

Digit Training in Noise Can Improve Cochlear Implant Users' Speech Understanding in Noise

Sandra I. Oba,¹ Qian-Jie Fu,^{1,2} and John J. Galvin III¹

Objective: While auditory training in quiet has been shown to improve cochlear implant (CI) users' speech understanding in quiet, it is unclear whether training in noise will benefit speech understanding in noise. The present study investigated whether auditory training could improve CI users' speech recognition in noise and whether training with familiar stimuli in an easy listening task (closed-set digit recognition) would improve recognition of unfamiliar stimuli in a more difficult task (open-set sentence recognition).

Design: CI users' speech understanding in noise was assessed before, during, and after auditory training with a closed-set recognition task (digits identification) in speech babble. Before training was begun, recognition of digits, Hearing in Noise Test (HINT) sentences, and IEEE sentences presented in steady speech-shaped noise or multitalker speech babble was repeatedly measured to establish a stable estimate of baseline performance. After completing baseline measures, participants trained at home on their personal computers using custom software for approximately 30 mins/day, 5 days/wk, for 4 wks, for a total of 10 hrs of training. Participants were trained only to identify random sequences of three digits presented in speech babble, using a closed-set task. During training, the signal-to-noise ratio was adjusted according to subject performance; auditory and visual feedback was provided. Recognition of digits, HINT sentences, and IEEE sentences in steady noise and speech babble was remeasured after the second and fourth week of training. Training was stopped after the fourth week, and subjects returned to the laboratory 1 mo later for follow-up testing to see whether any training benefits had been retained.

Results: Mean results showed that the digit training in babble significantly improved digit recognition in babble (which was trained) and in steady noise (which was not trained). The training benefit generalized to improved HINT and IEEE sentence recognition in both types of noise. Training benefits were largely retained in follow-up measures made 1 mo after training was stopped.

Conclusions: The results demonstrated that auditory training in noise significantly improved CI users' speech performance in noise, and that training with simple stimuli using an easy closed-set listening task improved performance with difficult stimuli and a difficult open-set listening task.

(*Ear & Hearing* 2011;32:573–581)

INTRODUCTION

Auditory training has been shown to improve cochlear implant (CI) users' speech understanding in quiet (Fu et al. 2005; Stacey et al. 2010). Unfortunately, CI users rarely experience quiet, optimal listening conditions in everyday life. Background noise is problematic for CI users (Skinner et al. 1994; Fu et al. 1998), especially dynamic noise (Nelson et al. 2003; Fu & Nogaki 2005). The relatively coarse spectral resolution of the CI is not sufficient to segregate speech from noise or one talker from another.

¹Division of Communication and Auditory Neuroscience, House Ear Institute; and ²Department of Biomedical Engineering and Neuroscience Program, University of Southern California, Los Angeles, California.

While training in quiet may direct listeners' attention to subtle auditory cues, it is unclear whether listeners can make use of these cues in noisy listening conditions. Indeed, there are few reports regarding the benefit of auditory training for speech understanding in noise for CI users. A pilot study by Fu and Galvin (2008) compared two training protocols (recognition of monosyllable words or sentences in noise) on CI users' speech understanding in noise. While recognition performance improved with both protocols, the sentence-in-noise training provided the greatest improvement in both test measures (phoneme or sentence recognition) and both types of noise (speech babble and steady noise).

Training outcomes for CI users and normal-hearing (NH) listeners attending to CI simulations can be affected by training frequency (Nogaki et al. 2007), as well as the training protocol and materials (Davis et al. 2005; Fu et al. 2005; Li & Fu 2007; Loebach & Pisoni 2008; Stacey & Summerfield 2007, 2008; Stacey et al. 2010). The time course and degree of adaptation can depend on the listening task or condition and can differ across individuals (Fu et al. 2005). Nogaki et al. (2007) found that, in a CI simulation study with NH listeners, the frequency of training sessions mattered less than the total amount of training completed. Fu et al. (2005) and Stacey and Summerfield (2007) suggested that multitalker training materials may be more effective than single-talker training stimuli. Stacey and Summerfield (2008) found that, in a CI simulation study with NH listeners, words and sentences were more effective training stimuli than phonemes. In other CI simulation studies with NH listeners, training outcomes were better with lexically meaningful training material and feedback than without (Davis et al. 2005; Li & Fu 2007). While these factors may affect training outcomes for speech in quiet, it is unclear how or whether they may affect speech understanding in noise. Most of these training studies focused listeners' attention to acoustic differences between stimuli. As such, these training methods might have targeted peripheral, "bottom-up" processes. While bottom-up processes are certainly important, central "top-down" processes contribute strongly to speech understanding in noise. It is possible that top-down processes should be targeted to improve CI users' performance in noise.

Auditory training for CI users has been receiving increased attention. Much research has been aimed at finding the most effective and efficient training methods. Computer-based auditory training has been used in CI studies (Fu et al. 2004, 2005; Galvin et al. 2007; Stacey et al. 2010) and in CI simulation studies with NH listeners (Nogaki et al. 2007; Stacey & Summerfield 2007, 2008; Loebach & Pisoni 2008). Computer-based auditory training allows for large numbers of training stimuli and listening tasks, with unlimited training time. Ideally, training for relatively short periods with a limited number of stimuli and a simple listening task can generalize to a larger number of stimuli and listening conditions. Some

TABLE 1. CI participant demographics

| Subject | Gender | Age (yrs) | Etiology | Age at HL Onset (yrs) | Duration of CI Use (yrs) | Device | Processor | Preprocessing |
|---------|--------|-----------|----------------|-----------------------|--------------------------|---------------------|-----------|----------------|
| S1 | F | 62 | Ototoxic drugs | 28 | 10 | Nucleus-24 (L) | Freedom | ADRO & Whisper |
| | | | | | 6 | Nucleus-24 (R) | Freedom | |
| S2 | F | 75 | Unknown | 37 | 8 | Nucleus-24 (R) | Freedom | ADRO & ASC |
| | | | | | | HA (L) | | |
| S3 | M | 67 | Unknown | 15 | 5 | Nucleus-24 (R) | Freedom | ADRO |
| | | | | | | HA (L) | | |
| S4 | F | 65 | Genetic | 4 | 5 | Nucleus-24 (R) | Freedom | ADRO |
| S5 | F | 70 | Genetic | 5 | 7 | Clarion-CII (L) | CII BTE | NA |
| | | | | | | HA (R) | | |
| S6 | F | 75 | Otosclerosis | ~20 | 20 | Nucleus-22 (L) | Freedom | ADRO |
| S7 | F | 46 | Unknown | 39 | 2 | Nucleus Freedom (R) | Freedom | ADRO |
| S8 | M | 78 | Noise | 40 | 14 | Nucleus-22 (L) | Freedom | ADRO & ASC |
| S9 | M | 69 | Otosclerosis | 25 | 1 | Nucleus Freedom (L) | Freedom | ADRO |
| | | | | | | HA (R) | | |
| S10 | M | 57 | Unknown | 37 | 15 | Nucleus-22 (L) | Freedom | ADRO |
| | | | | | 1 | Nucleus Freedom (R) | Freedom | |

