Adaptation to novel foreign-accented speech and retention of benefit following training: Influence of aging and hearing loss

Rebecca E. Bieber[a] and Sandra Gordon-Salant

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742, USA

Adaptation to speech with a foreign accent is possible through prior exposure to talkers with that same accent. For young listeners with normal hearing, short term, accent-independent adaptation to a novel foreign accent is also facilitated through exposure training with multiple foreign accents. In the present study, accent-independent adaptation is examined in younger and older listeners with normal hearing and older listeners with hearing loss. Retention of training benefit is additionally explored. Stimuli for testing and training were HINT sentences recorded by talkers with nine distinctly different accents. Following two training sessions, all listener groups showed a similar increase in speech perception for a novel foreign accent. While no group retained this benefit at one week post-training, results of a secondary reaction time task revealed a decrease in reaction time following training, suggesting reduced listening effort. Examination of listeners' cognitive skills reveals a positive relationship between working memory and speech recognition ability. The present findings indicate that, while this no-feedback training paradigm for foreign-accented English is successful in promoting short term adaptation for listeners, this paradigm is not sufficient in facilitation of perceptual learning with lasting benefits for younger or older listeners.
[http://dx.doi.org/10.1121/1.4980063]

I. INTRODUCTION

A. Recognition, generalization, and adaptation

During daily communication, all listeners must be able to recognize familiar speech stimuli, generalize this recognition to similar speech, and adapt to novel speech stimuli (Kleinschmidt and Jaeger, 2015). Listeners must navigate many levels of information included in any spoken utterance, including acoustic, phonological, lexical, and semantic. This can be negatively impacted by speech that is distorted in any way, including the presence of competing background noise or unfamiliar foreign accent. One group that appears to be particularly vulnerable to challenges in perceiving accented speech is older listeners, especially those with age-related hearing loss. Older listeners with normal hearing and with hearing loss often show greater difficulty understanding speech in noise than younger listeners with comparable hearing sensitivity (Dubno *et al.*, 1984; Ferguson *et al.*, 2010). These difficulties can be exacerbated in more challenging listening circumstances, such as listening to foreign-accented speech in the presence of noise (Gordon-Salant *et al.*, 2010b). It has been suggested that this poorer performance may be related to cognitive decline associated with aging, as well as to auditory temporal processing deficits (Cristia *et al.*, 2012).

For all three speech comprehension tasks (recognition, generalization, and adaptation), exposure to similar stimuli can promote improvements in performance (Adank and Janse, 2010; Bradlow and Bent, 2008; Clarke and Garrett, 2004; Nygaard and Pisoni, 1998; Samuel and Kraljic, 2009; Sidaras *et al.*, 2009). There is strong evidence that familiarity with a specific talker can improve perception of that talker's production of novel stimuli. Nygaard and Pisoni (1998) showed that listeners who are familiar with a specific talker have higher recognition of novel words and sentences produced by that talker following training. They suggest that listeners are able to take advantage of familiarity with the talker's individual idiosyncrasies in acoustic/phonemic production as well as their more global characteristics such as prosody in adapting to the novel stimuli.

In addition to the body of work on talker familiarity, there is growing evidence that exposures to multiple talkers may further facilitate generalization and adaptation to novel stimuli. It has been shown that prior exposure to a foreign accent can improve individuals' perception of any speech produced with that specific accent by facilitating a generalization to the variations in English pronunciation that are characteristic of that accent (Clarke and Garrett, 2004; Bradlow and Bent, 2008; Sidaras *et al.*, 2009).

In the case of talker-independent, accent-specific learning, the foreign accent (for example, Spanish-accented English) arises out of interactions between the phonologic characteristics of the target language (i.e., English) and the inherent phonologic constraints of the native language (i.e., Spanish). Therefore, talkers from a similar language background will have similar distortions in their production of the target language. Exposure to these characteristic distortions facilitates understanding of a novel talker by establishing a base of familiarity with the phonologic aspects of the accent.

Additional work has shown that exposure to numerous foreign accents may facilitate adaptation to English speech

---

[a]Electronic mail: rbieber@umd.edu.

produced with a novel foreign accent (Baese-Berk *et al.*, 2013). Baese-Berk *et al.* (2013) suggest that a generalization of adaptation occurs as a result of exposure to systematic variation in the possible productions of English language phonemes inherent to foreign-accented speech. While accent-specific adaptation is facilitated by prior knowledge of the expected variations within that given accent, sufficient exposure to multiple foreign accents is intended to facilitate accent-independent adaptation by providing the listener with a knowledge of more global sources of variation in accented speech (i.e., overall rate and prosody, or changes in vowel pronunciation). A successful multiple-talker training paradigm should promote the listener's flexibility in perception, rather than the learning of a specific accent's characteristics.

It may be that relationships between acoustic and lexical information influence adaptation. Guediche *et al.* (2014) showed that listeners were more likely to accurately perceive a sentence when it was immediately preceded by a non-degraded sentence that was either identical to or conceptually related to the degraded token, than when preceded by an unrelated sentence. When the acoustic characteristics are familiar, listeners can direct their attention to perception of the lexical content; this is seen in studies of talker familiarity and accent-specific generalization. In a study such as that of Baese-Berk *et al.* (2013), the lexical content of the training materials becomes familiar, and listeners are able to expand their flexibility of perception through exposure to the acoustic variation.

B. Auditory training paradigms

Speech perception training may be completed through any of a number of paradigms, which may vary along several parameters. Training may be completed within one session, or across two or more sessions. Wright and Sabin (2007) examined the effects of number of training days and length of daily training for psychoacoustic tasks and found that there is likely a critical number of trials per session required for across-session improvement, and that any additional training past this critical value may be superfluous. They suggest that this critical value may differ across tasks. The optimal number of training days and trials-per-day may vary with the difficulty of the training task.

Listener engagement with the task may be high (i.e., interactive paradigms with provision of feedback, native stimulus modeling, and so on) or limited (i.e., exposure-only with no listener response, or listener response with no feedback). Wright *et al.* (2015) have recently suggested that a combination of paradigms with high and low engagement may be most effective. Training may focus on phoneme-level, word-level, or sentence-level contrasts, each of which may involve different cognitive processes and facilitate different types of learning. For example, Nygaard and Pisoni (1998) showed that when listeners are trained to recognize voices using sentence length stimuli, this learning did not transfer to accurate identification of talkers based on presentation of isolated words.

Adank and Janse (2010) evaluated a series of training paradigms involving various levels of engagement with a novel foreign accent. Different listener groups were asked to respond to the stimulus in one of several ways: simply think about the stimulus sentence, repeat the sentence, transcribe the sentence, or repeat the sentence while imitating the accent. Listeners who imitated the accent were the only group who showed improvements in recognition of the accented speech following training. Here, improved understanding was best facilitated by the highest level of immersion – listeners who were asked to imitate the accent needed to comprehensively engage with the variations in the English production in order to reproduce them.

The paradigm implemented by Baese-Berk *et al.* (2013) and explored in this study is a design with limited listener engagement (i.e., no feedback), which takes place over the course of two training sessions and utilizes sentence-length stimuli. In their study, participants were limited to young adults. Participants in this and many previous studies were also limited to those with hearing thresholds within normal limits, and stimuli were presented in quiet or in steady state noise (Bradlow and Bent, 2008; Clarke and Garrett, 2004; Kraljic and Samuel, 2006; Sidaras *et al.*, 2009). While Baese-Berk *et al.* (2013) and others (Clarke and Garrett, 2004; Gordon-Salant *et al.*, 2010c; Adank and Janse, 2010) have documented rapid, short term adaptation to foreign accented speech, these prior investigations have not included any measures of retention. An understanding of the listener's retention of any short-term training benefit is critical for designing and implementing a successful training paradigm.

