Auditory training strategies for adult users of cochlear implants

PAULA STACEY[1,*] AND QUENTIN SUMMERFIELD[2]

[1] *Division of Psychology, Nottingham Trent University, Nottingham, England*

[2] *Department of Psychology, University of York, York. England*

There has been growing interest recently in whether computer-based training can improve speech perception among users of cochlear implants (Fu *et al*., 2005; Oba *et al*., 2011; Ingvalson *et al*., 2013). This paper reports a series of experiments which first evaluated the effectiveness of different training strategies with normal-hearing participants who listened to noise-vocoded speech, before conducting a small-scale study with users of cochlear implants. Our vocoder studies revealed (1) that 'High-Variability' training led to greater generalisation to new talkers than training with a single talker, and (2) that word- and sentence-based training materials led to greater improvements than an approach based on phonemes in nonsense syllables. Informed by these findings, we evaluated the effectiveness of a computer-based training package that included word- and sentence-based tasks, with materials recorded by 20 talkers. We found good compliance with the training protocol, with 8 out of the 11 participants completing 15 hours of training as instructed. Following training, there was a significant improvement on a consonant test, but in general the improvements were small, highly variable, and not statistically significant. A large-scale randomised controlled trial is needed before we can be confident that computer-based auditory training is worthwhile for users of cochlear implants.

INTRODUCTION

Cochlear implantation is highly effective at improving speech perception among adults with severe to profound hearing impairment. Developments in cochlear implant technology are constantly being made, with improved speech processing strategies and electrode design enhancing outcomes for adults who receive cochlear implants. In addition to these developments, auditory training is another way in which outcomes for cochlear implant users can possibly be maximised. By the mid-1990s, most hospital-based cochlear-implant programmes in the UK had ceased to provide an extensive amount of face-to-face aural rehabilitation to adult patients because there was little evidence in support of its effectiveness (e.g., Gagne *et al*., 1991). However, interest in auditory training was revived by studies in US by Gfeller *et al*. (2002) and Fu *et al*. (2005) which suggested that intensive computer-based auditory training can improve the timbre-recognition and speech-perception

*Corresponding author: paula.stacey@ntu.ac.uk

Proceedings of ISAAR 2013: Auditory Plasticity – Listening with the Brain. 4th symposium on Auditory and Audiological Research. August 2013, Nyborg, Denmark. Edited by T. Dau, S. Santurette, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. ISBN: 978-87-990013-4-7. The Danavox Jubilee Foundation, 2014.

skills of cochlear-implant users. Further research has added to these findings in recent years, and suggests that computer-based training may be valuable for adults who use cochlear implants (Ingvalson *et al.*, 2013; Oba *et al.*, 2011; Zhang *et al.*, 2012) and hearing aids (Sweetow and Sabes, 2006; Olson *et al.*, 2013).

An important consideration when evaluating the effectiveness of training regimes is to separate training-related learning from 'incidental' learning. Incidental learning refers to improvements that occur independent of the auditory training task, through procedural learning of task demands (Robinson and Summerfield, 1996), or perceptual learning resulting from repeated exposure to test materials. A further type of incidental learning has also been documented. Amitay *et al.* (2006) reported larger improvements in frequency discrimination for control participants who played a purely visual computer game (Tetris) between successive tests than for control participants who did not engage in an intervening task. These results suggest that maintaining attention and arousal, without explicit training, may be sufficient to lead to improvements on perceptual tasks. In order to evaluate the extent to which a training task has contributed to improvements in performance, it is therefore important to factor out improvements related to 'incidental learning'.

Although many of the studies into the effectiveness of computer-based training for users of cochlear implants have produced positive results, there is no consensus about how training materials should be structured to achieve maximal effectiveness. In this paper, we will (1) summarise results from a series of simulation studies which evaluated the effectiveness of different training strategies, before (2) assessing whether there is any evidence that our training materials are effective for users of cochlear implants.

SIMULATION STUDIES: HOW SHOULD TRAINING MATERIALS BE DESIGNED?

