Effects of electrode deactivation on speech recognition in multichannel cochlear implant recipients

Kara C. Schvartz-Leyzac[1,2], **Teresa A. Zwolan**[2], and **Bryan E. Pfingst**[1]

[1]Kresge Hearing Research Institute, University of Michigan, Department of Otolaryngology, 1150 West Medical Center Drive, Ann Arbor, MI, 48109, USA

[2]Hearing Rehabilitation Center, University of Michigan, Department of Otolaryngology, 450 W. Market Place, Building 1, Suite A, Ann Arbor MI, 48108, USA

Abstract

Objectives—The objective of the current study is to evaluate how speech recognition performance is affected by the number of active electrodes that are turned off in multichannel cochlear implants. Several recent studies have demonstrated positive effects of deactivating stimulation sites based on an objective measure in cochlear implant processing strategies. Previous studies using an analysis of variance have shown that, on average, cochlear implant listeners' performance does not improve beyond eight active electrodes. We hypothesized that using a generalized linear mixed model (GLMM) would allow for better examination of this question.

Methods—Seven peri- and post-lingual adult cochlear implant users (eight ears) were tested on speech recognition tasks using experimental MAPs which contained either 8, 12, 16 or 20 active electrodes. Speech recognition tests included CUNY sentences in speech-shaped noise, TIMIT sentences in quiet as well as vowel (CVC) and consonant (CV) stimuli presented in quiet and in signal-to-noise ratios of 0 and +10 dB.

Results—The speech recognition threshold in noise (dB SNR) significantly worsened by approximately 2 dB on average as the number of active electrodes was decreased from 20 to 8. Likewise, sentence recognition scores in quiet significantly decreased by an average of approximately 12%.

Discussion/Conclusion—Cochlear implant recipients can utilize and benefit from using more than eight spectral channels when listening to complex sentences or sentences in background noise. The results of the current study suggest a conservative approach for turning off stimulation sites is best when using site-selection procedures.

Kara C. Schvartz-Leyzac (corresponding author): Dr. Schvartz-Leyzac is a clinical instructor and holds dual appointments at the University of Michigan Hearing Rehabilitation Center and Kresge Hearing Research Institute., kleyzac@med.umich.edu; Phone: 734-763-5159.

Teresa A Zwolan: Dr. Zwolan is a Professor in the Department of Otolaryngology and Director of the Hearing Rehabilitation Center., Zwolan@med.umich.edu; Phone: 734-998-8119

Bryan E. Pfingst: Dr. Pfingst is a Professor in the Department of Otolaryngology and is a faculty member at Kresge Hearing Research Institute., bpfingst@umich.edu; Phone: 734-963-2292

Introduction

For more than 30 years, multichannel cochlear implants have proved to be an effective means of auditory rehabilitation for severely hearing-impaired individuals. Despite the success of cochlear implants and significant improvements in technology, there continues to be a wide range of performance across recipients (Beyea et al., 2016; Cullington & Zeng, 2008; Holden et al., 2013). It is hypothesized that variability in conditions near individual electrodes in the implanted cochlea contributes, at least partially, to variability observed among cochlear implant listeners (see Pfingst et al., 2015 for review) (Pfingst et al., 2015). Thus, some stimulation sites within the implanted cochlea may contribute more than other sites to the patient's overall performance. It has been demonstrated that within-subject speech recognition can be improved by deactivating poorer performing or otherwise problematic stimulation sites in a subject's speech processor MAP (Boëx, Kós, & Pelizzone, 2003; Garadat, Zwolan, & Pfingst, 2012, 2013; Noble, Labadie, Gifford, & Dawant, 2013; Zwolan, Collins, & Wakefield, 1997). We refer to this procedure as "site selection". "MAP" in this context refers to the pattern of electrodes chosen for activation and the stimuli assigned to the active electrodes.

Review of site-selection studies

Zwolan and colleagues (Zwolan et al., 1997) demonstrated improved speech understanding, on average across subjects, with experimental processor MAPs in which certain stimulation-sites were deactivated based on poor electrode discrimination. Individual benefit varied, and some listeners did not demonstrate improvements. Similarly, Garadat and colleagues (Garadat et al., 2013) demonstrated improved speech understanding for all 12 subjects when electrodes with the poorest masked modulation detection thresholds (MDTs) were deactivated from each tonotopic section of the electrode array. Other studies showed improved performance among cochlear implant listeners when electrode sites were selected for removal based on results of CT imaging (Bierer, Shea-Brown, & Bierer, 2015; Noble et al., 2013). For some site-selection strategies, as many as 14 channels were removed to create the modified MAP (Noble, Gifford, Hedley-Williams, Dawant, & Labadie, 2014).

It has been argued that cochlear implant listeners' performance does not significantly improve beyond 8 or 10 channels on average (Fishman, Shannon, & Slattery, 1997; Friesen, Shannon, Baskent, & Wang, 2001; Friesen, Shannon, & Cruz, 2005; Shannon, Cruz, & Galvin, 2011). This theoretically implies that for a 22-channel electrode array, as many as 12–14 electrodes could be deactivated using site-selection strategies while maintaining or improving performance. However, it is important to consider potential disadvantages of turning off stimulation sites. When stimulation sites are deactivated, the frequency allocation of the remaining sites is broadened, which reduces spectral resolution and might reduce or negate any potential benefit of the site-selection strategies.

Speech perception as a function of number of spectral channels

It is believed that performance with a CI is limited by poor spectral resolution caused by (1) a reduced cochlear neural population with respect to that observed in normal-hearing (NH) ears, and (2) overlap in the neural populations stimulated by adjacent electrodes. Friesen and

colleagues (2001) compared the speech recognition skills of cochlear implant recipients to that of NH individuals listening to noise-band vocoded speech (e.g., cochlear implant simulations). In brief, they found that cochlear implant users' average performance did not improve beyond 8–10 channels in any condition, whereas the performance of NH listeners continued to systematically improve with an increasing number of channels. It was hypothesized that these results were attributable to the significant channel interaction present in cochlear implant users. These results are consistent with other reports that speech recognition does not typically improve beyond 8–10 spectral channels for cochlear implant recipients (Fishman et al., 1997; Friesen et al., 2005; Shannon et al., 2011).