The listeners in the study by Baese-Berk *et al.*(2013) showed better speech recognition for a novel foreign accent following exposure to speakers with a variety of foreign accents, when compared to performance by groups in prior studies by their lab, who had been exposed to either a single foreign accent or to no foreign accent. Their protocol did not include a pre-training condition, which precludes any measurement of the magnitude of training benefit for their listeners.

In the present study, the training paradigm has been modified from that of Baese-Berk *et al.* (2013) to include both pre-testing as well as a retention testing visit, with no training session immediately preceding the retention testing. One purpose of the present investigation is to determine if older listeners will derive a similar magnitude of benefit as younger listeners from exposure to multiple foreign-accented talkers in perceiving a novel foreign accent, in realistic listening situations that include multiple background talkers.

C. Listening effort and cognition

The benefit of training with foreign-accented speech is typically quantified using a metric of percent correct in speech recognition scores. However, measurements other than percent correct speech recognition performance may prove valuable to probe the benefit of training. It is possible that individuals who show similar performance scores on word recognition measures expend differing amounts of effort in completing the task (Bourland-Hicks and Tharpe, 2002). There are a number of behavioral measures, both

subjective and objective, that have been used to assess the amount of cognitive effort associated with performance on a given task. Subjective measures may include participant rankings of perceived task difficulty, and behavioral objective measures often utilize relative performance on a secondary, or dual, task. However, Gosselin and Gagné (2010) note that subjective measures of cognitive load are often inconsistent with more objective measures, particularly in the case of older adults.

A dual-task paradigm can be used as an objective measure of cognitive load for speech recognition (Gosselin and Gagné, 2010). This strategy is grounded in the capacity theory of attention (Kahneman, 1973), which suggests that any individual has a limited amount of cognitive resources at any given time, and must prioritize the use of these resources in completing one or more perceptual tasks. In completing a task, an individual is thought to utilize a portion of their limited pool of cognitive resources. If a second competing task is introduced, the participant must appropriately allocate their cognitive resources, dividing their attention between the two tasks. The primary task is understood to take up the majority of the cognitive resources, with the remainder allocated to performance on the secondary task. Dual-task performance thus varies in relation to the difficulty of the primary task. Cognitive effort is extrapolated through observing the changes in the secondary task performance throughout completion of the primary task.

Various secondary tasks have been used to measure cognitive effort during speech recognition tasks, including response to a probe light, word/digit recall, or visual tracking (Downs, 1982; Bourland-Hicks and Tharpe, 2002; Tun et al., 2009; Gosselin and Gagné, 2010). In some cases, reaction time (RT) is used as a measure of cognitive effort. Bourland-Hicks and Tharpe (2002) showed that children with hearing loss had significantly longer reaction times for a secondary visual probe task compared to normal hearing peers, although both groups performed at levels considered "good" for clinical purposes on the primary word recognition task. For the current study, a secondary task involving response time to a visual probe is implemented in order to examine benefits of training for both speech recognition and associated listening effort.

Another common finding in studies evaluating the effects of age and hearing loss on auditory and psychoacoustic tasks is greater variability within groups of older listeners compared to younger listeners (Gordon-Salant et al., 2010a). While individual performance may vary on a primary training or adaptation task, it may not always be clear what the underlying cause of these differences may be. It is well established that cognition plays a large role in individual performance in auditory tasks, even when controlling for the effects of peripheral hearing sensitivity (Füllgrabe et al., 2015). Janse and Adank (2012) investigated a number of potential factors thought to predict adaptation capacity in older listeners. They found that listeners with poorer selective attention capacity showed less adaptation on a speech task, and that poor attention-switching control was correlated with poor selective attention. In the present study, participants complete a number of cognitive tasks examining

working memory capacity, selective attention and attention-switching control, with the goal of identifying which of these cognitive domains (if any) may be related to benefit of training with exposure to foreign accents.

D. Research questions and hypotheses

The two principal aims of the present study are the following: (1) to determine whether a training paradigm with no feedback, including multiple talkers with various foreign accents, can facilitate adaptation to a novel foreign accent for both younger listeners with normal hearing and older listeners with and without hearing loss, and (2) to determine whether there is retention of the training benefit. Benefit of training is examined through measures of speech recognition performance and listening effort. An additional goal of this study is to examine the relationship between speech recognition performance and cognitive resources across the course of training.

It is anticipated that the young normal hearing listeners will demonstrate adaptation to accented speech, and high speech recognition performance following training, consistent with previous literature (Baese-Berk et al., 2013). Varying results have been reported regarding the relative benefit of training for different age groups (Cristia et al., 2012). For this experiment, it is predicted that younger adults will show higher overall speech recognition scores than older adults. All groups are expected to show an overall increase in speech recognition performance following training, but younger adults are expected to show the greatest amount of improvement following the completion of the full training protocol (Adank and Janse, 2010). It is additionally predicted that all groups will show a reduction in cognitive effort as demonstrated through the dual task as training progresses. Finally, it is anticipated that cognitive measures of selective attention, attention-switching, and working memory will correlate with higher performance on the speech recognition task, similar to the findings of Janse and Adank (2012).

II. METHOD

A. Participants

Participants for this experiment included three listener groups: 15 young listeners aged 18–28 years (mean = 22.5 years) with normal hearing (YNH) as defined by pure-tone air and bone conduction thresholds less than 25 dB hearing level (HL) from 250 to 4000 Hz; 13 older listeners aged 65–76 years (mean = 69.7 years) with normal hearing (ONH) as defined by pure-tone air and bone conduction thresholds less than 25 dB HL from 250 to 4000 Hz; and 15 older listeners aged 70–82 years (mean = 74.5 years) with mild to moderate sloping sensorineural hearing loss (OHI). Participants all passed preliminary cognitive screening using the Montreal Cognitive Assessment (Nasreddine et al., 2005). All participants were native monolingual speakers of English, with little to no prior or current exposure to the native (L1) languages of the accented talkers.

Rebecca E. Bieber and Sandra Gordon-Salant

TABLE I. Experimental paradigm over three visits. Order of training speakers and novel (test) speakers is randomized across groups and participants.

Visit 1			Visit 2			Visit 3		
Task	Accent (rating)	List	Task	Accent (rating)	List	Task	Accent (rating)	List
Familiarization	English *(1.04)*	1	Training 2	Japanese *(7.64)*	5	*Retention*	**Greek (4.48)**	7
Pre-Test	**Korean (6.56)**	2		Portuguese *(5.72)*	5			
Training 1	Japanese *(7.64)*	3		Hebrew *(3.76)*	5			
	Portuguese *(5.72)*	3		Russian *(5.68)*	5			
	Hebrew *(3.76)*	3		Cantonese *(5.44)*	5			
	Russian *(5.68)*	3	*Post-Test*	**Gisu (6.28)**	6			
	Cantonese *(5.44)*	3						
Midway Test	**Turkish (5.36)**	4						

