Age, Hearing, and the Perceptual Learning of Rapid Speech

Maayan Manheim[1], Limor Lavie[1], and Karen Banai[1] ⓘ

Abstract

The effects of aging and age-related hearing loss on the ability to learn degraded speech are not well understood. This study was designed to compare the perceptual learning of time-compressed speech and its generalization to natural-fast speech across young adults with normal hearing, older adults with normal hearing, and older adults with age-related hearing loss. Early learning (following brief exposure to time-compressed speech) and later learning (following further training) were compared across groups. Age and age-related hearing loss were both associated with declines in early learning. Although the two groups of older adults improved during the training session, when compared to untrained control groups (matched for age and hearing), learning was weaker in older than in young adults. Especially, the transfer of learning to untrained time-compressed sentences was reduced in both groups of older adults. Transfer of learning to natural-fast speech occurred regardless of age and hearing, but it was limited to sentences encountered during training. Findings are discussed within the framework of dynamic models of speech perception and learning. Based on this framework, we tentatively suggest that age-related declines in learning may stem from age differences in the use of high- and low-level speech cues. These age differences result in weaker early learning in older adults, which may further contribute to the difficulty to perceive speech in daily conversational settings in this population.

Keywords

aging, hearing loss, adaptation, perceptual adjustment, compressed speech, rapid learning

Date received: 24 December 2017; accepted: 26 April 2018

Introduction

Perceptual learning allows listeners to adjust to natural variations in speech (e.g., accents, speech rates) and to learn new forms of speech (e.g., time-compressed and vocoded speech; for reviews see Davis & Johnsrude, 2007; Samuel, 2011; Samuel & Kraljic, 2009). Nevertheless, perceptual learning was rarely considered in the context of age- and hearing-related declines in speech perception. Rather, these declines were usually interpreted as outcomes of sensory and cognitive declines (e.g., Divenyi & Haupt, 1997; Gordon-Salant, Fitzgibbons, & Friedman, 2007; Goy, Pelletier, Coletta, & Pichora-Fuller, 2013; Humes & Christopherson, 1991; Humes & Roberts, 1990; Martin & Jerger, 2005; Pichora-Fuller, Schneider, & Daneman, 1995; Schneider & Pichora-Fuller, 2001; Wingfield, Tun, Koh, & Rosen, 1999; Wingfield & Tun, 2001). Therefore, the theoretical basis for using perceptual learning to support hearing rehabilitation in older adults (e.g., Burk, Humes, Amos, & Strauser, 2006; Henderson Sabes & Sweetow,

2007; Pichora-Fuller & Levitt, 2012) is not sufficiently formulated.[1] The goal of this study was therefore to study the effects of age and age-related hearing loss on the perceptual learning of time-compressed speech.

Following Gibson (1963), we define perceptual learning as an increase in the ability to extract information from stimulus arrays (in our case speech) due to experience. Perceptual learning often exhibits two phases and both were characterized in this study: an *early learning* phase characterized by rapid learning following brief experience with new stimuli and a *later* phase that follows longer experiences in which learning is usually slower and more gradual (Green, Banai, Lu, &

[1]Department of Communication Sciences and Disorders, University of Haifa, Israel

Corresponding author:
Karen Banai, University of Haifa, 199 Aba Khoushy Ave., Mount Carmel, Haifa 3498838, Israel.
Email: kbanai@research.haifa.ac.il

Bevalier, 2018). Although the literature is not decisive on the theoretical differences between these two phases, the distinction is important when speech recognition is considered. Although ongoing recognition of conversational speech and rapid adaptation to acoustic challenges depend on *early* learning, stable and long-lasting changes in speech representations require further learning (Mattys, Davis, Bradlow, & Scott, 2012). The latter is required if training is to play a meaningful role in rehabilitation. Here, we asked whether the effects of early learning and later learning following longer training differed as a function of age and hearing loss, and whether the transfer of learning, defined as changes in the recognition of stimuli different from those presented in training, is influenced by age and age-related hearing loss.

We used time-compressed speech as an analogue of natural-fast speech because older adults, even with normal hearing, often complain that conversational speech is too rapid for them to follow (Gordon-Salant & Fitzgibbons, 2001; Vaughan & Letowski, 1997; Wingfield, Poon, Lombardi, & Lowe, 1985). Moreover, current hearing aid technologies offer no solutions for this issue. Previous studies suggest that the recognition of time-compressed speech is amenable to both *early* (e.g., Dupoux & Green, 1997; Pallier, Sebastian-Galles, Dupoux, Christophe, & Mehler, 1998), and *later* (Banai & Lavner, 2014; Karawani, Bitan, Attias, & Banai, 2016; Song, Skoe, Banai, & Kraus, 2012) learning. Furthermore, in young adults learning on time-compressed speech might transfer to natural-fast speech (Adank & Janse, 2009), and time-compressed speech is used in rehabilitative training programs such as listening and communication enhancement (LACE; Henderson Sabes & Sweetow, 2007). However, if the transfer of learning declines with age, this might not be a vaiable option for older adults. As described below, studies on the effects of sensory load and linguistic complexity on the perception of time-compressed speech are consistent with a dynamic model of speech processing which is especially relevant in the context of perceptual learning and with the Reverse Hirarchy Theory (RHT) of perceptual learning (Ahissar, Nahum, Nelken, & Hochstein, 2009). These models are used in the final section of the introduction to derive the hypotheses for this study.

Speech Processing Declines in the Aging Auditory System: A Dynamic View

Age and hearing loss are both detrimental to time-compressed speech perception (e.g., Gordon-Salant & Fitzgibbons, 2001; Stine, Wingfield, & Poon, 1986; Tun, Wingfield, Stine, & Mecsas, 1992; Wingfield et al., 1999). Age effects are larger when sensory load is increased (e.g., by selective compression of consonants compared to compression of vowels or silent intervals, Gordon-Salant & Fitzgibbons, 2001). Furthermore, age effects are smaller when multiple linguistic cues are present (Stine & Wingfield, 1987; Stine et al., 1986) and when silent pauses are inserted between words to provide older adults more processing time (Wingfield et al., 1999). Findings such as the latter have been taken to suggest that age-related changes in the rate of speech processing play an important role in age-associated declines in speech perception (Stine et al., 1986; Wingfield et al., 1985; Wingfield et al., 1999). Hearing loss was found to have an independent contribution to the perception of time-compressed speech, with greater difficulties in older adults with age-related hearing loss than in those with age-normal hearing (Gordon-Salant & Fitzgibbons, 2001; Gordon-Salant et al., 2007; Stine et al., 1986; Wingfield et al., 1999).

The findings that sensory and linguistic factors both modulate the effects of age on time-compressed speech recognition can be parsimoniously interpreted within the framework of dynamic model of speech processing which is conceptually similar to the RHT model of perceptual learning described below. By this account, speech perception depends on the dynamic interaction between bottom-up processes of speech encoding, and nonsensory top-down processes (Davis & Johnsrude, 2007; Guediche, Blumstein, Fiez, & Holt, 2014; Mattys, Davis, Bradlow, & Scott, 2009; Mattys, Carroll, Li, & Chan, 2010; Nahum, Nelken, & Ahissar, 2008; Sohoglu, Peelle, Carlyon, & Davis, 2012). This interaction is evident when prior information such as lexical knowledge (Sohoglu et al., 2012; Sohoglu, Peelle, Carlyon, & Davis, 2014) and task demands (Mattys et al., 2009; Mattys & Wiget, 2011) influence speech recognition. However, growing evidence suggests that this fine interaction shifts with age (Alain, McDonald, Ostroff, & Schneider, 2004; Alain & Snyder, 2008; Eckert et al., 2008; Erb & Obleser, 2013; Goy et al., 2013; Mattys & Scharenborg, 2014; Peelle & Wingfield, 2016). Young adults flexibly use low-level acoustic–phonetic cues and higher level lexical cues depending on the situation, with greater weight assigned to low-level cues in the presence of sensory load (Mattys et al., 2009; Mattys & Wiget, 2011). Older adults on the other hand are not as flexible and tend to rely on higher level processes in compensation for low-level encoding difficulties (Alain et al., 2004; Du, Buchsbaum, Grady, & Alain, 2016). In line with the dynamic model, older adults had more difficulties than younger adults in ignoring lexical competition from similarly sounding words during speech recognition (Sommers, 1996; Sommers & Danielson, 1999) and phoneme categorization (Mattys & Scharenborg, 2014), but their speech recognition benefited from semantic context to a greater extent (Sommers & Danielson, 1999).

Changes in Learning of Perceptually Difficult Speech with Age and Age-Related Hearing Loss

In young adults, experience with perceptually difficult (e.g., rapid, accented, or noisy) speech yields perceptual learning consistent with the two phases described earlier. First, brief exposure to as few as 10 to 20 sentences of hard-to-recognize speech results in rapid improvements in recognition (i.e., *early learning;* Adank & Janse, 2009; Bradlow & Bent, 2008; Clarke & Garrett, 2004; Dupoux & Green, 1997; Reinisch & Holt, 2014). Second, further training, often several sessions long, yields additional improvements (i.e., *later learning*), but these are often more specific to the trained materials, with only partial transfer to untrained sentences (Banai & Lavner, 2014; Song et al., 2012). The Reverse Hierarchy Theory (RHT) of perceptual learning (Ahissar & Hochstein, 1993; Ahissar et al., 2009; Hochstein & Ahissar, 2002; Nahum et al., 2008) suggests that with training, young adults gradually learn to shift their perception to rely on the speech cues relevant to the task at hand. Although initially they rely on the high-level information (the default mode of perception in daily and familiar conditions) even when faced with acoustically challenging stimuli (e.g., time-compressed speech), learning enables relaiance on the low-level cues which are required to accurately percieve acoustically challenging stimuli. According to the RHT, learning is guided by task demands and thus depends on both low-level representations and on the cognitive processes required to locate these representations and make them accessible (Ahissar et al., 2009; Nahum, Nelken, & Ahissar, 2010). Because learning starts in response to the trained stimuli, it is often specific to these stimuli. However, because learning is triggered at high-levels of processing, transfer is afforded across stimuli that share high-level representations (e.g., the same word produced by different talkers). Findings on the perceptual learning of acoustically challenging speech in young adults are consistent with RHT predictions (Banai & Lavner, 2016; Neger, Rietveld, & Janse, 2014). Specifically, as it predicts, the learning of time-compressed speech was found to generalize across stimuli that share high-level representations, and further training led to further generalization to novel time-compressed sentences that do not share high-level representations with the trained ones (Banai & Lavner, 2012, 2014).