In this section, we will discuss the results from some of our studies with normal-hearing participants who listened to speech processed by a noise-excited vocoder (Shannon *et al.*, 1995). A noise-excited vocoder can be used to mimic the speech processing that occurs in a cochlear implant system, and it allows the spectral and temporal information that is transmitted to the listener to be manipulated. These studies allowed us to carry out controlled experiments that compared the effectiveness of different training strategies, in a way which would have been difficult with users of cochlear implants themselves. Specific issues that we addressed were (1) whether variability is needed in the speech materials, and (2) which training strategies appear to be most effective.

Variability in training materials

One issue in designing training materials for use by cochlear-implant users is whether it is important to incorporate variability in the training materials. Previous research from Lively and colleagues (Lively *et al.*, 1993) conducted with Japanese Americans learning to distinguish between /r/ and /l/ has suggested that training with several talkers is more effective than training with a single talker. For example, Lively *et al*. (1993) found that when training materials were recorded by five talkers there was significant generalisation to new talkers, but this was not the case when training materials were recorded by a single talker. It is important to know whether variability is an important consideration for cochlear-implant users, who listen to speech with reduced spectral and temporal cues which are important in differentiating between talkers (Fu *et al*., 2004).

Our study

We (Stacey and Summerfield, 2007) undertook a study to investigate whether variability is important. Experiment 1 included sixteen participants who listened to speech processed by an 8-channel noise-excited vocoder (frequency range 350 to 5500 Hz). To make the speech more difficult to understand, and following Rosen *et al*. (1999), signals were spectrally-shifted upwards to simulate a tonotopic misalignment of 6 mm according to the Greenwood (1990) function. This meant that the centre frequencies of the 8 bands were shifted upwards by approximately one and a half octaves.

Auditory training was provided by a 2-alternative forced-choice task. At the start of each trial, two words were presented orthographically on the left and right of the touch screen. The target word was then presented acoustically. Participants responded by touching the word corresponding to the target. Visual feedback on accuracy was given. In the 'High-Variability' (H-V) conditions, materials were recorded by 10 talkers, and in the 'Single-Talker' (S-T) conditions, materials were recorded by a single male talker. A visual control task was also implemented in order to separate training-related learning from 'incidental' learning. The visual control task was based on the same 2-alternative forced-choice task, but in this case participants responded to visually presented targets which were degraded by visual noise.

The experiment took place over the course of three days. During Test sessions 1, 2, and 3, participants completed tests of speech perception. There were 4 groups of participants. Groups 1 and 2 received H-V training, while groups 3 and 4 received S-T training. A cross-over design was used in which groups 1 and 3 received auditory training between Test Sessions 1 and 2, whilst groups 2 and 4 received auditory training between Test Sessions 2 and 3 (Table 1).

Figure 1 shows the results from this experiment. We can see that larger improvements in sentence perception were found following the auditory training task than following the visual control task. This difference was statistically significant (auditory training mean improvement: 7.98%, SD: 3.95%, visual control task mean improvement: 2.02%, SD: 4.46%, $t_{15} = 3.13$, $p < 0.01$).

Group	Day 1	Day 2		Day 3	
1	Test Session 1	H-V train	Test Session 2	Control	Test Session 3
2	Test Session 1	Control	Test Session 2	H-V train	Test Session 3
3	Test Session 1	S-T train	Test Session 2	Control	Test Session 3
4	Test Session 1	Control	Test Session 2	S-T train	Test Session 3

Table 1: Study design of Experiments 1 and 2 in Stacey and Summerfield (2007). H-V train = High-Variability auditory training, S-T train = Single-Talker auditory training.

However, Experiment 1 found no advantage for High-Variability training over Single-Talker training (High-Variability mean improvement: 7.16%, SD: 4.71%, Single-Talker mean improvement: 8.8%, SD: 3.09%, t_{14} = 0.83). We reasoned that this may have been because a spectral shift of 6 mm was too extreme to allow participants to differentiate between talkers. Therefore, Experiment 2 replicated Experiment 1, but simulated a tonotopic misalignment of 3 mm (signals were shifted upwards by approximately 3/4 octave). Thirty-two participants were recruited.

