Research Article

JARO
Journal of the Association for Research in Otolaryngology

Check for updates

Rate Discrimination Training May Partially Restore Temporal Processing Abilities from Age-Related Deficits

Samira Anderson[1] , Lindsay DeVries[1], Edward Smith[1], Matthew J. Goupell[1], and Sandra Gordon-Salant[1]

[1] *Department of Hearing and Speech Sciences, University of Maryland, College Park 20742, USA*

ABSTRACT

The ability to understand speech in complex environments depends on the brain's ability to preserve the precise timing characteristics of the speech signal. Age-related declines in temporal processing may contribute to the older adult's experience of communication difficulty in challenging listening conditions. This study's purpose was to evaluate the effects of rate discrimination training on auditory temporal processing. A double-blind, randomized control design assigned 77 young normal-hearing, older normal-hearing, and older hearing-impaired listeners to one of two treatment groups: experimental (rate discrimination for 100- and 300-Hz pulse trains) and active control (tone detection in noise). All listeners were evaluated during pre- and post-training sessions using perceptual rate discrimination of 100-, 200-, 300-, and 400-Hz band-limited pulse trains and auditory steady-state responses (ASSRs) to the same stimuli. Training generalization was evaluated using several temporal processing measures and sentence recognition tests that included time-compressed and reverberant speech stimuli. Results demonstrated a session × training group interaction for perceptual and ASSR testing to the trained frequencies (100 and 300 Hz), driven by greater improvements in the training group than in the active control group. Further, post-test rate discrimination of the older listeners reached levels that were equivalent to those of the younger listeners at pre-test. Generalization was observed in significant improvement in rate discrimination of untrained frequencies (200 and 400 Hz) and in correlations between performance changes in rate discrimination and sentence recognition of reverberant speech. Further, non-auditory inhibition/attention performance predicted training-related improvement in rate discrimination. Overall, the results demonstrate the potential for auditory training to partially restore temporal processing in older listeners and highlight the role of cognitive function in these gains.

INTRODUCTION

The brain's ability to process the temporal characteristics of auditory stimuli is an integral component of speech understanding, particularly in complex environments that reduce the redundancy of the speech signal. For example, the ability to discriminate between changes in temporal rate contributes to the listener's ability to discriminate fundamental frequency, which serves as a cue to speaker and gender identification. Thus, temporal processing is an important factor that supports speech segregation and speech understanding in noise (Zaltz and Kishon-Rabin 2022). In addition, non-speech measures of temporal processing such as gap detection can predict speech recognition in reverberation in normal-hearing listeners (Dreschler and Leeuw 1990; Gordon-Salant and Fitzgibbons 1993; Irwin and McAuley 1987), and pulse-rate discrimination can predict speech recognition in noise in cochlear-implant listeners (Zhou et al. 2019).

Listeners can discriminate small changes in the amplitude modulation rate of complex signals (e.g., pulse trains). Abilities are as good as 3 to 7% for rates of ~ 100 Hz, but performance declines rapidly for rates > 200–300 Hz

Correspondence to: Samira Anderson · Department of Hearing and Speech Sciences · University of Maryland · College Park, 20742, USA. email: sander22@umd.edu

(e.g., Carlyon and Deeks 2002; Carlyon et al. 2008; Kong et al. 2009; Macherey and Carlyon 2014). The physiological basis for this limitation on rate discrimination may arise from both peripheral and central sources (e.g., Ihlefeld et al. 2015; Johnson et al. 2021). Previous studies have demonstrated age-related declines in rate discrimination (DeVries et al. 2022; Gaskins et al. 2019) and in other temporal processing tasks, including gap detection (Snell 1997), duration discrimination (Fitzgibbons and Gordon-Salant 1995), and tempo discrimination (Fitzgibbons and Gordon-Salant 2001). Moreover, links between basic measures of auditory temporal processing and age-related deficits in speech recognition have been observed. Therefore, temporal processing deficits may underlie older adults' reported difficulties when understanding speech characterized by altered timing cues (i.e., time-compressed speech and reverberant speech), and the question remains whether these age-related deficits can be improved through targeted auditory training.

Previous studies have demonstrated the efficacy of training in improving temporal processing abilities. For example, training improves temporal rate discrimination thresholds in cochlear-implant listeners across a wide range of ages (Bissmeyer et al. 2020; Goldsworthy and Shannon 2014), and pitch discrimination can be improved through musical training in normal-hearing listeners (Bianchi et al. 2019; Micheyl et al. 2006). Furthermore, animal and human studies suggest that the brain retains some plasticity into older age. Age-related decreases in rat temporal coding and cortical firing synchrony can largely be reversed by training on a frequency discrimination auditory training paradigm (de Villers-Sidani et al. 2010). A cross-species study including mice and humans found that adaptive training on signal-in-noise detection in a closed-loop paradigm led to improvements in signal detection in both species and generalization to speech-in-noise performance in human listeners (Whitton et al. 2017). Older normal-hearing and hearing-impaired human listeners experience reductions in frequency-following response latencies, an indication of improved temporal precision, from training that adaptively increased or decreased both consonant-transition durations and auditory memory load (Anderson et al. 2013). Overall, these studies demonstrate the potential for training-related neuroplasticity in older listeners. It is, however, currently unknown whether auditory training targeted at auditory temporal processing can improve temporal rate discrimination ability in older normal-hearing listeners or hearing-impaired listeners and whether improvement in temporal rate discrimination generalizes to performance on speech-understanding measures.

Therefore, the current study was designed to (1) determine whether rate discrimination training can improve auditory temporal processing in older and younger listeners in both perceptual and neural responses; (2) determine the extent to which perceptual learning on rate discrimination generalizes to other temporal processing tasks and measures of speech understanding; and (3) investigate the neural and cognitive variables that are associated with training-related improvements in perception. Based on previous animal and human studies, we hypothesized that perceptual training would partially restore temporal processing abilities in older listeners. In addition, we hypothesized that neural responses to the trained pulse trains (auditory steady-state responses, ASSR) and cognitive ability would relate to changes in perception. Finally, given that previous studies have not shown significant effects of hearing loss on temporal processing tasks (Fitzgibbons and Gordon-Salant 1996; Roque et al. 2019a), we hypothesized a similar training benefit regardless of hearing status.

MATERIALS AND METHODS

Listeners

We recruited 301 listeners for a double-blind randomized controlled clinical trial and determined if they met the following age and audiometric criteria for these groups: young normal hearing (YNH, age 18–30 years), older normal hearing (ONH, age 65–85 years), and older hearing impaired (OHI, age 68–85 years). Normal hearing was defined as pure-tone thresholds ≤ 25 dB HL (re: ANSI 2018) from 125 to 4000 Hz in the right ear. Impaired hearing was defined by a high-frequency pure-tone average (average thresholds at 1, 2, and 4 kHz) > 30 dB HL and thresholds at 2 and 4 kHz < 70 dB HL (to ensure signal audibility). Hearing thresholds were required to be symmetrical (no interaural differences > 10 dB at any frequency) for all listeners, and there were no air-bone gaps > 10 dB at any frequency. Word recognition scores were > 70% for a single 25-word lists of the NU-6 test (Tillman and Carhart 1966) presented bilaterally at 75 dB HL in quiet. Middle ear function was normal bilaterally based on average values for tympanometric peak pressure, peak admittance, tympanometric width, and equivalent volume. Acoustic reflexes were present from 500 to 2000 Hz, elicited ipsilaterally and contralaterally. Finally, auditory brainstem responses (ABRs) were recorded, and Wave V latencies were < 6.8 ms with no interaural asymmetries > 0.2 ms. Additional criteria included the following: A passing score of ≥ 26 on the Montreal Cognitive Assessment (MoCA; Nasreddine et al. 2005), a negative history of neurological disease, a passing score on the Snellen vision screening chart ≤ 20/50 (Hetherington 1954), being a native English speaker, and earning a high school diploma. All procedures were reviewed and

approved by the Institutional Review Board (IRB) at the University of Maryland, College Park. Listeners provided informed consent and were monetarily compensated for their time.

The 125 listeners who met the study criteria were randomly assigned to one of two training groups: experimental and active control. Of these, 48 listeners did not complete the study. Seventeen listeners were dismissed due to: non-compliance with training (3), poor quality data (7), an adverse event (1), and excessive time delay associated with COVID-19 (6). Twenty-six listeners withdrew from the study due to medical or transportation issues. Eleven listeners were lost to follow-up. The final numbers of listeners in each training group were 40 experimental (14 YNH, 16 ONH, and 10 OHI; 30 females) and 37 active control (15 YNH, 14 ONH, and 8 OHI; 28 females). See Table 1 for additional demographic characteristics. Note that 1% of listener data (31 of 2618 measurements) are missing because of isolated issues during data collection or because of anomalous data that did not converge.

Pre- and Post-Testing

Both training groups were tested using the same battery of electrophysiological and behavioral measures prior to the onset and after completion of training. ASSRs were recorded to 100-, 200-, 300- and 400-Hz bandpass-filtered click trains, and behavioral pulse-rate discrimination was measured to the same stimuli. The behavioral test battery also included generalization measures: gap detection, gap duration discrimination, tempo discrimination, and several speech recognition measures. These measures will be described in more detail below. The duration of pre- and post-testing was approximately 1½ h for ASSR recording and 2 ½ h for behavioral testing.