As explained earlier, the dynamic use of high- and low-level speech cues changes with age. Therefore, age should influence only the outcomes of learning that are dependent on the ability to shift from high-level, top-down, operations to low-level ones. Age-related hearing loss is expected to have an additional effect due to the degradation of low-level acoustic–phonetic representations. Specifically, increased reliance on high-level speech cues should interfere with learning by making older adults less able to locate or attend to the relevant low-level cues.[2]

The evidence for the decline of auditory learning with age and hearing loss is mixed. A few studies suggest that the learning associated with brief exposure to time-compressed (Golomb, Peelle, & Wingfield, 2007; Peelle & Wingfield, 2005) and to accented (Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010) speech does not decline with age. For example, after taking into account age-related differences in accuracy, the rate and amount of adaptation to time-compressed speech were similar in younger and in older adults (Peelle & Wingfield, 2005). However, consistent with our suggestion that learning should decline with age when it depends on low-level acoustic cues, following exposure to one rate of time-compressed speech, older adults exhibited less transfer of learning to an untrained speech rate (i.e., a change in low-level cues) than young adults, despite similar amounts of learning (Peelle & Wingfield, 2005). Furthermore, the perceptual learning of auditory categories appears to decline with age (Maddox, Chandrasekaran, Smayda, & Yi, 2013; Scharenborg & Janse, 2013). Finally, other learning studies in older adults did not include a younger comparison group (Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Humes, Burk, Strauser, & Kinney, 2009; Janse & Adank, 2012; Karawani et al., 2016; Scharenborg, Weber, & Janse, 2015).

A few studies also reported relatively well-preserved auditory learning in older adults with age-related hearing loss (Burk et al., 2006; Karawani et al., 2016; Lavie, Attias, & Karni, 2013; Walden, Erdman, Montgomery, Schwartz, & Prosek, 1981). In one study, dichotic listening was compared between two groups of novice hearing-aid users, one of which received seven sessions of individual listening training (Lavie et al., 2013). Improvements were greater in the trained group, suggesting preserved learning in listeners with age-related hearing loss. In another study, older adults with hearing impairment needed longer training to improve their speech recognition in noise compared with normal-hearing young adults (Burk et al., 2006). However, as these studies did not include comparison groups of older adults with normal hearing, it is hard to determine whether hearing loss had an effect on learning outcomes. Finally, multisession learning curves were compared between older adults with normal hearing and older adults with age-related hearing loss on three speech recognition tasks (speech in noise, competing talker and time-compressed speech; Karawani et al., 2016). Group differences in learning were found for speech in noise but not for the other two conditions.

Specific Questions and Hypotheses for This Study

Our goal was to study the effects of age and hearing on several aspects of time-compressed speech learning. To disentangle the effects of age and hearing on learning from their already known effects on speech perception, accuracy was tested using speech rates that yielded equivalent levels of baseline recognition accuracy of time-compressed speech across groups. In addition, an adaptive training procedure which adjusts speech rates individually was used to equate levels of accuracy across groups during training (Leek, 2001; Levitt, 1971; for further discussion see Peelle & Wingfield, 2005). Different patterns of learning across groups should therefore reflect either an independent effect of high-level factors such as cognitive effects or an independent sensory effect on the learning process itself.

We asked whether the learning of time-compressed speech and its transfer to untrained sentences and to natural-fast speech are influenced by age and age-related hearing loss. Based on the aforementioned literature review, we assumed that on encounter with time-compressed speech, older adults should be less able than younger adults to switch from the default mode of relying on high-level cues to using low-level cues. If this is the case, older adults should exhibit less *early* learning than younger adults, because on initial encounter, highly time-compressed speech deviates substantially from normally encountered speech, thus requiring greater than normal reliance on low-level acoustic cues. On the other hand, over the course of adaptive training, during which difficulty is "tailored" to individual performance, and with a training task (semantic verification) that requires utilization of high-level information, no differences are expected as a function of age. The overall effects of learning and transfer were also assessed by comparing trained and untrained participants (matched for age and hearing). In these comparisons, better performance in trained than in untrained listeners is taken to indicate *later* learning or transfer. We expected that improvements in speech recognition from pretest to posttest should be smaller in older than in younger adults. We also hypothesized that transfer from time-compressed sentences to sentences repeated in natural-fast speech should not differ across groups because these stimuli share high-level representations with the trained stimuli. On the other hand, transfer to nonrepeated sentences was expected to decrease with age, because this transfer is based on listeners having learned (with training) to increase their reliance on low-level cues when processing perceptually difficult speech. The ability to do so is expected to decrease with age. Finally, differences between older adults with normal hearing and those with age-related hearing loss will provide evidence for an additional contribution of lower level sensory processing to learning.

Methods

Participants

One hundred fifty-seven potential participants were recruited through advertisements at academic institutions (young adults), retirement communities, and community centers (older adults). Prospective participants were screened to determine whether they met the inclusion criteria described in the next paragraph and grouped into three groups—young adults with normal hearing (YA), older adults with age-matching normal hearing (ONH), and older adults with age-related mild to moderate hearing loss (OHI).

Healthy young adults and older adults with healthy aging (with the possible exception of age-related hearing loss) were targeted. Therefore, the following inclusion criteria were used: (a) independent, community-dwelling adults with no prior experience with time-compressed speech; (b) high-school education or higher; (c) Hebrew as first or primary language based on self-report that each participant was using Hebrew as primary language (oral and written) for personal, familial, and professional needs and has been doing so for decades. The experimenters also judged the Hebrew proficiency of all participants as good (or higher); (d) normal neurological status (by self-report); (e) Mini-Mental State Examination Scores (Folstein, Folstein, & McHugh, 1975) $\geqslant 24$ (older adults only); (f) normal hearing or mild-to-moderate hearing loss. Hearing level was defined based on 4-frequency (500, 1000, 2000, and 4000 Hz) pure tone average (PTAs) thresholds; (f) suprathreshold speech recognition scores $\geqslant 60\%$; and (g) no experience (prior or current) with using hearing aids.

Six potential participants failed to meet the inclusion criteria and were excluded from the study: Three older adults were excluded for poor Hebrew proficiency, two for reporting additional neurological issues, and one for having more severe hearing loss. Two other participants—a young adult and an OHI participant completed only the pretest; their data were thus excluded from all analyses. The final study sample included 149 participants.

Normal hearing was defined as PTA thresholds (at 0.5, 1, 2, and 4 kHz) of 25 dB-HL or better (Duthey, 2013; Goman & Lin, 2016). By this criterion, all young adults (YA) had normal hearing. Fifty-six older adults with age-matching normal hearing (ONH) and 36 older adults with age-related mild to moderate hearing loss (OHI). Consistent with epidemiological studies (Gopinath et al., 2009; Roth, Hanebuth, & Probst, 2011), 39% of older adults in the current sample had

hearing loss. The characteristics of the three groups are shown in Table 1.

The relative proportions of male and female participants were similar across the three groups ($\chi^2(2) = 2.26$, $p = .32$). Mean ages did not differ significantly between the two groups of older adults, $t(90) = -1.73$, $p = .088$, but as expected, PTAs in both ears were lower in the ONH group, Right ear: $t(90) = -8.85$, $p < .001$; Left ear: $t(90) = -9.45$, $p < .001$. In the YA group, all suprathreshold recognition scores were 92% or higher. Among older adults, suprathreshold recognition scores were significantly higher in the ONH group, right ear: $t(90) = 4.11$, $p < .001$; left ear: $t(90) = 3.30$, $p = .001$.

Table 1. Characteristics of the Three Study Groups.

	YA (n = 57)	ONH (n = 56)	OHI (n = 36)
Age (years)[a]	27 ± 4	70 ± 5	72 ± 7
Age range	20–38	65–86	65–91
Gender distribution[b]	33:24	40:16	23:13
PTA[a]—right ear (dB HL)	10 ± 5	21 ± 6	*34 ± 8*
PTA[a]—left ear (dB HL)	10 ± 5	22 ± 6	*35 ± 7*
WRS[a]—right ear (%)	98 ± 3	96 ± 4	*91 ± 7*
WRS[a]—left ear (%)	99 ± 1	97 ± 4	*93 ± 8*
MMSE[a] (score)	–	29 ± 1	29 ± 1

Note. Italicized values indicate significantly poorer performance in the OHI group, see text for details. PTA = pure tone average across octave frequencies from 500 Hz to 4000 Hz; WRS = suprathreshold speech recognition scores; MMSE = Mini Mental State Examination scores; YA = young adults; ONH = older adults with age-normal hearing; OHI = older adults with mild to moderate hearing loss.
[a]Group means ± one standard deviation. [b]Numbers of female and male participants.

In the ONH group, the lowest recognition score was 88% and 47/56 participants had scores of 92% or higher; in the OHI group, the lowest score was 70%, all but two participants had scores of 80% or higher and 22/36 participants had scores of 92% or higher.

Overview of Design and General Procedures

Each participant completed two test sessions on consecutive days. On the first session, participants completed the pretest phase and were assigned at random to a training or to a no-training control group. On the second session, the training group completed the training phase described below and all participants completed the posttest phase. The ethics committee of the Faculty of Social Welfare and Health Sciences at the University of Haifa (protocol 199/12) approved all aspects of this study. Participants received a small compensation for participation in the study.