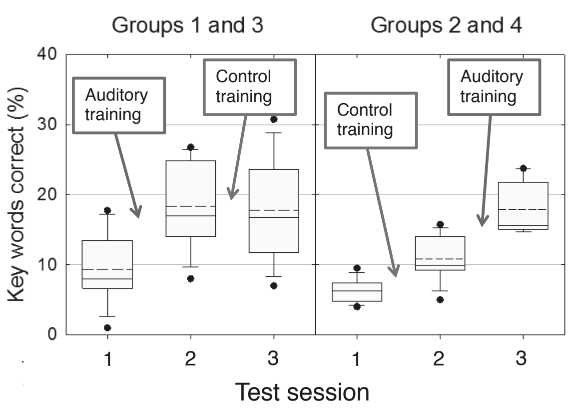

Fig. 1: Results from Stacey and Summerfield (2007), Experiment 1. Percentage of key words correctly identified in IEEE sentences according to test session and training group. The mean value is represented by the dashed line in the box, the median by the solid line. The box spans the interquartile range. Outliers are plotted as dots beyond the 10th-90th percentile whiskers.

Experiment 2 again found evidence that auditory training led to greater improvements in sentence perception than the control task, but it additionally found an advantage for High-Variability training over Single-Talker training ($t_{30} = 2.38$, $p < 0.05$, Fig. 2). The IEEE sentence test we used was recorded by 10 talkers, 5 of whom also recorded the training materials ('old' talkers), and 5 of whom did not ('new' talkers). The advantage of High-Variability over Single-Talker training was greater for the 'new' talkers (High-Variability mean improvement: 13.1%, Single-Talker mean improvement: 8.6%, $t_{30} = 2.03$, $p = 0.051$) than for the 'old' talkers (High-Variability mean improvement: 9.9%, Single-Talker mean improvement: 5.9%; $t_{30} = 1.31$, ns), thereby suggesting that High-Variability training leads to greater generalisation to new talkers than Single-Talker training.

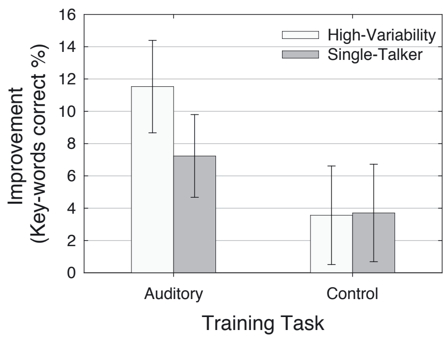

Fig. 2: Results from Stacey and Summerfield (2007), Experiment 2. Advantage of High-Variability training over Single-Talker training. Error bars indicate 95% confidence intervals.

This study therefore shows that there appear to be advantages for including materials recorded by several talkers, even when speech is noise-vocoded and spectrally shifted.

Different training strategies

Although we have found evidence that variability may be important when creating training materials for cochlear implant users, there remains uncertainty about what type of materials should be used for training. There is a long standing argument about whether training materials should be 'analytic' or 'synthetic'. Those who argue that training works best with analytic (bottom-up) approaches based on isolated phonemes claim that if people are trained with the basic building blocks of language, then greater generalisation to new materials can be obtained (e.g., Moore *et al*., 2005). However, other studies which have found that lexical information plays an important role in the perceptual learning of speech (Norris *et al*., 2003; Davis *et* *al*., 2005) suggest that synthetic (top-down) approaches using word and sentence materials will work better.

These differing views on the best approach to training are reflected in the diversity of training materials used in recent studies into the effectiveness of computer-based training for users of hearing aids and cochlear implants. For example, the study by Fu *et al*. (2005) created materials based on phonemes in words, Oba *et al*. (2011) used digits in noise, and Stecker *et al*. (2006) based their training materials on nonsense syllables.

