

Research Forum

Auditory Training With Frequent Communication Partners

Nancy Tye-Murray,^a Brent Spehar,^a Mitchell Sommers,^b and Joe Barcroft^b

Purpose: Individuals with hearing loss engage in auditory training to improve their speech recognition. They typically practice listening to utterances spoken by unfamiliar talkers but never to utterances spoken by their most frequent communication partner (FCP)—speech they most likely desire to recognize—under the assumption that familiarity with the FCP’s speech limits potential gains. This study determined whether auditory training with the speech of an individual’s FCP, in this case their spouse, would lead to enhanced recognition of their spouse’s speech.

Method: Ten couples completed a 6-week computerized auditory training program in which the spouse recorded the stimuli and the participant (partner with hearing loss) completed auditory training that presented recordings of their spouse.

Results: Training led participants to better discriminate their FCP’s speech. Responses on the Client Oriented Scale of Improvement (Dillon, James, & Ginis, 1997) indicated subjectively that training reduced participants’ communication difficulties. Performance on a word identification task did not change.

Conclusions: Results suggest that auditory training might improve the ability of older participants with hearing loss to recognize the speech of their spouse and might improve communication interactions between couples. The results support a task-appropriate processing framework of learning, which assumes that human learning depends on the degree of similarity between training tasks and desired outcomes.

Persons with hearing loss often receive auditory training to improve their listening performance. Training may lead to modest gains in recognition of speech produced by those particular talkers but may not generalize to the same words in a different context (Burk, Humes, Amos, & Strauser, 2006). As such, the value of traditional auditory training to an individual’s everyday communication interactions remains a critical but underinvestigated issue (e.g., Brouns & Price, 2011; Henshaw & Ferguson, 2013; Sweetow & Palmer, 2005).

Despite the relatively modest amount of research, one of the most robust findings concerns the relationship between talkers and tasks used during training and assessment. A number of studies, including meta-analyses of studies on auditory training (Brouns & Price, 2011; Henshaw & Ferguson, 2013; Sweetow & Palmer, 2005), indicate that benefits of auditory training are tied to what is transfer appropriate: Benefits are most likely to accrue when using the same talker during both training and assessment or

when using the same tasks during both training and assessment. In addition, when different tasks or talkers are used in training compared with assessment, the greater the overlap between the talkers (e.g., one out of six talkers overlapping as opposed to zero out of six talkers overlapping) and task type, the greater the benefit (e.g., Barcroft et al., 2011). A striking example of this pattern appears in Barcroft et al. (2011), who found that adults receiving auditory training with stimuli spoken by a single talker improved more on a single-talker assessment than on a multitalker assessment, whereas adults receiving auditory training with stimuli spoken by six talkers improved more on a multitalker assessment. These findings are compatible with transfer-appropriate processing (TAP) theory (Morris, Bransford, & Franks, 1977). According to TAP theory, human learning depends on the degree of compatibility between the tasks that one performs at the times of study and testing. The more similar tasks are at these two times, the better the learning.

The commonsense conclusion from this research and from TAP theory is that clinicians should tailor training to meet the communication demands of a particular patient. However, in practice, this is rarely if ever done. Instead, clinicians often use packaged computerized programs such as the Listening and Communication Enhancement (Sweetow & Sabes, 2006) or the Computer-Assisted Speech Perception Sentence (Boothroyd, 1987) to provide auditory

^aWashington University in St. Louis, MO

^bWashington University in St. Louis School of Medicine, MO

Correspondence to Nancy Tye-Murray: murrayn@wustl.edu

Editor: Rhea Paul

Associate Editor: Richard Dowell

Received May 15, 2015

Revision received September 16, 2015

Accepted November 25, 2015

DOI: 10.1044/2016_JSLHR-H-15-0171

Disclosure: The authors have declared that no competing interests existed at the time of publication.

training, or they provide live-voice training with their own voice—a voice that is likely highly intelligible to patients because clinicians spend their days talking to persons with hearing loss. In either scenario, clinicians adopt an implicit policy of “one size fits all” and assume that training with an unfamiliar communication partner or a communication partner whose speech is atypically intelligible will result in a patient hearing his or her everyday communication partners better in the real world.

Perhaps one reason why auditory training is not provided with stimuli spoken by a frequent communication partner (FCP) is an underlying assumption that if a patient is well practiced with a talker’s speech, then there is no potential for improvement to accrue as a result of training. Recent research has revealed a familiarity effect, which refers to how older persons with hearing loss can understand the speech of their spouse better than that of strangers (Souza, Gehani, Wright, & McCloy, 2013). Because they are already familiarized with their spouse’s speech, targeted auditory training with an FCP might yield minimal benefit.