B. Stimuli and noise

Stimuli were obtained from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSSTAR) corpus through Northwestern University's Speech Communication Research Group (Bradlow *et al.*, 2011). This corpus includes both native (L1) and non-native (L2) English recordings from over 120 talkers of a variety of speech materials. Selected stimuli for the present experiment include 120 sentences from the Hearing in Noise Test (HINT) (Nilsson *et al.*, 1994), spoken by ten different male talkers. Stimuli were root-mean-square (RMS) equalized using Adobe Audition CS5.5 and Cool Edit Pro V2 and stored in electronic data files on a personal computer (PC). Stimuli were rated by college-age native English speaking adults on a scale of 1 to 9, with 1 indicating no accent, and 9 indicating very strong accent. Five talkers with a range of accent strength—one mild (3.74), three moderate (5.44, 5.68, 5.72), and one strong (7.64)—were selected for use in the training sessions to ensure listeners' exposure to highly variable accented English pronunciation. Stimuli were also selected such that each speaker's native language (L1) came from a different language family, with the goal of exposing listeners to different patterns of phonological alterations associated with different foreign accents. For example, the five different L1s used as training stimuli (Japanese, Russian, Portuguese, Hebrew, and Cantonese) are Japonic, Slavic, Romance, Afro-Asiatic, and Sino-Tibetan languages, respectively. L1s used for test stimuli (Gisu, Greek, Korean, Turkish) all belong to unique language families (Niger-Congo, Indo-European, Koreanic, Turkic, respectively). This selection of language families was designed to rule out any possibility of accent-specific or language-family specific learning, and ensure that benefit of training was truly accent-independent. Stimuli were presented in six-talker babble comprised of both male and female native speakers of English. Calibration tones that were equivalent in RMS levels of the speech stimuli and background six-talker babble were created.

C. Procedures

Participants completed the study over the course of three visits, each separated by 7–10 days. For all tasks, participants were seated in a sound-attenuating booth, with short breaks provided as needed. Stimulus presentation was controlled using E-Prime (Psychology Software Tools, Inc). Sentences and noise were presented from a PC through an external sound card (ASUS Xonar Essence One) and routed to a monaural insert earphone (Etymotic ER-3A) to the participant's better hearing or preferred ear. Monaural presentation is used routinely in speech perception and psychoacoustic studies with older listeners to avoid potential effects of inter-aural asymmetries and/or binaural interference which is known to occur in some older listeners (Jerger *et al.*, 1993).

Stimuli were presented at 85 dB sound pressure level at a signal-to-noise ratio (SNR) of +5 dB SNR, which was established to provide equivalent performance across groups in pilot testing. Following presentation of the auditory stimulus, participants were asked to repeat what they heard, and accuracy was calculated based on keyword correct repetition. Similar to the procedure described by Baese-Berk *et al.* (2013), no feedback was given during training.

The order of tasks for each visit is shown in Table I. At the first visit, participants completed preliminary audiometric testing, task familiarization, and pre-testing. Preliminary audiometric testing included pure-tone threshold testing to confirm listener group criteria. Once participants passed the criteria for enrollment in the study, they began training at the first visit. For task familiarization, listeners were presented a half list (eight sentences, ~25 keywords) spoken by a native English talker. For pre-testing, listeners heard a full novel list (16 sentences, or ~50 keywords), spoken by an accented test speaker. In the training, listeners heard one full novel list (~50 keywords) repeated five times, each time presented by a different training speaker. Order of training speaker presentation was randomized across participants and groups. Following this training, listeners completed one additional novel list with a novel test speaker as a midway test, which provided a comparison point for participants' progress after half of the training.

At the second visit, participants completed the training portion of the study as well as a post-test. Participants heard one novel list (~50 keywords) repeated five times, presented by the same five training speakers with accents listed above (Japanese, Russian, Portuguese, Hebrew, and Cantonese). Order of speaker presentation was again randomized. Post-testing consisted of a novel list, presented by a novel test speaker with a novel accent.

The third visit consisted of a retention test as well as cognitive testing. Participants heard one final novel list with

a novel test speaker, to evaluate retention of the training effects. For all test lists (pre-test, midway, post-test, and retention), order of test talker accent was randomized across groups and participants to eliminate any potential influence of talker intelligibility. Cognitive measures were also completed at the third visit, including the Trail-Making Task A and B Test (assesses attention-switching; Reitan, 1958), the NIH Toolbox Flanker Inhibitory Control and Attention Test (assesses selective attention and inhibitory control; Eriksen and Eriksen, 1974; Gershon *et al.*, 2013), and the Listening Span test. The Listening Span test is a modified version of the Reading Span test (assesses working memory; Daneman and Carpenter, 1980), and was administered using the protocol described by Gordon-Salant and Cole (2016).

For assessment of cognitive load throughout task familiarization, training, and testing, participants completed a dual task requiring response to a visual probe. During auditory stimulus presentation, participants were seated facing a computer screen which presented visual prompts for the auditory task. For 12 out of 16 (75%) of the sentences in each list, a green dot appeared in a randomized position in the screen, at a randomized interval during primary task presentation. Participants were asked to press the space bar whenever the dot appeared. Reaction time to the dot was collected, with a maximum response time of 2000 ms (Downs and Crum, 1978; Downs, 1982; Feuerstein, 1992). Preceding the tasks, a familiarization/baseline trial consisting only of the secondary task was completed (Downs and Crum, 1978). For each of the subsequent trials, reaction time to the green dot was used as a measure of cognitive effort (Downs 1982; Bourland-Hicks and Tharpe 2002).

D. Data analysis

Data were analyzed through creation of linear mixed-effects models, using the lme4 software package in R (Bates *et al.*, 2014). To examine the benefit of training, speech recognition performance and reaction times were compared across test conditions (pre-test, midway test, post-test, retention test). Listener group (YNH, ONH, OHI) and test condition were included as fixed effects, and participants and items were included as random effects. The inclusion of participants as a random effect creates a model which takes into account the inherent independent variability of each participant within and across groups (Baayen, 2008). Models which included participants and/or items as random slopes were unable to converge and therefore no random slopes are included in these analyses.

Speech recognition performance and reaction times were used as dependent variables. For analyses, best fit models were determined to be those that accounted for the greatest amount of variance. Following determination of significant fixed effects, analysis of variance (ANOVA) was run on two models with and without the interaction term, to examine whether inclusion of the interaction increased the power of the model. Re-ordered analyses using dummy coding were run to explore all comparisons among the multi-level variables.

Speech recognition scores, collected as proportion of keywords correct per sentence, were arc-sin transformed for analyses. For the secondary task, reaction times for each listener were calculated at all test and training conditions, and were log-transformed for mixed-effects analysis in order to create a normal distribution of data (Kirk, 1995). While the mixed effects analysis includes trial-by-trial data for construction of the models, both speech scores and RT data are plotted by list means, for ease of visualization. Additionally, the RT scores are plotted as a normalized relative reaction time [(RT – BaselineRT)/BaselineRT], which decreases the variability due to factors such as age-related slowing in the raw scores (Hornsby *et al.*, 2013).

Scoring for the cognitive tasks was completed as follows. For the Flanker task, reaction time scores and age-adjusted scale scores were calculated for each listener. L-SPAN scores were calculated based on the level (i.e., number of sentences to recall) at which the participant performed correctly on two out of three sentence sets. The Trail-Making task contains two components, Trail-Making A and Trail-Making B. Trail-Making B tasks executive control, while Trail-Making A is primarily a task of speed. To generate a Trail-Making score, a normalized derived measure ([B-A]/A) was calculated for each participant. This normalized measure accounts for any age-related slowing in the older listener groups (Füllgrabe *et al.*, 2015).