All of the studies above included a 4-channel condition and all used an analysis of variance (ANOVA) to analyze results. In the present study, we analyzed how speech recognition is affected when the number of active electrodes was decreased from 20 to 8. We chose a linear mixed model rather than an ANOVA, to address this question. Generalized linear mixed models (GLMMs) are a type of regression analysis and are an extension of an ANOVA that allows for better control of hierarchical constraints. A GLMM approach was most appropriate for this study for several reasons. First, this study employed a within-subject repeated measures design that included several conditions (e.g., number of channels) with correlated errors. These correlated errors violate the assumption of independent observation that is required for ANOVA. Second, a mixed model approach is better designed to handle missing data values than an ANOVA as all available data points can be included for all subjects, even in the absence of a complete data set for any given subject. Lastly, and perhaps most important, a mixed model approach better accommodates a highly variable data set, such as that often observed with cochlear implant listeners. For example, in the current study we were interested in understanding how speech recognition was affected by the removal of active electrodes in a speech processor map. We analyzed clustered data points for each subject that were nested within a higher order variable of overall speech recognition abilities. In other words, each listener had varied speech recognition abilities regardless of the condition tested (e.g., number of channels). Therefore, the nested variable of condition was inherently affected by the overarching variable of speech recognition abilities (Simpson's paradox) (Simpson, 1951). Recently, Winn et al., (Winn, Edwards, & Litovsky, 2015) demonstrated the appropriateness of such linear mixed models and employed a similar analysis when measuring how spectral degradation affects listening effort as measured through pupil dilation. Their results show a systematic decline in listening effort with increasing spectral resolution.

In the currently study, we were particularly interested in knowing how performance is affected as the number of channels or stimulation sites varied between 8 and 20. Previous studies in CI listeners and NH individuals listening to noise-band vocoded speech, showed that performance on several speech recognition tasks certainly improved as the number of active channels and/or stimulation sites increases between 1 and 8 or 10 (Shannon et al., 2011; Shannon, Fu, & Galvin, 2004). However, there is little evidence to date to support the theory that providing more than 10 active electrodes in a listeners' MAP improves performance (Fishman et al., 1997; Friesen et al., 2001; Friesen et al., 2005).

In summary, there are now several studies showing the possible benefits of site-selection strategies for speech processor MAPs. However, for judicious application of site-selection procedures, it is important to re-evaluate the potential deleterious consequences that may occur when channels are removed from maps being used by recipients utilizing contemporary processing strategies. The aim of the current study was to determine the effects on speech recognition of randomly deactivating stimulation sites from the CI user's full clinical map. We hypothesized that random electrode site deactivation would have deleterious effects that were proportional to the number of sites turned off.

Subjects and methods

Subjects

Subjects included seven cochlear implant users with peri- or post-lingual sensorineural hearing loss. For one bilaterally implanted subject (S60) both ears were tested unilaterally and data from both ears was used in the current study. Four other subjects were bilaterally implanted, but testing was only completed in one ear. None of the subjects (unilaterally or bilaterally implanted) had residual hearing in their implanted ear; two subjects did have residual hearing in the opposite ear, but all unaided thresholds were within the severe to profound range. The ears were implanted with Nucleus CI24R(CA) or CI24RE(CA) devices, and all subjects used the ACE speech-processing strategy in their everyday/clinic MAP. All subjects used Nucleus 5 (CP810) speech processors. Demographic information for the subjects (eight ears) is shown in Table I. All subjects were native speakers of American English. The use of human subjects in the study was reviewed and approved by the University of Michigan Medical School Institutional Review Board.

Speech processor mapping procedures

Experimental speech processor MAPs were created using commercially available programming software, CustomSound®4.0 or 4.1. Threshold (T) and comfortable loudness (C) levels were assessed on each electrode prior to testing in order to ensure subjects were using an up-to-date speech processor MAP. For the current study, only the 20 most apical electrodes were used to create experimental speech processor MAPs; electrodes 1 and 2 were not used for any of the experimental MAPs because many of the subjects had these electrodes deactivated in their everyday/clinical MAP. None of the subjects who participated in our study had any more than two of the most basal electrodes (electrodes 1 or 2) deactivated in their everyday/clinical speech processor MAPs.

All speech processor MAPs were created using the default settings with the exception of manipulation of the number of active electrodes: ACE processing strategy, 900 pps stimulation rate per electrode, 8 maxima, 8 µs interphase gap, and MP1+2 stimulation mode. The overall pulse rate remained constant as the number of channels was decreased since the maxima was set to 8 for all conditions. The pulse width used for all but one subject was 25 µs/phase on each electrode; a 37 µs/phase pulse width was used for S60 which was consistent with the subject's everyday/clinical speech processor MAP.

T and C levels—The method used to set T (near the threshold for hearing) and C (loud but comfortable) stimulus levels was similar to that used in a typical cochlear-implant clinic. T levels were measured on each electrode across the array using a bracketing procedure similar to the standard clinical procedure. The T levels were then confirmed by stimulating five sequential electrodes (at the T level determined by the bracketing procedure), ensuring that the subject heard all five beeps. If T levels were not heard at these levels they were adjusted slightly until stimuli at all five electrodes were audible. C levels were set using a visual loudness scale in which the level of stimulation was increased until the subject reported that the stimulus was "loud but comfortable". C levels for each electrode were loudness balanced to the immediately apical adjacent electrode to ensure equal loudness across the array. Settings for the 20-channel MAP were verified via live voice, ensuring that each subject deemed speech comfortable and of comparable sound quality when compared with their clinical MAP. In all cases, subjects reported that the 20-channel MAP sounded similar to their everyday/clinical MAP.

Experimental speech processor MAPs—In total, 11 experimental speech processor MAPs were created for each subject (Table II). These MAPs consisted of 8, 12, 16 or 20 active channels. For all experimental MAPs, active electrodes were spaced as evenly as possible along the length of the electrode array in order to minimize distortion of the tonotopic map. As electrodes were deactivated, the frequency-to-electrode allocation was redistributed using the default setting in the clinical software and bandwidths for each electrode were increased to maintain the full frequency representation across the entire electrode array.

In order to avoid inadvertently selecting a better-functioning set of electrodes in one condition relative to another, the locations of the active electrodes were roved within each condition. Thus, we created various versions of the 8-, 12-, and 16 –channel MAPs which differed in the specific electrodes chosen for activation. Each of the speech recognition test conditions listed below was tested four separate times for each listener. For the 8 and 12 channel MAPs, each roved electrode configuration was tested once. For the 16 channel condition, each roved electrode configuration was tested twice since there were only two 16-channel configurations possible. For the 20 channel condition, the same 20-channel MAP was used for all four repetitions. Table II lists all of the electrode configurations for each of the experimental MAPs.

Some subjects perceived an overall decrease in loudness when using MAPs with fewer active electrodes. In order to help counter loudness differences between MAPs, subjects were allowed to increase the volume setting on the speech processor if needed. The experimenter then tested the subjective loudness of each experimental MAP via live voice and ensured that all MAPs were of appropriate loudness for each subject before writing the programs to the speech processor.