The study included three phases (see Figure 1): (a) A pretest phase during which participants were screened based on the inclusion criteria described earlier. Subsequently participants went through a baseline assessment of time-compressed and natural-fast speech recognition with three 20-sentence blocks of the recognition task described below; (b) a training phase in which half of the participants in each group (YA, ONH, OHI, and randomly assigned) completed 300 trials of adaptive training with time-compressed speech. Training was self-paced and took 40 to 55 min to complete; (c) a posttest phase in which all participants were tested again on the recognition of time-compressed speech and natural-fast speech. This phase was identical to the pretest. The same sentences were included in the pretest and posttest. The order of the three blocks of sentences presented
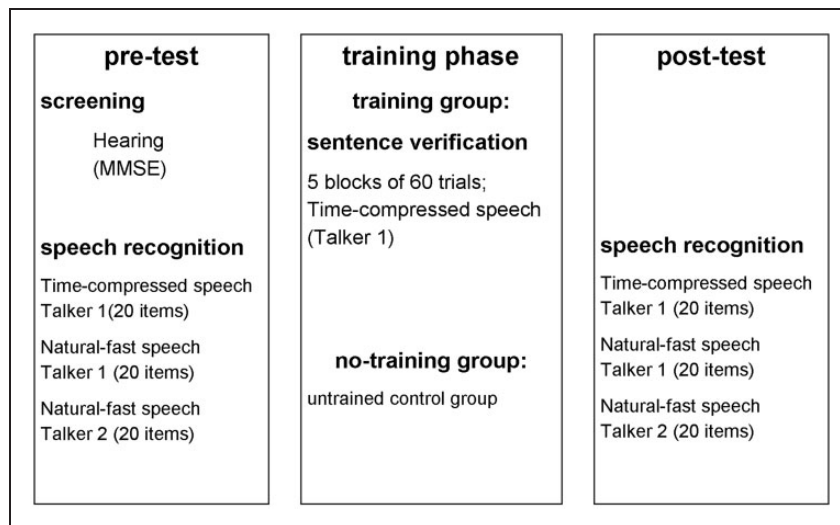

Figure 1. Study design.

during the pretest and posttest phases (see Figure 1) was counterbalanced across participants. Training and posttest were conducted on the same day to minimize participant attrition over the course of multiple sessions. Previous studies in young adults reported minimal differences in outcome between immediate (Gabay, Karni, & Banai, 2017) and delayed (Banai & Lavner, 2014) testing. Note that whereas during the pretest and posttest participants transcribed the sentences and recognition accuracy was scored based on the number of correctly reported words, during training listeners had to judge the semantic plausibility of the sentences with a yes or no task and the level of compression required to maintain a fixed level of performance was estimated as explained below.

A trained clinical audiologist (author M. M.) and a senior year audiology student who was blind to the goals of the study recruited the participants and collected the data. Hearing screening was conducted with MAICO MA 52 Audiometer.

Stimuli

Stimuli were 130 sentences in Hebrew (Prior & Bentin, 2006). Sentences were five to six words long in canonical Hebrew sentence structure (subject–verb–object). Each sentence had five content words and a maximum of one function word. Half of the sentences were semantically plausible (e.g., "the municipal museum purchased an impressionistic painting") and half were semantically implausible (e.g., "the comic book opened the back door"). Sentences were recorded by two female native-Hebrew speakers. Each of them recorded the sentences at a natural rate (Talker 1: 103 and Talker 2: 106 words/min) and at a natural-fast rate (Talker 1: 166 and Talker 2: 162 words/min). To obtain the natural-fast recordings, talkers were instructed to speak as rapidly as possible without omitting word parts. Each sentence was recorded 3 times and the clearest version was selected by two native Hebrew speakers who listened to all of the recordings. All sentences were recorded and sampled at 44 kHz using a standard microphone and PC soundcard. Time-compressed speech was generated by processing the natural-rate stimuli recorded by Talker 1 with a WSOLA algorithm (Verhelst & Roelands, 1993) in Matlab (see below for compression rates during the different phases).

As shown in Figure 2, of the 130 sentences, 100 sentences (50 plausible and 50 implausible) were randomly selected for the training phase during which they were presented in highly time-compressed format. Of these sentences, the recognition of 30 (15 plausible) sentences was assessed during the pretest and posttest phases (10 in time-compressed format and 20 in natural-fast format, 10 sentences from each talker). Overall, each of these sentences was presented once during the pretest,
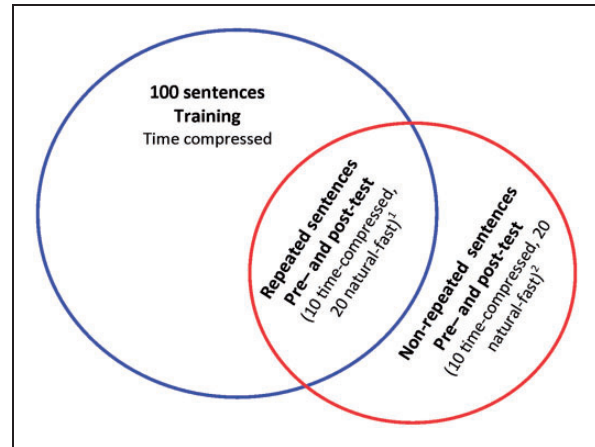
Figure 2. Schematic description of the stimuli.
[1]These sentences were presented during training and also during the test phases. [2]These sentences were not presented during the training phase.

3 times during training, and once during the posttest. The remaining 30 sentences were not included in the training set. Each was presented once during the pretest and once during posttest phases (10 in time-compressed format, 20 in natural-fast format, 10 from each talker).

Tasks

Speech recognition during the (nonadaptive) pretest and posttest phases. During the pretest and posttest phases, all participants completed three blocks of speech recognition tasks in counterbalanced order. After hearing each sentence once, listeners were asked to write it down as accurately as possible. Listeners could then initiate the next sentence. Each sentence could be played only once. No feedback was provided. Each block comprised 10 sentences taken from the training set (dubbed trained tokens) and 10 additional sentences (dubbed untrained tokens) as follows:

1. Time-compressed speech: 20 sentences produced by Talker 1 were compressed to 30% to 40% (see below) of their duration as recorded in natural speech rate.
2. Natural-fast 1: 20 sentences produced by Talker 1 at a natural-fast rate of 166 words/min.
3. Natural-fast 2: 20 sentences produced by Talker 2 at a natural-fast rate of 162 words/min.

These same blocks were repeated in the pretest and posttest phases. The order of sentences within a block was randomized, but the same random order was used for all participants to facilitate scoring by the experimenters. A different random order was used at the pretest and the posttest phases. The order of blocks was counterbalanced across participants. Natural-fast stimuli from

two talkers were included to allow us to test the transfer of learning across different talkers. However, as we did not have any differential hypotheses regarding the effects of age and hearing on cross-talker transfer, data from the two blocks were collapsed for simplicity. We note that talker did not moderate any of the effects reported in this manuscript.

Selection of compression rates. To disentangle the unique effects of age and age-related hearing loss on learning from their already documented effects on the perception of time-compressed speech, baseline performance on the time-compressed speech task was equated across groups. Because we were interested in rapid learning, we were concerned that the relatively large number of trials required by the individual adjustment method proposed by Peelle and Wingfield (2005) could mask this learning. We therefore decided to use a predetermined level of time-compression for each group for the pretest and posttest phases. As in previous studies (Banai & Lavner, 2014, 2016), young adults received the time-compressed sentences at 30% of their natural (un-speeded) duration. Sentences were presented at 40% of their natural duration for the OHI group and at 35% to 40% of their natural duration for the ONH group (half of the ONH listeners were randomly assigned to listen at 35% of natural duration, and the remaining half were assigned to listen at 40%). These rates were based on a small-scale pilot study in which 15 older adults were presented with 20 sentences compressed to 35% to 40% of their natural rate. These data suggested that over the course of 20 sentences, these compression rates yielded similar ranges of performance in older adults as in young adults presented with identical sentences compressed to 30% of their natural duration.

Scoring. The percentage of correctly transcribed words across sentences of a given type was used as an index of performance. Because we were interested in spoken language recognition, words with homophonic spelling errors were accepted as correct, consistent with our previous studies (Banai & Lavner, 2014, 2016). All words, including function words, were counted for scoring.

Speech verification during the (adaptive) training phase. Training was conducted with an adaptive speech verification task. Overall, 300 training trials were delivered, divided to five adaptive blocks of 60 trials each. On each trial, a sentence was randomly selected from the pool of 100 sentences. Once the pool was exhausted, sentences were resampled. The only restriction was that on each block of 60 sentences, an equal number of plausible and implausible sentences were to be selected. Prior to the start of the first block, participants were instructed that they were going to hear sentences spoken quite rapidly and that they are to determine, for each sentence, whether its content makes sense or not. They were then given an example of a plausible and an implausible sentence at a conversational rate. During each trial, participants heard one time-compressed sentence and had to indicate their plausibility judgment by selecting from the true (plausible) or false (implausible) options visually presented on screen. Visual feedback was provided in the form of a smiley face for correct responses and a sad face for incorrect ones. The level of time compression was varied adaptively in a 2-down or 1-up staircase procedure. Compression increased after each sequence of two correct responses and decreased after each error. The staircase comprised 25 logarithmically equal steps between 0.65 and 0.2 (e.g., compression dropped from 0.65 to 0.52, then to 0.44; for further details, see Banai & Lavner, 2014; Gabay et al., 2017). To minimize the effects of fatigue, participants were encouraged to take "refreshment breaks" as needed between the blocks. This phase of the study thus took 40 to 55 min to complete. All trained participants took a 10- to 15-min break between the end of training and the posttest.