Our study

We (Stacey and Summerfield, 2008) compared the effectiveness of three different approaches to training, based on words, sentences, and phonemes (see below).

i. Word-based training

Two-alternative forced-choice task (as described previously). Materials were recorded by a single male talker.

ii. Sentence-based training

The sentence-based training task was designed as a computer-based Connected Discourse Tracking procedure (CDT, De Fillipo and Scott, 1978; Rosen *et al*., 1999). The training task used IEEE sentences recorded by a single male talker. In this task, participants heard the target sentence, and their task was to decide which 3 out of 6 words displayed orthographically on the computer screen were in the sentence they had just heard.

iii. Phoneme-based training

For phoneme training we used Phonomena (Mindweavers, 2003; Moore *et al*., 2005). We used 10 sets of sounds, each of which consisted of a continuum which ranged from one nonsense syllable to another (e.g., ‘i’ to ‘e’, ‘va’ to ‘wa’, ‘sa’ to ‘sha’). At the extremes of each continuum was a synthesized example of that sound, and these examples were acoustically warped into one another so that each continuum consisted of 96 stimuli. The training task consisted of an XAB two-alternative forced-choice procedure, in which participants heard a target sound (X) and were asked to decide which of two following sounds (A or B) was the same as the target. The training package increased or decreased the separation between stimuli adaptively in order to track 71% correct performance (see Stacey and Summerfield (2008) for more detail).

Eighteen participants took part in the study, and speech was processed by a vocoder which simulated a 6-mm shift. Training was provided during nine 20-minute sessions. Tests of speech perception were administered at the beginning of the study and following every hour (3 sessions) of training thereafter.

The main results are shown in Fig. 3. This figure shows the average overall improvement following the word-, sentence-, and phoneme-based training approaches on each of the tests of speech perception. Overall, we can see that there were larger improvements following word- and sentence-based training than following phoneme-based training. The word- and sentence-based approaches led to significant improvements on the BKB and IEEE sentence tests and the consonant test, while the only significant improvement following phonetic training was found on the consonant test.

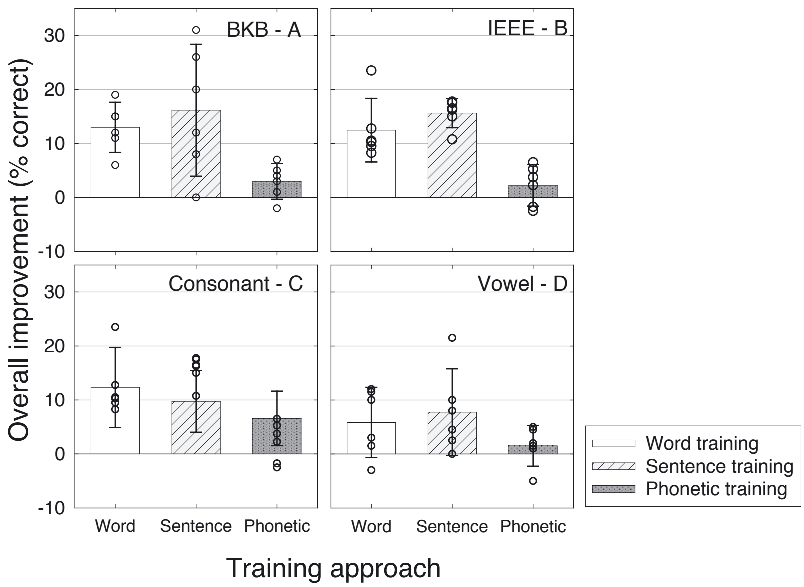

Fig. 3: Results from Stacey and Summerfield (2008). Overall improvements (following 3 hours of training) on the BKB sentence test (Panel A), the IEEE sentence test (Panel B), the consonant test (Panel C), and the vowel test (Panel D). Error bars indicate 95% confidence intervals, and improvements for individual participants are overlaid on the plots.

The approach to controlling for incidental learning we took in this experiment was to repeatedly administer these tests of speech perception at baseline, until an asymptote in performance was reached (see Fu *et al*., 2005). Participants' baseline level of performance was taken to be their highest level of performance during any of the baseline tests. Figure 4 shows the consequences of exercising this control, rather than taking baseline performance to be the first time participants completed the speech tests. We can see that if no control was exercised over the effects of repeated testing we would (1) find that improvements were much larger in magnitude, and (2) that many of the improvements would reach statistical significance.