III. RESULTS

A. Benefits of training on speech recognition

Speech recognition scores across all training and test conditions for the three listener groups are shown in Fig. 1. The three visits are separated by solid vertical lines; training sessions within visits are separated between dashed vertical lines. Test lists (pre-test, midway test, post-test, retention test) are
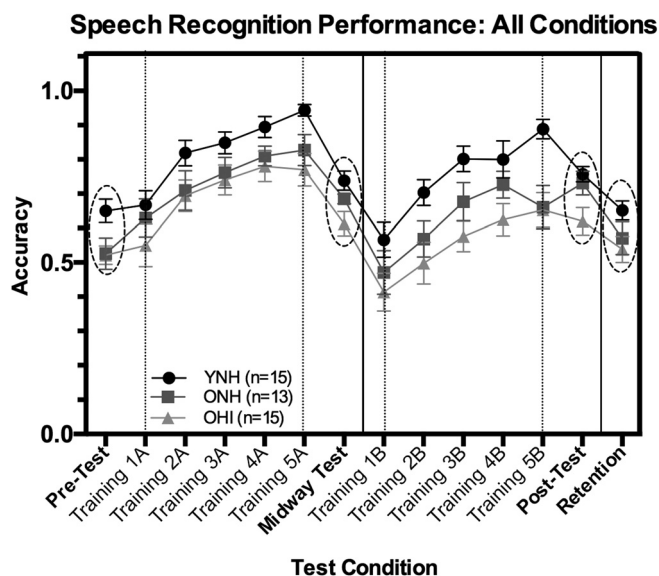

FIG. 1. Speech recognition performance and standard errors across training and test lists for three listener groups. Solid vertical lines indicate separate visit days, and dashed lines indicate training sessions. Test conditions (Pre-test, midway, post-test, retention test) are circled. Error bars represent standard error of measurement.

Rebecca E. Bieber and Sandra Gordon-Salant

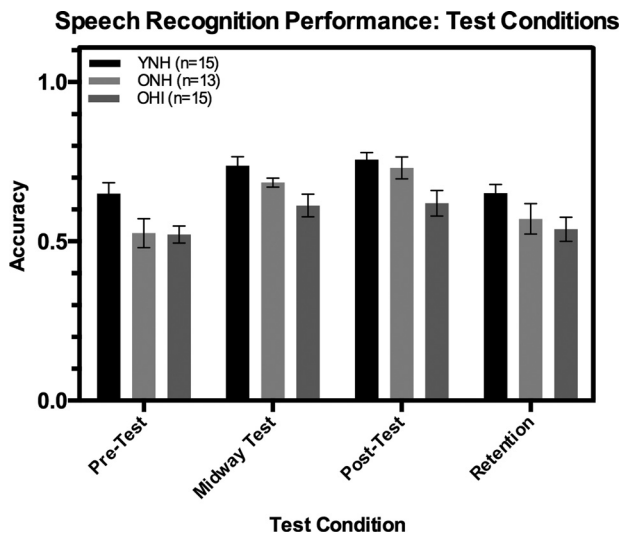**Speech Recognition Performance: Test Conditions**

FIG. 2. Mean speech recognition performance and standard errors across test lists (Pre-test, midway, post-test, retention test) for three listener groups. Error bars represent standard error of measurement.
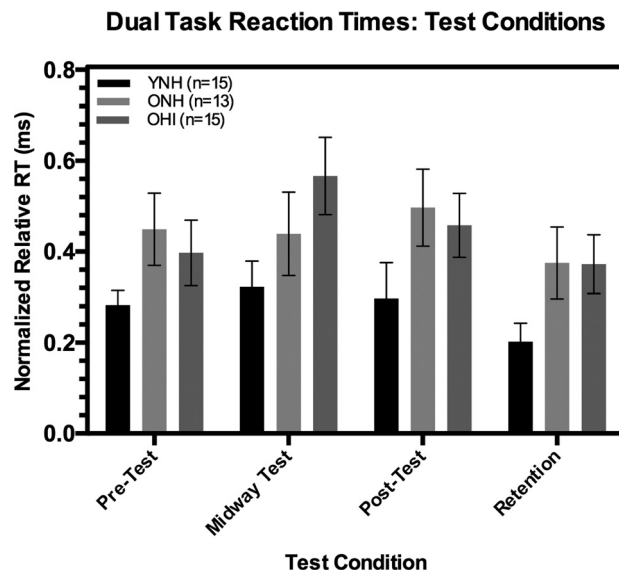
Dual Task Reaction Times: Test Conditions

FIG. 3. Mean normalized relative reaction times and standard errors across test lists (Pre-test, midway, post-test, retention test) for three listener groups. Error bars represent standard error of measurement.

circled. Figure 2 contains speech recognition scores for the four test conditions only. For all of these test conditions, YNH listeners demonstrated higher performance than either the ONH ($\text{ß} = -0.24$, $t = -2.21$, $p < 0.05$) or OHI ($\text{ß} = -0.4$, $t = -3.74$, $p < 0.005$) listeners. There was no difference in performance between the two older listener groups in the four test conditions ($\text{ß} = 0.15$, $t = 1.39$, $p = 0.16$).

For all three groups, there was a significant increase in performance from pre-test to midway test ($\text{ß} = 0.37$, $t = 6.06$, $p < 0.001$), demonstrating a fairly rapid short-term benefit of training. However, there were no significant differences in performance between the midway test and the post-test ($\text{ß} = 0.07$, $t = 1.17$, $p = 0.24$); listeners derived no additional benefit from a second day of training. Performance at the retention test, however, was significantly poorer than that at both the midway ($\text{ß} = -0.26$, $t = -4.34$, $p < 0.001$) and the post-test ($\text{ß} = -0.33$, $t = -5.55$, $p < 0.001$), indicating that on average, all three groups did not retain any benefit of this training. In fact, performance at the retention test did not differ significantly from the pre-test ($\text{ß} = 0.01$, $t = 1.71$, $p = 0.09$). No significant interactions between test condition and listener group were revealed, suggesting that all listener groups demonstrated similar patterns of improvement and decline across the course of training.

A further examination of the data revealed significant decreases in performance from the midway test to the first list of training day two ($\text{ß} = -0.56$, $t = -9.49$, $p < 0.005$). While this is not a true measure of retention of the generalization benefit seen on day one, it does suggest that any talker-specific adaptation from day one was not retained; training talkers are the same on the two training days.

B. Benefits of training on dual-task performance

The relative reaction time data for test lists for the three listener groups are shown in Fig. 3, with RT values listed in Table II. Listeners were asked to respond to a visual probe appearing at random intervals throughout each list with a button push. Lower RT values indicate faster reaction times, and are interpreted to reflect decreased cognitive effort associated with the primary, speech recognition task. YNH listeners demonstrated faster relative reaction times than either the ONH ($\text{ß} = 0.09$, $t = 2.84$, $p < 0.005$) or OHI ($\text{ß} = 0.13$, $t = 4.43$, $p < 0.005$) listeners. There was no difference in relative reaction time between the two older listener groups. ($\text{ß} = -0.05$, $t = -1.43$, $p = 0.15$), and no significant interactions between test condition and group were evident.

Overall, there was no significant difference in relative reaction times from pre-test to post-test ($\text{ß} = 0.005$, $t = 0.60$, $p = 0.55$). An initial increase in relative reaction times from pre-test to midway test ($\text{ß} = 0.02$, $t = 2.87$, $p < 0.005$) was followed by a decrease in relative reaction times from midway to post-test ($\text{ß} = -0.02$, $t = -2.18$, $p < 0.05$). However, the pattern of decreasing relative reaction times continued, with all listener groups demonstrating significantly faster relative reaction times at retention than at post-test ($\text{ß} = 0.01$, $t = 3.9$, $p < 0.001$).

