Research paper

Improving speech perception in noise with current focusing in cochlear implant users

Arthi G. Srinivasan [a,b,*], Monica Padilla [a], Robert V. Shannon [a,b], David M. Landsberger [a]

[a] Department of Communication and Auditory Neuroscience, House Research Institute, 2100 West 3rd Street, Los Angeles, CA 90057, USA
[b] Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089, USA

ARTICLE INFO

ABSTRACT

Cochlear implant (CI) users typically have excellent speech recognition in quiet but struggle with understanding speech in noise. It is thought that broad current spread from stimulating electrodes causes adjacent electrodes to activate overlapping populations of neurons which results in interactions across adjacent channels. Current focusing has been studied as a way to reduce spread of excitation, and therefore, reduce channel interactions. In particular, partial tripolar stimulation has been shown to reduce spread of excitation relative to monopolar stimulation. However, the crucial question is whether this benefit translates to improvements in speech perception. In this study, we compared speech perception in noise with experimental monopolar and partial tripolar speech processing strategies. The two strategies were matched in terms of number of active electrodes, microphone, filterbanks, stimulation rate and loudness (although both strategies used a lower stimulation rate than typical clinical strategies). The results of this study showed a significant improvement in speech perception in noise with partial tripolar stimulation. All subjects benefited from the current focused speech processing strategy. There was a mean improvement in speech recognition threshold of 2.7 dB in a digits in noise task and a mean improvement of 3 dB in a sentences in noise task with partial tripolar stimulation relative to monopolar stimulation. Although the experimental monopolar strategy was worse than the clinical, presumably due to different microphones, frequency allocations and stimulation rates, the experimental partial-tripolar strategy, which had the same changes, showed no acute deficit relative to the clinical.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Cochlear implant (CI) users tend to have good speech recognition in quiet but have difficulty understanding speech in more difficult listening environments. Understanding speech in quiet requires only 4 spectral channels (Shannon et al., 2004), but speech perception in more difficult listening conditions or perception of music requires more spectral channels. For example, understanding speech in noise requires at least 8–10 spectral channels of independent information (e.g. Friesen et al., 2001). Recognition of simple musical melodies requires 15 independent spectral channels (Burns et al., 2001), and recognition of more complex musical melodies requires more than 48 spectral channels (Smith et al., 2002) It is possible that musical melody recognition is improved with more spectral channels due to better formant resolution. It is worth noting, however, that increasing the number of functional spectral channels alone may not improve harmonic relationships, as these relationships may not depend solely on spectral resolution but also on accurate spectral coding of each harmonic.

Although modern CIs have between 12 and 22 physical electrodes, CI listeners perform as if they are only receiving between 4 and 8 independent channels of information (Friesen et al., 2001), that is, not enough spectral channels to understand speech in noise. The limited spectral information is thought to be due to channel interactions across stimulated electrodes, which arise from overlapping populations of activated neurons (Fu et al., 1998; Fu and Nogaki, 2005). Reducing the spread of excitation from a stimulated electrode would narrow the population of activated neurons and would therefore presumably reduce channel interactions

across electrodes. Reducing electric and neural interactions across electrodes should in theory improve spectral resolution, which should in turn improve speech perception in noise (e.g. Henry et al., 2005; Litvak et al., 2007b) and music perception (e.g. Burns et al., 2001; Smith et al., 2002).

Tripolar (TP) stimulation has been studied to reduce current spread in the cochlea. While monopolar (MP) stimulation (which is used in most speech processing strategies today) utilizes an extra-cochlear electrode as a ground, TP stimulation uses 2 intra-cochlear electrodes (adjacent to the active stimulating electrode) as grounds (see Fig. 1). By keeping the current loop entirely within the cochlea, the intra-scalar voltage gradient is steeper than for MP stimulation, possibly leading to more focused neural excitation. One potential problem with TP stimulation is that larger current amplitudes are required to reach adequate loudness levels. These amplitudes might not be achievable physically due to the compliance limits of the device (e.g. Litvak et al., 2007a). To address this issue, partial tripolar (PTP) stimulation has been implemented (e.g. Wilson et al., 1995; Bierer, 2007; Litvak et al., 2007a; Landsberger et al., 2012). With PTP stimulation, a fraction (σ) of the current is returned via the intracochlear electrodes (with $\sigma/2$ returned on each of the flanking electrodes), and the remainder ($1 - \sigma$) is returned via the extracochlear electrode, similarly to MP stimulation (see Fig. 1).

TP/PTP stimulation has been shown to sharpen intra-scalar voltage gradients (Mens and Berenstein, 2005), and there is some evidence that TP/PTP stimulation can reduce spread of excitation as shown physiologically (e.g. Bierer and Middlebrooks, 2002; Snyder et al., 2004), computationally (Litvak et al., 2007a) and, for at least some electrodes and subjects, psychophysically (Bierer and Faulkner, 2010; Landsberger et al., 2012). One of the issues with TP/PTP stimulation is that TP/PTP stimulation requires larger current amplitudes relative to MP stimulation to produce equivalent loudness level. However, as spread of excitation increases with current amplitude (e.g. Chatterjee and Shannon, 1998), it is possible that TP/PTP stimulation might not be narrower than MP at equal loudness levels. Landsberger et al. (2012) addressed this issue and showed that PTP stimulation reduces spread of excitation (measured with a psychophysical forward masking task) relative to MP stimulation at an equal loudness level for half of the subjects. For the other half of the subjects, the spread of excitation was similar for both MP and PTP stimulation. These studies have shown that TP/PTP stimulation can be beneficial in a single-channel context. However, as speech processing strategies stimulate multiple channels i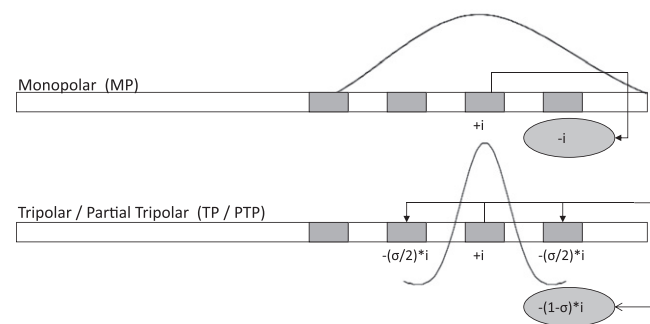n quick succession, single-channel metrics do not incorporate channel interactions across neighboring electrodes. In a test addressing these channel interactions, Berenstein et al. (2008) showed that PTP stimulation improves spectral ripple resolution (a multi-channel stimulus), when $\sigma = 0.75$. These results indicate that TP/PTP stimulation might be a promising tool for increasing spectral resolution and therefore improve speech perception in noise.