Speech recognition testing procedure

All speech recognition tests were administered in a double-walled sound-attenuated booth (Acoustic Systems Model RE 242 S; Applied Acoustic Systems, Montreal, Quebec). Tests

were administered via a graphic user interface programmed in MATLAB (Mathworks, Natick, MA, USA). Speech materials were delivered from the computer to a Rane ME60 graphic equalizer, a Rolls RA235 35W power amplifier, and presented via a loudspeaker positioned 1 m from subjects at 0° azimuth. All stimuli were calibrated at 60 dB (A) SPL with a sound-level meter (type 2231; B & K, Naerum, Denmark) using the slow time setting. For speech stimuli presented in the presence of background noise, the mixed signal (speech + noise) was presented at a constant level of 60 dB (A) SPL. During calibration sessions, the sound-level meter was positioned 1 m away from the loudspeaker at 0° azimuth. Speech recognition was measured while the subject used a laboratory owned CP810 speech processor. CustomSound® mapping features were the same as those used in each subject's everyday/clinical MAP. For all subjects, experimental MAPs were programmed to use pre-processing settings of "Everyday" (ADRO + ASC).

Four repetitions of each number-of-channel condition were presented in total. A block-randomized design was used and was consistent across all subjects. Specifically, subjects were presented with two full repetitions of each number of channel condition, starting with 20 channels. Two more repetitions were completed with 16 channels, and this procedure was repeated for 12 and 8 channel conditions. Then, two more repetitions were repeated for 8 channels, then increasing from 12 to 20 in the opposite order as presented above. The roved experimental MAP order of the 8-, 12- and 16-channel MAPs differed across subjects in a random fashion. For example, in the first repetition of the 8-channel condition, one subject may have listened with one version of the 8-channel map, while another subject may have listened with another version of the 8-channel map. Thus, all versions of roved electrode maps shown in Table II were presented to all subjects but in different random orders.

Speech stimuli

Sentence recognition in noise—CUNY sentences (Boothroyd, Hanin, & Hnath, 1985) were presented in a steady-state, speech-weighted background noise using a custom program created in MATLAB (MATLAB, 2010a). The CUNY sentences are meaningful utterances with contextual cues spoken by a male speaker. Four sentence lists (each containing 12 sentences) randomly chosen from a total of 72 lists were used for measuring one SRT. For this task, the background noise was presented alone for 1.5 seconds before the target, during presentation of the target, and for 0.5 second after the target sentence was completed. Raised cosine ramps were applied at the onset and offset of the noise stimulus with the onset and offset each measuring 5% of the entire stimulus length. Signal-to-noise ratio (SNR) was calculated for the time period when the target and noise overlapped. The mixed signal (target + noise) was normalized to its peak amplitude. Therefore, the level of the masker plus sentence was similar from trial to trial. SNR started at 20 dB at the beginning of the test and was adapted in a one-down one up procedure using a step size of 2 dB. The sentence was presented in the noise background one time and the subject was instructed to repeat the sentence to the experimenter. The experimenter lowered the SNR by 2 dB if the subject repeated all words in the sentence correctly, or increased the SNR by 2 dB for an incorrect response was provided. The one-down one-up procedure estimated a 50% correct point on the psychometric function. The SRT was taken as the mean of the SNRs at the last six reversals out of a total of 12 reversals.

Sentence recognition in quiet—The TIMIT acoustic-phonetic speech database (Lamel, Kassel, & Seneff, 1986) consisting of 34 lists of 20 sentences was used to assess speech recognition in quiet. The TIMIT sentences include various speakers and speaking patterns. A single run consisted of one TIMIT list. This database was chosen for two reasons: 1) the sentences are phonetically diverse, complex sentences are consistent with casual conversation that cochlear implant listeners might experience every day, and 2) they are unfamiliar to most cochlear implant recipients because they are not typically used in the clinic setting.

Vowel recognition—Vowel stimuli were taken from materials recorded by Hillenbrand et al. (1995) (Hillenbrand, Getty, Clark, & Wheeler, 1995) and were presented to the listeners using custom software written in MATLAB (MATLAB, 2010a). Four presentations (two females, two males) of each vowel were tested for each condition (Quiet, +10 dB SNR, 0 dB SNR). For conditions presented in background noise, a steady-state, speech weighted background noise was used. Each of the 12 medial vowels (i, $ɔ$, $ɛ$, u, $ɪ$, $ʊ$, $ʌ$, $æ$, $ɝ$, o, $ɑ$, e) was presented in a /h/-vowel-/d/ context [heed, hawed, head, who'd, hid, hood, hud, had, heard, hoed, hod, and haid]. Chance level was equal to 8.33% correct. Each calculated percent correct for one repetition for each condition tested was based on discrimination of 48 speech tokens (4 speakers × 12 speech tokens). One subject (S103) did not complete the +0 dB SNR condition due to time constraints.

Consonant recognition—Consonant stimuli were taken from Shannon et al (Shannon, Jensvold, Padilla, Robert, & Wang, 1999) and were presented with custom software written in (MATLAB, 2010a). Four presentations (2 females, 2 males) of each consonant were tested for each condition (Quiet, +10 dB SNR, +0 dB SNR). For conditions presented in background noise, a steady-state, speech weighted background noise was used. Stimuli consisted of 20 consonants, /b, d, g, p, t, k, m, n, f, s, ʃ, v, z, θ, l, dʒ, tʃ, w, j, r/, presented in a consonant-/a/- context (ba, da, ga, pa, etc). Chance level was equal to 5%. Each calculated percent correct for one repetition for each condition tested was based on discrimination of 80 speech tokens (4 speakers × 20 speech tokens). One subject (S103) did not complete the +0 dB SNR condition due to time constraints.

Statistical analysis

All data were analyzed using SPSS Version 22 (SPSS, 2013). A linear mixed-effects model approach was first used to examine how speech understanding is affected by the number of active electrodes removed from the speech processor MAP. For a review of this statistical method, please see the description and rationale provided in the introduction. For each speech test, the best model fit was determined by examining Schwarz's Bayesian Criterion (SBC). The SBC is a goodness-of-fit measure similar to the log-likelihood value used for logistic regression analysis but is specifically designed when examining multiple parameters. There are other goodness-of-fit measures that can be used when fitting GLMMs (e.g., Akaike's information criterion), but the SBC method is one that is commonly used and it is a conservative option compared to other approaches. For all models, the number of channels and slope intercept were entered as fixed values. In some cases the best model also included a random slope or intercept, or both. The results below reflect the best model fit for each

type of speech test based on SBC; a lower value indicated a better fit of the model. Table III provides coefficient and standard error values for each predictor variable entered in each significant model. Data were additionally analyzed using an ANOVA in order to replicate the statistical method used in previous studies (Friesen et al., 2001; Friesen et al., 2005).

Results

Linear Mixed Model Approach

Sentence recognition in background noise—Figure 1 shows the relationship between performance on the CUNY sentence recognition test in noise (SRT dB SNR, y-axis) as a function of number of active electrodes (x-axis). The regression line (thick dashed line) is shown. In general, the model reveals a gradual increase in SRTs (i.e., decrement in performance) as a function of the number of active electrodes removed. This significant result means that the modeled slope was significantly different from zero. The best-fit linear model included four predictors total: number of channels and intercept entered as both fixed and random effects. This is in keeping with the raw data shown in Figure 1 which demonstrates variability in both the y-intercept values and slope functions for each subject tested. Based on the fixed slope coefficient, these results suggest an average increase in SRT (decrease in performance) of 0.19 dB SNR for each stimulation site deactivated in the speech processor MAP. Thus, as the number of active channels decreased from 20 to 8, the average increase in SRT (decrement in speech recognition in noise) was 2.12 dB SNR. The range of dB SNR change observed across subjects varied from −5.79 (improvement) to 3.24 (decrement). The mean performance for each listener for the 8 and 20 channel conditions and the difference in performance between these two conditions are provided for each listener in Table IV.