C. Efficacy of training paradigm

Individual training curves and overall group means for the two training sessions are plotted in Fig. 4. Within each training session, listeners heard the same list of sentences five times, each repetition by one of five different talkers. Theoretically, listeners who are able to take optimal advantage of this training paradigm should show near-monotonic increases in speech recognition performance across these five lists: as they hear the same list five times, increased speech recognition is expected. The present results reveal that the actual trajectory of the training list performance varied greatly across individual listeners.

To examine the group patterns of training curves, linear mixed effects models were generated with speech recognition performance as the dependent variable. Listener group (YNH, ONH, OHI), training day (day 1 and 2), talker accent

TABLE II. Raw reaction times (ms) for all listener groups in the following conditions: RT-Only (secondary task familiarization); RT+Speech (dual-task familiarization); Pre-Test; Midway; Post-Test; Retention.

	YNH (n = 15)				ONH (n = 13)				OHI (n = 15)			
	Mean	*Minimum*	*Maximum*	SD[a]	Mean	*Minimum*	*Maximum*	SD	Mean	*Minimum*	*Maximum*	SD
RT-Only	460.02	*375.5*	*578.17*	69.83	513.01	*18.03*	*155.29*	21.86	570.03	*445.50*	*732.5*	95.24
RT+Speech	701.06	*397.83*	*1190.8*	225.37	846.48	*473.0*	*1375.92*	283.24	866.65	*642.33*	*1151.88*	156.14
Pre-Test	586.15	*474.33*	*823.33*	89.19	751.26	*453.50*	*1232.05*	234.05	784.04	*594.25*	*1098.67*	154.70
Midway	601.24	*460.5*	*814.00*	99.17	743.26	*442.83*	*1344.75*	242.51	881.99	*599.36*	*1261.75*	180.52
Post-Test	596.52	*474.08*	*1081.78*	175.56	774.35	*504.25*	*1402.33*	246.46	819.12	*644.17*	*1142.5*	151.06
Retention	546.91	*440.25*	*688.25*	72.62	710.15	*490.06*	*1303.31*	222.79	776.04	*602.58*	*1128.22*	161.69

[a]Standard deviation (SD).

rating (continuous variable on a scale of 0–9) and training list order (lists 1 through 5) were included as fixed effects, and participants and items were included as random effects. Significant main effects of listener group ($p < 0.05$), training day ($p < 0.05$), accent rating ($p < 0.001$) and order ($p < 0.05$) were found. The main effect of training list order is expected, and reflects learning of the repeated stimuli. The main effect of accent rating reflects a decrease in speech perception performance with an increase in accent rating (i.e., stronger accent). Significant interactions between listener group and accent rating were additionally seen. To examine the main effects and interactions, data were separated by listener group and training day for further analysis.

All three listener groups showed a significant effect of accent rating on both training days [YNH (ß = –1.45, $t = –4.5$, $p < 0.005$); ONH (ß = –0.38, $t = –9.5$, $p < 0.000$); OHI (ß = –0.44, $t = –11.5$, $p < 0.000$)]. While YNH listeners showed no differences in speech recognition across the first and second days of training, significantly higher performance was seen on training day one for both the ONH (ß = –0.45, $t = –9.92$, $p < 0.000$) and the OHI (ß = –0.49, $t = –11.3$, $p < 0.000$) groups when compared to training day two.

YNH listeners showed higher performance than OHI listeners as talkers' accent rating increased (ß = –0.09, $t = –3.91$, $p < 0.001$). A similar trend was seen when comparing the YNH and ONH listener groups, though this difference did not reach significance (ß = –0.05, $t = –1.95$, p = 0.05). No such differences were seen when comparing the two older listener groups (ß = 0.05, $t = 1.79$, p = 0.07).

Visual inspection of the individual training curves illustrates these differences in susceptibility to accent strength throughout the training sessions. Many listeners show large decreases in performance in the second half of the training sessions (See Fig. 3; lists 3–5), suggesting that despite the lexical familiarity of the training stimuli, the accentedness of the speech intruded significantly, causing mostly transient decreases in speech recognition performance throughout the training sessions. These listeners may not have been able to draw on their knowledge of the semantic content of the training sentences to match the incoming acoustic information from the strongly accented talkers.

D. Speech performance and cognition

All listeners completed measures of attention-switching (Trail-Making Task), selective attention (Flanker), and working memory (L-SPAN). Performance on these measures was compared to speech recognition performance to determine if cognitive abilities were related to training performance. Pearson product-moment correlations were conducted separately between each cognitive measure and speech recognition performance for all listeners combined. Of the three measures, only the L-SPAN correlated with absolute speech recognition
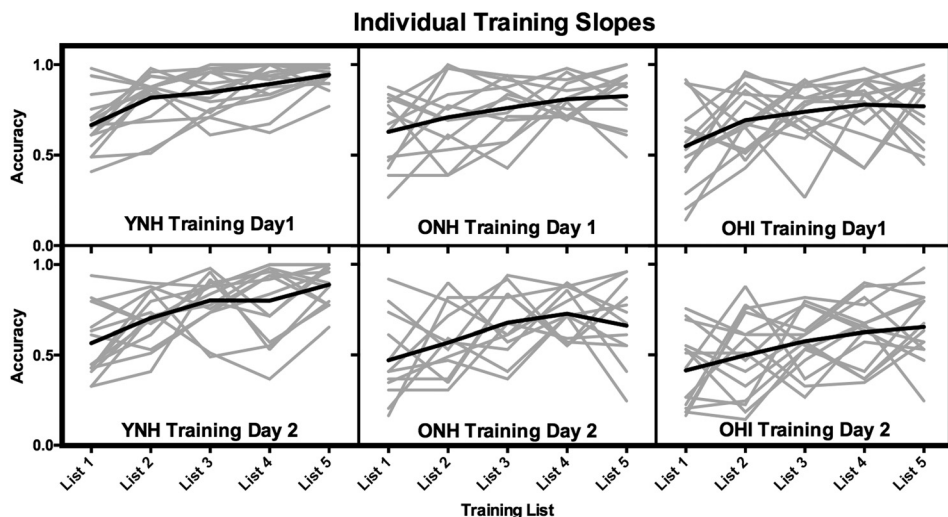

FIG. 4. Patterns of learning during training sessions, by listener group. Grey lines represent individual training curves; black lines represent the group mean. Training materials include identical lists spoken by talkers with different L1 accent.

Rebecca E. Bieber and Sandra Gordon-Salant

performance. Figure 5 presents the scatterplots of L-SPAN scores and speech recognition scores for the three listener groups in four test intervals (pre-test, midway test, post-test, and retention test). The correlation analyses revealed that working memory was significantly correlated with performance at pre-test ($r = 0.425$, $p < 0.01$), midway test ($r = 0.453$, $p < 0.01$), post-test ($r = 0.328$, $p < 0.05$), and retention test ($r = 0.402$, $p < 0.01$). A one-way ANOVA was conducted on the L-SPAN scores and revealed a significant main effect of listener group [$F(2, 42) = 14.451$, $p < 0.001$]; YNH listeners had significantly higher LSPAN scores than either the ONH listeners ($p < 0.001$) or the OHI listeners ($p < 0.001$). The two older listener groups did not differ in their L-SPAN performance ($p = 0.98$). Additional analyses explored the correlations between the magnitude of the training benefit (i.e., change in speech recognition performance) and the various cognitive measures; no significant relationships were found.