Scoring. Verification thresholds were calculated for each trained participant and each trained block by averaging the compression rates from the last five reversals on each block. Reversals are trials on which the direction of the adaptive track changes. Thus, trials on which a second correct response occurs after an error and trials on which an error occurs after a correct response are defined as reversal. These thresholds reflect 71% correct on the psychometric function (Leek, 2001; Levitt, 1971) linking the level of compression to performance accuracy. The slopes of the resulting five-block learning curves were calculated and used in statistical analysis, with negative slopes taken as an indication of training-induced learning.

Data Analysis

Data analysis focused on the following aspects of the data:

1. *Baseline performance* was defined as mean performance on each type of sentences presented during the pretest phase, calculated separately for time-compressed speech and natural-fast speech.
2. *Early learning* of time-compressed speech was defined as the amount of improvement over the course of listening to and transcribing 20 time-compressed sentences during the pretest. Learning was calculated for each listener as the difference between mean accuracy during the first and final "mini-blocks" of five time-compressed sentences presented during the time-compressed phase of the pretest.

3. *Changes in performance during the training phase* were evaluated for trained participants only, by calculating the slopes of the individual learning curves.
4. *Training effects* were defined as the accuracy differences between pretest and posttest performance with the time-compressed sentences that were included in the training phase.
5. *Transfer effects* were defined as the accuracy differences between pretest and posttest performance with materials that differed from those included in training. Three types of transfer were assessed: to time-compressed sentences that were not included in training, to natural-fast sentences that were included in training as time-compressed speech (and thus differed from the trained sentences in acoustic structure—a low-level feature), and to natural-fast sentences that were not included in the training phase (and thus differed from the trained materials in both high-level and low-level information).

For pretest and training phase data, variances were not homogenous across groups (Levene statistic > 13, $p < .001$), leading us to use Kruskal–Wallis analyses of variance (ANOVAs) to compare mean ranks across the three groups. To isolate the effects of age and hearing loss, these were followed by planned Mann–Whitney tests between the YA and the ONH and the ONH and OHI groups. Training and transfer effects were analyzed parametrically with ANOVAs. Differences on planned comparisons between the YA and ONH groups were attributed to age, whereas performance differences between the ONH and OHI groups were attributed to age-related hearing loss.

Results

Baseline Performance

Baseline performance on time-compressed speech was similar in the three groups (see Table 2 and Figure 3,

left), suggesting that the choice of compression rates yielded groups with similar performance distributions (Kruskal–Wallis $\chi^2(2) = 2.64$, $p = .27$). On the other hand, consistent with previous literature, the two groups of older adults recognized natural-fast speech less accurately than young adults (see Figure 3, right panel; Kruskal–Wallis $\chi^2(2) = 65.88$, $p < .001$;). These group differences stem from an age effect (YA vs. ONH: Mann–Whitney $U = 503$, $Z = -6.28$, $p < .001$) as well as from an effect of age-related hearing loss (ONH vs. OHI: Mann–Whitney $U = 701$, $Z = -2.46$, $p = .014$).
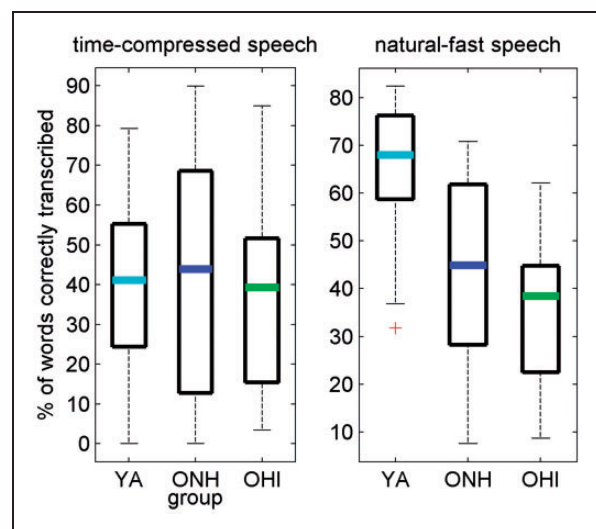
Figure 3. Pretest (baseline) performance by group. Left panel: time-compressed speech, right panel: natural-fast speech (collapsed across the two talkers). Each panel shows the median percentage of words correctly transcribed (thick line within each box) of young adults (YA, left), older adults with normal hearing (ONH, middle) and older adults with ARHL (OHI, right). Box edges mark the 25th and 75th percentiles of the group data; whiskers are 1.5 the interquartile range.

Table 2. Top: Pretest Recognition Performance by Group (Mean ± Standard Deviation [95% Confidence Interval]); Bottom: Pretest Performance by Subgroup (Mean ± Standard Deviation).

% Correct	Group					
	YA		ONH		OHI	
Time-compressed speech	39 ± 20 [33, 45]		44 ± 29 [36, 52]		36 ± 21 [28, 43]	
Natural-fast speech	66 ± 13 [63, 70]		43 ± 19 [38, 48]		34 ± 15 [29, 40]	
Subgroups	Trained	Control	Trained	Control	Trained	Control
Time-compressed speech	41 ± 18	38 ± 22	42 ± 25	47 ± 32	37 ± 21	34 ± 22
Natural-fast speech	67 ± 11	65 ± 14	42 ± 18	45 ± 20	35 ± 15	33 ± 15

Note: YA = young adults; ONH = older adults with age-normal hearing; OHI = older adults with mild to moderate hearing loss.
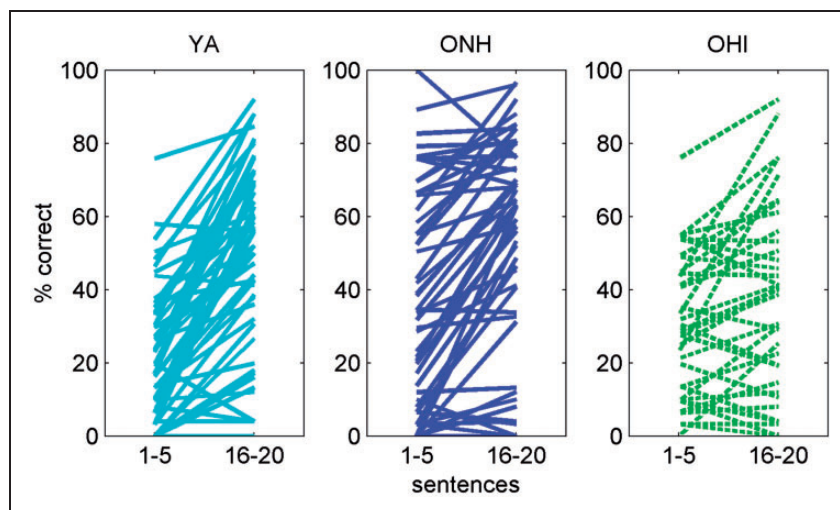
Early Learning of Time-Compressed Speech

We hypothesized that on first encounter with highly compressed speech early learning will be diminished in older adults due to decreased reliance on low-level speech cues. As shown in Figure 4, the recognition of time-compressed speech improved among most of the participants across the three groups over the course of listening to and transcribing 20 highly time-compressed sentences. Changes over the course of the pretest were quantified by calculating average performance for the first five sentences and final five sentences encountered in this block. Had improvements been random, roughly equal proportions of individuals from each group would have been expected to improve and to worsen between the two sets of sentences. Although this was not the case for the YA (50/57, 88% improved, binomial $p < .001$) and ONH (40/52, 77% improved, binomial $p < .001$) groups, in the OHI group, only 22/36 of the participants (61%, binomial $p = .24$) exhibited more accurate recognition in the final than in the first set of sentences. Thus, over the course of the pretest it seems that early learning occurred in significant proportion of participants only in the two normal hearing groups.

Although early learning was observed in young adults and in older adults with normal hearing, the magnitude of learning was nevertheless influenced by both age and hearing loss (Kruskal–Wallis $\chi^2(2) = 24.13$, $p < .001$ with mean ranks of 95, 70, and 51 in the YA, ONH, and OHI groups, respectively). As shown in Figure 5, early learning (operationally defined by the increase in recognition accuracy from the first five time-compressed sentences to the final five) decreased with both age and hearing loss.

Although more than 75% of young adults improved by more than 10% (with a median improvement of 30%) and half of the ONH and OHI participants improved by less than 10%. Planned comparisons suggest that the ONH group improved significantly less than the YA group (Mann–Whitney $U = 1,064$, $Z = -3.06$, $p = .002$) and that the OHI group improved less than the ONH group (Mann–Whitney $U = 747.5$, $Z = -2.08$, $p = .037$). Together, these data suggest that in contrast to earlier studies (Gordon-Salant et al., 2010; Peelle & Wingfield, 2005), age and age-related hearing loss both influenced
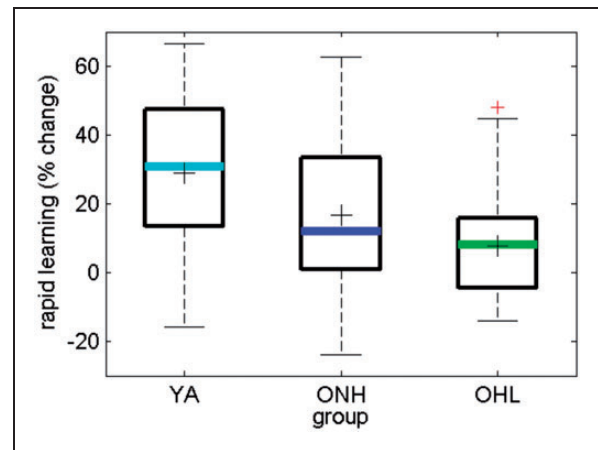
Figure 5. Early learning as a function of group. For each group, the interquartile range (25th–75th percentile) of learning scores is plotted with a black box. The thick line within each box marks the group's median. Whiskers are 1.5 the interquartile range. The black + signs within each box mark the group means; red + signs mark outliers.
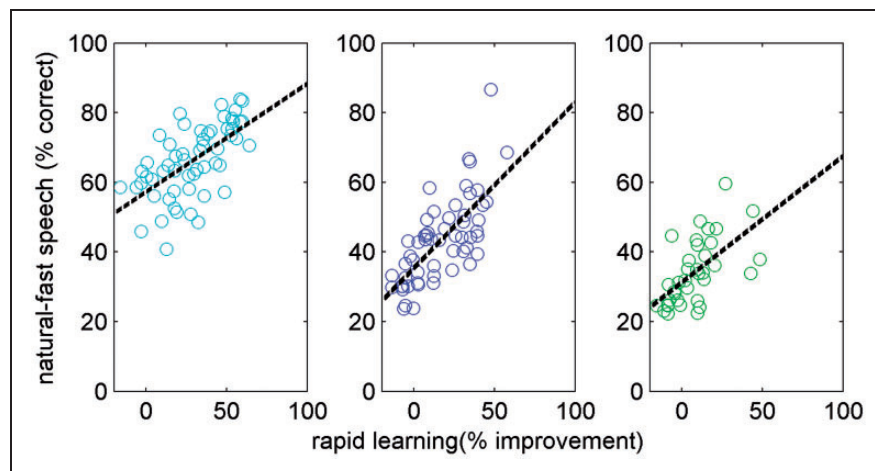