HA, hearing aid; ADRO, Adaptive Dynamic Range Optimization; ASC, Autosensitivity Control.

previous CI training studies have shown such generalizations. Fu et al. (2005) found that closed-set phonetic contrast training with monosyllable words generalized to improved vowel, consonant, and sentence recognition. Galvin et al. (2007) found that closed-set melodic contour identification (MCI) training generalized to improved MCI performance for untrained pitch ranges and improved familiar melody recognition. Loebach and Pisoni (2008) found that, in a CI simulation study with NH listeners, environmental sound recognition training significantly improved speech recognition; however, speech recognition training did not improve environmental sound recognition. All of the aforementioned training benefits and generalizations were observed for optimal listening environments (e.g., clean speech in quiet); it is unclear whether these might occur for noisy listening conditions. It is possible that a simple training task with simple stimuli (i.e., minimal acoustic variability) may target top-down processes that may be useful in developing listening strategies to compensate for noise.

The aims of this study were (1) to determine whether CI users' speech recognition in noise would improve with auditory training and (2) to determine whether training with familiar stimuli, in a closed-set task, would yield improvements in the more difficult task of open-set speech recognition in noise. In this study, 10 CI users were trained to identify digits presented in competing speech babble using a closed-set listening task. Speech recognition was measured for digits, Hearing in Noise Test (HINT) sentences (Nilsson et al. 1994), and IEEE sentences (1969) presented in two types of noise (steady state, speech-shaped noise and multitalker speech babble). "Pretraining" baseline performance for all speech tests was repeatedly measured for a minimum of four test sessions or until performance asymptoted. Participants were then trained only to identify digits in babble using custom training software loaded onto their home computers or loaner laptops. Participants trained for ~30 mins per day, 5 days per week for 4 wks, for a total of approximately 10 hrs of training. During training, the signal-to-noise ratio (SNR) was adjusted from trial to trial according to the correctness of the listener's response. Auditory and visual feedback was provided. Recognition of digits,

HINT sentences, and IEEE sentences in steady noise and babble was remeasured 2 and 4 wks after training was begun ("post-training measures"), as well 4 wks after training was stopped ("follow-up measures"). We hypothesized that the digit-in-babble training would significantly improve speech understanding in both types of noise and for all speech measures.

PARTICIPANTS AND METHODS

Participants

Ten postlingually deafened adult CI users participated in this experiment. Relevant demographic information is shown in Table 1. All participants were native speakers of American English. Seven participants had previous experience in speech recognition experiments. Three participants (S3, S7, and S9) had no prior experience with CI research. All provided informed consent before participating (in compliance with the local Institutional Review Board protocol) and all were reimbursed for their time and expenses associated with testing in the laboratory and training at home (equivalent to approximately \$15.00/hr).

Participants were tested using their clinical speech processors, set to their preferred "everyday" program and volume control settings. If a participant normally wore a hearing aid (HA) in combination with the CI, then testing and training were performed with the CI and HA set to the everyday settings. Participants were instructed to use the same CI (and HA) settings for all testing and training sessions. Participants with the Nucleus Freedom processor used preprocessing technology with their everyday programs for all testing and training. Some research (Müller-Deile et al. 2008; Wolfe et al. 2009; Gifford & Revit 2010) has shown that preprocessing strategies may improve speech perception in noise. Note that although the preprocessing may have influenced performance, we were interested in the change in performance after training for each subject.

General Testing and Training Timeline

Because of the high variability in performance among CI users, which might overwhelm any training benefits, a “within-subject” control procedure (instead of an experimental control group) was used, with each participant serving as their own experimental control. The within-subject control procedure required extensive baseline performance measures (see below) before training was begun. The within-subject control procedure also allowed procedural learning effects to be estimated independently for each participant. Baseline performance was repeatedly measured in the laboratory once a week for 4 wks or until achieving asymptotic performance. For each subject, asymptotic performance was determined using a one-way analysis of variance (ANOVA), with test session as a factor. Data from the digit and HINT sentence recognition tests in both types of noise were pooled within the ANOVA to obtain a broad estimate of asymptotic performance. Post hoc Bonferroni pairwise comparisons were performed to determine the test session at which performance began to asymptote. Data from this test session forward were used to calculate baseline performance within each test and noise condition.

After completing baseline “pretraining” measures, participants trained at home on their personal computers for approximately 30 mins per day, 5 days a week, for 4 wks. “Post-training” performance for all speech tests and noise types was measured after completing the second and fourth week of training. Training was stopped after the fourth week, and participants returned to the laboratory 1 mo later for “follow-up” performance measures for all speech tests and noise types. Post-training and follow-up performance was compared with baseline performance for each subject.

Test Materials and Methods

Speech performance was measured in steady speech-shaped noise (1000-Hz cutoff frequency, -12 dB/octave) and in six-talker speech babble. The SNR was calculated according to the long-term RMS of the speech and noise stimuli. Pretraining, post-training, and follow-up performances were assessed using three speech measures: digits in noise, HINT sentences in noise, and IEEE sentences in noise.

For digit-in-noise testing, speech recognition thresholds (SRTs; Plomp & Mimpen 1979) were measured using a closed-set, adaptive (one-up/one-down) procedure converging on the SNR that produced 50% correct (Levitt 1971). SRTs were defined as the SNR needed to produce 50% correct recognition of a three-digit sequence. Stimuli included digits “zero” through “nine” produced by a single male talker. During testing, three digits were randomly selected and presented in sequence (e.g., “three-zero-eight”) in the presence of background noise. The listener responded by clicking on response boxes (labeled “zero” through “nine”) shown on the computer screen or by typing in the numbers on the keyboard. Listeners were allowed to repeat the stimulus up to three times. If the listener correctly identified the entire sequence, the SNR was reduced by 2 dB; if the entire sequence was not correctly identified, the SNR was increased by 2 dB. Each test run was 25 trials; all reversals in SNR were used to calculate the SRT. Three runs in steady noise and in speech babble were measured at each test session.