Procedure

Listeners were seated in a double-walled sound-attenuating booth. The stimuli were presented to listeners through a single insert earphone (ER-2, Etymotic, Elk Grove Village, IL). Stimulus presentation and event timing were controlled from a laptop computer and a custom MATLAB script.

Perceptual and Neural Responses to Pulse Trains

Stimuli.

The stimuli were band-limited pulse trains (300-ms duration) having rates of 100, 200, 300, and 400 Hz. The pulses had a 1-kHz bandwidth arithmetically centered around 4 kHz, created using forward–backward Butterworth filters (5th order) (DeVries et al. 2022). Raised cosine Hanning windows with a 10-ms rise-fall time were applied to the stimuli to avoid filter-related onset and offset transients. The stimuli were presented monaurally to the right ear at 75 dBA for all electrophysiological and non-speech behavioral measures described below. For perceptual testing only, a low-frequency masking noise was mixed with the pulse train stimuli to eliminate the use of low-frequency distortion products to perform the task. Wideband masking noise was low-pass filtered using a 200-Hz cutoff with a − 3 dB/octave filter and presented at an overall level of 61 dB SPL.

Perceptual Rate Discrimination

Rate discrimination for each reference pulse rate was assessed by measuring pulse-rate difference limens (DLs) using a three-interval, two-alternative forced choice (3I-2AFC) procedure. Each rate (100, 200, 300, and 400 Hz) was tested with three blocks of 60 trials for a total of 720 trials across blocks. The order of reference pulse rate was randomized.

TABLE 1

Demographic characteristics of experimental and active control groups including sex, age, and pure-tone average (PTA). *YNH* young normal hearing, *ONH* older normal hearing, *OHI* older hearing impaired, *F* female, *M* mean, and *S.D.* standard deviation

Group	Experimental			Active Control		
	YNH (n = 14)	*ONH (n = 16)*	*OHI (n = 10)*	*YNH (n = 15)*	*ONH (n = 14)*	*OHI (n = 8)*
Sex	7 F	15 F	8 F	9 F	12 F	7 F
Age	21.1 (M) 2.2 (S.D)	69.9 (M) 4.0 (S.D)	74.0(M) 6.4 (S.D)	21.0 (M) 2.0 (S.D)	70.0 (M) 4.5 (S.D)	74.4(M) 6.4 (S.D)
HF PTA	5.5 (M) 2.6 (S.D)	14.1 (M) 3.0 (S.D.)	38.2 (M) 4.4 (S.D.)	6.2 (M) 2.9 (S.D)	13.9 (M) 4.3 (S.D.)	36.8 (M) 7.1 (S.D.)

Stimulus presentation was self-paced throughout the experiment. The listeners viewed a monitor that displayed four boxes. They were asked to click the box containing "Begin Trial" and then heard a sequence of three stimuli, with the presentation of each stimulus synchronized to a flash in the corresponding visual block in the sequence. The first stimulus was always the reference stimulus. The target stimulus with the higher rate was in the second or third interval, randomly chosen with a 50% *a priori* probability.

The listeners received the following instructions: "You will hear three brief sounds that sound like a buzz. The first one is the 'standard.' One of the other sounds has a slightly higher pitch that sounds different from the standard sound. Please select the sound, 2 or 3, that contains the higher pitch (or sounds different from the standard sound). If you are not sure, take a guess."

After each listener response, correct answer feedback was provided by flashing a green light at the box corresponding to the correct interval. A two-down-one-up adaptive procedure was employed to target 70.7% correct on the psychometric function (Levitt 1971). The initial rate difference between the reference and target stimulus was set at 40%. The maximum allowable rate difference was 40%. The rate difference after two correct responses or one incorrect response was changed by a factor of 2 (e.g., for a 100-Hz reference pulse rate, the 140-Hz target rate was changed to 120 Hz after correct answers on the first two trials). After three reversals of the adaptive procedure, the changes in step size were decreased by a factor of √2. The tracking ended after reaching the fixed number of 60 trials.

Analysis.

Perceptual responses were recorded in MATLAB. The pulse rate difference limen (DL) in percent for an individual adaptive track was found by calculating the geometric mean over all of the reversals in the adaptive procedure except the first two. The arithmetic mean of the second and third tracks was used to calculate the final DL for each listener and condition. The first track was omitted to decrease the possible impact of learning effects from the first track. The DLs were log-transformed due to a negative skew in the data prior to conducting the statistical analysis.

ASSR

Recording

The 300-ms pulse trains were presented at a rate of 1.66 Hz using the Intelligent Hearing Systems Continuous Acquisition Model (IHS SEPCAM, Miami, FL) through electromagnetically shielded insert ER-3 earphones (IHS) in an electrically shielded double-walled sound-attenuating booth. A three-electrode vertical montage was used (Cz active, right ear lobe reference, low forehead ground). Responses were recorded with a 10-kHz sampling rate and were filtered from 1 to 5 kHz on-line. A minimum of 1024 artifact-free sweeps (≤30 μV) were obtained for each condition. The listeners watched their movie of choice, muted with subtitles, to facilitate a relaxed but awake state.

Data Analysis

Responses were imported into MATLAB format using custom scripts and filtered from 50 to 500 Hz. An individual average response was created with the first 1000 artifact-free sweeps. Phase-locking factor (PLF) was assessed in a manner similar to that employed in previous studies (Jenkins et al. 2018; Roque et al. 2019b), using Morlet wavelets to decompose the signal from 50 to 500 Hz (Tallon-Baudry et al. 1996). The PLF value was then calculated for the response time region of 10–310 ms and around a 20-Hz frequency bin corresponding to the pulse rate of each condition. The PLFs were log-transformed due to a negative skew in the data.

Mid Generalization Measures

It was hypothesized that training to improve temporal processing on one measure (pulse-rate discrimination) would generalize to improvement on other non-speech auditory tasks that rely on accurate temporal processing. Gap detection, gap duration discrimination, and temporal interval discrimination measures were chosen because previous studies have demonstrated age-related deficits on these tasks (Fitzgibbons and Gordon-Salant 2001; Harris et al. 2010; Kumar 2011; Ross et al. 2010; Snell 1997).

Gap Detection

Gap detection thresholds were measured using target stimuli that were 250-ms wideband Gaussian noise bursts that had a silent gap temporally centered in the stimulus. Cosine squared windows with a 1-ms rise-fall time were applied to the stimuli to avoid transients.

A 3I-2AFC procedure was used. The first interval was the standard, with no gap. The target stimulus with the silent gap was in the second or third interval, randomly chosen with a 50 % *a priori* probability.

The listeners received the following instructions, "This is the 'standard' and is a continuous noise. One of the other noise bursts, 2 or 3, has a very brief pause or interruption that sounds different from the standard noise burst. Please select the noise burst, 2 or 3, that contains the brief pause (or sounds different from the standard noise burst). If you are not sure, take a guess."

After each listener response, correct answer feedback was provided. Then the gap duration was adapted according to the two-down-one-up adaptive rule,

targeting 70.7% correct discrimination. The initial gap duration was 25 ms. The maximum gap duration was 100 ms, and the minimum gap duration was 1 ms. The initial step size in the adaptive procedure was 5 ms. After two reversals, the step size was changed to 1 ms. The adaptive track continued until there were eight reversals. Threshold was defined as the arithmetic mean of the last six reversals. Three adaptive tracks were conducted. The arithmetic mean of the second and third tracks was used to calculate the gap detection threshold for each listener.

Gap Duration Discrimination

Gap duration discrimination was measured using 250-ms 1000-Hz tone pairs separated by a silent interval (Fitzgibbons and Gordon-Salant 1994). Cosine squared windows with a 5-ms rise-fall time were applied to the stimuli to avoid transients.

The listener received the following instruction: "Please select the tone pair, 2 or 3, that contains the longer silent interval (or sounds different from the standard tone pair). If you are not sure, take a guess."

After each listener response, correct answer feedback was provided. Then the gap duration was adapted according to the two-down-one-up adaptive rule. The initial gap duration for the target was 350 ms (i.e., 40% larger than the reference gap of 250 ms). The maximum gap duration was 450 ms, and the minimum gap duration was 252 ms. The initial step size in the adaptive procedure was 10 ms. After two reversals, the step size was reduced to 2 ms. The adaptive track continued until there were eight reversals. The relative gap duration discrimination DL in percent (based on the 250-ms reference) was calculated from the arithmetic mean of the last six reversals. Three adaptive tracks were measured. The arithmetic mean of the second and third tracks was used to calculate the gap duration discrimination DL for each listener.

Tempo (Rhythm) Discrimination

Discrimination DLs were measured for inter-onset intervals (IOIs) in isochronous sequences of five brief 50-ms 1000-Hz tones (see Fitzgibbons and Gordon-Salant 2001). The IOI is defined as the duration between the onset of one tone in the sequence and the onset of the subsequent tone. Cosine squared windows with a 5-ms rise-fall time were applied to the stimuli to avoid transients.