Figure 4. Early learning of time-compressed speech. Each line shows mean individual performance during the first and final blocks of five sentences encountered during the pretest in (left to right) young adults, older adults with normal hearing, and older adults with hearing impairment.

the ability of listeners to rapidly adjust to distorted speech.

Finally, previous studies suggested that the recognition of time-compressed and natural-fast speech are correlated, and the current pretest data replicate this observation (Pearson correlations between recognition accuracy in the first five time-compressed sentences and all natural-fast sentences: YA: $r = .58$, $p < .001$; ONH: $r = .76$, $p < .001$; and OHI: $r = .78$, $p < .001$). In addition, we now ask whether the magnitude of rapid learning of time-compressed speech also accounts for any variance in recognition of natural-fast speech. To this end, for each group, a regression model was constructed with the baseline recognition of natural-fast speech (averaged across the two talkers) as a dependent variable and two predictors: recognition of the first time-compressed sentences and the magnitude of learning over 20 time-compressed sentences. Overall, these models accounted for significant proportions of the variance across groups, YA: $R^2 = .59$, $F(2, 56) = 38.51$, $p < .001$; ONH: $R^2 = .80$, $F(2, 55) = 103.46$, $p < .001$; OHI: $R^2 = .75$, $F(2, 35) = 49.12$, $p < .001$. Importantly, early learning accounted for unique variance in all three groups with contributions of 25%, 22%, and 14% in the YA, ONH, and OHI groups, respectively (YA: $\beta = .51$, $t = 5.77$, $p < .001$; ONH: $\beta = .48$, $t = 7.51$, $p < .001$; and OHI: $\beta = .38$, $t = 4.30$, $p < .001$). This contribution is illustrated in Figure 6 after accounting for the first-order correlations between baseline performance on time-compressed speech and the two other model variables.

## Learning of Time-Compressed Speech During the Training Phase

Due to the structure of training, age effects were not hypothesized for this phase. Recognition thresholds decreased (improved) with training (see Figure 7), as expected from previous studies. Despite large individual differences among older adults, a substantial proportion of individuals within the two older groups nevertheless appeared to improve during training. To determine whether those decreases differed between the three study groups, the slopes of the individual learning curves were calculated (as the least-square fit of threshold to training block) and compared across groups. The mean rank (41, 43, and 22 in the YA, ONH, and OHI groups, respectively; lower ranks indicate smaller, more negative slopes) differed across groups (Kruskal–Wallis $\chi^2(2) = 11.90$, $p = .003$). This difference stems from a hearing effect, as the mean ranks of the OHI group was significantly smaller than that of the ONH group (Mann–Whitney $U = 123$, $Z = -3.02$, $p = .003$). No difference was found between the YA and ONH groups (Mann–Whitney $U = 367$, $Z = -.40$, $p = .688$). Note that although these data suggest that more learning was observed in the OHI group, this could simply reflect their poorer starting performance. The critical question is thus whether this experience resulted in greater pretest to posttest changes in trained OHI individuals than in untrained ones, which will be addressed in the following sections.

Figure 6. Recognition of natural-fast speech as a function of early (rapid) learning of time-compressed speech. Left to right: YA, ONH, and OHI. The values of the two variables were adjusted with baseline recognition of time-compressed speech as a covariate (using the formula $Y_i' = Y_i - B \times (X_i - E(X))$). Y' is the adjusted value, Y is the original value, X is the covariate, $E(X)$ is the mean of X, and B is the slope of the Y vs. X regression line). The adjustment was made to illustrate the relationship between the degree of rapid learning on time-compressed speech and the recognition of natural-fast speech. The dashed lines show the linear regression of adjusted natural-fast speech recognition and adjusted rapid learning.
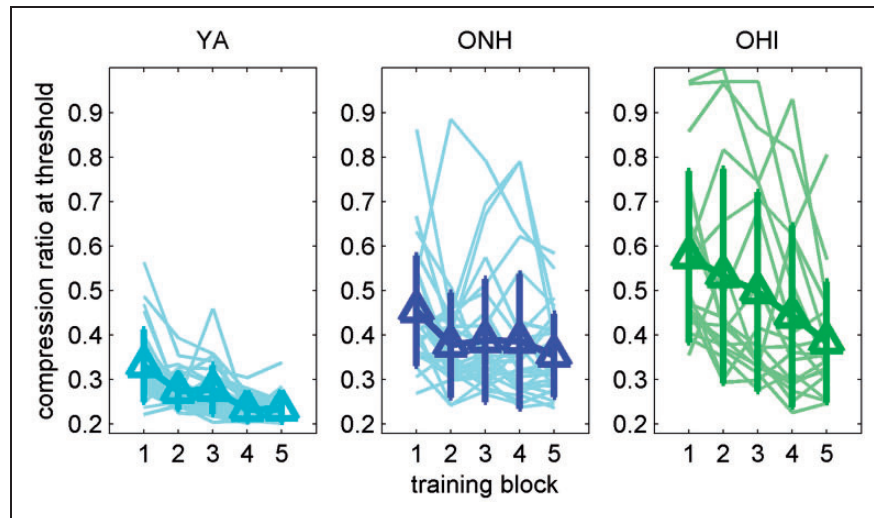
Figure 7. Verification thresholds during training as a function of training block. Thin lines denote individual data; triangles show group means over blocks of 60 trials. Error bars are ± I standard deviation.
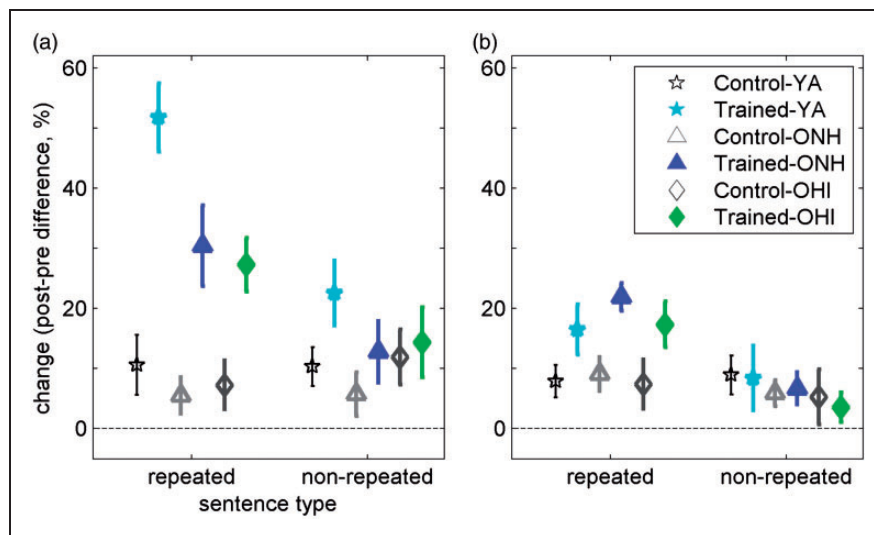
Figure 8. Pre-to-post-test changes in (a) time-compressed and (b) natural-fast speech recognition as a function of age, hearing, training, and sentence type. Group means and 95% confidence intervals of trained (filled symbols) and untrained control (empty symbols) groups are shown. On repeated time-compressed sentences that appeared during the training phase, all trained groups improved (on average) more than control groups between the pretest and the posttest, but the effect was smaller in older adults. Transfer to nonrepeated sentences was, however, limited to trained young adults. When sentences were repeated from the training phase as natural-fast speech, significant transfer of learning was observed across trained groups. However, no transfer to nonrepeated sentences was observed in any of the groups.

Learning and Transfer Effects: Time-Compressed Speech

Learning and transfer are defined as greater pretest to posttest improvements in trained than in untrained (control) participants. It was hypothesized that both learning of time-compressed sentences and transfer to new time-compressed sentences should decrease with age.

To determine whether these gains differed as a function of age and hearing loss, difference scores were calculated for corresponding blocks of sentences in the pretest and posttest phases. These difference scores, shown in Figure 8(a), were then submitted to a mixed-model ANOVA with one within-subject factor (sentence type: repeated from training or nonrepeated from training) and two between-subject factors—group (YA, ONH,

Table 3. Omnibus Analysis of Variance Outcomes for Time-Compressed Speech Learning as a Function of Group, Training, and Sentence Type.