For HINT sentences in noise, SRTs were measured using an open-set, adaptive (one-up/one-down) procedure (Levitt 1971).

HINT stimuli included 260 sentences produced by a single male talker, which were of easy to moderate difficulty in terms of vocabulary and syntactic complexity (e.g., “Strawberry jam is sweet.”). During each test run, a sentence was randomly selected (without replacement) from the 260-sentence stimulus set and presented in noise. Listeners were asked to repeat what they heard as accurately as possible. If the entire sentence was repeated correctly, the SNR was reduced by 2 dB; if not, the SNR was increased by 2 dB. On average, 15 sentences were presented within each test run, and the mean of the final six reversals in SNR was recorded as the SRT; a minimum of eight reversals in SNR was required for a valid test run. Results for HINT recognition in noise were reported in terms of the SRT, i.e., the SNR required to produce 50% correct whole sentence recognition. Three runs in steady noise and speech babble were measured at each test session. Note that sentences were not repeated within a test run. However, it is possible that some sentences were repeated across runs.

IEEE sentence recognition was measured at two SNRs that corresponded to moderately and extremely difficult. The SNRs were determined for each subject according to performance at the first test session. The HINT SRT in steady noise was used to estimate the moderately difficult SNR (i.e., the SNR that would produce approximately 50% IEEE sentence recognition in steady noise). Typically, the moderately difficult SNR for IEEE sentences was 2 to 3 dB higher than the HINT SRT. The extremely difficult SNR was 5 dB less than the moderately difficult SNR. Thus, if a subject’s HINT SRT was 8 dB, IEEE sentence recognition was measured +10 dB and +5 dB SNRs. IEEE stimuli included 720 sentences divided into 72 lists (10 sentences per list), produced by one male and one female talker. IEEE sentences are moderately difficult and much more difficult than HINT sentences in terms of semantic and syntactic complexity and word predictability (e.g., “Plead to the council to free the poor thief.”). During testing, a list was randomly chosen (without replacement) and a sentence was randomly chosen (without replacement) from the list and presented at the target SNR. Listeners were asked to repeat back what they heard as accurately as possible. Results were reported in terms of the percent of correctly identified words. Within a test session, performance was measured independently for each talker, noise type, and SNR. The test order for the different talker, noise type, and SNR conditions was randomized across test sessions and across subjects. In each test session, performance for each condition was evaluated using one sentence list. Thus, no IEEE sentences were repeated within or across test runs, as a different list was chosen for each condition at each test session.

Both HINT and IEEE sentences were used because they provide somewhat different measures of performance in noise. HINT SRTs (corresponding to 50% correct whole sentence recognition in noise) were adaptively measured, i.e., the SNR was adjusted from trial to trial according to the correctness of response, thereby avoiding floor and ceiling performance effects. IEEE sentence recognition was measured at two SNRs (moderate and difficult) that were fixed for each subject. While the target SNRs were selected to accommodate different performance levels across subjects, floor and ceiling effects may have influenced performance in some cases.

Testing was conducted in sound field in a sound-treated booth (IAC). Combined speech and noise were delivered via

single loudspeaker (Tannoy Reveal) at a fixed output (65 dBA) to avoid severe peak clipping by the CI speech processor. The target SNR was calculated according to the long-term RMS of the speech and noise; the level of the combined speech and noise was then adjusted to achieve the target output (65 dBA). Listeners were seated directly facing the loudspeaker. No training or trial-by-trial feedback was provided during test sessions. The same performance measures (digit recognition, HINT sentences, and IEEE sentences), noise types (steady state and speech babble), and test methods were used for all test sessions (pretraining, post-training, and follow-up).

Training Materials and Methods

After completing pretraining baseline measures, training was begun. Training was conducted at home, using participants' personal computers or loaner laptops loaded with custom training software (Sound Express). Participants trained while listening to stimuli play back via computer speakers. Participants were extensively trained on how to install and use the software, how to set up the computer speakers, and how set the listening level. Participants were also instructed to immediately contact the experimenters if there was any problem with the home training. The software also allowed for remote monitoring of training frequency and performance; the software logged each participant's training sessions, including performance and the total time spent training. Participants were instructed to train using the same CI (and HA) settings used for testing. Participants were asked to train for 30 mins/day, 5 days/week, for 4 wks, for a total of 10 training hours.

The same stimuli used for digit-in-noise testing were used for digit-in-noise training. However, participants trained only in competing speech babble; note that subjects were tested in both speech babble and steady noise. The training was computer administered. Similar to testing, participants were trained using an adaptive (one-up/one-down) SRT procedure. In each trial, three digits were presented in random sequence in the presence of competing speech babble; participants were allowed to repeat each stimulus presentation up to three times. Participants responded by clicking on numbers shown on the computer screen or by typing the numbers on the keyboard. If the participant answered correctly, visual feedback was provided, a new three-digit sequence was presented, and the SNR was reduced by 2 dB. If the participant answered incorrectly, audio and visual feedback was provided, in which the incorrect response and the correct response were played back and repeated. Then, the next digit sequence was presented and the SNR was increased by 2 dB. Each training exercise consisted of 25 three-digit sequences. Similar to the digit-in-noise testing, the SRT for digit-in-noise training was calculated across all reversals in SNR at the end of each run.

RESULTS

All participants completed the specified training; the total time spent training ranged from 583 to 767 mins with a mean of 647 mins (10.75 hrs). After training, performance improved for all subjects on some or all test measures; however, there was considerable variability in the amount of improvement.

Figure 1 shows individual and mean pre- and post-training (4 wks) SRTs for digit recognition. In steady noise, mean SRTs significantly improved from -2.2 to -5.0 dB