Mens and Berenstein (2005) published the first data comparing MP and TP/PTP speech processing tested chronically. They tested both a PTP strategy with $\sigma = 0.5$ and a strategy (called "flat TP + 2") with two flanking electrodes on each side (four flanking electrodes total) separated by one inactive electrode from the active stimulating electrode, with all of the current returning intra-cochlearly. Mens and Berenstein (2005) found no significant benefit with focused stimulation on monosyllabic word perception in quiet or in noise, although there was a trend of subjects preferring and performing better with the $\sigma = 0.5$ PTP strategy. Berenstein et al. (2008) performed a follow-up study, which compared speech perception with MP and PTP speech processing strategies in steady and fluctuating noise and found no benefit with PTP stimulation in speech perception. However, there may have been a potential problem with the analysis of the data in that study. Out of the 9 subjects tested, a subset of 5 subjects were given a strategy using $\sigma = 0.75$ (referred to as "RCF25" in Berenstein et al., (2008)) and the other 4 subjects were given a strategy using $\sigma = 0.25$ (referred to as "RCF75" in Berenstein et al., (2008)). The two subsets were collapsed into one group using TP/PTP stimulation in the speech perception analysis. However, studies (e.g. Bierer and Middlebrooks, 2002; Bonham and Litvak, 2008; Landsberger et al., 2012) have shown that $\sigma <= 0.5$ provides effectively the same current spread as MP stimulation. Therefore, it seems likely that the 4 subjects using $\sigma = 0.25$ would receive no benefit relative to MP stimulation as the electrical activation patterns of the neurons would be almost identical in the two conditions. Given that only 5 of the 9 subjects used a degree of current focusing shown to be potentially different from MP stimulation, it is possible that no difference between MP and PTP stimulation would be found because spectral improvements would be provided to only 5 of the 9 subjects.

In the present study, we compared speech perception in noise with a PTP strategy (using $\sigma = 0.75$ for all subjects) with matched (in terms of number of active electrodes, phase duration, stimulation rate, filters, hardware, loudness contributions of each electrode and experience) MP strategies. We expected to find that the degree of benefit with current focusing in the speech perception tasks would be related to the degree of reduction in spread of excitation with current focusing. Landsberger et al. (2012) proposed an adjective scaling task that correlated with the degree of spread of excitation reduction (measured with a psychophysical forward masking task) with PTP stimulation. Therefore, in this study, we used the Landsberger et al. (2012) adjective scaling task to estimate the degree of spread of excitation reduction with PTP stimulation for several electrodes across the array and correlated the results to the degree of speech perception benefit with current focusing.

2. Methods

2.1. Subjects

Six post-lingually deafened CI subjects implanted with the Advanced Bionics CII or HR90K device (none inserted with positioners) participated in this experiment. Five (C1, C3, C7, C8, and C9) of the subjects participated in the original adjective scaling task presented in Landsberger et al. (2012); part of that data is used in this study as well. Relevant subject demographics are shown in Table 1. All subjects were compensated for their participation and

Fig. 1. Illustration of Monopolar (MP) and Partial Tripolar (PTP) stimulation modes for a subset of electrodes in the array. "i" indicates the amount of stimulation current applied to the active electrode. For PTP stimulation, σ indicates the fraction of current returned by the intracochlear electrodes (i.e. the degree of current focusing). The oval beneath each diagram indicates the extra-cochlear return electrode. Note that only the first phase of an anodic-first biphasic pulse is shown. Theoretical regions of excitation for each stimulation mode are shown above each diagram.

Table 1
Relevant CI subject demographics.

Subject	Gender	Age	Device electrode type strategy	Duration of deafness prior to implantation (yrs)/duration of CI use (yrs)	HINT SRT with clinical strategy (in dB)
C1	M	78	CII HiFocus1J Fidelity 120	1/8	7.9
C3	F	54	HR90K HiFocus1J Fidelity120	11/5	7
C7	F	61	HR90K HiFocus1J Fidelity 120	45/5	1.7
C8	F	63	HR90K HiFocus1J Fidelity 120	1/3	7.4
C9	M	68	CII HiFocus1J HiRes	58/9	No open-set recognition
C14	M	45	HR90K HiFocus1J Fidelity 120	1/6	14.9

all provided informed consent in accordance with the local Institutional Review Board.

2.2. Experimental strategies

Experimental MP and PTP strategies (named Exp-MP and Exp-TP) were created for each subject. For both strategies, phase duration and per-channel stimulation rate were fixed at 140 μsec and 255 pps respectively. In a purely PTP strategy, electrodes 1 and 16 in the Advanced Bionics device cannot be used for active stimulation, as they are at the edges of the electrode array and do not have adjacent electrodes on both sides which can be used as grounds. For the purpose of fairly comparing the two strategies, both Exp-MP and Exp-TP strategies used electrodes 2–15 for active stimulation. Electrodes 1 and 16 were used exclusively for grounding in the Exp-TP strategy and provided no stimulation in the Exp-MP strategy. (It is unlikely that the grounding electrodes 1 and 16 in the Exp-TP strategy activated a substantially larger population of fibers near electrodes 1 and 16 relative to the Exp-MP strategy as studies (e.g. Landsberger et al., 2012) have shown that spread of excitation is similarly broad for TP, if not narrower.) A frequency allocation of 350–5500 Hz was used across the 14 channels. Strategies were programmed to the Platinum Series Processor (PSP) body worn processor using the Bionic Ear Programming System (BEPS). Impedance levels were measured prior to fitting to ensure stimulation levels were within compliance (Current levels in microamperes were kept below $7300/[R_{accessK\Omega} + (PhaseDur_{\mu sec}*0.01)]$, where $Ra_{ccessK\Omega}$ is the access resistance and $PhaseDur_{\mu sec}$ is the pulse phase duration – from Litvak (2012)).

Fitting parameters for the strategies were measured using the HRStream interface (Advanced Bionics Corp.) and a custom-built MATLAB front-end. All electrodes in both MP and PTP stimulation modes were loudness-balanced to PTP electrode 8, with the reference amplitude set at a "most comfortable level". Loudness balancing was performed by alternating stimulation on two electrodes and asking the subject to adjust the loudness of one relative to the other. By loudness balancing all electrodes across stimulation modes, we ensured that each electrode contributed an equivalent loudness to the overall strategy, e.g. the spectral shape was similar for both strategies. Each MP electrode was then compared to the same PTP electrode to verify the loudness balance for each

electrode across strategies – at this state, only very small adjustments (if any) were needed to achieve equivalent loudness as the loudness balancing from the previous step tended to be sufficient. Lastly, all electrodes were swept at the amplitude measured to be of equivalent "most comfortable" loudness to provide a final verification that all provided a similar, comfortable loudness.