Effect	F	df	p	partial η^2
Sentence type	67.5	1,142	<.001	.322
Group	18.5	2,142	<.001	.206
Training	130.1	1,142	<.001	.478
Sentence Type × Group	7.3	2,142	.001	.093
Sentence Type × Training	91.4	1,142	<.001	.392
Group × Training	8.6	2,142	<.001	.108
Sentence Type × Group × Training	3.1	2,142	.048	.042

Table 4. Outcomes for Planned Comparison of Time-Compressed Speech Learning as a Function of Age, Training, and Sentence Type.

Effect	F	df	p	partial η^2
Age-group	30.3	1,108	<.001	.22
Training	132.4	1,108	<.001	.55
Sentence type	79.2	1,108	<.001	.42
Age × Training	8.4	1,108	.005	.07
Sentence Type × Training	78.6	1,108	<.001	.42
Age × Sentence Type × Training	4.47	1,108	.037	.04

and OHI) and training (trained and control). The effects of this ANOVA are summarized in Table 3. Consistent with Figure 8(a), the main effects and all interaction terms were significant, suggesting that pre-to-post-test gains were influenced by group, training, and sentence type. To unpack the effects of age and age-related hearing loss, this omnibus ANOVA was followed by two planned comparisons—between the YA and ONH groups and between the ONH and the OHI groups. Each of these comparisons involved a two sentence types by two groups by two training statuses ANOVA.

Age effects (YA vs. ONH): As in the omnibus ANOVA, all main effects were significant (see Table 4). All interaction terms that involve training were also significant (see Table 4). Together these effects suggest that whereas across ages, gains between the pretest and posttest were greater in participants who went through training, these gains were larger among young adults. The interaction of sentence type and training suggests that although transfer of learning to untrained items was observed, this transfer was only partial because the effect of training was larger for sentences that were repeated from the training phase than for nonrepeated sentences. Finally, the highest order interaction of age, training, and sentence type suggests that whereas the effects of training were overall weaker in older adults, transfer to nonrepeated sentences was especially reduced in this group (see Figure 8(a), right section).

Hearing effects (ONH vs. OHI groups): In contrast to the age effects, none of the effects that involved hearing status was significant, hearing: $F(1, 88) = 0.71$, $p = .40$, partial $\eta^2 = .008$; Hearing × Training: $F(1, 88) = 1.58$, $p = .21$, partial $\eta^2 = .018$; Sentence Type × Hearing: $F(1, 88) = 3.16$, $p = .079$, partial $\eta^2 = .035$; Sentence Type × Training × Hearing: $F(1, 88) = .002$, $p = .96$, partial $\eta^2 = 0$.

Together, these data are consistent with the interpretation that the effects of training were similar in the two groups of older adults, regardless of their hearing status.

Although significant gains were observed on sentences that were encountered during training, it seems that older adults fail to achieve even the partial transfer to untrained sentences that is seen in younger adults.

Transfer to Natural-Fast Speech

Pre-to-post-test changes in the recognition of natural-fast speech conditions are shown in Figure 8(b). These were also hypothesized to diminish with age. Similar to the case of time-compressed speech, data were submitted to an omnibus ANOVA with sentence type (repeated and nonrepeated) as a within-subject factor and two between-subject factors: training (training, no-training) and group (YA, ONH, and OHI). The main effect of training and the Training × Sentence Type interaction were both significant, $F(1, 142) = 36.11$, $p < .001$, partial $\eta^2 = 0.203$ and $F(1, 142) = 83.6$, $p < .001$, partial $\eta^2 = 0.203$, respectively. Neither of the other interactions involving training (Group × Training and Group × Training × Sentence Type) approached significance. These outcomes are consistent with the idea that whereas training on time-compressed speech yielded significant transfer to natural-fast speech (on which there was no training), this transfer was similar across the three groups of listeners and limited to the sentences that were encountered during training. This finding is consistent with the idea that transfer occurred across stimuli that shared high-level representations with the sentences that were presented during training in time-compressed format, despite the fact that these sentences did not share low-level, acoustic representations (time-compressed vs. natural-fast speech).

The planned comparisons between the YA and ONH groups and between the ONH and OHI groups also support the conclusions from the omnibus ANOVA. First, comparing the YA and ONH groups suggested that the effect of training and the Training × Sentence Type interaction were both significant, $F(1, 108) = 33.75$, $p < .001$, partial $\eta^2 = 0.238$ and $F(1, 108) = 63.34$, $p < .001$, partial $\eta^2 = 0.370$, respectively. The effects of age,

Age × Training and Age × Training × Sentence Type did not approach significance ($F < 2$, $p > 0.14$), suggesting that the lack of transfer to unrepeated natural-fast sentences was not age dependent. Second, comparing the two groups of older adults also suggests significant training and Training × Sentence Type effects, $F(1, 88) = 27.19$, $p < .001$, partial $\eta^2 = 0.236$ and $F(1, 88) = 58.06$, $p < .001$, partial $\eta^2 = 0.397$, respectively. Although the effects of hearing and sentence type were both significant, $F(1, 88) = 5.80$, $p = .018$, partial $\eta^2 = 0.062$ and $F(1, 88) = 121.33$, $p < .001$, partial $\eta^2 = 0.579$, respectively, none of the interactions involving hearing status approached significance ($F < 1.68$, $p > .19$). Therefore, there is no indication that age-related hearing loss interfered with the transfer of learning to natural-fast speech (limited as it was to sentences repeated from training).

Discussion

The effects of age and hearing loss on the perceptual learning of rapid speech were investigated. Both had a negative influence on the early learning of time-compressed speech (see Figure 5). Further training induced significant learning in all groups (see Figure 7), but both learning and transfer to unrepeated time-compressed sentences were weaker in older than in younger adults (see Figure 8(a)). Age and hearing had no effect on the transfer of learning to repeated sentences presented as natural-fast speech (see Figure 8(b)). Although hearing loss had a negative effect on early learning, it had no additional effects on learning and transfer following longer training.

Early Learning of Time-Compressed Speech

Interactive models of speech perception and the RHT suggest that on initial encounter with time-compressed speech, the unfamiliar acoustics require greater than normal reliance on lower level speech cues (Ahissar et al., 2009; Mattys et al., 2009; Mattys et al., 2012). The literature reviewed in the Introduction also suggests that older adults should be less able than younger adults to switch from the "default" mode of relying on high-level speech cues to the use of lower level acoustic representations. Based on these, we hypothesized that early learning of time-compressed speech decreases with age. Consistent with this hypothesis, we found that age has a negative influence on the early learning of time-compressed speech. A smaller proportion of older adults exhibited any amount of early learning compared with young adults, and the magnitude of improvement declined with age, as shown in Figure 5. Although 75% of young adults improved their performance by 13% or more, only approximately 50% of the older adults with normal hearing improved by similar amounts. Because baseline

performance was equated across groups, we argue that these group differences are attributable to real differences in the ability to rapidly learn perceptually difficult speech. Although further studies are needed to establish causation, the high correlations between early learning of time-compressed speech and the perception of natural-fast speech are also consistent with this proposal.

While older adults can rapidly learn to access lower level speech representations that are needed to process rapid speech, the current findings suggest that this process is less effective than in younger adults, likely due to cognitive slowing (Wingfield et al., 1999) or other cognitive factors. That hearing loss has an additional effect on early learning (see Figure 5) is also consistent with the interpretation that learning acoustically challenging speech could depend on the availability of low-level speech cues. Alternatively, age-related declines in learning could stem from differences in the processing of sentence context, but this seems unlikely because the processing of time-compressed speech in older adults tends to benefit from increased linguistic context (Stine & Wingfield, 1987; Stine et al., 1986; Wingfield et al., 1985).

The present findings appear inconsistent with the outcomes of previous studies on early learning of perceptually difficult speech in older adults (Golomb et al., 2007; Gordon-Salant et al., 2010; Peelle & Wingfield, 2005). Gordon-Salant et al. (2010) investigated adaptation to accented speech over the course of four lists of either 40 monosyllabic words or 40 low-context sentences. They found equivalent learning effects in young adults and in older adults with hearing impairment. One explanation for the disparity between the outcomes of the two studies is that aging and age-related hearing loss have different effects on learning time-compressed and accented speech. Alternatively, listening to artificially compressed speech creates a greater sensory load than listening to naturally accented speech. In this case, listeners would need to rely to a greater extent on low-level cues when listening to time-compressed speech than when listening to accented speech. If older adults give more weight to high-level information and less weight to sensory information (Mattys & Scharenborg, 2014), age-related differences in learning are expected to be greater in the case of time-compressed than in the case of accented speech.

Earlier studies on early learning of time-compressed speech also reported equivalent levels of learning in older and younger adults (Golomb et al., 2007; Peelle & Wingfield, 2005). We believe that the process of individual adjustment of compression rates used by Peelle and Wingfield may have masked age differences at the onset of learning. Their own findings of age-related differences in the first learning trials as well as of reduced transfer to a new compression rate in older adults are nevertheless

consistent with the idea that age diminishes the capacity to rely on low-level cues to resolve acoustic challenges such as time compression. Differences between the current findings and those of Golomb et al. (2007) are harder to reconcile, but we note two possible explanations. First, Golomb et al. administered four adaptation conditions per participant. As they did not report data from the first condition separately, their report may have masked age-related group differences in the early stages of learning. Second, half of the sentences in the early learning phase of this study were semantically anomalous. If, as suggested earlier, older adults pay more attention to high-level cues at the expense of lower level ones, or have difficulty inhibiting competing high-level information (Mattys & Scharenborg, 2014; Sommers & Danielson, 1999), this could have put them in a learning disadvantage at this study. Testing the validity of this interpretation is left for future studies because we do not have sufficient data to analyze normal and anomalous sentences separately.