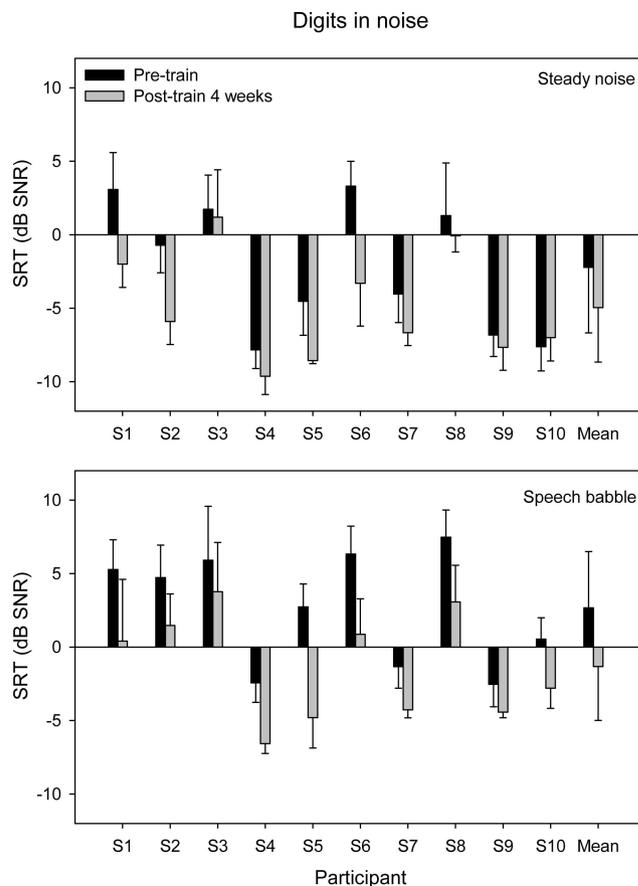


Fig. 1. Speech reception thresholds (SRTs) for digit recognition in speech-shaped steady noise (top panel) and multitalker speech babble (bottom panel). Mean data (across trials) are shown for each participant for pretraining performance and performance after 4 wks of training. Mean data (across participants) are shown at the far right. The error bars show 1SD of the mean.

after training ($t_9 = 3.66$, $p = 0.005$). Post-training SRTs worsened by 0.6 dB for subject S10 but improved for subjects S1, S2, S3, S4, S5, S6, S7, S8, and S9 (range = 0.5 to 6.6 dB). In speech babble, mean SRTs significantly improved from 2.7 to -1.3 dB after training ($t_9 = 7.50$, $p < 0.001$). Post-training SRTs improved for all 10 subjects (range = 1.9 to 7.5 dB).

Figure 2 shows individual and mean pre- and post-training (4 wks) SRTs for HINT sentence recognition. In steady noise, mean SRTs improved from 6.6 to 5.6 dB after training; however, the improvement was not significant ($t_9 = 1.83$, $p = 0.093$). Post-training SRTs worsened for subjects S4, S6, and S8, (range = 0.5 to 1.4 dB), improved for subjects S1, S2, S3, S7, S9, and S10, (range = 0.6 to 4.5 dB), and were unchanged for subject S5. In speech babble, mean SRTs significantly improved from 11.2 to 8.3 dB after training ($t_9 = 4.037$, $p = 0.003$). Post-training SRTs worsened by 0.5 dB for subject S8 and improved for subjects S1, S2, S3, S4, S5, S6, S7, S9, and S10 (range = 1.4 to 7.6 dB).

Figure 3 shows individual and mean pre- and post-training (4 wks) IEEE sentence recognition scores at the moderate SNR. In steady noise, mean performance significantly improved from 58.5 to 66.0% correct after training ($t_9 = -2.68$, $p = 0.025$). Post-training performance worsened by 10.1 points

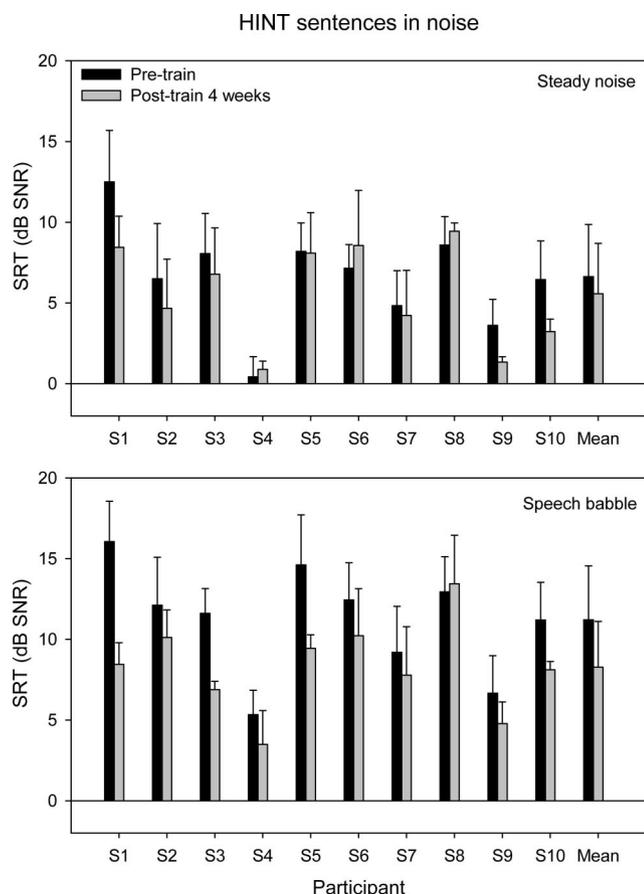


Fig. 2. Similar to Figure 1 but for HINT sentence recognition in speech-shaped steady noise (top panel) and multitalker speech babble (bottom panel).

for subject S4, improved for subjects S1, S2, S6, S7, S8, S9, and S10 (range = 6.1 to 23.5 points), and was unchanged for subjects S3 and S5. In speech babble, mean performance significantly improved from 32.5 to 41.7% correct after training ($t_0 = -3.12$, $p = 0.012$). Post-training performance worsened for subjects S2 and S9 (range = 5.0 to 5.8 points), improved for subjects S1, S3, S6, S7, S8, and S10 (range = 11.4 to 26.0 points), and was unchanged for subjects S4 and S5.

Figure 4 shows individual and mean pre- and post-training (4 wks) IEEE sentence recognition scores at the difficult SNR. In steady noise, mean performance significantly improved from 36.0 to 44.7% correct after training ($t_0 = -2.54$, $p = 0.032$). Post-training performance worsened by 8.8 points for subject S9, improved for subjects S1, S2, S4, S5, S6, S8, and S10 (range = 5.9 to 19.8 points), and was unchanged for subjects S3 and S7. In speech babble, mean performance improved from 14.3 to 17.7% correct after training; however, the improvement was not significant ($t_0 = -1.88$, $p = 0.093$). Because of potential floor performance effects, the scores for the difficult SNR were transformed into rationalized arcsine units (Studebaker 1985). Even after this transformation, there was no significant difference between pre- and post-training performance ($t_0 = -2.051$, $p = 0.071$). Post-training performance worsened by 6.3 points for subject S7, improved for subjects S3, S5, S6, S8, S9, and S10 (range = 3.8 to 13.2 points), and was unchanged for subjects S1, S2, and S4.