In the present investigation, we did not follow the traditional route of providing auditory training with the goal of improving participants’ abilities to recognize the speech of a variety of talkers that they might encounter. Rather, we were concerned solely with enhancing their abilities to recognize the speech of a particular FCP. We determined whether auditory training with the speech of an individual’s FCP—in this case their spouse—would lead to enhanced speech recognition. Because a spouse’s voice is probably the voice they hear most often and in the greatest array of listening environments, training with the FCP might yield little if any benefit. Addressing this issue directly, the present auditory training program provided structured practice to older listeners using stimuli that had been recorded by each participant’s spouse.

Method

Participants

Ten older married adult couples participated in the study (mean age = 73.2 years, $SD = 5.7$). Each had been a couple for at least 14 years ($M = 45.8$ years, $SD = 14.7$). On the basis of the .85 effect size observed in Barcroft et al. (2011), it was determined that a sample size of 10 couples provides power of .77 to find an effect size as large or larger using a within-subject design. Within each couple, the partner with hearing loss (henceforth referred to as *participant*) received auditory training, and the participant’s spouse recorded the stimuli for the participant’s auditory training experience. Each participant had at least a moderate hearing loss (pure-tone average: left ear, $M = 40.3$ dB HL, $SD = 13.3$; right ear, $M = 38.3$ dB HL, $SD = 9.9$) and had used a hearing aid for at least 3 months ($M = 10.7$ years, $SD = 14.0$). Nine participants were men and one was a woman, and there were no same-sex couples. Couples were recruited through the Volunteers for Health program at Washington University School of Medicine. Each member

of the pair was paid a participant fee of \$10/hr. Approval from Washington University School of Medicine’s institutional review board was obtained, and research was conducted in accordance with the provisions of the World Medical Association Declaration of Helsinki.

The spouse of each participant recorded the stimuli, which were then inserted into the Customized Listening: Exercises for Aural Rehabilitation (cLEAR, formerly I Hear What You Mean; Tye-Murray et al., 2012) training software. Recording occurred 3 to 4 weeks prior to the participant’s receipt of auditory training and was completed over the course of three visits using a teleprompter and microphone. Before insertion into the program software, the stimuli were edited and leveled to have equal root-mean-square amplitudes.

In preparation for this investigation, we first determined that individuals would not improve on a four-alternative forced-choice (4AFC) test (Barcroft et al., 2011) when their spouse was the talker simply by virtue of having taken the test twice and hence having learned the test task. The 4AFC test is described in the Assessment section. Five older married adult couples (mean age = 73.6 years, $SD = 4.2$) participated in this preliminary study. Each had been a couple for at least 11 years ($M = 46.0$ years, $SD = 20.1$). The spouse with hearing loss (two men, three women) received the same assessment intervals as participants in the main study but no computerized training. The participant’s spouse recorded the assessment material approximately 2 weeks prior to the first test session. Recording was completed in one session using a teleprompter and microphone. All stimuli were edited and leveled to have equal root-mean-square amplitudes. Each participant had a sensorineural hearing loss (mean pure-tone average in the better ear = 45.8, $SD = 28.8$) and had used a hearing aid for at least 3 months ($M = 17.8$ years, $SD = 23.1$). Participants in the preliminary study did not significantly improve on the 4AFC test from the pretest to the posttest interval, changing by an average of only 4.4 percentage points.

Training Program

The subset of the cLEAR training program used in this experiment includes five activities varying in linguistic and contextual content (for a detailed description, see Tye-Murray et al., 2012) and is anchored in a meaning-based approach as opposed to a nonsense syllable or drill approach to auditory skills learning. The program includes word, sentence, and paragraph stimuli. Of particular relevance to the present investigation is the four-choice discrimination task (Activity 2), which presents spoken word pairs such as *bat–cat* in the presence of six-talker babble. The response screen for this pair shows four picture pairs: *bat–bat*, *bat–cat*, *cat–bat*, and *cat–cat*. The participant’s task is to touch the correct picture pair displayed on the computer touchscreen monitor. Following a wrong response, the stimuli are presented again at an easier signal-to-noise ratio (SNR). Following a correct response, the word pair is presented auditorily in quiet and orthographically on the