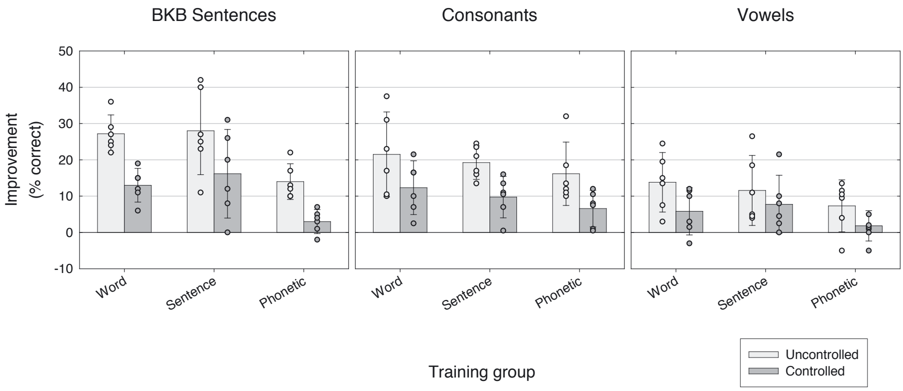

Fig. 4: Results from Stacey and Summerfield (2008). Improvements in performance on the BKB sentence test, the consonant test, and the vowel test. The white bars (uncontrolled) show the overall level of improvement between the first time tests were completed in the baseline session and the final testing session. The light grey bars (controlled) show the level of improvement between the 'highest baseline' and the final testing session. Error bars denote 95% confidence intervals.

The results from this study support the view that lexical information is important in the perceptual learning of speech (Norris *et al*., 2003; Davis *et al*., 2005). Our results are also consistent with the results from Faulkner *et al*. (2005) who found that training generalised best to similar tests, and that analytic training approaches did not lead to improvements in sentence perception.

EFFECTIVENESS OF TRAINING FOR COCHLEAR-IMPLANT USERS

Our study

Informed by our earlier results, we created a training package for use by adult users of cochlear implants (Stacey *et al*., 2010). The training package was designed to be user-friendly for people with limited experience of computers, and minimal computer skills were required. The training package contained the word- and sentence-based training materials, and materials were recorded by 20 talkers.

Eleven people who had used cochlear implants for over 3 years took part in the study (average duration of use: 5.73 years, SD: 2.69). They were aged between 23 and 71 years (average age: 58.82 years, SD: 18.89), and their latest score on the BKB sentence test recorded by the implant teams ranged from 34-81% correct (average BKB score: 59.82%, SD: 18.82).

Participants were asked to complete one hour of training a day, 5 days per week, over a period of three weeks (totalling 15 hours). Training was self-administered in participants' own homes, and they were visited by the first author at the beginning of the study who administered tests of speech perception repeatedly until an asymptote in performance was reached. Participants were then visited again following every week of training, when they completed further tests of speech perception and a questionnaire (the Glasgow Benefit Inventory, Robinson *et al*., 1996) which measured whether training had benefitted participants in their everyday lives.

We found good compliance with the training protocol, with eight of the eleven participants completing 14-16 hours of training as instructed. The following analyses were limited to these 8 participants who completed the required amount of training.

The average overall improvements following three hours of training are shown in Fig. 5. We can see that there is a significant improvement consonant perception following training (mean improvement: 8.06%, SD: 6.90, $t_7 = 3.31$, $p < 0.05$), but the improvements did not reach significance for the other speech tests and there was a great deal of variability between participants.

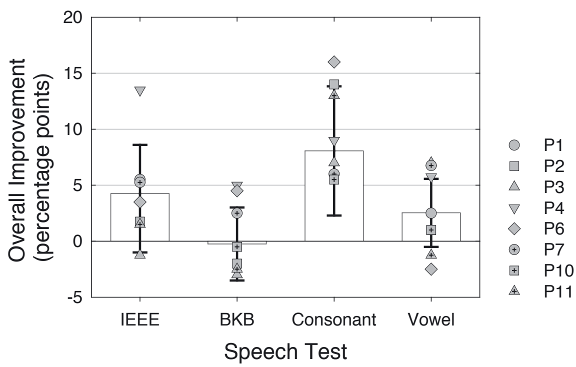

Fig. 5: Results from Stacey *et al*. (2010). Overall improvements following three hours of training on each test of speech perception. Error bars indicate 95% confidence intervals, and improvements by individual participants are overlaid on the plots.