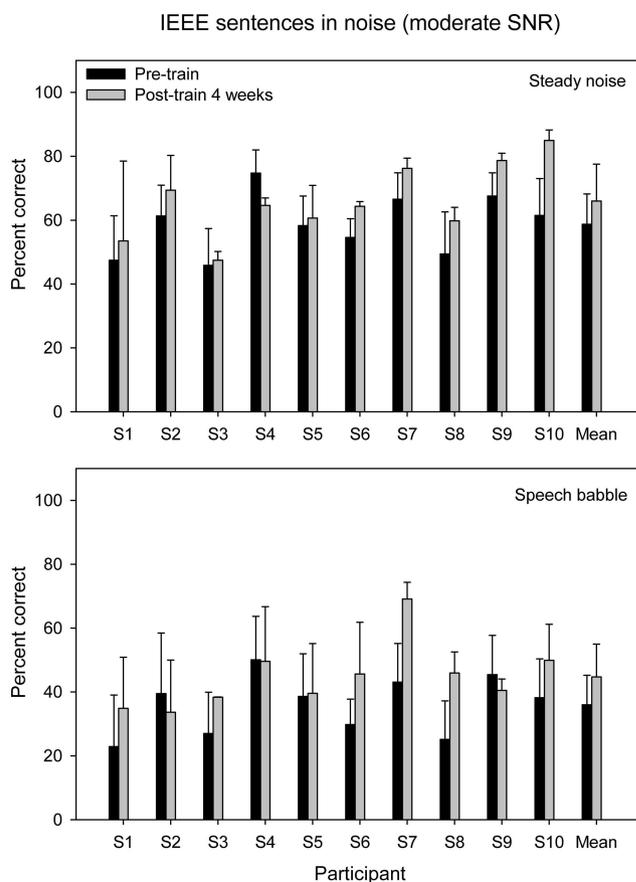


Fig. 3. Similar to Figure 1 but for IEEE sentence recognition in speech-shaped steady noise (top panel) and multitalker speech babble (bottom panel). Data are shown for the moderate SNR condition.

Figure 5 shows mean performance (across participants) for digit recognition in noise (top left), HINT sentence recognition in noise (top right), IEEE sentence recognition at the moderate SNR (bottom left), and IEEE sentence recognition at the difficult SNR (bottom right), as a function of noise type. Mean data are shown for pretraining, 2 wks post-training, 4 wks post-training, and follow-up measures.

A two-way repeated-measures ANOVA, with test session (pretraining, 2 wks post-training, 4 wks post-training, and follow-up) and noise type (steady state or speech babble), was performed for each of the speech measures shown in Figure 5. The results are shown in Table 2. For all measures, performance in steady noise was significantly better than that in speech babble, as would be expected. Post-training performance (4 wks after training was begun) and follow-up performance (4 wks after training was stopped) were significantly better than pretraining performance for all test measures. There was no significant difference between post-training and follow-up measures, suggesting that the training benefits were largely retained 1 mo after training was stopped.

Data were also analyzed to see whether there was a difference in post-training performance gains across test and noise conditions. Two-way repeated-measures ANOVAs were performed on the performance difference data (4 wks post-training – pretraining baseline), with test type (digit or HINT; IEEE moderate or difficult SNR) and noise type (steady state or speech babble) as factors. The two-way ANOVAs were per-

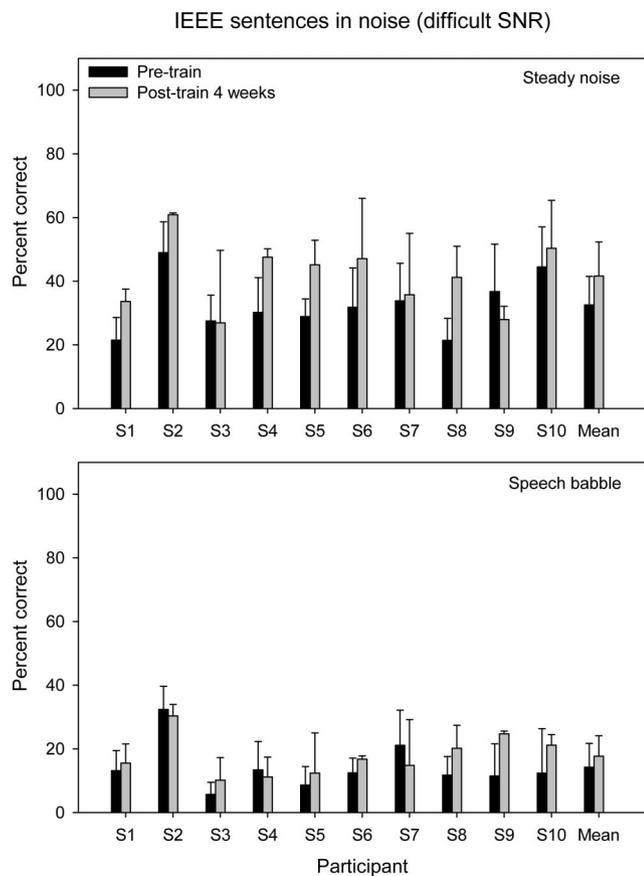


Fig. 4. Similar to Figure 3 but for the difficult SNR condition.

formed independently for the SRT measures (difference in SNR) and the IEEE measures (difference in percent correct). The results are shown in Table 3. There was no significant difference in post-training performance gains between the digit recognition in noise or HINT sentence recognition in noise tests or between the IEEE sentence recognition in noise at the moderate or difficult SNRs. There was a significant difference in post-training performance gains between noise types only for the HINT sentence recognition in noise, with a greater improvement observed in babble than in steady noise.

The digit-in-babble training was very similar to the digit-in-babble testing, in that the task was exactly the same, the stimuli were exactly the same, and SRTs were calculated in exactly the same way (the mean of all reversals in SNR across all 25 trials). The digit-in-babble training differed from the digit-in-babble testing only in terms of the trial-by-trial feedback and the listening environment. Participants trained at home on their own personal computers or laptops that were loaned to them; the sound quality and potential for distraction may have varied greatly among subjects. In contrast, all subjects were tested in the laboratory in a sound-treated room using high-quality audio equipment; thus, the sound quality and test environment was the same for all participants. Participants' digit-in-babble SRTs were compared between training and testing; the results are shown in Figure 6. Testing SRTs were highly correlated with training SRTs, suggesting that improvements in the home training environment were reflected in the more controlled laboratory testing environment. In

general, the digit-in-babble SRT measurements were robust to different listening conditions.