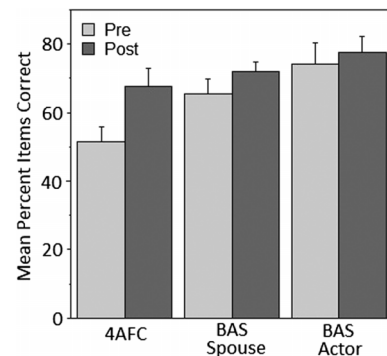
screen. This task was included in both training and assessment (without feedback) because it requires participants to map form and meaning using minimal pairs (e.g., *bat–cat*), in contrast to more traditional phonemic discrimination drill tasks that divorce form from meaning (e.g., listening to /ba/ and /ka/ and attempting to determine whether the two segments are the same or different). Activity 1 required participants to determine whether a target sound occurred in the initial, medial, or final position; Activity 3 required them to complete a sentence with one of four word choices, where each word differed by only one sound; Activity 4 required them to guess the next sentence from a choice of three after hearing a preceding sentence, where all three word choices might be plausible given how much of the preceding sentence was recognized; and Activity 5 required them to answer multiple-choice questions after listening to a paragraph.

Participants received 12 hr of computer-based auditory training over the course of 6 weeks (twice a week in 1-hr sessions). During training, a participant sat in a sound-treated room and listened to speech stimuli in approximately 62 dB SPL of background speech babble. Audio levels of the stimuli were adapted using a two-down, one-up procedure designed to keep performance at approximately 79% items correct (Levitt, 1971). This level was chosen to keep participants engaged and focused while avoiding any frustration that occurs when tasks are too difficult.

Assessment

Pre- and posttraining assessments included two speech-in-noise tests—the Build-a-Sentence Test (BAS; Tye-Murray et al., 2008) and the 4AFC test (Barcroft et al., 2011)—and the Client Oriented Scale of Improvement (COSI) questionnaire (Dillon, James, & Ginis, 1997). The BAS is a closed-set words-in-sentence context test in which a set of 36 words is randomly presented in a consistent sentence context (*The boys and the dog watched the mouse*). Two versions of the BAS were administered, one with the stimuli recorded by the participant’s spouse and a second with the stimuli recorded by a professional actor. Pre- and post-BAS testing for participants included counterbalanced and blocked presentations from their spouse and from the professional actor. Sentences were presented in babble at randomly varying SNRs. Here we report the pre- and post-training results from test items presented in background babble at 62 dB SPL with the target speech presented at 0, +5, and +10 dB SNR. The 4AFC test utilizes a format similar to that of Activity 2 (described above) but presents items with a constant (+3 dB SNR) rather than varying SNR and does not provide feedback. Test items were spoken by the participant’s spouse. The COSI assesses subjective changes in listening performance. Before training, participants listed three everyday listening situations in which they desired to improve their listening performance. After training, participants rated how much, if any, improvement resulted.

Figure 1. Results for the four-alternative forced-choice test and the Build-a-Sentence Test with test items spoken by the spouse and test items spoken by an actor.



Results

The results for the pre- and posttraining speech tests appear in Figure 1. The amount of improvement after training with the spouse for the 4AFC test is estimated to be 16.1 percentage points, 95% confidence interval (CI) of the difference [7.5, 24.7]. This indicated a notable improvement from training, $t(9) = 4.2$, $p < .010$. This degree of improvement is comparable to that which persons with hearing loss realize when they train with an unfamiliar talker and are tested before and after training with the same talker. Barcroft et al. (2011) reported gains of 17.1 percentage points on the 4AFC test for participants who trained for 12 hr with the cLEAR auditory training program, in which the talker was one of six trained actors.

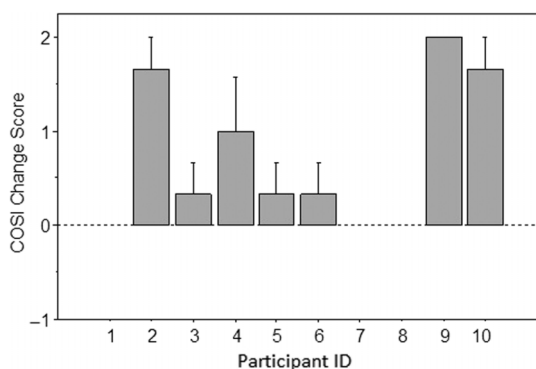
The results for the BAS for either talker in Figure 1 are averaged for the three SNRs. On average, participants improved by 6.3 percentage points after training when the spouse spoke the BAS test items, 95% CI of the difference [1.1, 11.5], and by 3.6 percentage points when the actor spoke the items, 95% CI of the difference [−4.2, 11.3]. Although the trend was toward improvement following training, especially when the spouse was the speaker, a two-way analysis of variance with the independent variables of interval and talker type revealed no effect of interval, $F(1, 9) = 3.38$, $p = .099$, $\eta_p^2 = .273$, or talker, $F(1, 9) = 3.18$, $p = .112$, $\eta_p^2 = .257$.