Loudness-balanced current amplitudes were then imported into the BEPS fitting system. The electric dynamic range was fixed at 20 dB for each electrode (as is standard for mapping of Advanced Bionics devices) and the input acoustic dynamic range was 60 dB. As we found that there was no significant difference in electric dynamic range between MP and PTP stimulation for any subject (data not shown), the loudness mapping of the Exp-MP and Exp-TP strategies were likely similar (similar findings regarding similar dynamic ranges for MP and PTP stimulation averaged across subjects have been shown in Berenstein et al., 2008; Zhu et al., 2011). Speech perception and overall strategy were tested by playing live speech through the microphone through a sound-field in a sound booth. If the map was too soft, the entire map (threshold and MCL levels for all electrodes) was raised by a fixed (in dB) step size to maintain the same overall spectral shape across the array. Subject C7 required an increase of 2 dB for both Exp-MP and Exp-TP strategies, and subject C9 required an increase of 4 dB for both Exp-MP and Exp-TP strategies (none of the other subjects required an increase in loudness for either strategy). Strategies were then programmed into the PSP and subjects were instructed to toggle back and forth between the Exp-MP and Exp-TP strategies to verify that that both strategies provided comprehensible and similarly loud speech. Subjects were not informed about the differences between the two strategies and were only told that they would be tested on two experimental strategies that would sound very different relative to their clinical map.

The experimental strategies tested differed from the clinical strategies in several ways. The output filter corresponding to electrode 1 in the clinical map was shifted to electrode 2 (because electrode 1 was not used for active stimulation in either experimental strategy), resulting in a spectral shift basalward of 1.1 mm (adjacent electrodes in the Advanced Bionics HiFocus electrode array are 1.1 mm apart). Both experimental strategies used a phase duration of 140 μsec and a stimulation rate of 255 pps/e while the clinical maps used phase durations of either 18 usec (subjects C1, C3, C8, C9, C14) or 24.2 usec (subject C7) and stimulation rates of either 3712 pps/e (subjects C1, C3, C8, C9, C14) or 2750 pps/e (subject C7). Additionally, the PSP research processor necessitated the use of the microphone on the coil, while all subjects used the "t-mic" on their clinical, behind-the-ear processors. Consequently, both experimental strategies were novel, but were equivalently novel during testing.

In order to allow for adaptation to the novel experimental strategies, subjects were asked to listen to an audiobook for 40 min total. Subjects first listened to one of the two experimental strategies (randomly selected) and were switched to the other experimental strategy after 20 min. Subjects were provided a physical copy of the book and were allowed to read along with the audiobook for the first 5 min to familiarize themselves with the character names; the book was subsequently taken away and subjects were asked to follow the story by only listening. Subjects all reported that they could follow the audiobook story even without the aid of the book.

2.3. Speech perception tests

All speech tests were conducted in a double-walled sound-treated booth. Speech tests were conducted using TigerSpeech Technology (SoundExpress and I-STAR: www.tigerspeech.com)

software. Exp-MP and Exp-TP strategies were alternated between runs (where each run consists of one Speech Recognition Threshold [SRT] measurement) for each test to preclude learning effects across runs. All speech stimuli were presented at 65 dB SPL, and noise level was adaptively changed for both tests. Presentation level was calibrated with a 1 kHz pure tone (all stimuli were normalized to have the same long-term RMS level). Speech was presented in multi-talker babble noise for all tests. The patients' clinical map was tested after the two experimental strategies to serve as a comparison. For both tests, the experimenter was not blinded as to which strategy was being tested.

Speech recognition thresholds (SRTs) were adaptively measured using sentences from the Hearing in Noise Test (HINT) presented in multi-talker babble. All speech stimuli were presented at a fixed level of 65 dB SPL and the noise level was adaptively changed. The stimulus set consisted of 260 HINT sentences (Nilsson et al., 1994), which are short and grammatically simple. SRTs were adaptively measured using a 1-up/1-down rule. During testing, a sentence was selected randomly without replacement from the stimulus set and presented at the target SNR; the initial SNR was 30 dB. If the subject repeated at least 50% of the "key" words (that is, words conveying the content of the sentence rather than prepositions, conjunctions, etc. This distinction was at the discretion of the experimenter) correctly, the SNR was decreased by 2 dB; if the subject repeated less than 50% of the words correctly, the SNR was increased by 2 dB (i.e. Rule 3 from Chan et al., (2008)). Therefore, the SRT converged on the signal-to-noise ratio (SNR) that produced 50% correct word in sentence recognition (Levitt, 1971). Note that sentences were never repeated during an SRT measurement run, although it is possible a sentence may have been repeated across runs. Only one experimenter determined how many words were repeated correctly. The mean of the 8 reversals was recorded as the SRT for that trial. Typically, 3 SRT measurements were obtained for each subject and each strategy. If one of the measurements differed from the others by more than 3 dB, an additional SRT measurement was made for both strategies (therefore, the number of SRT measurements for each strategy was the same). The mean of all SRT measurements was recorded as the SRT for that strategy. HINT SRTs were measured for all subjects, with the exception of C9, who was not capable of open-set speech recognition at a 30 dB SNR with either of the experimental strategies or with his clinical processor.

SRTs for digits in multi-talker babble noise were also measured, using the same procedure as that used in Oba et al. (2011). In this task, 3 digits (from 0 to 9) were spoken in the presence of multi-talker babble. Speech stimuli were presented at a fixed level of 65 dB SPL and the noise level was adaptively changed. If all 3 digits were identified correctly, the SNR was lowered. Conversely, if a mistake was made, the SNR was raised. The SNR was adapted in a 1 up-1 down procedure with a step size of 2 dB. The number of stimulus presentations in each run was fixed at 25, and the average of all reversals was taken as the SRT for that run. Typically, 3 SRT measurements were made for each subject and each strategy. If one of the measurements differed from the others by more than 3 dB, an additional SRT measurement was made for each strategy. The mean of all SRT measurements was recorded as the SRT for that strategy.