Learning and Transfer Following Training

According to the RHT, learning and its transfer to new tokens depend on the degree of similarity between the materials used in training and the materials used to assess training outcomes. Learning is expected to transfer across stimuli that share the same representations with the trained materials (Ahissar et al., 2009). We thus hypothesized that the effects of age on learning and transfer should increase with increasing discrepancies between the representations of the trained and the untrained materials. Specifically, we hypothesized that age effects should be minimal for time-compressed stimuli that were encountered in training but increase as the overlap between trained and untrained stimuli decreases. As discussed in the following paragraphs, the current data are partially consistent with this hypothesis (Figure 8(a)) in showing robust learning but reduced transfer in older adults with either normal or impaired hearing. In contrast to early learning, age-related hearing loss did not interfere with either learning or transfer beyond the age effect, suggesting that impoverished sensory input does not in and of itself interfere with learning. Consistent with previous findings from multiday training (Karawani et al., 2016), in this study, learning of time-compressed tokens repeated from training was quite robust in older adults with and without hearing impairment. However, although robust, this learning was nevertheless weaker than learning in young adults. Thus, what this study adds, is that not only are older adults at a learning disadvantage on initial encounter with rapid speech (early/rapid learning), but additional training results in smaller gains in older adults. Sadly, this suggests that training alone is not likely to resolve

speech perception deficits in older adults, and that additional forms of interventions should be considered (Kricos & Holmes, 1996).

Consistent with our initial hypothesis, the transfer of learning to nonrepeated sentences presented in time-compressed speech was altogether absent in older adults. These nonrepeated stimuli do not share high-level representations with the trained stimuli. They do, however, share the same acoustic structure with the trained materials (both are time compressed). As discussed earlier, the use of these cues appears to alter with age, putting older adults at a disadvantage (Mattys & Scharenborg, 2014; Sommers, 1996; Sommers & Danielson, 1999). Similarly, reduced generalization to an untrained compression rate was observed by Peelle and Wingfield (2005) following brief adaptation. We have suggested that training makes the low-level speech cues, that are required for the successful perception of time-compressed speech, more accessible, and that after training young adults can utilize these low-level (sub-lexical) representations even with new sentences (Banai & Lavner, 2014). If older adults give lesser weight to these low-level representations (Mattys & Scharenborg, 2014; Sommers & Danielson, 1999), or if their ability to attend to these cues is reduced due to either sensory or attention-switching difficulties (Scharenborg et al., 2015), one could expect reduced generalization compared with young adults.

On the other hand, and consistent with the hypothesis that transfer requires some representational similarity across stimuli, no transfer of learning occurred to nonrepeated natural-fast sentences in either younger or older adults. This finding is inconsistent with findings in young adults that initial training with time-compressed speech does generalize to novel natural-fast items (Adank & Janse, 2009). It is, however, consistent with the outcomes of multiple other training studies that reported limited generalization to novel items (Banai & Lavner, 2014; Burk et al., 2006; Karawani et al., 2016). For example, although speech in noise learning was not specific to the trained stimuli, Karawani et al. (2016) found that it was highly specific to the trained task. As we have previously found that transfer to untrained time-compressed sentences increased with longer training (Banai & Lavner, 2014), it is also possible that with more training listeners could become able to exploit the similarities between time-compressed and natural-fast speech. Testing this idea requires further studies, perhaps with matched rates of time-compressed and natural-fast speech.

Interestingly, whereas age and hearing had significant negative effects on early learning, they did not diminish the amount of improvement during the adaptive training session (Figure 7). One possible reason for this difference is that learning could be simply slower in older adults and the longer duration of the training session allowed

them to catch up. Alternatively, adaptively changing speech rates during training could support learning in older adults by allowing gradual adaptation to increasing speech rates (see Gabay et al., 2017 for comparison of adaptive and nonadaptive training protocols). Finally, the semantic verification task used during training could be less taxing for older adults than the sentence-writing task used to estimate early learning. Testing these alternatives is beyond the scope of this study. That learning during training was not negatively influenced by either age or hearing loss suggests that although the training task was probably more effortful for older than for younger participants (verification thresholds were consistently higher in the older groups), this greater effort did not prevent the recruitment of the resources needed to initiate learning. Furthermore, the learning curves provide no indication of increased fatigue in the older groups toward the end of the session. Although participants with age-related hearing loss improved more than the other two groups during training, their overall performance remained poorer, consistent with the additional sensory deficits characteristic of this group.

We note three limitations of this study. First, training and testing occurred on the same day. Thus, it is not clear whether the robust learning observed in older adults is retained over time. Findings from studies with longer training protocols or multiple test sessions (Burk et al., 2006; Golomb et al., 2007; Karawani et al., 2016) show that learning in older adults certainly accumulates over training sessions and retains past the end of training, regardless of hearing status. Further studies are required to determine whether this is the case for a single dose of training, as administered in this study. Second, an "active" control group was not included in this study. Thus, training effects may have stemmed from the greater time trained participants invested in the study, rather than from training itself. We find this unlikely because substantial training effects were found even when control participants spent the time between tests sitting quietly in the lab (Gabay et al., 2017) or chatting with the experimenters (Karawani, Bitan, et al., 2017). Third, the participants in the two groups of older adults had above median education (for their cohort) and high Mini-Mental State Examination scores. Hence, our findings may not be representative of all older adults. Nevertheless, as age effects emerged even with this well-selected sample, it seems likely that our findings even under represent the true extent of the effect of aging on perceptual learning of time-compressed speech.

Conclusions and Implications

Three indices of perceptual learning were explored, and age interfered with all three. First, age and age-related hearing loss both had a negative effect on the early learning of time-compressed speech. Furthermore, across groups, the amount of early learning was a significant predictor of the ability to recognize natural-fast speech. Although further studies are needed to determine the causal direction, we tentatively suggest that age-related declines in learning might contribute to the perceptual difficulties of older adults with and without hearing loss. Second, although older adults with and without hearing loss learned substantially during training, when compared to age- and hearing-matched untrained participants, gains were significantly smaller in older adults with either normal or impaired hearing. Third, in older adults, the transfer of learning following more intensive training followed a specific pattern, which is consistent with the notion that age-related declines in speech processing are associated with age differences in the weights assigned to the top-down and bottom-up processes. This suggests that even after training, initial encounter with acoustically difficult speech, which is acoustically similar to the trained items is still arduous for older adults, regardless of hearing status. Combined with earlier findings of limited transfer of learning following perceptual training of different types, our findings suggest that attempts to improve speech perception in ecological conditions with training will be successful only if training can bolster the rapid learning of newly encountered speech.

The use of hearing aids can potentially improve speech perception in complex listening situations and augment the effects of auditory training (Lavie et al., 2013; Lavie, Banai, Karni, & Attias, 2015). It remains to be seen whether this augmentation makes the low-level cues that are presumably amplified by the hearing aids more useful for older adults in the context of rapid speech perception. Another issue left for future studies is whether the effect of hearing loss on early learning in older adults with presbycusis is also characteristic of groups with hearing impairments of different etiologies such as young adults and children.

Notes

1. Perception and learning are intimately linked (see Ahissar et al., 2009 for review), thus preserved learning could to some extent offset age-related sensory declines, but declines in learning are likely to exacerbate perceptual declines.
2. This is a more specific prediction than what would have been predicted based on theories of age-related perceptual or cognitive slowing, because if older adults take longer than younger adults to complete each speech-related computation, the cumulative effect of the slowing will result in a breakdown in speech processing (Wingfield et al., 1999) and interfere with all aspects of learning.

ORCID iD

Karen Banai http://orcid.org/0000-0002-2990-0470

References

Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *Journal of the Acoustical Society of America*, *126*, 2649–2659. doi:10.1121/1.3216914

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, *90*, 5718–5722.

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *364*, 285–299. doi:10.1098/rstb.2008.0253

Alain, C., McDonald, K. L., Ostroff, J. M., & Schneider, B. (2004). Aging: A switch from automatic to controlled processing of sounds? *Psychology and Aging*, *19*, 125–133. doi:10.1037/0882-7974.19.1.125

Alain, C., & Snyder, J. S. (2008). Age-related differences in auditory evoked responses during rapid perceptual learning. *Clinical Neurophysiology*, *119*, 356–366. doi:10.1016/j.clinph.2007.10.024

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 4357–4362. doi:10.1073/pnas.1213555110

Banai, K., & Lavner, Y. (2012). Perceptual learning of time-compressed speech: More than rapid adaptation. *PloS One*, *7*, e47099. doi:10.1371/journal.pone.0047099

Banai, K., & Lavner, Y. (2014). The effects of training length on the perceptual learning of time-compressed speech and its generalization. *Journal of the Acoustical Society of America*, *136*, 1908. doi:10.1121/1.4895684

Banai, K., & Lavner, Y. (2016). The effects of exposure and training on the perception of time-compressed speech in native versus nonnative listeners. *Journal of the Acoustical Society of America*, *140*, 1686. doi:10.1121/1.4962499

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729. doi:10.1016/j.cognition.2007.04.005

Burk, M. H., Humes, L. E., Amos, N. E., & Strauser, L. E. (2006). Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear and Hearing*, *27*, 263–278. doi:10.1097/01.aud.0000215980.21158.a2

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*, 3647–3658. doi:10.1121/1.1815131

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, *229*, 132–147. doi:10.1016/j.heares.2007.01.014

Divenyi, P. L., & Haupt, K. M. (1997). Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects. *Ear and Hearing*, *18*, 42–61.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature communications*, *7*, 12241. doi:10.1038/ncomms12241

Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology-Human Perception and Performance*, *23*, 914–927. doi:10.1037/0096-1523.23.3.914

Duthey, B. (2013). *Background paper 6.21 hearing loss*. Geneva, Switzerland: WHO.