A 3I-2AFC procedure was used. The reference intervals had a fixed IOI, either 100 ms (fast reference) or 600 ms (slow reference). The target stimulus with the relatively slower tone sequence was in the second or third interval, randomly chosen with a 50% *a priori* probability.

The listeners received the following instructions: "You will hear three sequences of 5 brief tones. The first sequence is the 'standard.' One of the other sequences, 2 or 3, sounds slower than the standard sequence. Please select the tone sequence, 2 or 3, that is a slower sequence (or sounds different from the standard sequence). If you are not sure, take a guess."

After each listener response, correct answer feedback was provided. Then, the IOI was adapted according to the two-down-one-up adaptive rule. The starting target IOI was 150 ms for the 100-ms reference IOI and 700 ms for the 600-ms reference IOI. The maximum target IOI was 200 ms, and the minimum target IOI was 101 ms for the 100-ms reference IOI; the maximum target IOI was 800 ms, and the minimum target IOI was 601 ms for the 600-ms reference IOI. The initial step size in the adaptive procedure was 10 ms. After two reversals, the step size decreased to 2 ms. The adaptive track continued until there were eight reversals. The DL for each IOI was calculated from the arithmetic mean of the last six reversals of each track. Three adaptive tracks were conducted for each reference IOI (i.e., there were six separate adaptive tracks). The arithmetic mean of the second and third tracks was used to calculate the relative IOI DL in percent (based on either the 100-ms or 600-ms IOI reference) for each listener.

Sentence Recognition

Sentence recognition in quiet was measured for sentences from the IEEE corpus (IEEE 1969) in five conditions: normal rate with no reverberation (i.e., clean speech), two levels of time compression (TC; 40% and 60%), and two levels of reverberation (REV; 0.6 s and 1.2 s). There were 10 sentences presented in each condition. Each sentence was preceded by a carrier phrase, "Number 1," "Number 2," etc. Listeners were instructed to repeat the sentence they heard. The experimenter scored which of the five keywords in each sentence were repeated correctly, and the percent correct keywords words out of 50 was calculated for each condition.

Training

Experimental

Listeners received in-lab perceptual rate discrimination training for two rates, 100 and 300 Hz, using a procedure similar to that described above for rate discrimination assessment. The training was blocked by rate, with four blocks of 60 trials for each rate, for a total of 480 trials. Correct-answer feedback was provided after each trial throughout the training sessions. Nine sessions of this training took place in the sound-attenuating booth over the course of 2 to 3 weeks. The duration of

training per session was 45 to 60 min, depending on the participant's speed of response.

Active Control

Listeners received in-lab training on tone-in-noise detection, using a 3I-2AFC procedure. A notched-noise paradigm and simultaneous masking were used to measure filter bandwidths (Desloge et al. 2012), using a 300-ms 1-kHz stimulus tone and a 500-ms white Gaussian noise (0.25–6 kHz). The target tone was temporally centered in the noise. Cosine squared windows with a 10-ms rise-fall time were applied to the noise and target tones to avoid transients. The noise level was fixed at 75 dBA, and the tone level varied adaptively to determine threshold in three notch bandwidths: 90, 120, and 150 Hz.

After each listener response, correct answer feedback was provided. Then, the tone level was adapted according to the two-down-one-up adaptive rule. The initial and maximum target tone level was 75 dBA, and the minimum target tone level was −20 dBA. The initial step size in the adaptive procedure was 3 dB. After three reversals, the step size decreased to 0.5 dB. Each of the three notch bandwidth conditions was presented in four blocks, with 40 trials per block, for a total of 480 trials; therefore, the procedure had the same number of trials when compared to the pulse-rate discrimination training, except that the task was different. Nine sessions of this training took place in the sound-attenuating booth over the course of 2 to 3 weeks. The masked threshold in dB for an individual adaptive track was found by calculating the arithmetic mean over the last four reversals in the adaptive track. The arithmetic mean of the second and third tracks was used to calculate the final masked threshold for each listener and condition. The duration of training per session was 45 to 60 min, depending on the participant's speed of response.

Cognitive Testing

Assessments from the National Institutes of Health (NIH) Cognition Toolbox (Weintraub et al. 2013) were used to determine if particular cognitive skills predicted perceptual training benefits. These tests included the List Sorting Working Memory Test, the Flanker Inhibitory Control and Attention Test, the Pattern Comparison Processing Speed Test, and the Dimensional Card Sort Test. The tests were administered using the NIH toolbox application on an Apple iPad (Apple, Inc., Cupertino, CA). The Uncorrected Standard scores (not age-corrected, mean = 100, SD = 15) were downloaded from the application.

Statistical Analysis

The data were analyzed by JASP (v.14.1, 2020) statistical software.

Pulse Rate Discrimination Improvement and Near Generalization

Separate four-way mixed analyses of variance (ANOVA) were conducted to evaluate the effects of training on perception and neural representation of the pulse trains, comparing pre-test and post-test measures. There were two between-subjects factors (listener group: YNH, ONH, OHI; training group: experimental, active control) and two within-subjects factors (rate: 100, 200, 300, 400 Hz; session: pre-test, post-test). For perceptual measurements, the dependent variable was pulse-rate DL. For the electrophysiological measurements, the dependent variable was the ASSR PLF. In addition, multivariate ANOVAs (MANOVAs) were conducted to assess differences between post-test rate discrimination in the older listeners with pre-testing rate discrimination in the YNH listeners to determine if training restores temporal processing deficits in the older listeners across rates. Bonferroni-corrected two-way ANOVAs, independent-samples *t* tests (assuming equal variance), and paired-samples *t* tests were used to perform *post hoc* analyses when significant main effects or interactions were observed.

Mid Generalization

Separate mixed ANOVAs were conducted to evaluate mid generalization to the other temporal processing measures using the same two between-subjects factors (listener group and training group) as for the pulse trains and the same within-subjects factor (session). For gap detection, the gap detection threshold was submitted to a three-way mixed ANOVA. For gap duration discrimination, the gap duration discrimination DL was submitted to a three-way mixed ANOVA. For tempo discrimination, the IOI discrimination DL was submitted to a four-way mixed ANOVA, because there was an additional within-subjects factor (reference IOI: 100, 600 ms).

Far Generalization

A mixed ANOVA was conducted to evaluate generalization to sentence recognition measures using the same between-subjects factors. The within-subjects factors were condition [clean speech, two levels of time compression (TC40, TC60), and two levels of reverberation
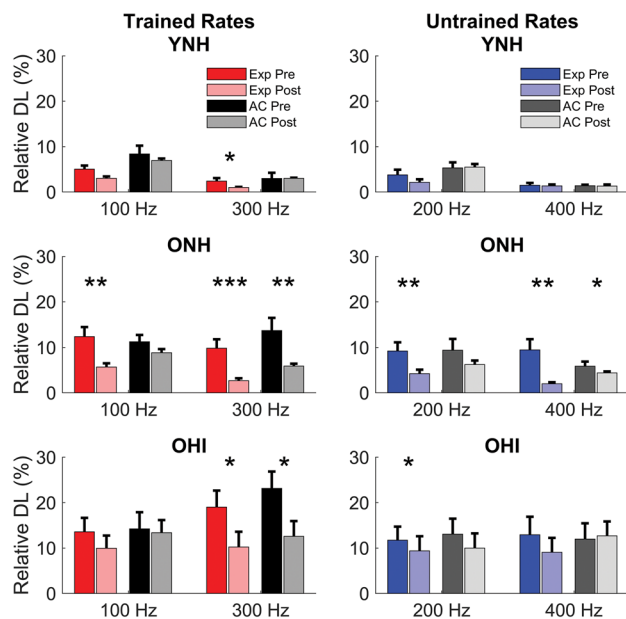
Fig. 1 Average rate discrimination difference limens (DLs) are displayed for 100 and 300 Hz (left panels) and 200 and 400 Hz (right panels) in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners who completed nine sessions of rate discrimination training (experimental group, EXP) or tone-in-noise detection training (active control group, AC). There were significant improvements in performance (smaller DLs) in the EXP group that were not observed in the AC group. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Error bars = ±1 S.E

(0.6-s RV, 1.2-s RV)] and session (pre-test, post-test). The dependent variable was the sentence recognition score. The percent correct scores were transformed using the rationalized arcsine unit (RAU) transform (Studebaker 1985) to avoid violation of the homogeneity of variance assumption required for an ANOVA.

Relationships Among Perceptual Temporal Processing Measures

To test our hypothesis that the behavioral measures share a common temporal processing mechanism, correlations were conducted to evaluate relationships among the pulse-rate discrimination DLs, gap detection thresholds, gap duration discrimination DLs, temporal interval discrimination DLs, and RAU-transformed sentence recognition scores for temporally altered stimuli. Spearman's rho was calculated because not all of the data were normally distributed. The False Discovery Rate was used to correct for multiple comparisons (Benjamini and Hochberg 1995).