Sentence recognition in quiet—Figure 2 shows the relationship between performance on the TIMIT sentence recognition test (percent correct, y -axis) as a function of number of active electrodes (x-axis). The best-fit linear model was significant and included three predictors total (number of channels entered as a fixed effect and intercept entered as a fixed and random effect). Based on the fixed slope coefficient, these results suggest an average decrease of 0.98% for each active electrode removed from the speech processor MAP. Thus, as the number of active channels was decreased from 20 to 8, the sentence recognition in quiet score decreased by an average of 10.58%. The range of percent correct change observed across subjects varied from −6.80 to −16.46% as the number of active channels was decreased from 20 to 8. The mean performance for each listener for the 8 and 20 channel conditions and the difference in performance between these two conditions are provided for each listener in Table IV.

Vowel recognition in quiet and in background noise—Figure 3 shows the relationship between performance on the vowel recognition task in quiet and at signal-to-noise ratios of +10 and +0 dB (percent correct, y-axis) as a function of number of active electrodes (x-axis). For vowels in quiet, the best-fit linear model included three predictors total (number of channels entered as a fixed effect and intercept entered as a fixed and random effect). Based on the fixed slope coefficient, these results suggest an average

decrease of 0.56% for each active electrode removed from a speech processor MAP. When the number of active channels was decreased from 20 to 8, the score for vowel recognition in quiet decreased by 5.97%. Change in performance observed across subjects varied from +4.73 to −29.1 % as the number of active channels was decreased from 20 to 8. The mean performance for each listener at for the 8 and 20 channel conditions and the difference in performance between these two conditions are provided for each listener in Table IV. For the results in background noise (+10 or +0 dB SNR) a significant model could not be derived for either vowel recognition condition. These results show that, performance on a vowel recognition in noise task did not significantly decrease as the number of active electrodes was decreased from 20 to 8.

Consonant recognition in quiet and in background noise—Figure 4 shows performance on the consonant recognition task in quiet and in a +10 and 0 dB SNR condition. A linear mixed model could not be derived for any of the consonant recognition tasks. These results show that performance on a consonant recognition task in quiet or noise did not significantly decrease as the number of active electrodes was decreased from 20 to 8.

Analysis of variance (ANOVA) approach—In order to make a more direct comparison between our findings and those reported in previous investigations, the data of the current study were also analyzed using an ANOVA model. We felt this was appropriate since subjects used in this study demonstrated pre-operative test scores similar to subjects used in the Friesen et al (2005) study (Table I). One-way ANOVAs with Tukey post hoc analyses were performed for TIMIT and CUNY sentences as well as for each of the vowel and consonant phoneme recognition conditions (Quiet, +10 dB SNR, +0 dB SNR). Using this analysis approach, the results were fairly similar to those reported in previous studies (Fishman et al., 1997; Friesen et al., 2001; Friesen et al., 2005). Recall that some of these previous studies included conditions with fewer numbers of channels than used in the current study (<8 channels). ANOVA analyses in the current study revealed that performance did not differ significantly when listening with 8, 12, 16 or 20 channels for all speech conditions with the exception of TIMIT sentences. Post-hoc analysis (Tukey HSD) revealed a significant difference between performance for 8- and 16- channel conditions (*p=0.024*) only. Statistics and results of the ANOVA are reported in Table V.

Discussion

Contrary to previous findings, the current investigation reveal that the speech recognition abilities of Nucleus CI users programmed with the ACE strategy tended to decrease as the number of channels was decreased from 20 to 8. Results for challenging sentences in quiet and sentences in noise showed an appreciable decline when the number of channels was decreased from 20 to 8, while the decline in performance for vowel and consonants was smaller or not significant. For example, for the GLMM analyses, the model for vowel and consonant stimuli could either 1) not be fit or 2) the fit was significant but predicted minimal per-channel decreases for a total decrement of less than 7% as the number of channels was decreased from 20 to 8. Decrements for sentence stimuli appear more meaningful. For contextually difficult and complex speech (TIMIT sentences) in quiet, the overall average decrement of nearly 12% was significant and is likely to be noticeable to a cochlear implant

recipient. Likewise, performance on the speech in noise task demonstrated a decrement in SRT of almost 2 dB SNR and was also significant and is also likely noticeable to the implant recipient. Some studies have cited that, in some cases, a 1 dB change in SNR can affect speech understanding by 10% or more (Chasin, 2013). However, Chasin (2013) also notes that the relationship between improvement in speech understanding and change in dB SNR depended on the stimulus used and also on the point along the psychometric function at which performance was measured.

Similar to many cochlear implant studies, we observed a large range of performance across subjects. For some subjects, performance decreased by nearly 30% in at least one of the conditions tested (TIMIT sentences). The individual results are reported in Table IV. The individual results are also displayed in each graph (dashed lines). It appears that the performance of some listeners was unaffected by the number of channels while others experienced a slight decrease in performance as the number of channels was increased. However, random variability in the data is expected and the comprehensive results of the model show a significant trend when data across all subjects is combined.

In the current study, we generally observed that subjects who would typically be classified as "excellent" performers tended to be more susceptible to the effects of channel deactivation and that slightly poorer performers appeared to be less susceptible to such effects. It is plausible that excellent users were able to achieve a high level of performance because they were able to benefit from more active channels when compared to listeners who perform more poorly. In fact, a similar trend was reported by Friesen and colleagues (Friesen et al. 2001) and statistical analyses showed this trend to be significant in 19 subjects. In the current study, we did attempt to analyze the relationship between overall performance levels and change in performance as the number of channels was manipulated. However, limited sample prevented us from drawing definitive conclusions. For example, we performed regression analyses to determine the extent to which performance with the 20-channel MAP predicted the amount to which overall performance changed as the number of channels was decreased from 20 to 8 channels. This analysis was performed for both CUNY and TIMIT sentences given that performance using these stimuli seemed to be most susceptible to changes in electrode configuration (see previous results). Results did show a significant relationship for CUNY sentences ($R^2 = 0.58$, $p = 0.03$), however these results cannot be interpreted with any certainty because further statistical analysis revealed autocorrelation of the residuals (Durbin-Watson statistic = 0.67). A similar trend was also found for TIMIT sentences, but this analysis was also confounded by autocorrelation of the residuals. Due to the limited sample size we were unable to perform further analyses to determine if other factors such as duration of hearing loss, implant type, subject age or years of implant use were related to the results reported here.