IV. DISCUSSION

A. Benefits of exposure on speech recognition

Following one session of exposure to systematic variation in accented English, all listeners show a significant improvement in speech recognition scores. These findings add further support to the notion that exposure to multiple accents in a short-term generalization paradigm improves recognition of a novel foreign accent, as found by Baese-Berk et al. (2013). The current results additionally expand on these prior findings by demonstrating that older listeners, with and without hearing loss, may also benefit from exposure to a broad range of accents for improving recognition of foreign-accented speech.

However, the benefit of this training paradigm appears limited in some respects. The current findings suggest that listeners do not retain the benefits of training at an interval of seven to ten days post-exposure, as measured by speech

FIG. 5. Scatterplots showing the relationship between L-SPAN sores and speech recognition accuracy for participants in the three listener groups, measured at the four test intervals (top left panel: pre-test; top right panel: midway test; bottom left panel: post test; bottom right panel: retention test). The solid line in each plot represents the overall correlation with L-SPAN scores at each test condition.

recognition performance. Listeners show significant decreases in performance across both training visit intervals. Scores are significantly lower from midway test (end visit 1) to training list 2A (start visit 2) and from post-test (end visit 2) to retention test (visit 3). Both talker-specific generalization (end visit 1 – start visit 2) and accent-independent adaptation (end visit 2 – visit 3) are not retained.

One limitation of the present study is that it did not include a true control group. In the study by Baese-Berk et al. (2013), listeners in their multiple-accent paradigm were compared to a group from a prior study who were training on a single accent or on no accent. Post-training scores for the multiple-accent group were higher than those of the other groups. Without a true control group such as this, the results of the present study may be influenced by practice effects. However, the inclusion of pre-testing and retention testing is an important addition to the protocol used by Baese-Berk et al. (2013). Their data reflect group comparisons, but do not provide information regarding the individual magnitude of training benefit. Another limitation of the present study is the relatively small sample sizes for each listener group. These small group sizes may contribute to some of the null findings documented here. Given the high variability within and across listener groups – particularly with older listeners, larger participant groups and inclusion of a true control group will be critical for future work.

The current findings hold important implications for design and implementation of future training paradigms. While many prior studies of short-term or rapid adaption have not included retention measures (Clarke and Garrett, 2004; Baese-Berk et al., 2013), there is evidence that listeners can retain learning for trained aspects of speech stimuli such as distinction of novel phonological contrasts (Lively et al., 1993; Flege et al., 1995). However, it should be noted that these prior studies utilized paradigms involving more direct engagement by the listeners. In training native Japanese listeners to distinguish between English /l/ and /r/ production, Logan et al. (1993) utilized a two-alternative-forced-choice task, with immediate feedback given to listeners during the training sessions. In a similar training paradigm, Flege et al. (1995) asked listeners to identify either a voiced or unvoiced stop consonant, or decide whether or not two stimuli contained the same stop consonant. Listeners were provided trial-by-trial feedback. This type of dynamic task, as well as the provision of feedback, creates a more active engagement on the part of the listener throughout training. It may be the case that a training paradigm including only exposure to the stimulus, rather than an active training variant, is not successful at providing long-term retention of training benefit, even for younger listeners with normal hearing.

B. Benefit of exposure on listening effort

Throughout training, listeners were asked to complete a secondary task involving reaction time to a visual probe. Reaction times were used to examine the effort associated with the speech recognition task before, during, and after training. Listening effort is known to be a critical component of language processing, and has been shown to increase as
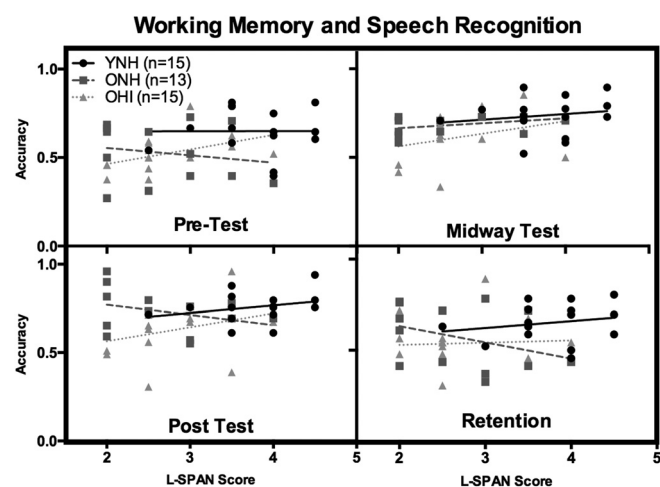
the listening environment grows more challenging (Zekveld *et al.*, 2011). Interestingly, indicators of increased listening effort can occur without any significant changes in performance on a listening task (Mackersie and Cones, 2011).

In the present study, the dual-task (RT) data do not reflect the training pattern observed on the speech recognition scores. In fact, the present results suggest that, following an initial increase, all listeners demonstrated consistent decreases in reaction time over the course of the entire experiment – resulting in a global decrease in listening effort. While speech recognition performance was equivalent at pre-test and retention test, listeners displayed evidence that the task had become significantly less effortful by the retention test. Mackersie and Cones (2011) examined behavioral and psychophysiological components of listening effort across two test sessions. They found that listeners demonstrated reduced ratings of mental effort and frustration on the second day of testing. The authors propose that participants may have become more comfortable in the test setting, and had decreased anxiety related to the task as the task became more familiar, though task demand remained high. A similar effect may be evident in the present study, suggesting that exposure training can help mitigate the effects of high listening demands during speech processing in daily communication settings.

The inclusion of a secondary task alters the nature of the training paradigm from that of Baese-Berk *et al.* (2013), who did not include any measure of listening effort. The current protocol, which tracks reaction times (and thereby probes listening effort) across the entirety of the protocol, did not include a primary task-only condition to rule out a decrease in primary task performance due to the inclusion of a secondary task. While it is possible that such a decrease may have occurred in this study, prior work has demonstrated that, with sufficient instruction, there is no detrimental influence of a simple visual probe reaction time task on a primary speech recognition task (Downs, 1982; Hornsby *et al.*, 2013; Picou and Ricketts, 2014).

C. Speech recognition performance and cognition

Listeners completed a number of cognitive measures probing working memory, attention-switching control, and selective attention. While it is well accepted that cognitive status plays an important role in speech perception, reports vary as to the relative contributions of individual cognitive domains. In the present study, working memory capacity was the only measure that was significantly correlated with higher levels of accuracy in recognition of foreign-accented speech. This finding is consistent with many previous reports of the strong relationship between working memory and speech perception, particularly in challenging listening conditions (Akeroyd *et al.*, 2008; Rudner *et al.*, 2011; Anderson *et al.*, 2013; Gordon-Salant and Cole, 2016). In contrast, the measures relating to attention were not significantly associated with speech recognition performance, consistent with findings reported by Janse and Adank (2012), who found no relationship between attention-switching control and accuracy in a study of foreign accent adaptation. Other investigations observed a significant relationship between speech recognition accuracy and measures of attention (Füllgrabe *et al.*, 2015; Zekveld *et al.*, 2011), however, the speech stimuli in those studies were not spoken with a foreign accent.

While absolute speech recognition scores were significantly correlated with attentional control, the lack of a significant relationship between attentional control and rate of change in accuracy in the current study is unexpected, and the explanation for this is unclear. It may be that a single global measure of executive control may be more effective than multiple measures of independent cognitive components in examining underlying causes for differences in speech recognition performance (Füllgrabe *et al.*, 2015); these relationships require further examination.