Performance Predictors

A step-wise multiple linear regression was conducted to identify the potential factors that contributed to changes

in pulse-rate discrimination performance for 100- and 300-Hz rates in the experimental group. The dependent variable was the average change in rate DL (post-test–pre-test) for 100- and 300-Hz reference rates. The Pattern Comparison Processing Speed Test scores were included as an independent variable due to its relationship to pre-test DLs (Gaskins et al. 2019). Additional cognitive measures included in the analyses were the List Sorting Working Memory Test, the Flanker Inhibitory Control and Attention Test, the Pattern Comparison Processing Speed Test, and the Dimensional Card Sort. The PTA in the right ear (500 to 4000 Hz) was also included to determine the contribution of audibility to performance. A log transform was applied to the skewed PTA distribution. Finally, to determine the contributions of subcortical neural processing to performance changes, the pre-test PLF and change in PLF averaged for 100- and 300-Hz rates were included.

RESULTS

Rate Discrimination

Figure 1 displays pre- and post-test performance for the 100- to 400-Hz reference rates in YNH, ONH, and OHI listeners. The mixed ANOVA showed a main effect of session ($F_{(1, 69)} = 52.62$, $P < 0.001$, $\eta^2 = 0.43$), such that DLs were lower (better) at the post-test compared to the pre-test. There was a significant training group × session interaction ($F_{(1, 69)} = 5.48$, $P = 0.005$, $\eta^2 = 0.15$), which was driven by the larger decrease in DL from post- to pre-test sessions in the experimental compared to the active control group ($T_{(73)} = 2.96$, $P = 0.004$). Furthermore, improvement occurred for the untrained rates in the experimental group but not in the active control group. Post hoc testing revealed that the experimental group exhibited significant effects of session for both the trained rates [two-way mixed ANOVA with factors session and rate (100 and 300 Hz); $F_{(1, 39)} = 41.96$, $P < 0.001$, $\eta^2 = 0.53$] and the untrained rates (200 and 400 Hz; $F_{(1, 39)} = 21.39$, $P < 0.001$, $\eta^2 = 0.35$), but the active control group showed a significant effect of session for only the trained rates (100 and 300 Hz; $F_{(1, 35)} = 17.89$, $P < 0.001$, $\eta^2 = 0.34$) and not the untrained rates (200 and 400 Hz; $F_{(1, 35)} = 3.41$, $P = 0.073$, $\eta^2 = 0.09$).

The training group × listener group × session interaction was not significant ($F_{(2, 69)} = 0.53$, $P = 0.592$, $\eta^2 = 0.02$), suggesting that training effects on rate discrimination did not differ significantly by listener group (YNH, ONH, OHI). In addition, there was no listener group × session interaction in either training group (all P values > 0.05).

A MANOVA was then used to compare the post-test DLs in the ONH and OHI listeners to the pre-test DLs in the YNH listeners in the experimental training group
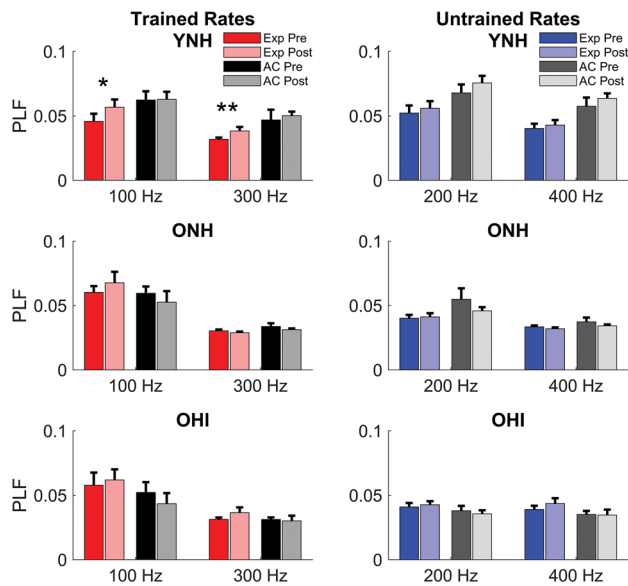
Fig. 2 Average pre- and post-training phase-locking factor (PLF) are displayed for 100 and 300 Hz (left panels) and 200 and 400 Hz (right panels) in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners in the experimental (EXP) and active control (AC) groups. There were significant increases in PLF in the training group, especially in the YNH listeners that were not observed in the active control group. *$P < 0.05$, **$P < 0.01$. Error bars $= \pm$ S.E

for the four different rates. At the pre-test, there was a main effect of listener group ($F_{(2, 36)} = 14.28$, $P < 0.001$, $\eta^2 = 0.44$); post hoc t tests showed that both groups of older listeners had higher (poorer) DLs than the YNH listeners ($P < 0.001$), but the older groups did not differ from each other ($P = 1$). A comparison of the pre-test DLs in the YNH listeners with the post-test DLs in ONH and OHI listeners showed a main effect of listener group ($F_{(2, 37)} = 8.29$, $P = 0.001$, $\eta^2 = 0.31$); post hoc t tests showed that the DLs of ONH listeners at post-test did not differ from those of YNH listeners at pre-test ($P = 0.426$). However, the OHI listeners had higher DLs than both the ONH listeners at post-test ($P = 0.025$) and the YNH listeners at pre-test ($P < .001$). There was also a rate × listener group interaction ($F_{(6, 111)} = 4.68$, $P < 0.001$, $\eta^2 = 0.13$). At the 100-Hz rate, there was no main effect of listener group ($P = 0.18$). At the 200-, 300-, and 400-Hz rates, there was no significant difference between the YNH and ONH listeners ($P > 0.05$ for all three comparisons), but the OHI listeners had higher DLs than the YNH listeners ($P < 0.05$ for all three comparisons). Given that pre-test DL differences existed between the ONH and YNH listeners ($P < 0.001$), these results demonstrate that training on rate discrimination at least partially restored temporal processing abilities on this measure in ONH listeners.
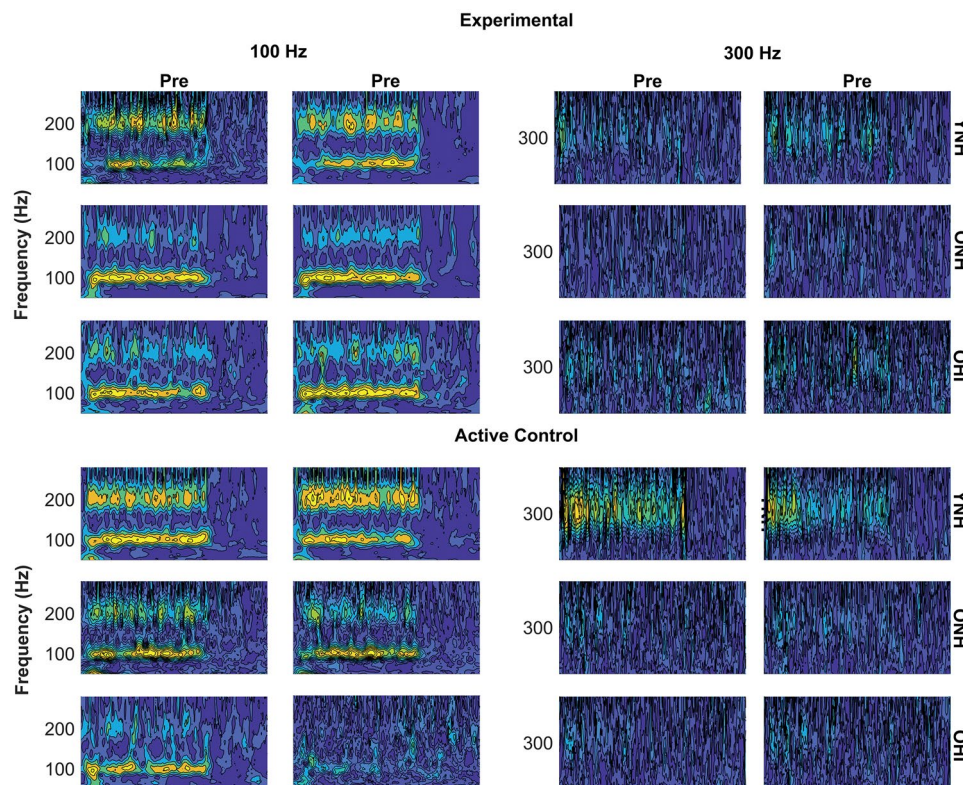

Fig. 3 Pre- and post-training phase-locking factor (PLF) for 100- and 300-Hz rates is displayed in the time–frequency domain for young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners in the experimental (top three panels) and active control (bottom three panels) groups
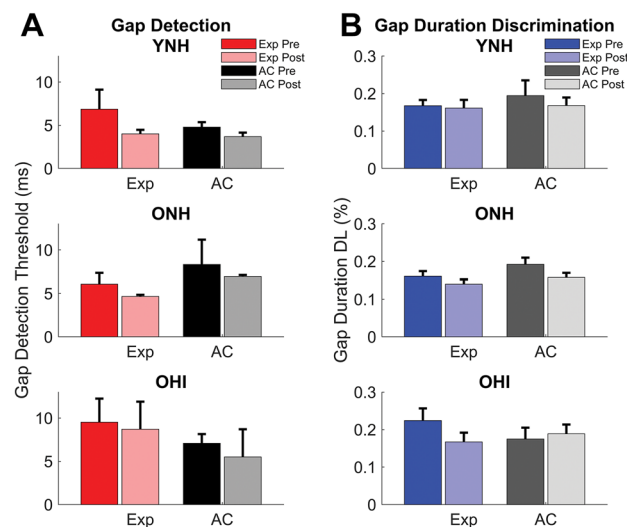
Fig. 4 Average pre- and post-training gap detection thresholds (**A**) and gap duration DLs (**B**) are displayed for young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners in the experimental (EXP) and active control (AC) groups. No changes in performance were noted from pre-test to post-test in any group. Error bars = ±S.E