Baseline levels of performance are shown in Table 2. There was large variability in participants' performance levels prior to training. For example, performance on the IEEE test ranged from 10% correct to 49%, and performance ranged from 20 to 61% correct on the consonant test. Given the diversity in the baseline level of performance of our participants, we tested whether there was an association between performance levels and levels of improvements (Table 2). These analyses revealed no significant associations, although it must be recognised that these analyses were limited by the small sample size.

Finally, there was no evidence that auditory training led to improvements in everyday life (as measured by the GBI).

		IEEE	*BKB*	*Consonant*	*Vowel*
Baseline performance	Average	22.20	48.69	43.84	47.47
	St. dev.	15.43	15.11	13.19	6.78
	Range	10 to 49	25 to 71	20 to 61	36 to 51
Correlations		$r_7 = 0.32$, $p = 0.48$	$r_8 = 0.63$, $p = 0.09$	$r_8 = 0.11$, $p = 0.80$	$r_8 = 0.50$, $p = 0.21$

Table 2: Average baseline levels of performance from Stacey and Summerfield (2010). The numbers indicate percentage correct, and baseline performance levels consisted of participants' average performance during the last two tests they completed in the first session. This table also shows the results of the correlation analyses.

Comparison with previous research

Overall, the results from this study do not give strong evidence that computer-based training is a worthwhile intervention for adult users of cochlear implants. This is in contrast to the generally positive outcomes reported by other studies (Fu *et al*., 2005; Oba *et al*., 2011; Zhang *et al*., 2012; Ingvalson *et al*., 2013). This disparity may be in part due to the highly variable levels of pre-training performance of the participants in our study, and more positive results may have been found if our participants had performed more poorly overall.

Evidence from other studies that auditory training can produce positive effects is encouraging, and these studies give useful insight into the type of training materials that might be effective for use by cochlear implant users. However, before we can be confident that auditory training is responsible for the improvements that have been found, a larger-scale randomised controlled study is needed (see Henshaw and Ferguson, 2013, for a review). The most common type of design used in this area to date has been a repeated measures design (Fu *et al*., 2005; Oba *et al*., 2011; Zhang *et al*., 2012; Ingvalson *et al*., 2013). Some studies have sought to control for the effects of incidental learning by employing the methodology we used here of repeatedly administering tests at baseline, so that participants act as their own controls (e.g., Fu *et al*., 2005; Ingvalson *et al*., 2013). However, this methodology makes the assumption that incidental learning is a short-term phenomenon, and is complete by the end of baseline testing. This methodology fails to take into account the fact that incidental learning might continue over the course of the study, and performance may improve because participants are engaged in *any* task, rather than one we would expect to improve speech perception. Ideally therefore, a control group who completes an alternative to auditory training should also be included.

CONCLUSION

Our simulation studies suggested that word- and sentence-based training materials recorded by several talkers will be most effective for use by cochlear-implant users. However, our small-scale study with cochlear-implant users found only minimal evidence of improvements following training. Although previous studies have found positive effects of training for users of cochlear implants, more robust evidence in the form of a large-scale randomised controlled trial is needed before we can be confident that computer-based auditory training is worthwhile for users of cochlear implants.

REFERENCES

Amitay, S., Irwin, A., and Moore, D.R. (**2006**). “Discrimination learning induced by training with identical stimuli,” Nat. Neurosci., **9**, 1446-1448.

Davis, M.H., Hervais-Adelman, A., Taylor, K., McGettigan, C., and Jonsrude, I.S. (**2005**). “Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences,” J. Exp. Psychol. Gen., **134**, 222-241.

De Filippo, C.L., and Scott, B.L. (**1978**). “A method for training and evaluating the reception of ongoing speech,” J. Acoust. Soc. Am., **63**, 1186-1192.