The results of the current study have important implications for site-selection strategies. Previous studies from our laboratory have removed 5 sites when using site-selection techniques (Garadat et al., 2013). Zwolan and colleagues removed 1–7 electrodes based on an electrode discrimination task (Zwolan et al., 1997), while methods employed elsewhere removed as many as 14 channels (Bierer et al., 2015; Noble et al., 2014; Noble et al., 2013). Due to the variable performance observed across subjects, the results of the current study do

not necessarily result in an absolute recommendation for how many electrodes can be removed during site-selection methods without causing detriments to performance. However, the results suggest that an experimenter and/or clinician should exercise caution when selecting sites for removal and when interpreting the results of site-selection experiments. Unfortunately, there is currently no efficient method to determine how each listener's performance may be affected by electrode-site removal. The optimal number of electrodes removed from a MAP may depend on the stimulation strategy or stimulation mode. Bierer and Litvak (2016) found improvements in phoneme recognition among some poorer performers when they used a partial tri-polar, as opposed to monopolar, stimulation strategy with either 1) a MAP with a full complement of electrodes (no site-selection) or 2) a site-selection program using monopolar stimulation (Bierer & Litvak, 2016). Grasmeder and colleagues (Grasmeder, Verschuur, & Batty, 2014) demonstrated that frequency-to-electrode assignment using a reduced-frequency MAP based on post-operative x-ray imaging and insertion angle can improve performance in some listeners. Results from Vickers and colleagues argue that site deactivation for *n-of-m* strategies (e.g., SPEAK, ACE) is not beneficial when choosing stimulation sites based on pitch-ranking tasks. It is hypothesized that, for such strategies, only a given number of sites (*n)* are stimulated in one cycle and therefore it may be possible to deactivate many sites without consequence (Vickers, Degun, Canas, Stainsby, & Vanpoucke, 2016). It is also important to consider that, for *n-of-m* strategies, a CIS strategy (8 channels, 8 maxima) will process a greater proportion of quiet sounds than a typical ACE strategy (20 channels, 8 maxima). The 8-channel ACE strategy conditions used in this study are essentially equivalent to an 8-channel CIS strategy. Taken together, these results demonstrate it is likely that site-selection benefits vary across listeners and are likely dependent on several factors, including spectral resolution/channel interaction across the electrode array, stimulation mode, and/or stimulation strategy.

The results of the current study should be interpreted with caution. First, the number of subjects tested was limited and the participants were generally good to excellent performers with peri- or post-lingual onset of hearing loss. It is uncertain how these results reflect the cochlear implant recipient population as a whole, and particularly those who are poorer performers than those who participated in the current study. Second, the current study measured the acute speech recognition abilities when subjects listened with experimental MAPs; we did not provide listening experience or training with each experimental MAP. It is known that performance with other types of experimental MAPs can improve if the recipient is allowed time for practice and training (Li, Galvin, & Fu, 2009). Although electrode site removal was spread along the length of the electrode array to maintain the frequency- to-place alignment in the cochlea, greater reduction in numbers of electrodes likely altered the frequency-to-place alignment when compared to the patients' everyday-clinical MAP. It is not certain if a performance decrease observed as the number of active electrodes decreases is a reflection of a) the reduced number of channels, b) frequency-to-place warping, c) reduced similarity to the subject's everyday map, or d) some weighted combination of these factors.

The results of the current study suggest that using appropriate statistical analyses can also have an impact on the outcomes of the study. The results of the current study are somewhat in contrast to previous reports that average performance of cochlear implant listeners does

not improve beyond 8 to 10 active electrodes. Cochlear implant users' performance tends to be highly variable and previous studies have examined data across subjects using an ANOVA approach. The current study primarily used an alternative statistical method (GLMM) that is more robust to variable data than an ANOVA. As outlined earlier, this approach has other statistical advantages when examining repeated measures and nested variables.

Of note, when data from the current study were analyzed using a traditional ANOVA approach, similar to the approach used in previous studies, our results were similar (Fishman et al., 1997; Friesen et al., 2001; Friesen et al., 2005; Shannon et al., 2011). However, in the current study we did find that for contextually difficult sentences (TIMIT sentences in quiet), the ANOVA analysis did show that performance decreased from 16 to 8 channels only. It is possible that these results differ from those obtained in previous studies due to the difficulty of the sentence corpus. TIMIT sentences are low-context, and often consist of short phrases. It is likely that listening with a reduced number of stimulation sites and/or channels becomes more difficult as the task complexity increases. This would support the hypotheses that a greater number of stimulation sites are needed as the listening effort increases (Winn et al., 2015).

While the overall results in the current study differ from those reported by Friesen and colleagues (Friesen et al., 2001) with regards to how many channels provide meaningful improvements for CI users, a similar pattern was noted across the studies related to stimulus type: both studies found that performance for sentence recognition tended to be more dependent on the number of active electrodes in a subject's MAP compared to performance for phoneme recognition. Sentence stimuli are spectro-temporally dynamic over a longer duration of time containing both vowels and consonants. In contrast, both vowel and consonant phonemes, while not completely spectro-temporally static, are more so than sentence stimuli and are also shorter in duration. Adding electrodes to a speech processor MAP might result in better resolution of spectrally dynamic sentences if the neural populations stimulated by individual stimulation sites are relatively independent. However, for stimuli which do not vary as much in spectral content over time (vowel and consonant phonemes), fewer numbers of channels may be required for optimal performance. These results, which are consistent across both studies, highlight the importance of speech recognition materials when assessing the number of channels required for adequate speech understanding.

Acknowledgments

References

Beyea JA, McMullen KP, Harris MS, Houston DM, Martin JM, Bolster VA, … Moberly AC. Cochlear Implants in Adults: Effects of Age and Duration of Deafness on Speech Recognition. Otol Neurotol. 2016; doi: 10.1097/mao.0000000000001162

Bierer JA, Litvak L. Reducing Channel Interaction Through Cochlear Implant Programming May Improve Speech Perception: Current Focusing and Channel Deactivation. Trends Hear. 2016; 20doi: 10.1177/2331216516653389

Bierer, JA., Shea-Brown, ET., Bierer, SM. Modeling the electrode-neuron interface to select cochlear implant channels for programming. Conference on Implantable Auditory Prostheses; Lake Tahoe, CA. 2015.

Boëx C, Kós MI, Pelizzone M. Forward masking in different cochlear implant systems. The Journal of the Acoustical Society of America. 2003; 114(4):2058–2065. http://dx.doi.org/10.1121/1.1610452. [PubMed: 14587605]

Boothroyd A, Hanin L, Hnath T. A sentence test of speech perception: Reliability, set equivalence, and short term learning. 1985

Chasin M. Slope of PI Function is not 10%-per-dB in Noise for All Noises and All Patients. Hearing Review. 2013; 20(11):1.