Examples of challenging everyday listening situations cited on the COSI included “I can’t hear my wife’s voice clearly when she is looking down,” “When she speaks to me from another room,” and “When she speaks with the TV or radio turned on.” As shown in Figure 2, seven of 10 participants reported improvement in their three designated situations. On average, the COSI improvement was 0.73 point on the 5-point scale, 95% CI [0.4, 1.1]. This indicated an improvement after the training, $t(29) = 4.63$, $p < .001$.

Discussion

The present experiment suggests that auditory training using the speech of a spouse was beneficial for older

Figure 2. Results from the Client Oriented Scale of Improvement for individual participants, displayed as a function of assigned participant identification number.



participants with hearing loss, despite long-term cohabitation. Training led to better word discrimination and led participants to rate their communication interactions as improved. However, participants did not significantly improve on the word recognition test (BAS) for either their spouse or the unfamiliar actor.

The results indicate the applicability of the TAP framework to auditory training, which posits that learning depends on the degree of compatibility between the tasks that one performs at the times of study and testing. The more similar tasks are at these two times, the better the learning. The present results support a TAP perspective because they demonstrate learning in a situation where training and test tasks were highly similar (i.e., same talker, same background noise). It should be noted that our three tests cover the range from an unnatural form of speech recognition (4AFC) to a more natural form of speech recognition (BAS) to a form of speech communication that is encountered in everyday life. We provided specific training for the 4AFC test but not the BAS. Significant improvement was found for the former but not the latter, consistent with TAP theory. In designing future training programs, stimuli might better mimic the types of speech communication that characterize everyday life. For example, training stimuli might include vocabulary that a patient is likely to encounter (e.g., work-related vocabulary) and common everyday phrases that occur in a patient's home setting.

The present results also suggest that exposure alone is not sufficient to improve speech perception. The FCPs likely have received all the benefit they can from talker familiarity, and the current findings suggest that further gains can be obtained only with some form of structured auditory training. Training gains may have occurred because participants received structured practice in the auditory streaming of their spouse's voice from a background of six competing talkers in a way that would never have happened in daily life. In his classic work, Bregman (1990) described auditory scene analysis as a process whereby the human auditory system organizes sound streams into meaningful perceptual units. An important component of

auditory scene analysis is auditory streaming, wherein an individual is able to follow distinct streams of sound patterns over time. Auditory streaming is evident in the "cocktail party" effect: Listeners can follow a particular voice even though other voices are speaking simultaneously (Cherry, 1953). Auditory streaming is highly developed in professionals such as conductors or music producers, who must listen to several sound sources, such as voices and instruments, simultaneously. This kind of expertise suggests that practice (as well as natural talent) may fine tune one's ability to perform auditory scene analysis. Auditory training may have taught participants to discern their spouse's voice from background babble and to integrate the unfolding speech signal into a unified and coherent percept. This led to better performance on the listening tests and to better subjective listening performance in everyday situations.

Auditory training with the FCP may have also led to the participants increasing their attention to the formal features of the speech of their FCPs. Often considered in the context of second language learning, sufficient attention to the formal features of language is needed for successful acquisition of the forms in question (see Schmidt, 2001, on the role of attention in second language acquisition). Participants in the present study were pushed to attend to the formal (phonetic and phonological) features of the speech of their FCPs in a focused manner and to a degree that (at least in all likelihood) surpasses what they experience in their day-to-day lives.

As indicated by the COSI, participants demonstrated improvements in their everyday communication difficulties. Although one possibility is that the auditory training was responsible for these improvements, other factors may have contributed as well. For example, participants may have been motivated to observe improvements, which may be an inherent problem in using a subjective evaluation to gauge efficacy regardless of the kind of audiological intervention (e.g., an expensive listening device) that is being assessed. Furthermore, improvements might be related to changes in the behavior of the FCPs. For example, the FCPs implicitly may have learned the technique of "clear speech" in the process of recording speech stimuli. As a result, they may have been motivated to use clear speech in the home setting.