2.4. Qualitative ratings of current focusing − clean/Dirty Index

Speech perception results were compared with a single-channel, qualitative metric of current focusing, specifically, the "Clean/Dirty" index used in Landsberger et al. (2012). In this task, subjects are presented with a pulse train on a single active electrode and are asked to rate how "clean" the stimulus is by clicking on a horizontal bar which corresponds to a continuum from "less

clean" to "more clean". Similarly, in separate trials, subjects are asked to rate how "dirty" the stimulus is, by clicking on a horizontal bar corresponding to a continuum from "less dirty" to "more dirty". The task is performed for stimuli with $\sigma = 0$ (MP) and $\sigma = 0.75$ (PTP) In Landsberger et al. (2012), data were collected only on electrode 9 for subjects C1, C3, C7, C8 and C9. In this study, we used the EL9 data from Landsberger et al. (2012) for the 5 subjects tested; additionally, data were collected for all subjects for electrodes 2, 4, 6, 8, 10, 12 and 14, using MP and PTP ($\sigma = 0.75$) stimuli.

As in Landsberger et al. (2012), a "Clean/Dirty Index" was computed for each electrode. The "Clean/Dirty Index" was defined as: |Clean Agreement Score $\sigma = 0$ − Dirty Agreement Score $\sigma = 0$| + |Clean Agreement Score $\sigma = 0.75$ − Dirty Agreement Score $\sigma = 0.75$|. The Clean/Dirty Index therefore quantifies the difference in how "clean" and "dirty" the stimuli sounded for MP and PTP stimulation (for more details, see Landsberger et al., 2012). For example, a Clean/Dirty Index near 0 indicates that both the MP and PTP stimulation on a given electrode sounded equivalently clean and dirty; by contrast, a larger Clean/Dirty Index (near 1) indicates that the PTP stimulus sounded more clean (and less dirty) compared to the MP stimulus on the same electrode. Landsberger et al., 2012 found a significant correlation between the Clean/Dirty Index and the degree of current focusing achieved with PTP ($\sigma = 0.75$) stimulation as measured with psychophysical forward masking. Therefore, in this study, we use the Clean/Dirty Index as a metric of degree of current focusing achieved on a given electrode.

All stimuli consisted of 300 msec biphasic pulse trains, with a phase duration of 226 μsec and a stimulation rate of 1000 pps. All measurements were repeated 15 times.

3. Results

The mean Exp-MP SRT for the HINT sentence test was 15.3 dB and the mean Exp-TP SRT was 12.29 dB. For the 5 subjects tested, all subjects received an improvement with PTP stimulation (range of improvement: 1.86 dB−4.2 dB), with a mean improvement of 3.01 dB. A two-tailed, paired t-test showed that the Exp-TP strategy yielded significantly better performance relative to the Exp-MP strategy ($t_4 = 6.26$, $p < 0.01$). Fig. 2a shows the SRTs for all subjects for Exp-MP and Exp-TP strategies (scatter plot).

The mean Exp-MP SRT for the digits in noise test was 3.79 dB and the mean Exp-TP SRT was 1.08 dB. 5 out of 6 subjects (all but subject C14) showed an improvement with PTP stimulation (range of improvement: −0.05 dB−4.98 dB), with a mean improvement of 2.71 dB. A two-tailed, paired t-test showed that this difference was statistically significant ($t_5 = 3.86$, $p < 0.05$). Fig. 2b shows the SRTs for all subjects for the Exp-MP and Exp-TP strategies.

The results for both the HINT test and the digits test show a large effect of stimulation mode. We found a Cohen's d of 2.8 for the HINT test and a d of 1.6 for the digits test, where a d of greater than 0.8 is considered a "large" effect. However, as relatively few subjects were used and to minimize the effect of possible outliers in the data, we repeated the statistical analyses with robust techniques. Percentile-t bootstrap analyses were used to compare the performance with the Exp-MP strategy with that of the Exp-TP strategy for both the HINT and digits tests. For all bootstrap analyses, each bootstrap distribution had the same number of samples as the original distribution (for more detail on the use of robust statistical techniques, see Aronoff et al., 2011). The percentile-t bootstrap analysis showed that the Exp-TP strategy yielded significantly better performance relative to the Exp-MP strategy for both the HINT task (CI: 1.35–4.73, $p < 0.05$) and for the digits task (CI: 0.77–4.67, $p < 0.05$). Therefore, both the paired t-test and the percentile-t bootstrap were in agreement, showing a significant effect of stimulation mode.
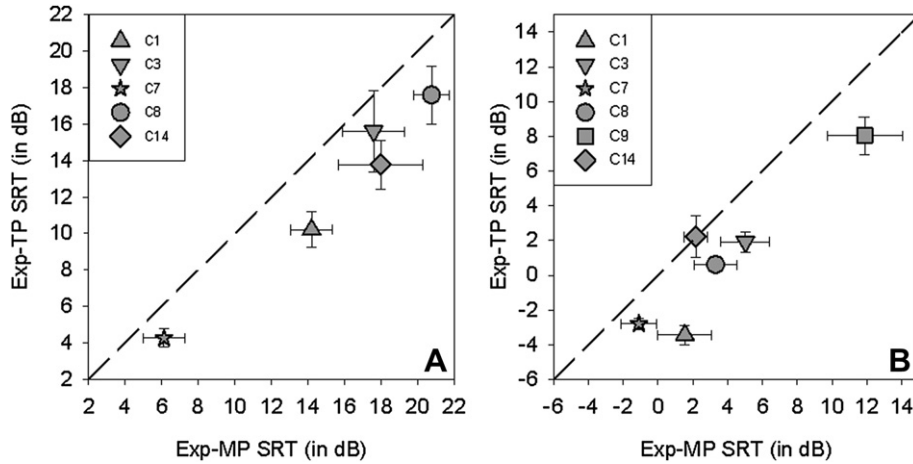
Fig. 2. Left (A) – a scatter plot showing the SRTs in dB for HINT sentences in babble noise for the Exp-TP strategy (*y*-axis) and the Exp-MP strategy (*x*-axis). Lower numbers indicate better performance. The dashed diagonal line is the line of equality, where Exp-TP performance is equal to Exp-MP performance. The Exp-TP strategy was significantly better than the Exp-MP strategy ($p < 0.05$). Error bars indicate 1 standard error of the mean. Right (B) – a similar scatter plot showing the SRTs in dB for the digits in noise task. The Exp-TP strategy was significantly better than the Exp-MP strategy ($p < 0.05$). Error bars indicate 1 standard error of the mean.