Cullington HE, Zeng FG. Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. J Acoust Soc Am. 2008; 123(1): 450–461. DOI: 10.1121/1.2805617 [PubMed: 18177173]

Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J Speech Lang Hear Res. 1997; 40(5):1201–1215. [PubMed: 9328890]

Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001; 110(2):1150–1163. [PubMed: 11519582]

Friesen LM, Shannon RV, Cruz RJ. Effects of stimulation rate on speech recognition with cochlear implants. Audiol Neurootol. 2005; 10(3):169–184. DOI: 10.1159/000084027 [PubMed: 15724088]

Garadat SN, Zwolan TA, Pfingst BE. Across-site patterns of modulation detection: Relation to speech recognitiona). The Journal of the Acoustical Society of America. 2012; 131(5):4030–4041. http://dx.doi.org/10.1121/1.3701879. [PubMed: 22559376]

Garadat SN, Zwolan TA, Pfingst BE. Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners. Audiol Neurootol. 2013; 18(4):247–260. DOI: 10.1159/000351302 [PubMed: 23881208]

Grasmeder ML, Verschuur CA, Batty VB. Optimizing frequency-to-electrode allocation for individual cochlear implant users. J Acoust Soc Am. 2014; 136(6):3313.doi: 10.1121/1.4900831 [PubMed: 25480076]

Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. J Acoust Soc Am. 1995; 97(5 Pt 1):3099–3111. [PubMed: 7759650]

Holden LK, Finley CC, Firszt JB, Holden TA, Brenner C, Potts LG, … Skinner MW. Factors affecting open-set word recognition in adults with cochlear implants. Ear Hear. 2013; 34(3):342–360. DOI: 10.1097/AUD.0b013e3182741aa7 [PubMed: 23348845]

Lamel FL, Kassel RH, Seneff S. Speech database development: design and analysis of the acoustic-phonetic corpus. DARPA Speech Recognition Workshop. 1986

Li T, Galvin JJ 3rd, Fu QJ. Interactions between unsupervised learning and the degree of spectral mismatch on short-term perceptual adaptation to spectrally shifted speech. Ear Hear. 2009; 30(2): 238–249. DOI: 10.1097/AUD.0b013e31819769ac [PubMed: 19194293]

MATLAB. MATLAB Natick. Massachusetts, United States: The MathWorks, Inc; 2010a.

Noble JH, Gifford RH, Hedley-Williams AJ, Dawant BM, Labadie RF. Clinical evaluation of an image-guided cochlear implant programming strategy. Audiol Neurootol. 2014; 19(6):400–411. DOI: 10.1159/000365273 [PubMed: 25402603]

Noble JH, Labadie RF, Gifford RH, Dawant BM. Image-guidance enables new methods for customizing cochlear implant stimulation strategies. IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society. 2013; 21(5):820–829. DOI: 10.1109/TNSRE.2013.2253333

Pfingst BE, Zhou N, Colesa DJ, Watts MM, Strahl SB, Garadat SN, … Zwolan TA. Importance of cochlear health for implant function. Hearing Research. 2015; 322:77–88. DOI: 10.1016/j.heares.2014.09.009 [PubMed: 25261772]

Shannon RV, Cruz RJ, Galvin JJ 3rd. Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise. Audiol Neurootol. 2011; 16(2):113–123. DOI: 10.1159/000315115 [PubMed: 20639631]

Shannon RV, Fu QJ, Galvin J 3rd. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. Acta Otolaryngol Suppl. 2004; (552):50–54.

Shannon RV, Jensvold A, Padilla M, Robert ME, Wang X. Consonant recordings for speech testing. J Acoust Soc Am. 1999; 106(6):L71–74. [PubMed: 10615713]

Simpson EH. The Interpretation of Interaction in Contingency Tables. Journal of the Royal Statistical Society Series B (Methodological). 1951; 13(2):238–241.

SPSS, I. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp; 2013.

Vickers D, Degun A, Canas A, Stainsby T, Vanpoucke F. Deactivating Cochlear Implant Electrodes Based on Pitch Information for Users of the ACE Strategy. Adv Exp Med Biol. 2016; 894:115–123. DOI: 10.1007/978-3-319-25474-6_13 [PubMed: 27080652]

Winn MB, Edwards JR, Litovsky RY. The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation. Ear Hear. 2015; 36(4):e153–165. DOI: 10.1097/aud. 0000000000000145 [PubMed: 25654299]

Zwolan TA, Collins LM, Wakefield GH. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. J Acoust Soc Am. 1997; 102(6):3673–3685. [PubMed: 9407659]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Summary