The present findings and corresponding methodology present two very practical implications for clinical practice. The first implication is that when individuals report a specific listening complaint, such as an inability to understand their FCP in noise, auditory training can be tailored to address their situation-specific need. Customized auditory training might also prepare an individual for an interaction. For example, a patient requiring round-the-clock nursing care might receive auditory training with the speech of a new caretaker. A grandparent who is anticipating a visit from a grandchild might receive auditory training that presents training items spoken by the child. The FCPs in this investigation had to return to the laboratory for multiple recording sessions in order to record the training materials. For each FCP, our laboratory staff then spent approximately

15 to 20 hr editing and leveling the stimuli and inserting the stimuli into the cLEAR program. The overall process was time consuming and one not likely to be implemented within the confines of a real-world audiological practice. For this reason, we have subsequently developed a more automated and streamlined methodology and have adopted different hardware and created new software. The hardware comprises a simple tablet, microphone, and headphones. The software incorporates a teleprompting system, an instantaneous playback accept/reject routine, a waveform editing and leveling algorithm, and an auditory training game and testing programs. The recordings populate the game and test software with the new recordings immediately after the FCP finishes recording the stimuli, thereby overcoming the time-consuming stimuli-preparation process.

The second implication of the present findings and methodology is that it is feasible to use the FCP's speech for routine audiological assessment of speech recognition. In this investigation, we used speech recognition tests that presented stimuli spoken by a participant's FCP in order to assess the efficacy of an aural rehabilitation intervention. In like fashion, audiologists can use test stimuli recorded by a patient's FCP to evaluate how well a particular patient performs in his or her everyday environment. Instead of using speech babble as background noise (as we did in this investigation) or in addition to, the audiologist might use homelike or work-related background noise (depending on the identity of the FCP and the kinds of everyday environments that are of interest) to mimic real-world listening conditions. For example, in the case of a hearing aid fitting, a patient might be tested with an FCP's voice both without and with the new hearing aid. Counseling would then include a review of the speech recognition benefit afforded by the new hearing aid. We will explore this possibility in future research.

Acknowledgments

This research was supported by National Institutes of Health Grant RO1DC008964. We thank Elizabeth Mauzé and Shannon Sides for their contributions in data collection.

References

- Barcroft, J., Sommers, M. S., Tye-Murray, N., Mauze, E., Schroy, C., & Spehar, B. (2011). Tailoring auditory training to patient needs with single and multiple talkers: Transfer-appropriate gains on a four-choice discrimination test. *International Journal of Audiology*, 50, 802–808.
- Boothroyd, A. (1987, June). *CASPER: Computer-Assisted speech perception evaluation and training*. Paper presented at the 10th Annual Conference on Rehabilitation Technology, Washington, DC.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brouns, K., & Price, H. (2011). Auditory training and adult rehabilitation: A critical review of the evidence. *Global Journal of Health Science*, 3, 49–63.
- Burk, M. H., Humes, L. E., Amos, N. E., & Strauser, L. E. (2006). Effect of training word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear & Hearing*, 27, 263–278.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *The Journal of the Acoustical Society of America*, 25, 975–979.
- Dillon, H., James, A., & Ginis, J. (1997). Client Oriented Scale of Improvement (COSI) and its relationship to several other measures of benefit and satisfaction provided by hearing aids. *Journal of the American Academy of Audiology*, 8, 27–43.
- Henshaw, H., & Ferguson, M. (2013). Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence. *PLoS ONE*, 8(5), e62836. doi:10.1371/journal.pone.0062836
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 33, 467–476.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519–533.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, United Kingdom: Cambridge University Press.
- Souza, P., Gehani, N., Wright, R., & McCloy, D. (2013). The advantage of knowing the talker. *Journal of the American Academy of Audiology*, 24, 689–700.
- Sweetow, R., & Palmer, C. V. (2005). Efficacy of individual auditory training in adults: A systematic review of the evidence. *Journal of the American Academy of Audiology*, 16, 494–504.
- Sweetow, R., & Sabes, J. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) program. *Journal of the American Academy of Audiology*, 17, 538–558.
- Tye-Murray, N., Sommers, M. S., Mauze, E., Schroy, C., Barcroft, J., & Spehar, B. (2012). Using patient perceptions of relative benefit and enjoyment to assess auditory training. *Journal of the American Academy of Audiology*, 23, 623–634.
- Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., Hale, S., & Rose, N. (2008). Auditory-visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *International Journal of Audiology*, 47(Suppl. 2), S31–S37.