ASSR

Figure 2 displays pre- and post-training PLF values for all four rates, and Fig. 3 displays pre- and post-training phase-locking spectra for the 100- and 300-Hz rates. The mixed ANOVA showed no main effect of session ($F_{(1, 69)} = 0.31$, $P = 0.582$, $\eta^2 = 0.00$), but there was a training group × session interaction ($F_{(1, 69)} = 4.61$, $P = 0.035$, $\eta^2 = 0.06$), driven by a significant increase in PLF in the experimental group (post hoc two-way repeated-measures ANOVA with factors rate and session; $F_{(1, 37)} = 4.99$, $P = 0.032$, $\eta^2 = 0.12$) that was not observed in the active control group ($F_{(1, 36)} = 0.87$, $P = 0.357$, $\eta^2 = 0.02$). The training group × listener group × session interaction was not significant ($F_{(2, 69)} = 0.08$, $P = 0.927$, $\eta^2 = 0.00$), suggesting that training effects on PLF did not differ by listener group.

Post hoc testing in the experimental group showed significant effects of session for the trained rates (two-way mixed ANOVA with factors session and rate (100 and 300 Hz); $F_{(1, 37)} = 5.44$, $P = 0.025$, $\eta^2 = 0.13$), but not the untrained rates (200 and 400 Hz; $F_{(1, 37)} = 2.88$, $P = 0.098$, $\eta^2 = 0.07$). The active control group did not show significant effects of session for either the trained rates (100 and 300 Hz; $F_{(1, 36)} = 1.66$, $P = 0.21$, $\eta^2 = 0.04$) or the untrained rates (200 and 400 Hz; $F_{(1, 36)} = 0.29$, $P = 0.594$, $\eta^2 = 0.01$).

The mixed ANOVA showed a significant effect of rate ($F_{(3, 216)} = 48.85$, $P < 0.001$, $\eta^2 = 0.40$) because there were higher PLFs for the lower rates compared to the higher rates. There was a main effect of listener group

($F_{(2, 72)} = 8.24$, $P < 0.001$, $\eta^2 = 0.19$). Post hoc t tests showed that the YNH group had higher PLFs than the ONH group ($P = 0.002$) and the OHI group ($P = 0.005$), but the group difference was not significant between the ONH and OHI groups ($P = 1.00$). There was a significant listener group × rate interaction ($F_{(6, 216)} = 6.42$, $P < 0.001$, $\eta^2 = 0.15$). In separate *post hoc* mixed ANOVAs with factors listener group and training group, there was no significant listener group difference for the 100-Hz PLF ($F_{(2, 73)} = 0.85$, $P = 0.431$, $\eta^2 = 0.02$), but there were significant group differences for the 200-Hz PLF ($F_{(2, 73)} = 13.65$, $P < 0.001$, $\eta^2 = 0.27$), 300-Hz PLF ($F_{(2, 72)} = 6.86$, $P = 0.002$, $\eta^2 = 0.16$), and 400-Hz PLF ($F_{(2, 72)} = 11.11$, $P < 0.001$, $\eta^2 = 0.24$). At the 200-, 300-, and 400-Hz rates, there was no significant difference between the ONH and OHI listeners ($P > 0.05$ for all three comparisons), but the YNH listeners had higher PLF than the either group of older listeners ($P < 0.05$ for all six comparisons).

Mid Generalization–Temporal Processing

Gap detection

Figure 4A displays pre- and post-training data for the gap detection tasks. The mixed ANOVA showed that there was a main effect of session ($F_{(1, 70)} = 5.41$, $P = 0.023$, $\eta^2 = 0.01$), but there was no training group × session interaction ($F_{(1, 70)} = 0.09$, $P = 0.77$, $\eta^2 = 0.00$). There was no main effect of listener group ($F_{(2, 70)} = 1.51$, $P = 0.29$, $\eta^2 = 0.03$).

Gap Duration Discrimination

Figure 4B displays pre- and post-training data for the gap duration discrimination tasks. The mixed ANOVA showed a main effect of session ($F_{(1, 69)} = 7.00$, $P = 0.01$, $\eta^2 = 0.01$), but there was no training group × session interaction ($F_{(1, 69)} = 0.75$, $P = 0.56$, $\eta^2 = 0.00$). The was no main effect of listener group ($F_{(2, 69)} = 0.53$, $P = 0.59$, $\eta^2 = 0.01$).

Tempo Discrimination

Figure 5 displays pre- and post-training data for relative DLs as a function of 100- and 600-ms IOIs. The mixed ANOVA showed neither a main effect of session ($F_{(1, 66)} = 1.10$, $P = 0.301$, $\eta^2 = 0.02$) nor a training group × session interaction ($F_{(1, 66)} = 0.02$, $P = 0.893$, $\eta^2 = 0.00$). The was no main effect of listener group ($F_{(2, 66)} = 0.36$, $P = 0.696$, $\eta^2 = 0.01$). There was a main effect of IOI ($F_{(1, 66)} = 23.66$, $P < 0.001$, $\eta^2 = 0.23$); the relative DLs were smaller for the 600-ms IOI than for the 100-ms IOI. No other interactions were significant.

Far Generalization–Speech Recognition

Figure 6 displays pre- and post-training speech recognition data in experimental and active control groups,
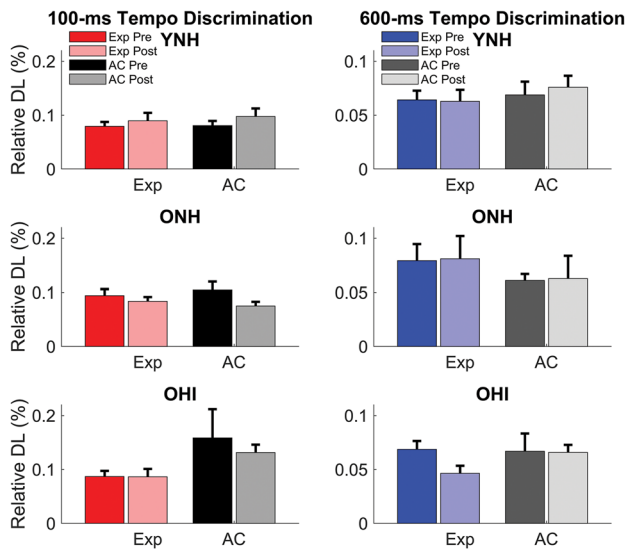
FIG. 5 Average relative difference limens (DL) are displayed as a function of 100- and 600-ms inter-onset intervals (IOIs) obtained in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners in the experimental (EXP) and active control (AC) groups. No changes in performance were noted in any group. Error bars = ±S.E

FIG. 6 Average percent correct sentence recognition scores are displayed for pre- and post-training for clean (undistorted) speech, 40 % time-compressed speech (40 % TC), 60 % time-compressed speech (60 % TC), and 0.6-s and 1.2-s reverberation time (0.6s REV, 1.2s REV, respectively) in young normal-hearing (YNH), older normal-hearing (ONH), and older hearing-impaired (OHI) listeners in the experimental (EXP) and active control (AC) groups. No changes in performance were noted in any listener group. Error bars = ±S.E

respectively. The mixed ANOVA showed neither a main effect of session ($F_{(1, 72)} = 1.10$, $P = 0.299$, $\eta^2 = 0.00$) nor a training group × session interaction ($F_{(1, 71)} = 0.77$, $P = 0.381$, $\eta^2 = 0.00$), suggesting that sentence recognition did not improve across groups. There was a main effect of listener group ($F_{(2, 71)} = 60.03$, $P < 0.001$, $\eta^2 = 0.29$). Post hoc testing showed that the OHI listeners had poorer overall performance than the YNH and ONH listeners ($P < 0.001$ for both), and ONH listeners had poorer overall performance than the YNH listeners ($P = 0.008$). There was a significant measure × listener group interaction ($F_{(8, 284)} = 44.82$, $P < 0.001$, $\eta^2 = 0.08$). The OHI listeners had lower scores than the YNH and ONH listeners on all scores (all $P < 0.001$) except clean speech ($P > 0.05$). The ONH listeners had lower scores than the YNH listeners for the TC60 condition ($P < 0.001$) but not for any other measure. Removal of the outlier in the OHI experimental group did not change these results.