DISCUSSION

The present results demonstrate that auditory training in noise can improve CI users' speech perception in noise. Training with the closed-set digit identification task using simple single-talker digit stimuli significantly improved speech understanding in noise for the untrained listening tasks (i.e., single-talker HINT sentence recognition in noise and multitalker IEEE sentence recognition in noise). Training with the more difficult noise type (speech babble) also improved performance with the untrained noise type (steady speech-shaped noise). Most important, performance gains were largely retained 1 mo after training was stopped.

While mean performance significantly improved with training, the amount of improvement varied across participants. Some participants (S2) benefited most strongly in the trained task (digit recognition in noise), and others (S1) benefited in all speech recognition tasks. Nevertheless, most participants' performance improved after training, for nearly all speech measures. All participants had at least 1 yr of experience with their CI and the majority had at least 5 yrs of implant experience. Thus, participants had a long "passive learning" experience with their device. Longitudinal studies have shown that the greatest improvements in speech recognition (typically measured in quiet) occurred within the first 3 to 9 mos postimplantation, but that performance may continue up to improve 5 yrs postimplantation (Spivak & Waltzman 1990; Loeb & Kessler 1995; Tyler et al. 1997). Before enrolling in this study, all participants reported that while they were pleased with their CI performance, background noise remained especially problematic, even after many years of implant use. The present results suggest that even experienced CI users can improve performance in noise with auditory training.

In this study, the mean age of participants was 66 yrs, and 8 of the 10 participants were older than 60 yrs. Background noise is especially difficult for older hearing-impaired listeners (Gordon-Salant & Fitzgibbons 1995; Frisina & Frisina 1997). In a CI simulation study with NH listeners, Schwartz et al. (2008) found significantly better phoneme recognition performance with younger than with older participants. Other studies have shown lower performance for some speech measures for elderly CI users, relative to younger CI users (Chatelin et al. 2004; Vermeire et al. 2005; Friedland et al. 2010). Given these deficits, auditory training may be especially important for older individuals. In this study, nearly all subjects improved performance on nearly all speech measures after the digit-in-noise training. Auditory training, especially for older adults, should not be overlooked as an important component of auditory rehabilitation. While the digits-in-noise training may have been effective, what did participants learn? Fu and Galvin (2008) presented pilot data for training in noise using a phonetic contrast protocol or a keyword-in-sentence protocol. They found that both a "bottom-up" approach (targeting acoustic contrasts with phoneme recognition training) and a "top-down" approach (targeting contextual cues and source streaming available with sentence training) were effective for improving speech understanding in noise, although the sentence training seemed to provide a greater overall bene-

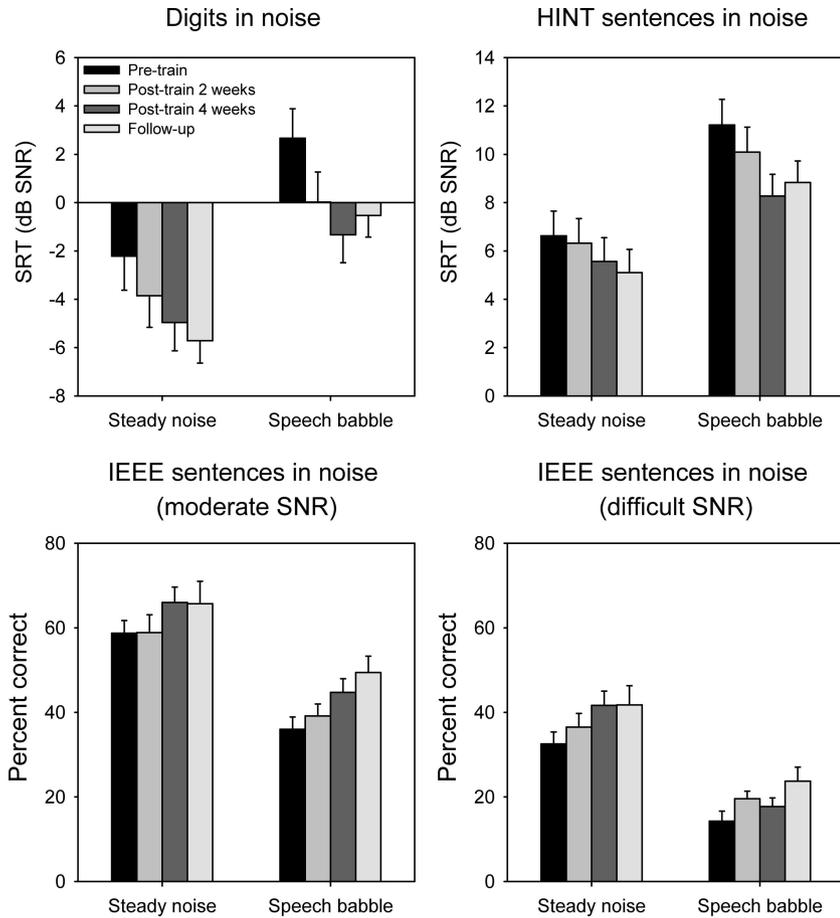


Fig. 5. Mean performance (across participants) for digits recognition in noise (top left panel), HINT sentence recognition in noise (top right panel), IEEE sentence recognition at the moderate SNR (bottom left panel), and IEEE sentence recognition at the difficult SNR (bottom right panel), as a function of noise type. Mean performance is shown for pretraining baseline, 2 wks post-training, 4 wks post-training, and follow-up measures (4 wks after training was stopped). The error bars show 1SE of the mean.

fit. The present digit-in-noise training included aspects of each of these approaches. There may have been some bottom-up learning (sensory component) in discriminating phonemic contrasts (e.g., the acoustic differences between the consonants in “five” and “nine”). The closed-set task and limited number of stimuli may have aided top-down processing, in which participants developed strategies for

segregating the target from the noise. Because the digits were speech stimuli, central pattern processing (i.e., top-down) most likely played a very strong role.