D. Effects of aging and hearing loss

While the multiple-accent exposure training paradigm has been documented as beneficial for younger listeners with normal hearing sensitivity (Baese-Berk *et al.*, 2013), this paradigm has not been explored previously in populations that are older, including older adults with and/or without hearing impairment. These groups are particularly important to study, because older people often present with a complaint of difficulty understanding accented English. The current findings show that as a group, older listeners derive benefit in the form of short-term adaptation to speech with a novel foreign accent, and this applies to older listeners with hearing loss as well. Moreover, the older listener groups (with and without hearing loss) demonstrated improvements in speech recognition performance following exposure to the training stimuli. The lack of interaction between age and test condition suggests that the magnitude of benefit was similar across groups, despite differences in absolute levels of performance. However, it should be noted that within groups, the individual older listeners showed much more variability in performance than younger listeners within the training sessions. A significant proportion of older listeners displayed non-monotonic patterns of performance across the training lists. These age-related patterns may be the result of greater susceptibility by older listener groups to talkers with stronger foreign accents.

E. Benefits of high variability training stimuli

Accent-independent adaptation to foreign-accented speech is understood to occur as a result of exposure to speech produced with a wide range of foreign accents (Baese-Berk *et al.*, 2013). A training paradigm such as this includes a sufficient range of variability in English pronunciation for listeners to expand their flexibility in perceiving unfamiliar speech stimuli and assign lexical meaning. Following this exposure, listeners should theoretically be able to access a mental store of the possible variants of English pronunciation and apply this internal flexibility to understanding a novel foreign accent.

In the study by Baese-Berk *et al.* (2013), training stimuli included recordings from five talkers with different accents (Thai, Korean, Hindi, Romanian, and Mandarin), with a

Rebecca E. Bieber and Sandra Gordon-Salant

similar "mid-range intelligibility." These languages are all from unique language families. In the present study, talkers were selected in an attempt to magnify any benefit of training seen in the Baese-Berk et al. (2013) study, that is, by selecting highly varied talkers in terms of L1 and degree of accent. Training and test talkers were selected, again, from different native language families and originating from different continents, as available. Talkers with a range of accent strength were selected for training talkers in order to increase the degree of variation from standard American English. These changes to the composition of the accented-ness of training stimuli appear not to have eliminated the benefit of training seen in the original study. It is difficult to evaluate whether the magnitude of the training benefit was similar to that of the prior study given the differences in study paradigms (inclusion of true control group and/or pre-test condition). However, listeners in the present study did improve in performance with a novel foreign accent following the high-variability training protocol; including a wider range of accent strengths appears to be beneficial.

Kleinschmidt and Jaeger (2015) propose a model for understanding listeners' abilities to recognize familiar speech, generalize this understanding to similar talkers, and adapt their perceptual generalization to talkers or listening situations that are entirely novel. This can often include understanding talkers with novel foreign accents. Kleinschmidt and Jaeger (2015) suggest that, both within and across language or accent groups, phonetic cues are distributed in a structured way. In their framework, a listener, or "ideal adaptor," is able to adapt to a novel speech stimulus by drawing on their internal knowledge of this distribution of phonetic cues, and continually updating this knowledge.

Naturally, this ideal adaptor framework cannot and is not designed to capture the immense variability associated with individual listeners and their performance. Individual listeners will vary in numerous characteristics, including age, hearing sensitivity, cognitive status, language background, motivation, and aptitude for learning. While some of these characteristics may be controlled in the experimental paradigm, they all likely influence the individual's ability to take advantage of any given training paradigm.

Perrachione et al. (2011) examined the benefit of including high variability training stimuli in a training paradigm designed to teach a novel phonological contrast. They found that listeners who had a high aptitude for learning phonological contrasts (assessed prior to training) benefitted from training stimuli with high variability, whereas low aptitude listeners were negatively impacted by high stimulus variability, and benefitted more from training paradigms where stimuli were blocked by talker. These authors suggest that learning depends on a critical interaction between the characteristics of the listener and the design of the training paradigm. Interestingly, the ideal adaptor framework proposed by Kleinschmidt and Jaeger (2015) is built on the assumption that the listener identifies that there is a need to adapt and is motivated to complete this process. The model also depends on the listener's prior experiences. Future studies may benefit from a more comprehensive evaluation of their listeners' aptitudes, motivations, and weaknesses, in order to develop highly specific training paradigms based on individual listener attributes.

High variability of training materials is only one aspect of the training paradigm described here. As described earlier, training paradigms may vary in duration, stimulus characteristics, blocking of stimuli, and participant engagement. Wright et al. (2015) examined a training paradigm that combines both active training and passive exposure-only conditions, and concluded that this hybrid paradigm is more beneficial than either type of training alone. These authors also suggested that effective training can occur only when a sufficient amount of daily practice is achieved, and that benefits of training may be greatest when such practice is maintained, stimulating sustained neural plasticity.

The design of the current study, with two training sessions separated in time by one week, may have precluded listeners from maintaining this state of plasticity. Previous work has demonstrated that short-term adaptation occurs fairly quickly within a training session and then plateaus (Clarke and Garrett, 2004). The ideal adaptor framework suggests that, in the case of selective adaptation, the adaptation may occur gradually, and build up over time (Kleinschmidt and Jaeger, 2015), which may be the culmination of several short-term training sessions.

In the present study, while an initial training session was beneficial, listeners clearly did not retain this benefit, or even a talker-specific adaptation, at a one week post-training interval. A future study might be designed to provide training on a number of consecutive days, and over a longer time course. Consecutive training days may be critical for retention of the training benefit within training sessions, and, ideally, may facilitate long term retention of the adaptation. It is possible that a longer-term training paradigm could be beneficial in maintaining adaptive skills.

V. CONCLUSION

The results of this study support a short-term benefit of exposure to multiple foreign accents in adapting to speech with a novel foreign accent. This finding builds on previous work by Baese-Berk et al. (2013) and documents a benefit not only for young, normal hearing listeners, but also for listeners who are older and who have either normal hearing or hearing impairment.

While older listeners show poorer overall speech recognition performance, there does not seem to be an effect of age in the magnitude of improvement in performance following a training session. However, a critical finding of this study is the lack of retention of any training benefit. These findings also suggest that repeated exposure to foreign-accented speech may decrease the overall cognitive load associated with speech perception. This finding warrants further investigation, considering the critical role of listening effort in perception of speech in challenging environments. Future studies of auditory training and adaptation to foreign-accented speech should be designed with consideration of individual variability in learning styles and cognitive status, as well as facilitation of long-term plasticity.

J. Acoust. Soc. Am. 141 (4), April 2017

Rebecca E. Bieber and Sandra Gordon-Salant   2809

Adank, P., and Janse, E. (**2010**). "Comprehension of a novel accent by young and older listeners," Psychol. Aging **25**(3), 736.

Akeroyd, M. A. (**2008**). "Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults," Internat. J. Audiol. **47**(2), S53–S71.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (**2013**). "A dynamic auditory-cognitive system supports speech-in-noise perception in older adults," Hear. Res. **300**, 18–32.

Baayen, R. H. (**2008**). *Analyzing linguistic data: A practical introduction to statistics using R* (Cambridge University Press, Cambridge, UK).

Baese-Berk, M. M., Bradlow, A. R., and Wright, B. A. (**2013**). "Accent-independent adaptation to foreign accented speech," J. Acoust. Soc. Am. **133**(3), EL174–EL180.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., and Grothendieck, G. (**2014**). "Package 'lme4'," R Foundation for Statistical Computing, Vienna.

