to the brain. How to adjust these parameters to achieve the best performance for individual cochlear implant users in the short-term as well as the long-term is still one of the biggest challenges in electric hearing. Three factors appear to be most important for speech performance in electric hearing: (a) number of electrodes, (b) the allocation of frequency regions to electrodes, and (c) the stimulation mode and location of electrodes. However, little is known about the interactive effects of these critical factors on speech performance. The current study addresses this issue.

METHODS

Three post-lingually deafened adults with the Nucleus-22 cochlear implant participated in this study. All had at least six months experience utilizing the SPEAK speech processing strategy and all were native speakers of American English. Vowel recognition was measured in a 12-alternative identification paradigm, including 10 monophthongs and 2 diphthongs, presented in a /h/-vowel/ d/ context, e.g., "heed", "hid", "hayed", "head", "had", "hod", "hawed", "hoed", "hood", "who d", "hud", "heard". The tokens for these closed-set tests were digitized natural productions from 5 men, 5 women, 3 boys, and 2 girls, drawn from the material collected by Hillenbrand et al. (4). A stimulus token was randomly chosen from all 180 tokens in vowel recognition and presented in random order. Each run included 180 trials and each data point included 2 or 3 runs. No feedback was provided. The subjects started the formal test without training and with no appreciable period of adjustment to the new processor. All signals were presented at comfortable audible levels through a custom interface (5). The implementation of the speech processor with CIS speech processing strategy in Nucleus-22 cochlear implant users was as follows. The signal was first pre-emphasized using a first-order Butterworth high-pass filter with a cutoff frequency of 1200 Hz, and then band-pass filtered into 4 broad frequency bands using eighth-order Butterworth filters. The corner frequencies of the bands were dependent on the simulated insertion depth. The corner frequencies of the frequency allocation set were 290, 585, 1081, 1913, and 3310 Hz, corresponding to the tonotopic locations at 7, 10.75, 14.5, 18.25, and 22 mm from the apex based on the Greenwood’s place-to-frequency function: f(x)=165.4*(10^0.064*x-0.88), where x is position on the basilar membrane (in mm) from the apex (3). The tonotopic width of each frequency band was a constant 3.75-mm, same as the spacing between adjacent electrode pairs, e.g. (20,22) vs. (15,17). The most apical tonotopic location was gradually shifted from 7-mm (set1) to 13.75-mm (set10) in steps of 0.75 mm for a total of 10 frequency sets. The envelope of the signal in each band was extracted by half-wave rectification and low-pass filtering (eight-order Butterworth) with a 160 Hz cutoff frequency. The acoustic amplitude was transformed into the electric amplitude by a power-law function with an exponent of 0.2 (2). Then this transformed amplitude was used to modulate the amplitude of a continuous, 500 pulse/sec biphasic pulse train with a 100
μs/phase pulse duration. Three electrode configurations were used, including an apical set with BP+1 mode [ABP1: (20,22), (15,17), (10,12), (5,7)], a basal set with BP+1 mode [BBP1: (16,18), (11,13), (6,8), (1,3)], and an apical set with BP+5 mode [ABP5: (16,22), (11,17), (6,12), (1,7)].

RESULTS AND DISCUSSION

Figure 1 shows the vowel recognition score as a function of the frequency allocations for three electrode configurations. There are several significant findings in the present study. First, consistent with the previous results (6), speech performance changed significantly when either frequency allocation or electrode configuration was changed. Results suggest that best performance may be achieved when the electrodes receive the speech information extracted from the “matched” frequency region. A spectral mismatch, especially a severe mismatch, between the speech information and the tonotopic location of the electrodes will result in poor speech recognition and will distort the accurate evaluation of how many electrodes (or channels) are necessary (1). Second, subjects with different insertion depths had the same recognition patterns in all electrode configurations except for a difference of overall performance, indicating that all these subjects already accommodated to the new patterns of speech information after a long time of exposure to their normal 20 electrode SPEAK processor. Third, a 3-mm shift in the tonotopic location of electrodes required the same amount of tonotopic shift in the filter frequency allocation to obtain the best performance, suggesting that a shifted place-to-frequency mapping may be still preserved in cochlear implant users. Fourth, all subjects had similar recognition patterns for electrode configurations ABPI and ABP5, which had the same return electrodes, indicating that pattern recognition of speech sounds in electric hearing depends primarily on the electrically evoked discharge patterns at the return electrode.

ACKNOWLEDGEMENTS

We wish to thank Professor James Hillenbrand for allowing us to use the multi-talker vowel test materials and the three implant listeners for their participation in the current experiments. The assistance of Xiaosong Wang in collecting data is also gratefully acknowledged. The research was supported by NIDCD.

REFERENCES