Correlations Among Perceptual Measures

Correlations were calculated among the pre-test pulse-rate discrimination DLs and the other non-speech and speech temporal processing measures. Table 2 displays the *R* values for these correlations. Correlations were generally high and significant within the groups of measures (near, mid, or far generalization). There were also significant correlations across groups of measures. The pulse-rate discrimination DLs were significantly correlated with most of these measures, including gap detection
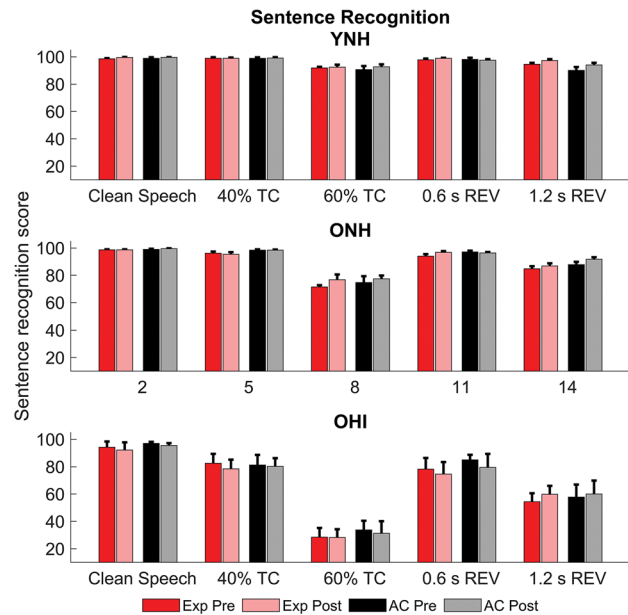
thresholds, temporal interval discrimination thresholds, and all temporally altered sentence recognition measures. The correlations between pulse-rate discrimination DLs and the mid generalization measures, as well as temporally altered sentence recognition measures and the mid generalization measures tended to be the smallest, often lacking significant correlations.

Correlations were also calculated for the improvements in measures (post-test minus pre-test change) for the 300-Hz DL and the PLF (the rate at which the greatest changes were observed across groups) and measures that were related to pre-test 300-Hz DLs in Table 2 (non-speech measures: gap detection and 100-ms tempo discrimination; speech measures: 60 % TC, 0.6-s RV, and 1.2-s RV). The analysis was restricted to the training group listeners who had scores less than 100 % on the pre-test measures ($N = 39$), in other words, the listeners that had the potential to improve. The sentence recognition score in the 0.6-s RV condition was negatively correlated with the 300-Hz DL ($\rho = -0.434$, $P = 0.008$) and was positively correlated with the 300-Hz PLF ($\rho = 0.394$, $P = 0.018$) after correcting for multiple comparisons using the False Discovery Rate procedure (Benjamini and Hochberg 1995). Figure 7 displays scatter plots for these relationships. No other correlations were significant.

TABLE 2

Correlations among pulse-rate discrimination difference limens and other non-speech and speech temporal processing measures. Spearman correlation ρ values are displayed for the following variables: difference limens (DLs) for discrimination of 100-, 200-, 300-, and 400-Hz pulse trains, gap detection thresholds (GAP DET), gap duration detection thresholds (GAP DUR), temporal interval discrimination thresholds for 100 and 600 ms (TEMPO 100 and TEMPO 600), and sentence recognition presented in 40% and 60% time compression ratio conditions (40% TC and 60% TC) and in 0.6-s reverberation time and 1.2-s reverberation time conditions (0.6-s REV and 1.2-s REV). Boldface font indicates values that are significant at an alpha level of 0.05 or better. *$P<0.05$, **$P<0.01$, ***$P<0.001$

|	DL 100	DL 200	DL 300	DL 400	GAP DET	GAP DUR	TEMPO 100	TEMPO 600	40% TC	60% TC	0.6-s RV	1.2-s RV
Pulse rate DLs (Near Generalization)												
DL 100												
DL 200	**.671*****											
DL 300	**.607*****	**.592*****										
DL 400	**.588*****	**.616*****	**.799*****									
Temporal Processing (Mid Generalization)												
GAP DET	**.425*****	**.433***	**.522*****	**.463*****								
GAP DUR	.196	.137	**.233***	**.231***	**.313****							
TEMPO 100	**.373****	**.343****	**.289***	**.358****	**.258***	**.341****						
TEMPO 600	**.274***	**.456*****	**.250***	**.289***	**.271***	**.297***	**.487*****					
Sentence Recognition (Far Generalization)												
40% TC	**-.331****	**-.361****	**-.587*****	**-.509*****	**-.458*****	-.214	-.141	-.201				
60% TC	**-.392*****	**-.411*****	**-.729*****	**-.611*****	**-.433****	**-.276***	-.163	-.171	**.613*****			
0.6-s RV	**-.363****	**-.363****	**-.610*****	**-.532*****	**-.298****	**-.251***	-.100	-.165	**.717*****	**.662*****		
1.2-s RV	**-.346****	**-.433*****	**-.581*****	**-.496*****	**-.267***	-.332	-.174	-.221	**.573*****	**.681*****	**.656*****	
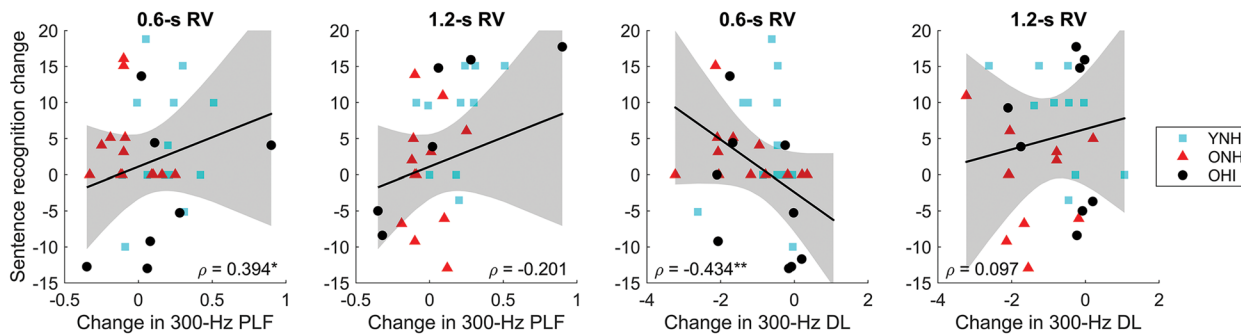
Fig. 7 Scatter plots demonstrating relationships among training-related changes in relative phase-locking factor (PLF) (left two panels) and differences limens (DLs right two panels) to the 300-Hz rate and sentence recognition in young normal-hearing (YNH, cyan squares), older normal-hearing (ONH, red triangles), and older hearing-impaired (OHI, black circles) listeners. Improvement in 300-Hz PLF and 300-Hz DLs was related to improvement in sentence recognition in the 0.6-s reverberation condition. *$P<0.05$, **$P<0.01$

Factors Contributing to Training-Induced Changes in Pulse Rate Discrimination

To identify the potential factors that contributed to changes in pulse-rate discrimination performance for 100- and 300-Hz rates in the experimental group, a step-wise multiple linear regression was conducted. The multiple linear regression collinearity diagnostics showed satisfactory tolerance (lowest 0.30) and variance inflation factor (highest 2.61) values, suggesting that the predictor variables were not highly correlated. One significant regression equation was returned; the Flanker score (attention) significantly predicted change in rate discrimination ($F_{(1, 35)} = 13.53$, $P<0.001$) with $R^2 = 0.29$. None of the other variables contributed significantly to the change in rate discrimination. This model is summarized in Table 3.

DISCUSSION

The overarching goal of this investigation was to determine the effect of rate discrimination training on temporal processing in older and younger listeners. The results showed training-related improvements in temporal rate discrimination DLs and phase locking. A larger degree of improvement in temporal rate discrimination DLs occurred for the experimental group compared to the active control group, suggesting perceptual learning for the experimental group and some procedural learning for both groups (Koziol and Budding 2012). The training × listening group interactions were not significant; therefore, training effects were not limited to a specific listener group (YNH, ONH, and OHI). Improved rate discrimination and phase locking related to higher sentence recognition scores in the condition with the shorter reverberation time (0.6 s), suggesting that generalization of training effects may potentially extend to real-world listening situations.

Effects of Aging and Hearing Loss on Training Benefits

Results showed training benefits across listener groups for both rate discrimination and phase locking. There was no significant listener × training group interaction, suggesting that training effects did not differ by age or hearing status (Fig. 1). These results appear to contrast with those of Sabin et al. (2013), who found improvement in spectrotemporal modulation thresholds in young listeners

---

TABLE 3

Summary of "stepwise" regression analysis for variables contributing to change in rate discrimination. Unstandardized (B) and standard error (S.E.) coefficients and standardized (β) coefficients in a model were automatically generated by evaluating the significance of each variable's contribution to the average change in 100- and 300-Hz rate discrimination. Only one model was generated, in which the Flanker score predicts significant variance in rate discrimination change. All other variables were excluded from the model (working memory, speed of processing, dimension card sort, pure-tone average, pre-training phase-locking factor, and change in phase-locking factor)

Variable	R^2	B	S.E	β	95% C.I. for B	p value
Model 1	0.35					<0.001
Flanker		0.26	0.06	0.59	0.14–0.39	<0.001

but not in older listeners. The older listeners in the Sabin et al. study had mild to moderate hearing loss (thresholds ranging from 15 to 70 dB HL from 0.5 to 4 kHz), which may have affected their ability to benefit from training on spectrotemporal modulation due to decreased spectral resolution associated with hearing loss. Our study focused on a measure of temporal processing, an acoustic dimension that is less affected by hearing loss (Fitzgibbons and Gordon-Salant 1996), and we did not find effects of hearing loss on pre-training rate discrimination. Bianchi et al. (2019) found that the extent of musical training benefit on F_0 discrimination was limited by the extent of hearing loss. Therefore, we examined the relationship between high-frequency PTA and change in the 100-Hz DL, and found a modest correlation among the older listeners ($\rho = 4.20$, $P = 0.03$), suggesting that the training benefit decreased with increased hearing threshold.