Bourland-Hicks, C., and Tharpe, A. M. (**2002**). "Listening effort and fatigue in school age children with and without hearing loss," J. Speech Hear. Res. **45**, 573–584.

Bradlow, A. R., and Bent, T. (**2008**). "Perceptual adaptation to non-native speech," Cognition **106**(2), 707–729.

Bradlow, A. R., L. Ackerman, L. A. Burchfield, L. Hesterberg, J. Luque, and K. Mok. (**2011**). "Language- and talker-dependent variation in global features of native and non-native speech," in *Proceedings of the 17th International Congress of Phonetic Sciences*, pp. 356–359.

Clarke, C. M., and Garrett, M. F. (**2004**). "Rapid adaptation to foreign-accented English," J. Acoust. Soc. Am. **116**(6), 3647–3658.

Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., and Floccia, C. (**2012**). "Linguistic processing of accented speech across the lifespan," Front. Psych. **3**, 479.

Daneman, M., and Carpenter, P. A. (**1980**). "Individual differences in working memory and reading," J. Verbal Learn Verbal Behav. **19**(4), 450–466.

Downs, D. W. (**1982**). "Effects of hearing aid use on speech discrimination and listening effort," J. Speech Hear. Dis. **47**(2), 189–193.

Downs, D. W., and Crum, M. A. (**1978**). "Processing demands during auditory learning under degraded listening conditions," J. Speech Hear. Res. **21**(4), 702–714.

Dubno, J. R., Dirks, D. D., and Morgan, D. E. (**1984**). "Effects of age and mild hearing loss on speech recognition in noise," J. Acoust. Soc. Am. **76**(1), 87–96.

Eriksen, B. A., and Eriksen, C. W. (**1974**). "Effects of noise letters upon the identification of a target letter in a nonsearch task," Attent. Percep. Psychophys. **16**(1), 143–149.

Ferguson, S. H., Jongman, A., Sereno, J. A., and Keum, K. (**2010**). "Intelligibility of foreign-accented speech for older adults with and without hearing loss," J. Am. Acad. Audiol. **21**(3), 153–162.

Feuerstein, J. F. (**1992**). "Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort," Ear Hear. **13**(2), 80–86.

Flege, J. E., Takagi, N., and Mann, V. (**1995**). "Japanese adults can learn to produce English/I/and/l/accurately," Lang. Speech **38**(1), 25–55.

Füllgrabe, C., Moore, B. C., and Stone, M. A. (**2015**). "Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition," Front. Aging Neurosci. **6**, 347.

Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., and Nowinski, C. J. (**2013**). "NIH toolbox for assessment of neurological and behavioral function," Neurology **80**, S2–S6.

Gordon-Salant, S., and Cole, S. S. (**2016**). "Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing," Ear Hear. **37**(5), 593–602.

Gordon-Salant, S., Frisina, R. D., Fay, R. R., and Popper, A. (Eds.). (**2010a**). *The Aging Auditory System* (Springer Science and Business Media, New York), Vol. 34, p. 131.

Gordon-Salant, S., Yeni-Komshian, G. H., and Fitzgibbons, P. J. (**2010b**). "Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss," J. Acoust. Soc. Am. **128**(1), 444–455.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Schurman, J. (**2010c**). "Short-term adaptation to accented English by younger and older adults," J. Acoust. Soc. Am. **128**(4), EL200–EL204.

Gosselin, P. A., and Gagné, J. P. (**2010**). "Utilisation d'un paradigme de double tâche pour mesurer l'attention auditive [Use of a dual-task paradigm to measure listening effort]," Inscript. Rép. **34**(1), 43.

Guediche, S., Reilly, M., and Blumstein, S. E. (**2014**). "Facilitating perception of speech in babble through conceptual relationships," J. Acoust. Soc. Am. **135**(4), 2257–2258.

Hornsby, B. W. (**2013**). "The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands," Ear Hear. **34**(5), 523–534.

Janse, E., and Adank, P. (**2012**). "Predicting foreign-accent adaptation in older adults," Q. J. Exp. Psychol. **65**(8), 1563–1585.

Jerger, J., Silman, S., Lew, H. L., and Chmiel, R. (**1993**). "Case studies in binaural interference: Converging evidence from behavioral and electrophysiologic measures," J. Am. Acad. Audiol. **4**(2), 122–131.

Kahneman, D. (**1973**). *Attention and Effort* (Prentice-Hall, Englewood Cliffs, NJ).

Kirk, R. E. (**1995**). *Experimental Design: Procedures for the Behavioral Sciences*, 3rd ed. (Brooks/Cole, Pacific Grove, CA) p. 105.

Kleinschmidt, D. F., and Jaeger, T. F. (**2015**). "Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel," Psychol. Rev. **122**(2), 148.

Kraljic, T., and Samuel, A. G. (**2006**). "Generalization in perceptual learning for speech," Psychonom. Bull. Rev. **13**(2), 262–268.

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**(3), 1242–1255.

Mackersie, C. L., and Cones, H. (**2011**). "Subjective and psychophysiological indexes of listening effort in a competing-talker task," J. Am. Acad. Audiol. **22**(2), 113–122.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., and Chertkow, H. (**2005**). "The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment," J. Am. Ger. Soc. **53**(4), 695–699.

Nilsson, M., Soli, S. D., and 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**(2), 1085–1099.

Nygaard, L. C., and Pisoni, D. B. (**1998**). "Talker-specific learning in speech perception," Attent. Percep. Psychophys. **60**(3), 355–376.

Perrachione, T. K., Lee, J., Ha, L. Y., and Wong, P. C. (**2011**). "Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design," J. Acoust. Soc. Am. **130**(1), 461–472.

Picou, E. M., and Ricketts, T. A. (**2014**). "The effect of changing the secondary task in dual-task paradigms for measuring listening effort," Ear Hear. **35**(6), 611–622.

Reitan, R. M. (**1958**). "Validity of the Trail Making Test as an indicator of organic brain damage," Percep. Motor Skills **8**(3), 271–276.

Rudner, M., Rönnberg, J., and Lunner, T. (**2011**). "Working memory supports listening in noise for persons with hearing impairment," J. Am. Acad. Audiol. **22**(3), 156–167.

Ryan, C. (**2013**). "Language Use in the United States: 2011," American Community Survey Reports, http://www.census.gov/prod/2013pubs/acs-22.pdf (Last viewed 4/9/2017).

Samuel, A. G., and Kraljic, T. (**2009**). "Perceptual learning for speech," Attent. Percep. Psychophys. **71**(6), 1207–1218.

Sidaras, S. K., Alexander, J. E., and Nygaard, L. C. (**2009**). "Perceptual learning of systematic variation in Spanish-accented speech," J. Acoust. Soc. Am. **125**(5), 3306–3316.

Tun, P. A., McCoy, S., and Wingfield, A. (**2009**). "Aging, hearing acuity, and the attentional costs of effortful listening," Psychol. Aging **24**(3), 761.

Wright, B. A., Baese-Berk, M. M., Marrone, N., and Bradlow, A. R. (**2015**). "Enhancing speech learning by combining task practice with periods of stimulus exposure without practice," J. Acoust. Soc. Am. **138**(2), 928–937.

Wright, B. A., and Sabin, A. T. (**2007**). "Perceptual learning: how much daily training is enough?," Exper. Brain Res. **180**(4), 727–736.

Zekveld, A. A., Kramer, S. E., and Festen, J. M. (**2011**). "Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response," Ear Hear. **32**(4), 498–510.