Alternatively, participants’ general attention may have improved after training. Amitay et al. (2006) found that NH listeners’ frequency discrimination improved after training in an unrelated visual task or after training with identical stimuli,

TABLE 2. Results of two-way repeated measures ANOVAs performed for each speech measure

| | Factor | dF, res | F-Ratio | p | Bonferroni (p < 0.05) |
|--------------------|---------------------------|---------|---------|------------------|-----------------------------|
| Digits | Training (Pre, T2, T4, F) | 3, 27 | 17.122 | <0.001 | T2, T4, F > Pre |
| | Noise (SS, babble) | 1, 27 | 86.476 | <0.001 | SS > babble |
| | Training × Noise | 3, 27 | 1.874 | 0.158 | |
| HINT | Training (Pre, T2, T4, F) | 3, 27 | 8.518 | <0.001 | T4, F > Pre |
| | Noise (SS, babble) | 1, 27 | 121.869 | <0.001 | SS > babble |
| | Training × Noise | 3, 27 | 2.697 | 0.066 | |
| IEEE moderate SNR | Training (Pre, T2, T4, F) | 3, 27 | 9.243 | <0.001 | T4 > Pre F > Pre, T2 |
| | Noise (SS, babble) | 1, 27 | 143.519 | <0.001 | SS > babble |
| | Training × Noise | 3, 27 | 1.012 | 0.403 | |
| IEEE difficult SNR | Training (Pre, T2, T4, F) | 3, 27 | 16.602 | <0.001 | T2, T4 > Pre F > T2, Pre |
| | Noise (SS, babble) | 1, 27 | 101.515 | <0.001 | SS > babble |
| | Training × Noise | 3, 27 | 1.334 | 0.284 | |

Factor levels for “training” included: Pre (pre-training), T2 (2 wks post-training), T4 (4 wks post-training), and F (follow-up 4 wks after training was stopped). Factor levels for “noise” included: SS (steady state speech-shaped) and babble (multitalker speech babble). The far right column shows significant differences (p < 0.05) between levels for post hoc Bonferroni pairwise comparisons. Values in bold are statistically significant.

TABLE 3. Results of two-way repeated measures ANOVAs performed on the post-training performance gain data (i.e., the difference in performance between 4 wks post-training and pretraining baseline)

| | Factor | dF, res | F-Ratio | <i>p</i> | Bonferroni (<i>p</i> < 0.05) |
|--------|---------------------------------|---------|---------|--------------|----------------------------------|
| T4-Pre | Test (digits, HINT) | 1, 9 | 2.980 | 0.118 | |
| | Noise (SS, babble) | 1, 9 | 12.510 | 0.006 | HINT: babble > SS |
| | Test × Noise | 1, 9 | 0.510 | 0.494 | |
| T4-Pre | Test-IEEE (moderate, difficult) | 1, 9 | 0.342 | 0.573 | |
| | Noise (SS, babble) | 1, 9 | 1.154 | 0.311 | |
| | Test × Noise | 1, 9 | 1.216 | 0.299 | |

Because of the differences in performance measures between the adaptive SRT (dB SNR) and the fixed SNR tests (percent correct), ANOVAs were performed separately for the two groups of tests. For both tests, factor levels for “noise” included SS (steady state) and babble (multitalker speech babble). Factor levels for “test” included digits (digit-in-noise SRT) and HINT (HINT sentence-in-noise SRT). The far right column shows significant differences (*p* < 0.05) between levels for post hoc Bonferroni pairwise comparisons. Value in bold is statistically significant.

suggesting that some benefits observed with auditory training may not be due to improved auditory processing but due to better overall attention. More recently, Moore et al. (2010) found that auditory training improved adult listeners’ ability to keep three stimuli in short-term memory in a three-alternative forced choice psychophysical task. In this study, we did not have a separate control group who performed an unrelated training task that might shed light on the effects of “attention” on learning effects. Instead, we used a within-subject control that allowed relative improvements with training to be observed, regardless of pretraining baseline performance. There was a wide range (0.4 to 12.5 dB) in pretraining baseline HINT SRTs in steady noise. It would be difficult to establish a comparable control group, and the range in baseline performance might overwhelm post-training changes in performance.

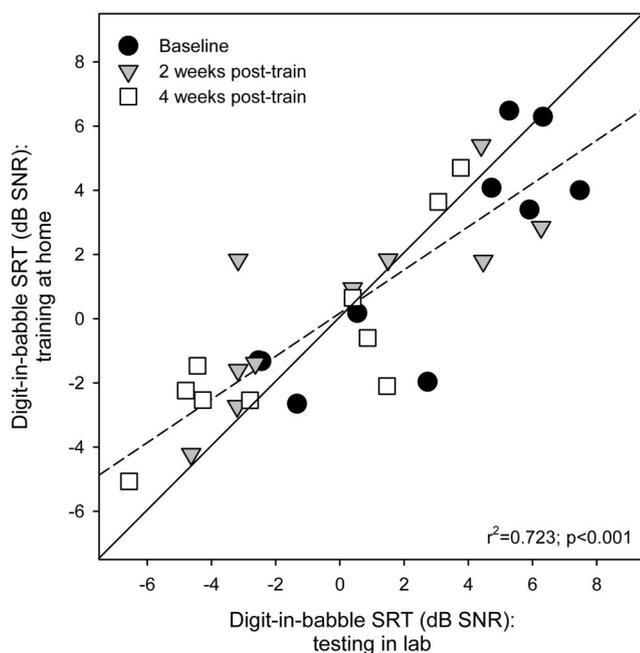


Fig. 6. Individual participants’ digit-in-babble SRTs (measured at home during training, with feedback) as a function of digit-in-babble SRTs (tested in the laboratory, no feedback). Performance is shown for pretraining baseline (black circles), 2 wks post-training (gray triangles), and 4 wks post-training (white squares). The solid diagonal line shows unity performance. The dashed diagonal line shows a linear regression fit across all data; r^2 value and p value for the regression are shown at lower right.

It is also difficult to completely rule out long-term procedural learning due to exposure to the stimuli, test environment, and methods. The extensive baseline testing was performed to reduce these procedural learning effects. However, because of the repeated testing (a minimum of seven test sessions throughout this study), there was a risk of familiarization and learning effects. This may be especially true for the HINT SRT measures, in which some sentences may have been repeated across test runs and sessions. However, increased familiarity with the test materials would tend to sharpen the slope of the performance-intensity function; noise would provide either complete interference or none at all. Given the easy level of difficulty for the HINT sentences, familiarity with the materials was most likely a minor issue. In general, better experimental controls will help to clarify auditory learning and demonstrate any benefits of auditory training.

Anecdotal reports suggest that participants benefited from the training outside of the laboratory. One participant reported better awareness of background noise after training. She no longer “tuned-out” when presented with background noise and instead tried to “listen beyond the noise” in more situations. Another participant commented that she was surprised how well she performed in seemingly large amounts of noise. Another participant commented that she used to be intimidated when listening to numbers over the telephone (an auditory-only listening task not unlike the present digit recognition testing); the digit recognition training had improved her confidence. CI users may gain confidence if they see they are making progress with auditory training. Indeed, some participants commented that the auditory training was motivating, as they were able to see their progress over time. Motivation is certainly an important factor in training outcomes. The participants in this study were highly motivated (hence, their willingness to participate in CI research), which may influence the training outcomes. With effective training tools and resources, CI users may receive the necessary “motivational boost” to get the most benefit from their implant.