The improvement in behavioral temporal processing with training partially reduced age-related deficits. The ONH listeners' post-training DLs decreased to levels that approached those of the YNH listeners' pre-training DLs. These results are consistent with animal models of neuroplasticity in auditory aging that have shown that perceptual training can reduce or eliminate age-related deficits in temporal processing (de Villers-Sidani et al. 2010). Although there was a modest decrease in DLs in the OHI listeners, their post-training DLs remained significantly different from the pre-training DLs of the YNH listeners.

We did not find a similar reduction of the age-related deficit in neural temporal processing. Significant group differences in the PLF (at rates > 100 Hz) at the pre-test session persisted at the post-test session. Our selection of rates was motivated to match testing between rate discrimination and the ASSR, and rates of 100–400 Hz arise from low to high brainstem sources (Herdman et al. 2002). The de Villers-Sidani et al. (2010) study found changes in temporal precision in the rat auditory cortex, and therefore, it is possible that a selection of a lower frequency rate (40 Hz or lower) that represents cortical sources would have shown an improvement in temporal precision.

Generalization

Generalization was evaluated by comparing pre- and post-test performance on untrained measures and by determining the extent to which changes in rate discrimination and phase locking correlated with changes in untrained measures. We found that relative DLs were lower for the untrained rates (200 and 400 Hz), but we did not find significant training-related changes in any of the other measures. We also found that improvements in rate discrimination and phase locking were related to improvement in recognition of reverberant sentences.

Rate Discrimination

The improvement in pulse-rate discrimination and the specificity of training effects for trained and untrained rates are consistent with previous studies that have demonstrated near generalization effects that were specific to the training task. For example, Fitzgerald and Wright (2011) trained YNH listeners to detect sinusoidal amplitude modulations and found that training generalized to untrained modulation rates but not to untrained carrier spectra or to rate discrimination using the trained rate and carrier spectrum. Similarly, YNH listeners trained to detect depth of either spectral, temporal, or spectrotemporal modulations did not show generalization of training effects to untrained modulations (Sabin et al. 2012).

The Fitzgerald and Wright (2011) and Sabin et al. (2012) studies only trained YNH listeners, but different learning patterns might be found in older listeners. For example, Sabin et al. (2013) found that training to detect spectral modulations generalized to an untrained spectral modulation frequency in ONH listeners but not YNH listeners. In the current study, there were significant improvements in rate discrimination for the untrained frequencies (200 and 400 Hz) in the ONH and OHI listeners but not in the YNH listeners (see Fig. 1). The reason for the lack of generalization in YNH listeners is currently unknown but may be due to the fact that performance in the YNH listener group was excellent at the pre-test across rates, thus limiting capacity for improvement.

ASSR

No near generalization was found for untrained rates (200 and 400 Hz). The absence of generalization suggests two points: (1) the lack of increased PLF to 200- and 400-Hz rates suggests that the increase to 100- and 300-Hz rates is due to effects of training rather than to the effects of repeated testing and (2) cortical neural processes may underlie generalization in perceptual performance, but the ASSR recordings in the current study targeted subcortical processing.

Generalization to other temporal processing and sentence recognition measures

When comparing pre- and post-test performance, no mid or far generalization was observed for any of the other temporal processing (non-rate discrimination) or sentence recognition measures. This is in contrast to other training studies employing temporally based training that have observed generalization to speech stimuli. For example, Lakshminarayanan and Tallal (2007) trained YNH listeners' perception of frequency-modulated (FM) sweeps that varied in direction of change, duration of FM sweep, and inter-stimulus interval between sweeps. They found that this training led to enhanced discrimination between syllables that differed in the onset of the second formant (/ba/ vs /da/), transition duration (/ba/ vs /wa/), and

silence duration (/sa/ vs /sta/). The transfer of temporally based training has also been observed in older listeners. Fostick et al. (2020) trained older listeners with normal to mild hearing loss levels on a spatial temporal order judgment task and found that improvement on this task generalized to recognition of word stimuli presented in quiet, narrowband noise, and wideband noise. They did not observe similar generalization for training on an intensity discrimination task.

In the current study, significant relationships among the pre-test measurements DLs and all of the temporally distorted sentence recognition measures, consistent with previous studies that have found that performance on non-speech temporal processing measures predicts sentence recognition in challenging listening environments (Gordon-Salant and Fitzgibbons 1993; Zhou et al. 2019). Although there was no significant effect of training on sentence recognition overall, there were relationships between training-related changes on the key measures of rate discrimination and phase locking and improvements in speech recognition performance from pre- to post-testing, specifically the 0.6-s REV condition (Fig. 7). These results and those of Fostick et al. (2020) support the hypothesis that improvements in temporal processing ability may lead better speech recognition in some situations. The strength of this result needs further investigation because a similar relationship was not also seen for the 1.2-s RT condition.

Other training studies employing speech stimuli have observed generalization, and these effects vary depending on training parameters (Banai and Lavner 2019; Burk and Humes 2008; Karawani et al. 2015). Banai and Lavner (2019) trained young listeners to recognize time-compressed sentences under several different listening protocols that varied by stimulus set size, training schedule (trials presented in one training session vs. several sessions), and training duration. They found that all protocols led to improvement on the trained task and generalization to untrained tasks (new talker or sentences), but training over several sessions was the only protocol that led to generalization to new untrained sentences. Banai and Lavner concluded that distributed training provides multiple opportunities to consolidate learning. The current study also implemented distributed training during ten sessions over the course of 2 to 3 weeks and found that training with non-speech stimuli may lead to improvements in recognition of reverberant speech stimuli.

Factors That Contribute to Perceptual Learning

The Flanker score was the only variable that contributed significantly to change in rate discrimination from pretest to post-test. Individuals with better response inhibition/attention experienced greater decreases in relative DLs following training. We had hypothesized that both cognitive and ASSR measures would relate to changes in rate discrimination. This hypothesis was based in part on the results of Gaskins et al. (2019), who found that both processing speed and ASSR spectral energy predicted 400-Hz rate discrimination. The current study found relationships among all of the cognitive variables and the pre-test relative DLs (R^2 values ranging from 0.14 to 0.37), but not among the pre-test ASSR PLFs and relative DLs (no R^2 value higher than 0.10). Overall, the current results suggest that cognitive function could potentially be important factor in the potential for improvement in temporal processing ability, at least with respect to rate discrimination. We note that the relatively high rates used in the current study arise from brainstem sources (Herdman et al. 2002). Perhaps the inclusion of a lower rate emanating from the cortex (e.g., ≤40 Hz) would reveal a relationship between ASSR PLF and perceptual change due to the likelihood that cortical sources may be more highly influenced by top-down cognitive influences.

CONCLUSION

The current results suggest that perceptual training improves rate discrimination across listeners and can partially restore behavioral auditory temporal processing deficits in older listeners. Neural phase locking also improves with training, but there was no relationship among behavioral and neural measurements with the tested rates. At least one measure of cognitive function, response inhibition/attention, accounts for significant variance in improvement in rate discrimination. Therefore, the paradigm used in the study protocol may be efficacious for individuals with average attention ability, but individuals with impaired attention or cognitive function may benefit from a different paradigm.

REFERENCES

ANDERSON S, WHITE-SCHWOCH T, PARBERY-CLARK A, KRAUS N (2013) Reversal of age-related neural timing delays with training. Proc Natl Acad Sci U S A 110:4357–4362. https://doi.org/10.1073/pnas.1213555110

ANSI (2018) ANSI/ASA S3.6-2018 Specification for audiometers. American National Standards Institute, New York

Banai K, Lavner Y (2019) Effects of stimulus repetition and training schedule on the perceptual learning of time-compressed speech and its transfer. Atten Percept Psychophys 81:2944–2955. https://doi.org/10.3758/s13414-019-01714-7

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing J R Stat Soc Series B Stat Methodol 289–300. https://www.jstor.org/stable/2346101

Bianchi F, Carney LH, Dau T, Santurette S (2019) Effects of musical training and hearing loss on fundamental frequency discrimination and temporal fine structure processing: psychophysics and modeling. J Assoc Res Otolaryngol 20:263–277. https://doi.org/10.1007/s10162-018-00710-2

Bissmeyer SRS, Hossain S, Goldsworthy RL (2020) Perceptual learning of pitch provided by cochlear implant stimulation rate. PLoS ONE 15:e0242842. https://doi.org/10.1371/journal.pone.0242842

Burk MH, Humes LE (2008) Effects of long-term training on aided speech-recognition performance in noise in older adults. J Speech Lang Hear Res 51:759–771. https://doi.org/10.1044/1092-4388(2008/054)