The subjects' clinical maps were better than both experimental strategies for all subjects in the HINT test, with a mean clinical SRT of 10.6 dB (relative to 15.3 dB for Exp-MP and 12.29 for Exp-TP). The difference was highly variable across subjects. For example, some subjects showed a fairly small difference between the clinical map and the experimental strategies (C1, C7), while other subjects showed a much larger difference (C3, C8). Fig. 3 shows the difference between subjects' clinical strategies and each experimental strategy for both the digits and sentences tasks. A paired *t*-test shows that the clinical strategy was better than the Exp-MP strategy ($t_4 = 3.93$, $p < 0.05$) but was not significantly different from the Exp-TP strategy ($t_4 = 2.13$, $p = 0.09$); however, with only 5 subjects, statistical power was low. The difference between the clinical maps

and the experimental strategies was much smaller for the digits in noise test, with a mean clinical SRT of 1.97 dB (relative to 3.79 dB for Exp-MP and 1.08 dB for Exp-TP), and the Exp-TP strategy was better on average than the clinical map. There was no statistically significant difference between the clinical strategy and either of the experimental strategies.

Most subjects received similar improvements with the Exp-TP strategy (relative to Exp-MP) for both the HINT sentence test and the digits in noise test. For example, subject C7, who received the smallest benefit (1.86 dB) in the HINT test received a comparable 1.7 dB improvement in the digits in noise test. Similarly, subject C1 who received a 3.9 dB improvement in the HINT test received a comparable 4.9 dB improvement in the digits in noise test. Fig. 4
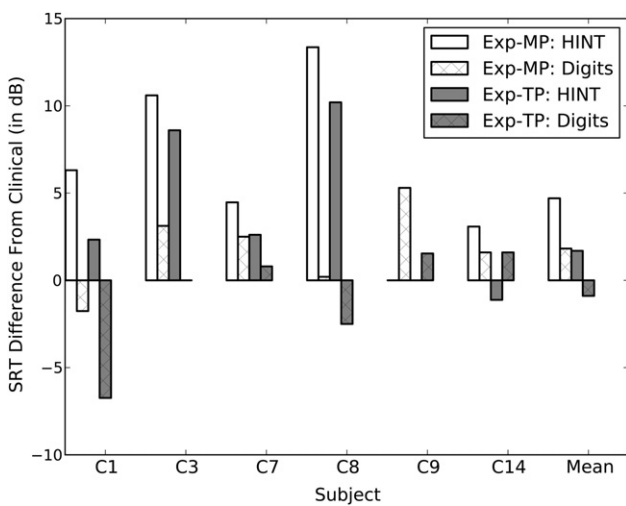

Fig. 3. The difference between each subjects' clinical strategy and the two experimental strategies for both the HINT and digits tests. Differences between the clinical strategy and the Exp-MP strategy are shown in the white bars for each subject and differences between the clinical strategy and the differences between the clinical strategy and the Exp-TP strategy are shown in the gray bars. Solid bars show differences between the clinical and experimental strategies for the HINT test, and hatched bars show differences for the digits test. Positive values indicate that the clinical strategy was better than the experimental. The bar for subject C3 for the Exp-TP strategy in the digits test is not visible as the difference was 0 dB. No bars are shown for subjects C9 for the HINT tests, as we were unable to measure HINT SRTs with any strategy (including clinical).
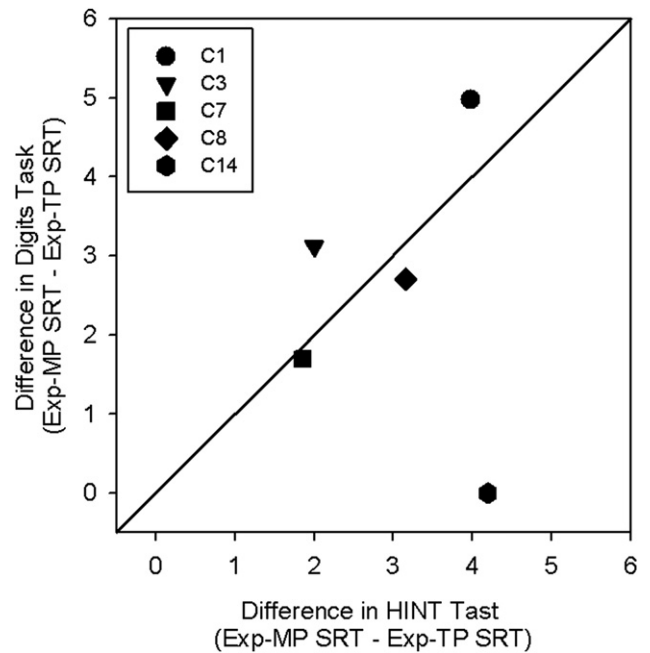

Fig. 4. A scatter plot comparing the improvement with the Exp-TP strategy (relative to the Exp-MP strategy) for the Digits (*y*-axis) and HINT (*x*-axis) tasks. Each symbol represents a different subject (no symbol is shown for subject C9 as we were unable to measure a HINT SRT). The solid line represents the line of equality, where there was equal improvement on both tests.

shows a scatter plot comparing the improvement with the Exp-TP strategy (relative to Exp-MP) in the digits test versus improvement in the HINT test (each point represents a subject); the line illustrates the line of equality, where improvement with the Exp-TP strategy was the same for both tests. For 4 of the 5 subjects shown (subject C9 did not have enough open-set speech recognition to get a HINT score, and is therefore not represented on this plot), the improvement for both tests is similar and the points cluster close to the line of equality. There was one exception to this pattern — subject C14 received the largest benefit from current focusing (4.2 dB) in the open-set sentence test but received no benefit in the closed set digits test. There was no statistically significant correlation between the improvement in the digits test and the improvement in the HINT test ($r^2 = 0.0009$, $p = 0.96$). Even if the data from C14 were removed, the correlation would still not be significant ($r^2 = 0.67$, $p = 0.18$).

The degree of improvement with the Exp-TP strategy did not seem to depend on the subject's absolute performance. For example, subject C7, who had the best performance for both tasks with the Exp-MP strategy (indicating that she was one of the best performers), received close to a 2 dB improvement in SRT with the Exp-TP strategy relative to the Exp-MP strategy on both tasks. On the other end of the performance spectrum, subject C9, who did not have enough open-set recognition to complete the HINT sentence task (indicating poorer overall performance, regardless of stimulation mode), received a 3.7 dB improvement with current focusing on the digits in noise task. Absolute monopolar performance (Exp-MP SRT) did not predict the degree of improvement from current focusing (Exp-MP SRT − Exp-TP SRT) for either the HINT test ($r^2 = 0.21$, $p = 0.73$) or for the digits in noise test ($r^2 = 0.12$, $p = 0.82$). Similarly, absolute performance with the Exp-TP strategy did not predict improvement for either the HINT test ($r^2 = 0.08$, $p = 0.65$) or the digits in noise test ($r^2 = 0.002$, $p = 0.93$), nor did absolute performance with the clinical maps ($r^2 = 0.63$, $p = 0.11$ and $r^2 = 0.34$, $p = 0.23$ for the HINT and digits tests respectively).