Based on this study and previous work, most CI users would seem to benefit from structured auditory training, due to improved performance and increased confidence. In this study, subjects improved regardless of pretraining performance levels. Speech performance improved with training for both elderly CI users and younger CI subjects who had many years of experience with their device. Future research will determine whether auditory training can accelerate adaptation for newly implanted CI patients. In this study, there was great intersubject

variability in terms of post-training performance gains. Future research may help to identify CI users who may benefit most from training, as well as to select training tasks tailored for individual CI users and skill levels.

CONCLUSIONS

The results of this study demonstrate that auditory training with the digit-in-noise task significantly improved CI users' ability to recognize digits in the presence of both steady noise and speech babble. These training benefits generalized to improve open-set HINT and IEEE sentence recognition in both steady noise and babble. The results suggest that auditory training with familiar stimuli using a simple, closed-set task may improve CI users' speech understanding in difficult listening conditions.

ACKNOWLEDGMENTS

The authors thank all of the CI users who participated in this study.

This work was supported by NIH-NIDCD grant R01-DC004792.

Address for correspondence: Sandra I. Oba, Division of Communication and Auditory Neuroscience, House Ear Institute, 2100 West Third Street, Los Angeles, CA 90057, USA. E-mail: soba@hei.org.

Received March 31, 2010; accepted January 4, 2011.

REFERENCES

- Amitay, S., Irwin, A., Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nat Neurosci*, *9*, 1446–1448.
- Chatelin, V., Kim, E. J., Driscoll, C., et al. (2004). Cochlear implant outcomes in the elderly. *Otol Neurotol*, *25*, 298–301.
- Davis, M. H., Hervais-Adelman, A., Taylor, K. et al. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J Exp Psych*, *134*, 222–241.
- Friedland, D. R., Runge-Samuels, C., Baig, H., et al. (2010). Case-control analysis of cochlear implant performance in elderly patients. *Arch Otolaryngol Head Neck Surg*, *136*, 432–438.
- Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hear Res*, *106*, 95–104.
- Fu, Q. J., & Galvin, J. J., III. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hear Res*, *242*, 198–208.
- 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., Galvin, J. J., III, Wang, X., et al. (2004). Effects of auditory training on adult cochlear implant patients: A preliminary report. *Cochlear Implants Int*, *5*, 84–90.
- Fu, Q. J., Galvin, J. J., III, Wang, X., et al. (2005). Moderate auditory training can improve speech performance of adult cochlear implant users. *J Acoust Soc Am*, *6(ARLO)*, 106–111.
- Fu, Q. J., Shannon, R. V., Wang, X. (1998). Effects of noise and number of channels on vowel and consonant recognition: Acoustic and electric hearing. *J Acoust Soc Am*, *104*, 3586–3596.
- Galvin, J. J., III, Fu, Q. J., Nogaki, G. (2007). Melodic contour identification by cochlear implant listeners. *Ear Hear*, *28*, 302–319.
- Gifford, R. H., & Revit, L. J. (2010). Speech perception for adult cochlear implant recipients in a realistic background noise: Effectiveness of preprocessing strategies and external options for improving speech recognition in noise. *J Am Acad Audiol*, *21*, 441–451.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1995). Recognition of multiply degraded speech by young and elderly listeners. *J Speech Hear Res*, *38*, 1150–1156.
- IEEE Subcommittee. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans Audio Electroacoust*, *AU-17*, 225–246.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am*, *49*, 467–476.
- Li, T. H., & Fu, Q. J. (2007). Perceptual adaptation to spectrally shifted vowels: The effects of lexical and non-lexical labeling. *J Assoc Res Otolaryngol*, *8*, 32–41.
- Loeb, G. E., & Kessler, D. K. (1995). Speech recognition performance over time with the Clarion cochlear prosthesis. *Ann Otol Rhinol Laryngol Suppl*, *166*, 290–292.
- Loebach, J. L., & Pisoni, D. B. (2008). Perceptual learning of spectrally degraded speech and environmental sounds. *J Acoust Soc Am*, *123*, 1126–1139.
- Moore, D., Amitay, S., Halliday, L. (2010). Changes in accuracy with interval order during auditory development and learning. Abstracts of the Association for Research in Otolaryngology 33rd Midwinter Meeting, 304.
- Müller-Deile, J., Kiefer, J., Wyss, J., et al. (2008). Performance benefits for adults using a cochlear implant with adaptive dynamic range optimization (ADRO): A comparative study. *Cochlear Implants Int*, *9*, 8–26.
- Nelson, P. B., Jin, S. H., Carney, A. E., et al. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *J Acoust Soc Am*, *113*, 961–968.
- 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.
- Nogaki, G., Fu, Q. J., Galvin, J. J., III. (2007). The effect of training rate on recognition of spectrally shifted speech. *Ear Hear*, *28*, 132–140.
- Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *J Acoust Soc Am*, *66*, 1333–1342.
- Schvartz, K. C., Chatterjee, M., Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *J Acoust Soc Am*, *124*, 3972–3988.
- Skinner, M. W., Clark, G. M., Whitford, L. A., et al. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus-22 channel cochlear implant system. *Am J Otol*, *15*, 15–27.
- Spivak, L. G., & Waltzman, S. B. (1990). Performance of cochlear implant patients as a function of time. *J Speech Hear Res*, *33*, 511–519.
- Stacey, P. C., Raine, C. H., O'Donoghue, G. M., et al. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *Int J Audiol*, *49*, 347–356.
- Stacey, P. C., & Summerfield, A. Q. (2007). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J Acoust Soc Am*, *121*, 2923–2935.
- Stacey, P. C., & Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *J Speech Lang Hear Res*, *51*, 526–538.
- Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J Speech Lang Hear Res*, *28*, 455–462.
- Tyler, R. S., Parkinson, A. J., Woodworth, G. G., et al. (1997). Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *J Acoust Soc Am*, *102*, 508–522.
- Vermeire, K., Brokx, J. P. L., Wuyts, F. L., et al. (2005). Quality-of-life benefit from cochlear implantation in the elderly. *Otol Neurotol*, *26*, 188–195.
- Wolfe, J., Schafer, E. C., Heldner, B., et al. (2009). Evaluation of speech recognition in noise with cochlear implants and dynamic FM. *J Am Acad Audiol*, *20*, 409–421.