Carlyon RP, Deeks JM (2002) Limitations on rate discrimination. J Acoust Soc Am 112:1009–1025. https://doi.org/10.1121/1.1496766

Carlyon RP, Long CJ, Deeks JM (2008) Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues. J Acoust Soc Am 123:2276–2286. https://doi.org/10.1121/1.2874796

de Villers-Sidani E, Alzghoul L, Zhou X, Simpson KL, Lin RCS, Merzenich MM (2010) Recovery of Functional and Structural Age-Related Changes in the Rat Primary Auditory Cortex with Operant Training Proceedings of the National Academuy of Sciences - USA 107:13900–13905. https://doi.org/10.1073/pnas.1007885107

Desloge JG, Reed CM, Braida LD, Perez ZD, Delhorne LA (2012) Auditory-Filter Characteristics for Listeners with Real and Simulated Hearing Impairment Trends Amplif 16:19–39. https://doi.org/10.1177/1084713812445510

DeVries L, Anderson S, Goupell MJ, Smith E, Gordon-Salant S (2022) Effects of aging and hearing loss on perceptual and electrophysiological measures of pulse-rate discrimination. J Acoust Soc Am 151:1639. https://doi.org/10.1121/10.0009399

Dreschler WA, Leeuw AR (1990) Speech reception in reverberation related to temporal resolution. J Speech Hear Res 33:181–187. https://doi.org/10.1044/jshr.3301.181

Fitzgerald MB, Wright BA (2011) Perceptual learning and generalization resulting from training on an auditory amplitude-modulation detection task. J Acoust Soc Am 129:898–906. https://doi.org/10.1121/1.3531841

Fitzgibbons PJ, Gordon-Salant S (1994) Age effects on measures of auditory duration discrimination J Speech Hear Res 37:662. https://doi.org/10.1044/jshr.3703.662

Fitzgibbons PJ, Gordon-Salant S (1996) Auditory temporal processing in elderly listeners. J Am Acad Audiol 7:183–189

Fitzgibbons PJ, Gordon-Salant S (2001) Aging and temporal discrimination in auditory sequences. J Acoust Soc Am 109:2955–2963. https://doi.org/10.1121/1.1371760

Fitzgibbons PJ, Gordon-Salant S (1995) Age effects on duration discrimination with simple and complex stimuli. J Acoust Soc Am 98:3140–3145. https://doi.org/10.1121/1.413803

Fostick L, Taitelbaum-Swead R, Kreitler S, Zokraut S, Billig M (2020) Auditory training to improve speech perception and self-efficacy in aging adults. J Speech Lang Hear Res 63:1270–1281. https://doi.org/10.1044/2019_jslhr-19-00355

Gaskins C, Jaekel BN, Gordon-Salant S, Goupell MJ, Anderson S (2019) Effects of aging on perceptual and electrophysiological responses to acoustic pulse trains as a function of rate. J Speech Lang Hear Res 62:1087–1098. https://doi.org/10.1044/2018_jslhr-h-ascc7-18-0133

Goldsworthy RL, Shannon RV (2014) Training improves cochlear implant rate discrimination on a psychophysical task. J Acoust Soc Am 135:334–341. https://doi.org/10.1121/1.4835735

Gordon-Salant S, Fitzgibbons PJ (1993) Temporal factors and speech recognition performance in young and elderly listeners. J Speech Hear Res 36:1276–1285. https://doi.org/10.1044/jshr.3606.1276

Harris KC, Eckert MA, Ahlstrom JB, Dubno JR (2010) Age-related differences in gap detection: effects of task difficulty and cognitive ability. Hear Res 264:21–29

Herdman AT, Picton TW, Stapells DR (2002) Place specificity of multiple auditory steady-state responses. J Acoust Soc Am 112:1569–1582. https://doi.org/10.1121/1.1506367

Hetherington R (1954) The Snellen Chart as a Test of Visual Acuity Psychol Forsch 24:349–357. https://doi.org/10.1007/BF00422033

IEEE (1969) IEEE recommended practice for speech quality measurements. https://doi.org/10.1109/IEEESTD.1969.7405210

Ihlefeld A, Carlyon RP, Kan A, Churchill TH, Litovsky RY (2015) Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners. J Assoc Res Otolaryngol 16:641–652. https://doi.org/10.1007/s10162-015-0527-7

Irwin RJ, McAuley SF (1987) Relations among temporal acuity, hearing loss, and the perception of speech distorted by noise and reverberation J Acoust Soc Am 81:1557–1565. https://doi.org/10.1121/1.394508

JASP Team (2020) JASP (Version 0.14.1[Computer software])

Jenkins KA, Fodor C, Presacco A, Anderson S (2018) Effects of amplification on neural phase locking, amplitude, and latency to a speech syllable. Ear Hear 39:810–824. https://doi.org/10.1097/aud.0000000000000538

Johnson KC, Xie Z, Shader MJ, Mayo PG, Goupell MJ (2021) Effect of Chronological Age on Pulse Rate Discrimination in Adult Cochlear-Implant Users Trends Hear 25:23312165211007370. https://doi.org/10.1177/23312165211007367

Karawani H, Bitan T, Attias J, Banai K (2015) Auditory Perceptual Learning in Adults with and without Age-Related Hearing Loss Front Psychol 6:2066. https://doi.org/10.3389/fpsyg.2015.02066

Kong YY, Deeks JM, Axon PR, Carlyon RP (2009) Limits of temporal pitch in cochlear implants. J Acoust Soc Am 125:1649–1657. https://doi.org/10.1121/1.3068457

Koziol LF, Budding DE (2012) Procedural Learning. In: Seel NM (ed) Encyclopedia of the Sciences of Learning. Springer US, Boston, MA, pp 2694–2696. https://doi.org/10.1007/978-1-4419-1428-6_670

Kumar U (2011) Temporal processing abilities across different age groups. J Am Acad Audiol 22:5–12

Lakshminarayanan K, Tallal P (2007) Generalization of non-linguistic auditory perceptual training to syllable discrimination. Restor Neurol Neurosci 25:263–272

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

Macherey O, Carlyon RP (2014) Re-examining the upper limit of temporal pitch. J Acoust Soc Am 136:3186. https://doi.org/10.1121/1.4900917

Micheyl C, Delhommeau K, Perrot X, Oxenham AJ (2006) Influence of musical and psychoacoustical training on pitch discrimination. Hear Res 219:36–47

Nasreddine ZS et al (2005) The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53:695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

ROQUE L, GASKINS C, GORDON-SALANT S, GOUPELL MJ, ANDERSON S (2019a) Age effects on neural representation and perception of silence duration cues in speech. J Speech Lang Hear Res 62:1099–1116. https://doi.org/10.1044/2018_jslhr-h-ascc7-18-0076

ROQUE L, KARAWANI H, GORDON-SALANT S, ANDERSON S (2019b) Effects of age, cognition, and neural encoding on the perception of temporal speech cues. Front Neurosci 13:749. https://doi.org/10.3389/fnins.2019b.00749

ROSS B, SCHNEIDER B, SNYDER JS, ALAIN C (2010) Biological markers of auditory gap detection in young, middle-aged, and older adults. Plos ONE 5:e10101

SABIN A, CLARK C, EDDINS D, WRIGHT B (2013) Different patterns of perceptual learning on spectral modulation detection between older hearing-impaired and younger normal-hearing adults. J Assoc Res Otolaryngol 14:283–294. https://doi.org/10.1007/s10162-012-0363-y

SABIN AT, EDDINS DA, WRIGHT BA (2012) Perceptual learning evidence for tuning to spectrotemporal modulation in the human auditory system. J Neurosci 32:6542–6549. https://doi.org/10.1523/jneurosci.5732-11.2012

SNELL KB (1997) Age-related changes in temporal gap detection. J Acoust Soc Am 101:2214–2220. https://doi.org/10.1121/1.418205

STUDEBAKER GA (1985) A "rationalized" arcsine transform. J Speech Hear Res 28:455–462. https://doi.org/10.1044/jshr.2803.455

TALLON-BAUDRY C, BERTRAND O, DELPUECH C, PERNIER J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. J Neurosci 16:4240–4249. https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996

TILLMAN TW, CARHART R (1966) An expanded test for speech discrimination utilizing CNC monosyllabic words. Northwestern University Auditory Test No. 6. SAM-TR-66–55 Tech Rep SAM-TR:1–12. https://doi.org/10.21236/ad0639638

WEINTRAUB S ET AL (2013) Cognition Assessment Using the NIH Toolbox Neurology 80:S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded

WHITTON JP, HANCOCK KE, SHANNON JM, POLLEY DB (2017) Audio-motor perceptual training enhances speech intelligibility in background noise. Curr Biol 27:3237-3247.e3236. https://doi.org/10.1016/j.cub.2017.09.014

ZALTZ Y, KISHON-RABIN L (2022) Difficulties Experienced by Older Listeners in Utilizing Voice Cues for Speaker Discrimination Front Psychol 13:797422. https://doi.org/10.3389/fpsyg.2022.797422

ZHOU N, MATHEWS J, DONG L (2019) Pulse-rate discrimination deficit in cochlear implant users: is the upper limit of pitch peripheral or central? Hear Res 371:1–10. https://doi.org/10.1016/j.heares.2018.10.018