Because we hoped that the improvement in speech perception with the Exp-TP strategy (relative to performance with the Exp-MP strategy) could be predicted by a reduction in spread of excitation, we measured Clean/Dirty indices across the electrode array in order to estimate the reduction in spread of excitation for each subject. Fig. 5 shows Clean/Dirty indices for all electrodes tested for each subject (a larger Clean/Dirty index indicates a larger difference between MP and PTP stimuli). There was a broad range of Clean/Dirty indices across subjects. For instance, subjects C8 and C9 showed Clean/Dirty indices near 0 (i.e. presumably little to no effect of current focusing) for all electrodes tested. Conversely, subject C1 showed large Clean/Dirty indices for nearly all electrodes tested.

It is unclear whether or not the reduction in spread of excitation is equally important for each cochlear location. For example, it is possible that a reduction in spread is most important in the apex (or middle or base) for an improvement in speech processing with current focusing. Similarly, it is possible that the cochlear region that shows the greatest reduction in spread might predict an improvement in performance. Alternatively, it could be that improvements across the entire electrode array are necessary for a sizeable improvement in performance. Therefore, we correlated the Clean/Dirty Index on each electrode individually, the electrode for each subject witch provided the largest Clean/Dirty Index, and the average Clean/Dirty Index for each subject to the improvements in speech perception with the Exp-TP strategy for both the digits task and the HINT task. Even without control for type I error, none of the correlations were significant.

4. Discussion

There was a mean improvement of 3 dB with the PTP speech processing strategy for HINT SRTs and a 2.7 dB mean improvement for digits in noise SRTs relative to the experimental MP strategy. An improvement of 3 dB in HINT SRT has been shown to translate to roughly a 30% improvement in sentence intelligibility (e.g. Aronoff et al., 2011; Chan et al., 2008). Furthermore, subjects' performance with their clinical strategy was highly variable, from 1.7 dB on the HINT SRT test to no open set recognition; however, all subjects showed an improvement regardless of absolute performance level. The results of this study show that, with all other parameters equalized, current focusing, specifically PTP stimulation, can improve speech perception in noise. While other studies have shown that current focusing can improve psychophysical metrics
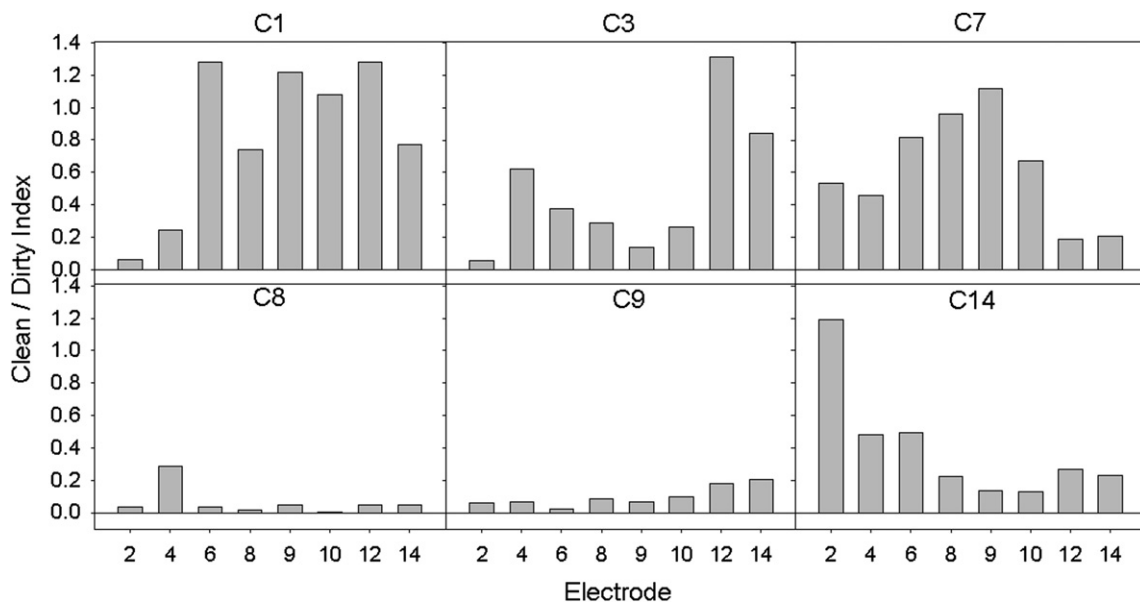

Fig. 5. Bar graphs showing Clean/Dirty Indices for each subject. The *x*-axis in each plot is electrode number. The *y*-axis indicates the Clean/Dirty index, where larger values indicate a larger difference between MP and PTP stimulation.

such as spread of excitation (e.g. Bierer and Faulkner, 2010; Srinivasan et al., 2010; Landsberger et al., 2012), virtual channel discrimination (Landsberger and Srinivasan, 2009; Srinivasan et al., 2012), and spectral ripple resolution (Berenstein et al., 2008), this study illustrates that the benefit of current focusing can translate to improved speech perception in noise. Although this study is limited by both the small number of subjects and that subjects were given only a short time to adapt to the tested experimental strategies, this study serves as a promising first step in showing the potential utility of current focusing in improving speech perception.

Although not reported in Oba et al. (2011), a significant correlation between digits in noise SRTs and HINT SRTs was found in that study (Oba, 2012). Therefore, we expected the size of the improvement with current focusing in the HINT task to be correlated with the size of the improvement with current focusing in the digits in noise task. However, although the size of the improvement was similar for all subjects other than C14 (see Fig. 4), we found no significant relationship. The lack of relationship may be due to the two tests measuring different underlying processes. The closed-set digits test had a relatively small set of stimuli, most of which could likely be discriminated by listening to only the vowel formants. Indeed, subjects reported particular difficulty discriminating "five" from "nine", which were two of the only tokens with a similar vowel sound. Although listening primarily to the vowel formants may have been sufficient to do the closed-set digits task, more information would be required to do the open-set HINT test. It is therefore possible that the lack of correlation in improvement between the two tests may be due to subjects using different cues for each task.