Results of the current study showed that, on average, cochlear implant listeners' sentence recognition performance degraded as fewer stimulation sites were included in the MAP, ranging from 20 to 8 channels. Individually, cochlear implant users demonstrated a wide range of results; some listeners' performance was not dependent on the number of channels whereas other listeners' performance decreased significantly as the number of active channels was decreased. These results have important implications for electrode site-selection strategies in multichannel cochlear implant users.
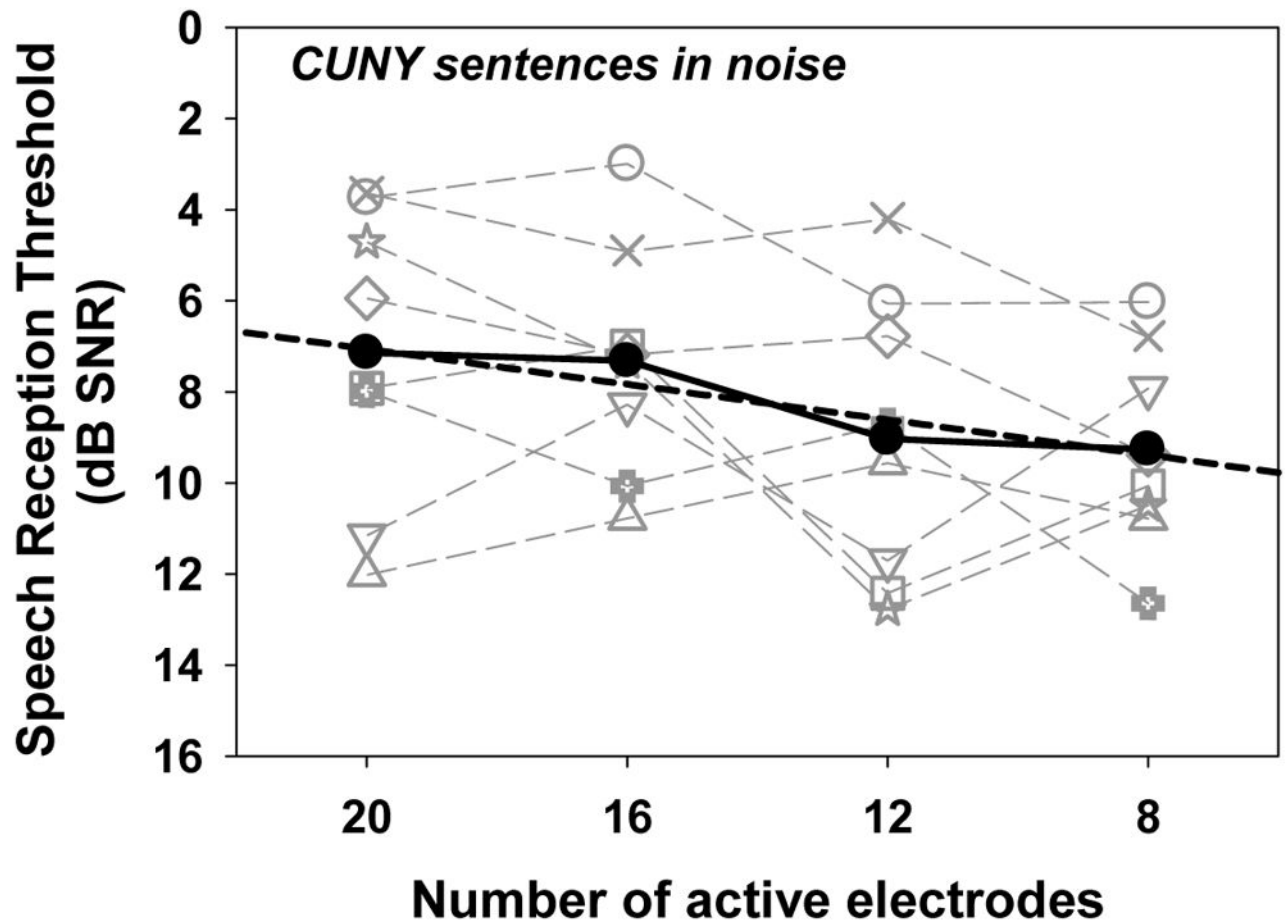
Figure 1.
Results for the CUNY sentence recognition in noise task. The number of active electrodes is shown along the x-axis, while the speech recognition threshold (SRT) in dB SNR is shown along the y-axis. Smaller SNR values indicate better performance: i.e. reception of sentence material at more challenging signal to noise ratios. The gray symbols and dashed lines indicate results for individual subjects, while the black filled circles and solid lines indicate average data across all subjects. The dashed black line represents the regression line.
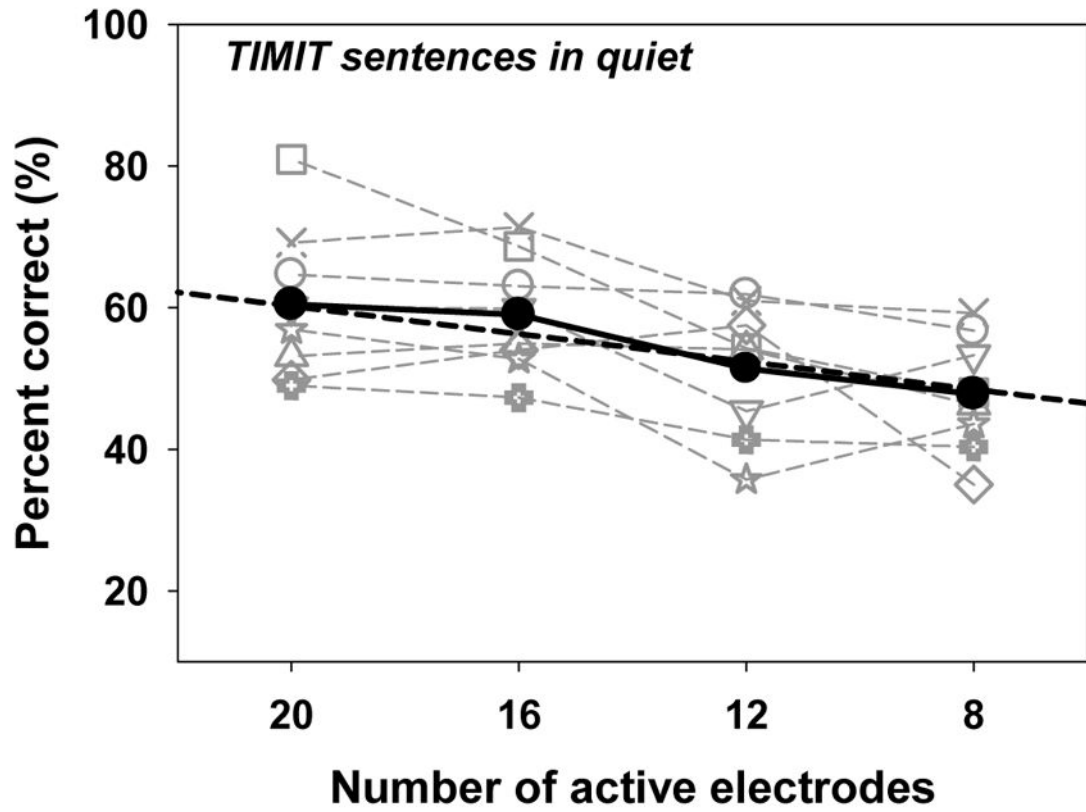
Figure 2.
Results for the TIMIT sentence recognition task in quiet. The number of active electrodes is shown along the x-axis, while the percent correct score is shown along the y-axis. Otherwise, the details of the graph as the same as noted in Figure 1.
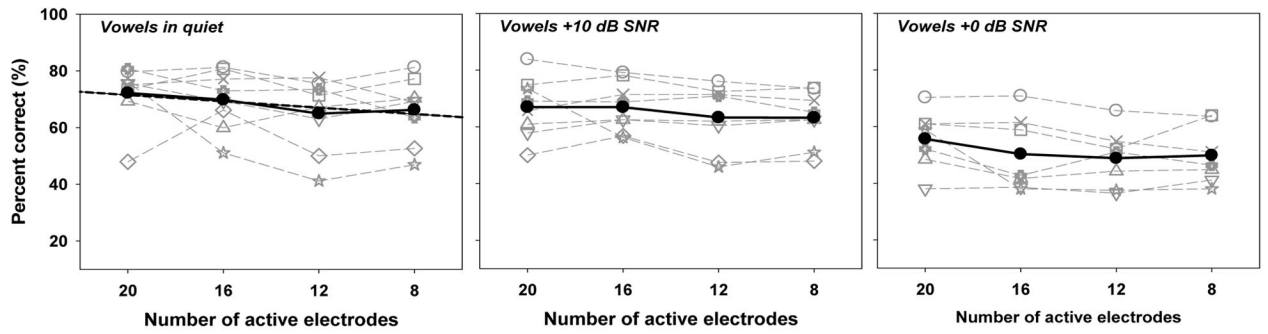
Figure 3.
Results for the vowel recognition tasks performed in quiet and at SNRs of +0 and +10 dB. Note that the middle and right panels lack the black-dashed regression lines because there were no model fits to the data.
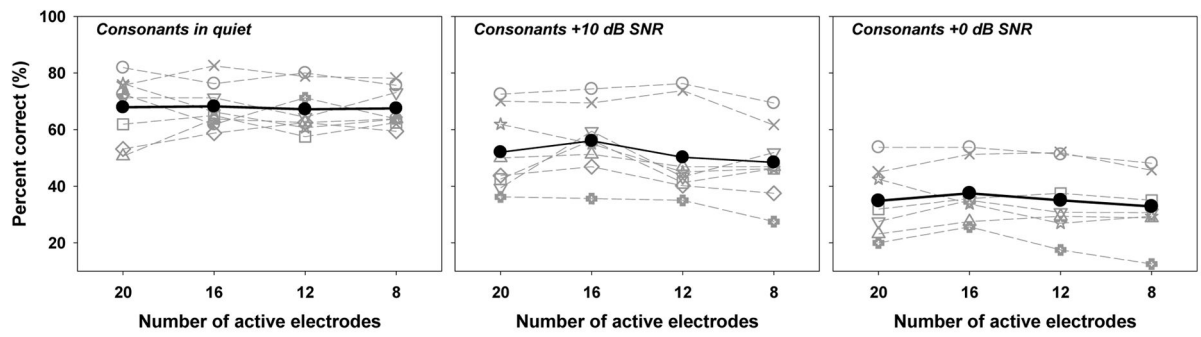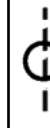
Figure 4.
Results for the consonant recognition tasks performed in quiet and at SNRs of +0 and +10 dB. Note that no regression lines are shown because there were no model fits to the data.