Faulkner, A., Rosen, S., and Jackson, A. (**2005**). “Relative effectiveness of training methods for shifted speech,” Poster presented at the British Society of Audiology short-papers meeting, Cardiff, UK.

Fu, Q.-J., Chinchilla, S., and Galvin, J.J. (**2004**). “The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users,” J. Assoc. Res. Oto., **5**, 253–260.

Fu, Q.-J., Galvin, J.J., Wang, X., and Nogaki, G. (**2005**). “Moderate auditory training can improve speech performance of adult cochlear implant patients,” Acoust. Res. Lett. Online, **6**, 106-111.

Gagne, J.P., Parnes, L.S., LaRocque, M., Hassan, R., and Vidas, S. (**1991**). “Effectiveness of an intensive speech perception training program for adult cochlear implant recipients,” Ann. Otol. Rhinol. Laryngol., **100**, 700-707.

Gfeller, K., Witt, S., Ademek, M., Mehr, M., Rogers, J., Stordahl, J., and Ringgenberg, S. (**2002**). “Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients,” J. Am. Acad. Audiol., **13**, 132-145.

Greenwood, D.D. (**1990**). “A cochlear frequency-position function for several species – 29 years later,” J. Acoust. Soc. Am., **87**, 2592-2605.

Henshaw, H., and Ferguson, M.A. (**2013**). “Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence,” PLoS One, **8**, e62836.

Ingvalson, E.M., Leea, B., Fiebigb, P., and Wonga, P.C.M. (**2013**). "The effects of short-term computerized speech-in-noise training on postlingually deafened adult cochlear implant recipients," J. Speech Lang. Hear. Res., **56**, 81-88.

Lively, S.E., Logan, J.S., and Pisoni, D.B. (**1993**). "Training Japanese listeners to identify English /r/ and /l/. II. The role of phonetic environment and talker variability in learning new perceptual categories," J. Acoust. Soc. Am., **94**, 1242-1255.

Moore, D.R., Rosenberg, J.F., and Coleman, J.S. (**2005**). "Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children," Brain Lang., **94**, 72-85.

Norris, D., McQueen, J.M., and Cutler, A. (**2003**). "Perceptual learning in speech," Cogn. Psychol., **47**, 204-238.

Oba S.I., Fu, Q.-J., and Galvin, J.J. (**2011**). "Digit training in noise can improve cochlear implant users' speech understanding in noise," Ear Hearing, **32**, 573-581.

Olson, A.D., Preminger, J.E., and Shinn, J.B. (**2013**). "The effect of LACE DVD training in new and experienced hearing aid users," J. Am. Acad. Audiol., **24**, 214-230.

Robinson, K., Gatehouse, S. and Browning, G.G. (**1996**). "Measuring patient benefit from otorhinolaryngological surgery and therapy," Ann. Otol. Rhinol. Laryngol., **105**, 415-422.

Robinson, K., and Summerfield, A.Q. (**1996**). "Adult auditory learning and training." Ear Hearing, **17**, 51S-65S.

Rosen, S., Faulkner, A., and Wilkinson, L. (**1999**). "Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants," J. Acoust. Soc. Am., **106**, 3629-3636.

Shannon, R.V., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primary temporal cues," Science, **270**, 303-304.

Stacey, P.C., and 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., and 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.

Stacey, P.C., Raine, C.H, O'Donoguhe, G.M., Tapper, L., Twomey, T., and Summerfield, A.Q. (**2010**). "Effectiveness of computer-based auditory training for adult users of cochlear implants," Int. J. Audiol., **49**, 347-356.

Stecker, G.C., Bowman, G.A., Yund, E.W., Herron, T.J., Roup, C.M., and Woods D.L. (**2006**). "Perceptual training improves syllable identification in new and experienced hearing aid users," J. Rehabil. Res. Dev., **43**, 537-552.

Sweetow, R.W., and Sabes, J.H. (**2006**). "The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program," J. Am. Acad. Audiol., **17**, 538-558.

Zhang, T., Dorman, M.F., Fu, Q.-J., and Spahr, A.J. (**2012**). "Auditory training in patients with unilateral cochlear implant and contralateral acoustic stimulation," Ear Hearing, **33**, e70–e79.