We had hoped to find a significant relationship between the degree of speech perception improvement from the Exp-TP strategy and the degree of current focusing achieved (as assessed by the Clean/Dirty Index) across the array. However, there did not appear to be any relationship between the two measures for either speech task. The lack of relationship between the Clean/Dirty Index and speech perception improvement is not surprising, given that all subjects showed improvements in speech perception with the Exp-TP strategy, but subjects C8 and C9 had Clean/Dirty indices near 0 for all electrodes tested across the array (suggesting little reduction in spread of excitation with PTP stimulation, based on Landsberger et al., (2012)). One possibility for the lack of relationship is that the Clean/Dirty Index is not an adequate predictor of the degree of current focusing. Although Landsberger et al. (2012) found a significant correlation between the Clean/Dirty Index and degree of current focusing (as measured by a psychophysical forward masking task), that study only used one electrode on each of 6 subjects. It is also possible that some subjects may have cognitive difficulty with the adjective scaling task – such a subject may show a reduction in spread of excitation measured with a psychophysical forward masking task, but may be unable to understand the adjective scaling task. This type of difficulty may not have been evident in the Landsberger et al. (2012) study as only one electrode was tested per subject. Further investigation (e.g. testing on more electrodes across the array) may reveal that the Clean/Dirty Index does not predict the degree of reduction in spread of excitation as fully as we had assumed. Perhaps a measure of spread of excitation at many locations along the array would provide a better predictor of performance than the Clean/Dirty index. Unfortunately, psychophysical forward-masked curves take much too long to measure to be clinically relevant. An alternative approach would be use eCAPS to measure spread of excitation, but pilot work in our lab has suggested that the long phase durations required to provide adequate loudness make artifact cancellation difficult. Another possibility is that the reduction in spread of excitation on a single channel is not a good predictor of speech perception. As other

studies (e.g. Hughes and Abbas, 2006; Zwolan et al., 1997) have shown no significant correlation between single-channel metrics (such as pitch ranking or electrode discrimination) and speech perception, it is possible that a single channel spread of excitation measure might not account for across–channel interactions that impact speech perception.

Subjects typically performed better with their clinical maps than with the experimental strategies. This was to be expected, as the experimental strategies were very different than the subjects' clinical map in several crucial ways. Differences include the microphone (on the coil attached to the side of the head for the experimental strategies vs. a "t-mic" hanging over the ear for the clinical maps), a reduction in the number of active electrodes (the experimental strategies used 14 channels vs. 16 physical channels or 120 virtual channels for the clinical), strategy (the experimental strategies used a CIS implementation vs. Fidelity 120 for most subjects), stimulation rate (255 pps/e for the experimental strategy vs. greater than 2700 pps/e for the clinical maps), and a spectral shift (the output of the most apical filterbank was sent to channel 2 in the experimental strategies, resulting in a basalward spectral shift of 1.1 mm). All of these changes are likely to result in an immediate drop in performance when acutely tested. In a study measuring the difference in performance between the Continuous Interleaved Sampling (CIS+) strategy (Wilson et al., 1991) and the Fine Structure Processing (FSP) strategy (Hochmair et al., 2006) (with new external hardware), subjects initially had an approximately 4 dB decrement in SRT performance when they were switched from the familiar CIS + strategy to the unfamiliar FSP strategy and were tested acutely (Vermeire et al., 2010). However, over the course of a year, subjects adapted to the different hardware and signal processing and performance with the FSP strategy improved to the point where performance with FSP (and with the new hardware) was better than the initial CIS + baseline. Similarly, Dorman and Loizou (1997) showed that the representation of vowels can take up to a month to stabilize when changing stimulation mode (from analog stimulation to pulsatile CIS, in that study). Other studies investigating the drop in performance from more severe spectral shifts than that imposed in this study have found that the acute drop is substantial, but that subjects adapt significantly over a period of days to weeks (e.g. Rosen et al., 1999; Fu and Galvin, 2003). We would therefore expect that CI users would adapt to both experimental strategies tested, had they been given a few weeks of chronic exposure prior to testing. However, due to limitations in the research processor equipment, the experimental strategies were only tested acutely in the lab over a period of a few hours, prohibiting full adaptation (although they received 40 min of adaptation). Subjects had much more practice with their own clinical processors, and were therefore expected to perform better with them. Allowing subjects to more fully adapt to the experimental strategies would enable more fair comparisons of the experimental strategies to the subjects' clinical processors in future experiments.

The present experiment used experimental speech processing strategies which were designed to isolate and only test the effect of current focusing. All parameters across the two experimental strategies (e.g. hardware, filterbank allocation, number of active electrodes, stimulation rate, etc) were kept the same, resulting in fairly simple strategies that were ideal for initial research testing as all of the parameters were tightly controlled. However, both the Exp-MP and Exp-TP strategies tested here were sub-optimal for clinical implementation in terms of the number of channels and stimulation rate. In the experimental strategies, electrodes 1 and 16 were not used for active stimulation (and were only used as grounds). In a clinical implementation of PTP stimulation, monopolar, bipolar, or phantom electrode stimulation (e.g. Saoji

and Litvak, 2010) could be used for electrodes 1 and 16, thereby allowing greater spatial coverage of the cochlea. Secondly, the experimental strategies tested both used a stimulation rate of 255 pulses per second per electrode (pps/e), which is lower than most clinical stimulation rates. However, this may not inherently be a problem, as studies have shown that there is no consistent benefit to higher stimulation rates across subjects (e.g. Vandali et al., 2000). Although higher stimulation rates may not be necessary, the stimulation rate of a PTP strategy could be increased through several methods. An n-of-m peak-picking algorithm (similar to the ACE strategy of Cochlear Corp.) would stimulate fewer channels in one sweep and would increase the overall stimulation rate (discussed in more detail in Landsberger and Srinivasan, 2009). Another option might be to use paired stimulation to stimulate two channels with a large spatial separation (e.g. ELs 1 and 9) simultaneously, effectively doubling the per-channel stimulation rate. Modeling results (Frijns, 2012) suggest that the region of neural excitation for paired and sequential stimulation is almost identical for PTP stimulation. Finally, the same benefits of current focusing to speech perception might be realized by only providing sharpening for a subset of electrodes (e.g. specific cochlear regions, such as the apical and middle regions, or in regions where subjects get a definite psychophysical benefit with current focusing) while using shorter duration MP pulses on the remainder of the electrodes. A mixed-mode strategy would effectively increase the per-channel stimulation rate across the array. Given that current focusing yields benefits to speech perception, a more thorough exploration of methods of optimizing strategies for clinical implementation is warranted.