Table I

Subject demographic table. Post-operative speech recognition scores (AzBio sentences and CNC words) reflect performance using everyday settings in clinical speech processor maps.

Subject ID	Age (years old)	Gender	Ear	Symbol	Age at onset of hearing loss	Years CI use	Implant	Pre-operative speech scores		Post-operative speech scores	
								HINT sentence score (quiet)	CNC words (quiet)	AzBio sentence score (quiet)	CNC words (quiet)
S60	74	M	R	✳	40–50 years old	5	CI24RE (CA)	0%	NA	93%	72%
			L	⊖		11	CI 24R (CS)	8%	NA	100%	100%
S85	66	F	R	▷	5 years old	9	CI24RE (CA)	15%	20%	NA	80%
S86	67	F	L	◁	5 years old	5	CI24RE (CA)	NA	NA	93%	72%
S98	78	M	L	⊡	Teenager	3	CI24RE (CA)	0%	0%	81%	80%
S99	47	F	R	☆	5 years old	2	CI24RE (CA)	20%	4%	94%	80%
S100	74	F	L	✦	3 years old	2	CI24RE (CA)	NA	NA	NA	NA
S103	72	M	L	◇	30 years old	1	CI24RE (CA)	0%	0%	86%	68%

NA = data not available. Test scores were obtained in the clinic at the most recent time point relative to the subject's participation in the study (within 1 year of participation).

Table II

Number-of-channel conditions and roving assignments. The roving assignments refer to the various electrode configurations that were used in each number-of-channels condition (e.g., there were four 'roved' conditions for the 8-channel map).

										Electrode number										
8-channel																				
MAP	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
1																				
2																				
3																				
4																				
12-channel																				
MAP	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
1																				
2																				
3																				
4																				
16-channel																				
MAP	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
1																				
2																				
20-channel																				
MAP	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
1																				

Gray/shaded squares indicate active channels for each condition. White/open squares indicate inactive channels for each condition.

Table III

Parameters for significant GLMM results reported in the study.

CUNY sentences			
	Predictor	**Coefficient**	**Standard Error**
Fixed	Number of channels	−.194	.066
	Intercept	10.94	1.01
Random	Number of channels	.017	.015
	Intercept	4.15	3.48
TIMIT sentences			
	Predictor	**Coefficient**	**Standard Error**
Fixed	Number of channels	.980	.177
	Intercept	40.62	3.62
Random	Intercept	50.28	29.59
Vowels in quiet			
	Predictor	**Coefficient**	**Standard Error**
Fixed	Number of channels	.569	.180
	Intercept	60.30	4.24
Random	Intercept	87.81	49.71

Table IV

Individual performance for each subject on CUNY Sentences (dB SNR re: 50% SRT), TIMIT Sentences (% correct) and Vowel recognition in quiet (% correct) for the 8 and 20 channel conditions. The difference between performance using the 8 and 20 channel maps ("Change") is also shown.

	CUNY Sentences (dB SNR)			TIMIT Sentences (%)			Vowels in quiet (%)		
	8-channel	20-channel	Change	8-channel	20-channel	Change	8-channel	20-channel	Change
S60R	6.85	3.42	+3.17	59.24	69.18	−9.94	68.75	75.00	−6.25
S60L	6.03	3.75	+2.29	56.75	64.68	−7.93	81.25	79.68	+1.57
S85	11.16	7.92	−3.24	53.29	60.09	−6.80	69.27	75.52	−6.25
S86	10.78	12.02	−1.24	46.31	53.15	−6.84	70.31	69.27	+1.04
S98	10.06	7.92	+2.14	51.71	68.18	−16.46	77.18	73.43	+3.74
S99	10.50	4.71	+5.79	43.60	56.91	−13.30	46.88	76.04	−29.16
S100	12.64	7.98	+4.66	40.36	48.99	−8.63	63.54	80.73	−17.19
S103	9.38	5.95	+3.43	35.04	49.81	−14.77	52.65	47.91	+4.73
AVG	**9.67**	**6.70**	**+2.12**	**48.28**	**58.87**	**−10.58**	**66.22**	**72.19**	**−5.97**
SD	**2.21**	**2.85**	**2.99**	**8.38**	**7.96**	**3.76**	**11.61**	**10.43**	**11.8**

Note that a decrement in performance as the number of channels was decreased from 20 to 8 is indicated with negative (−) change for TIMIT and Vowel stimuli but a positive (+) change for CUNY sentences. Average (AVG) and standard deviations (SD) for the group of subjects are provided for each condition, for each speech test.

Table V

Results of One-way ANOVA analyses for all speech stimuli conditions. *F*-values (followed by degrees of freedom) and *p*-values are reported for each condition.

	CUNY sentences
F-value	1.252 (3,28)
p-value	0.310
	TIMIT sentences
F-value	3.369 (3,28)
p-value	0.032*
	Vowels in quiet
F-value	0.677 (3,28)
p-value	0.574
	Vowels +10 dB SNR
F-value	0.349 (3,28)
p-value	0.790
	Vowels +0 dB SNR
F-value	0.530 (3,28)
p-value	0.666
	Consonants in quiet
F-value	0.020 (3,28)
p-value	0.996
	Consonants +10 dB SNR
F-value	0.441 (3,28)
p-value	0.725
	Consonants +0 dB SNR
F-value	0.174 (3,24)
p-value	0.913

Significant results (*p<0.05)* are indicated with an asterisk (*).