Acknowledgements

References

Aronoff, J.M., Freed, D.J., Fisher, L.M., Pal, I., Soli, S.D., 2011. The effect of different cochlear implant microphones on acoustic hearing individuals' binaural benefits for speech perception in noise. Ear Hear. 32, 468–484.

Berenstein, C.K., Mens, L.H.M., Mulder, J.J.S., Vanpouke, F.J., 2008. Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar, and virtual channel electrode configurations. Ear Hear. 29, 250–260.

Bierer, J.A., 2007. Threshold and channel interaction in cochlear implant users: evaluation of the tripolar electrode configuration. J. Acoust. Soc. Am. 121, 1642–1653.

Bierer, J.A., Faulkner, K.F., 2010. Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves. Ear Hear. 31, 247–258.

Bierer, J.A., Middlebrooks, J.C., 2002. Auditory cortical images of cochlear-implant stimuli: dependence on electrode configuration. J. Neurophysiol. 87, 478–492.

Bonham, B.H., Litvak, L.M., 2008. Current focusing and steering: modeling, physiology, and psychophysics. Hear. Res. 242, 141–153.

Burns, E.M., Sanborn, E.S., Shannon, R.V., Fu, Q.J, 2001. Perception of familiar melodies by implant users, Conference of Implantable Auditory Prosthesis, Vol. 81, Pacific Grove, CA.

Chan, J.C., Freed, D.J., Vermiglio, A.J., Soli, S.D., 2008. Evaluation of binaural functions in bilateral cochlear implant users. Int. J. Audiol. 47, 296–310.

Chatterjee, M., Shannon, R.V., 1998. Forward masked excitation patterns in multi-electrode electrical stimulation. J. Acoust. Soc. Am. 103, 2565–2572.

Dorman, M.F., Loizou, P.C., 1997. Mechanisms of vowel recognition for Ineraid patients fit with continuous interleaved sampling processors. J. Acoust. Soc. Am. 102, 581–587.

Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J. Acoust. Soc. Am. 110, 1150–1163.

Frijns, 2012. Personal communication.

Fu, Q.J., Galvin 3rd, J.J., 2003. The effects of short-term training for spectrally mismatched noise-band speech. J. Acoust. Soc. Am. 113, 1065–1072.

Fu, Q.J., Nogaki, G., 2005. Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. J. Assoc. Res. Otolaryngol. 6, 19–27.

Fu, Q.J., Shannon, R.V., Wang, X., 1998. Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. J. Acoust. Soc. Am. 104, 3586–3596.

Henry, B.A., Turner, C.W., Behrens, A., 2005. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. J. Acoust. Soc. Am. 118, 1111–1121.

Hughes, M.L., Abbas, P.J., 2006. The relation between electrophysiologic channel interaction and electrode pitch ranking in cochlear implant recipients. J. Acoust. Soc. Am. 119, 1527–1537.

Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schosser, H., Garnham, C., Anderson, I., 2006. Med-EL cochlear implants: state of the art and a glimpse into the future. Trends Amplif. 10, 201–219.

Landsberger, D.M., Padilla, M., Srinivasan, A.G., 2012. Reducing current spread using current focusing in cochlear implant users. Hear. Res. 284, 16–24.

Landsberger, D.M., Srinivasan, A.G., 2009. Virtual channel discrimination is improved by current focusing in cochlear implant recipients. Hear. Res. 254, 34–41.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49 (Suppl 2), 467.

Litvak, L.M., Spahr, A.J., Emadi, G., 2007a. Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. J. Acoust. Soc. Am. 122, 967–981.

Litvak, L.M., Spahr, A.J., Saoji, A.A., Fridman, G.Y., 2007b. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. J. Acoust. Soc. Am. 122, 982–991.

Litvak, 2012. Personal communication.

Mens, L.H., Berenstein, C.K., 2005. Speech perception with mono- and quadrupolar electrode configurations: a crossover study. Otol Neurotol. 26, 957–964.

Nilsson, M., Soli, S.D., Sullivan, J.A., 1994. Development of the Hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am. 95, 1085–1099.

Oba, S.I., Fu, Q.J., Galvin 3rd, J.J., 2011. Digit training in noise can improve cochlear implant users' speech understanding in noise. Ear Hear. 32, 573–581.

Oba, 2012. Personal communication.

Rosen, S., Faulkner, A., Wilkinson, L., 1999. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. J. Acoust. Soc. Am. 106, 3629–3636.

Saoji, A.A., Litvak, L.M., 2010. Use of "phantom electrode" technique to extend the range of pitches available through a cochlear implant. Ear Hear. 31, 693–701.

Shannon, R.V., Fu, Q.J., Galvin 3rd, J., 2004. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. Acta Otolaryngol. Suppl., 50–54.

Smith, Z.M., Delgutte, B., Oxenham, A.J., 2002. Chimaeric sounds reveal dichotomies in auditory perception. Nature 416, 87–90.

Snyder, R.L., Bierer, J.A., Middlebrooks, J.C., 2004. Topographic spread of inferior colliculus activation in response to acoustic and intracochlear electric stimulation. J. Assoc. Res. Otolaryngol. 5, 305–322.

Srinivasan, A.G., Landsberger, D.M., Shannon, R.V., 2010. Current focusing sharpens local peaks of excitation in cochlear implant stimulation. Hear. Res. 270, 89–100.

Srinivasan, A.G., Shannon, R.V., Landsberger, D.M., 2012. Improving virtual channel discrimination in a multi-channel context. Hear. Res. 286, 19–29.

Vandali, A.E., Whitford, L.A., Plant, K.L., Clark, G.M., 2000. Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. Ear Hear. 21, 608–624.

Vermeire, K., Punte, A.K., Van de Heyning, P., 2010. Better speech recognition in noise with the fine structure processing coding strategy. ORL J. Otorhinolaryngol. Relat. Spec. 72, 305–311.

Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D.K., Rabinowitz, W.M., 1991. Better speech recognition with cochlear implants. Nature 352, 236–238.

Wilson, B., Lawson, D., Zerbi, M., 1995. Speech Processors for Auditory Prostheses. Final Report of contract # N01-DC-2-2401. National Institutes of Health, Bethesda, MD.

Zhu, Z., Tang, Q., Zeng, F.G., Guan, T., Ye, D., 2011. Cochlear-implant spatial selectivity with monopolar, bipolar and tripolar stimulation. Hear. Res. 283, 45–58.

Zwolan, T.A., Collins, L.M., Wakefield, G.H., 1997. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. J. Acoust. Soc. Am. 102, 3673–3685.