Eckert, M. A., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2008). Age-related effects on word recognition: Reliance on cognitive control systems with structural declines in speech-responsive cortex. *Journal of the Association of Research in Otolaryngology*, *9*, 252–259. doi:10.1007/s10162-008-0113-3

Erb, J., & Obleser, J. (2013). Upregulation of cognitive control networks in older adults' speech comprehension. *Frontiers in Systems Neuroscience*, *7*, 116. doi:10.3389/fnsys.2013.00116

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. doi:10.1016/0022-3956(75)90026-6

Gabay, Y., Karni, A., & Banai, K. (2017). The perceptual learning of time-compressed speech: A comparison of training protocols with different levels of difficulty. *PloS One*, *12*, e0176488. doi:10.1371/journal.pone.0176488

Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology*, *14*, 29–56. doi:10.1146/annurev.ps.14.020163.000333

Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *Journal of the Acoustical Society of America*, *121*, 1701–1708. doi:10.1121/1.2436635

Goman, A. M., & Lin, F. R. (2016). Prevalence of hearing loss by severity in the United States. *American Journal of Public Health*, *106*, 1820–1822. doi:10.2105/AJPH.2016.303299

Gopinath, B., Rochtchina, E., Wang, J. J., Schneider, J., Leeder, S. R., & Mitchell, P. (2009). Prevalence of age-related hearing loss in older adults: Blue Mountains Study. *Archives of Internal Medicine*, *169*, 415–416. doi:10.1001/archinternmed.2008.597

Gordon-Salant, S., & Fitzgibbons, P. J. (2001). Sources of age-related recognition difficulty for time-compressed speech.

Journal of Speech, Language, and Hearing Research, *44*, 709–719. doi:10.1044/1092-4388

Gordon-Salant, S., Fitzgibbons, P. J., & Friedman, S. A. (2007). Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, *50*, 1181–1193. doi:10.1044/1092-4388 (2007/082)

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010). Short-term adaptation to accented English by younger and older adults. *Journal of the Acoustical Society of America*, *128*, EL200–204. doi:10.1121/1.3486199

Goy, H., Pelletier, M., Coletta, M., & Pichora-Fuller, M. K. (2013). The effects of semantic context and the type and amount of acoustic distortion on lexical decision by younger and older adults. *Journal of Speech, Language, and Hearing Research*, *56*, 1715–1732. doi:10.1044/1092-4388(2013/12-0053)

Green, C. S., Banai, K., Lu, Z. L., & Bevalier, D. (2018). Perceptual learning. In J. T. Wixted, & J. Serences (Eds.), *Stevens' handbook of experimental psychology and cognitive neuroscience, sensation, perception, and attention.* (Vol 2, pp. 755–802, 4th ed.). Hoboken, NJ: John Wiley & Sons. doi:10.1002/9781119170174

Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: Insights from behavioral, computational, and neuroscience research. *Frontiers in Systems Neuroscience*, *7*, 126. doi:10.3389/fnsys.2013.00126

Henderson Sabes, J., & Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (LACE) training. *International Journal of Audiology*, *46*, 374–383. doi:10.1080/14992020701297565

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804. doi:10.1016/S0896-6273(02)01091-7

Humes, L. E., Burk, M. H., Strauser, L. E., & Kinney, D. L. (2009). Development and efficacy of a frequent-word auditory training protocol for older adults with impaired hearing. *Ear and Hearing*, *30*, 613–627. doi:10.1097/AUD.0b013e3181b00d90

Humes, L. E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *Journal of Speech and Hearing Research*, *34*, 686–693. doi:10.1044/jshr.3403.686

Humes, L. E., & Roberts, L. (1990). Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility. *Journal of Speech and Hearing Research*, *33*, 726–735. doi:10.1044/jshr.3304.726

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology*, *65*, 1563–1585. doi:10.1080/17470218.2012.658822

Karawani, H., Bitan, T., Attias, J., & Banai, K. (2016). Auditory perceptual learning in adults with and without age-related hearing loss. *Frontiers in Psychology*, *6*, 2066. doi:10.3389/fpsyg.2015.02066

Karawani, H., Lavie, L., & Banai, K. (2017). Short-term auditory learning in older and younger adults. *Proceedings of the International Symposium on Auditory and Audiological Research*, *6*, 1–8. Retrieved from https://proceedings.isaar.eu/index.php/isaarproc/article/view/2017-01.

Kricos, P. B., & Holmes, A. E. (1996). Efficacy of audiologic rehabilitation for older adults. *Journal of the American Academy of Audiology*, *7*, 219–229.

Lavie, L., Attias, J., & Karni, A. (2013). Semi-structured listening experience (listening training) in hearing aid fitting: Influence on dichotic listening. *American Journal of Audiology*, *22*, 347–350. doi:10.1044/1059-0889(2013/12-0083)

Lavie, L., Banai, K., Karni, A., & Attias, J. (2015). Hearing aid-induced plasticity in the auditory system of older adults: Evidence from speech perception. *Journal of Speech, Language, and Hearing Research*, *58*, 1601–1610. doi:10.1044/2015_JSLHR-H-14-0225

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception and Psychophysics*, *63*, 1279–1292. doi:10.3758/BF03194543

Levitt, H. (1971). Transformed up-down methods in psycho-acoustics. *Journal of the Acoustical Society of America*, *49*, 467–477. doi:10.1121/1.1912375

Maddox, W. T., Chandrasekaran, B., Smayda, K., & Yi, H. G. (2013). Dual systems of speech category learning across the lifespan. *Psychology and Aging*, *28*, 1042–1056. doi:10.1037/a0034969

Martin, J. S., & Jerger, J. F. (2005). Some effects of aging on central auditory processing. *Journal of Rehabilitation Research and Development*, *42*, 25–44. doi:10.1682/JRRD.2004.12.0164

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, *59*, 203–243. doi:10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Carroll, L. M., Li, C. K. W., & Chan, S. L. Y. (2010). Effects of energetic and informational masking on speech segmentation by native and non-native speakers. *Speech Communication*, *52*, 887–899. doi:10.1016/j.specom.2010.01.005

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978. doi:10.1080/01690965.2012.705006

Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. *Psychology and Aging*, *29*, 150–162. doi:10.1037/a0035387

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*, 145–160. doi:10.1016/j.jml.2011.04.004

Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: The case of speech in noise. *PLoS Biology*, *6*, e126. doi:10.1371/journal.pbio.0060126

Nahum, M., Nelken, I., & Ahissar, M. (2010). Stimulus uncertainty and perceptual learning: Similar principles govern auditory and visual learning. *Vision Research*, *50*, 391–401. doi:10.1016/j.visres.2009.09.004

Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience*, 8, 628. doi:10.3389/fnhum.2014.00628

Pallier, C., Sebastian-Galles, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: A cross-linguistic study. *Memory and Cognition*, 26, 844–851. doi:10.3758/BF03211403

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1315–1330. doi:10.1037/0096-1523.31.6.1315

Peelle, J. E., & Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends in Neurosciences*, 39, 486–497. doi:10.1016/j.tins.2016.05.001

Pichora-Fuller, M. K., & Levitt, H. (2012). Speech comprehension training and auditory and cognitive processing in older adults. *American Journal of Audiology*, 21, 351–357. doi:10.1044/1059-0889(2012/12-0025)

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, 97, 593–608. doi:10.1121/1.412282

Prior, A., & Bentin, S. (2006). Differential integration efforts of mandatory and optional sentence constituents. *Psychophysiology*, 43, 440–449. doi:10.1111/j.1469-8986.2006.00426.x

Reinisch, E., & Holt, L. L. (2014). Lexically guided phonetic retuning of foreign-accented speech and its generalization. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 539–555. doi:10.1037/a0034409

Roth, T. N., Hanebuth, D., & Probst, R. (2011). Prevalence of age-related hearing loss in Europe: A review. *European Archives of Oto-Rhino-Laryngology*, 268, 1101–1107. doi:10.1007/s00405-011-1597-8

Samuel, A. G. (2011). Speech perception. *Annual Review of Psychology*, 62, 49–72. doi:10.1146/annurev.psych.121208.131643

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics*, 71, 1207–1218. doi:10.3758/APP.71.6.1207

Scharenborg, O., & Janse, E. (2013). Comparing lexically guided perceptual learning in younger and older listeners. *Attention Perception and Psychophysics*, 75, 525–536. doi:10.3758/s13414-013-0422-4

Scharenborg, O., Weber, A., & Janse, E. (2015). The role of attentional abilities in lexically guided perceptual learning by older listeners. *Attention Perception and Psychophysics*, 77, 493–507. doi:10.3758/s13414-014-0792-2

Schneider, B. A., & Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: Implications for speech perception. *Seminars in Hearing*, 22, 227–240. doi:10.1055/s-2001-15628

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, 32, 8443–8453. doi:10.1523/JNEUROSCI.5069-11.2012

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2014). Top-down influences of written text on perceived clarity of degraded speech. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 186–199. doi:10.1037/a0033206

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, 11, 333–341. doi:10.1037/0882-7974.11.2.333

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, 14, 458–472. doi:10.1037/0882-7974.14.3.458

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex*, 22, 1180–1190. doi:10.1093/cercor/bhr196

Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, 2, 272–279. doi:10.1037/0882-7974.2.3.272

Stine, E. L., Wingfield, A., & Poon, L. W. (1986). How much and how fast: Rapid processing of spoken language in later adulthood. *Psychology and Aging*, 1, 303–311. doi:10.1037/0882-7974.1.4.303

Tun, P. A., Wingfield, A., Stine, E. A., & Mecsas, C. (1992). Rapid speech processing and divided attention: Processing rate versus processing resources as an explanation of age effects. *Psychology and Aging*, 7, 546–550. doi:10.1037/0882-7974.7.4.546

Vaughan, N. E., & Letowski, T. (1997). Effects of age, speech rate, and type of test on temporal auditory processing. *Journal of Speech, Language, and Hearing Research*, 40, 1192–1200. doi:10.1044/jslhr.4005.1192

Verhelst, W., & Roelands, M. (1993). *An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech*. In IEEE international conference on acoustics, speech, and signal processing, Minneapolis, MN (pp. 554–557). Minneapolis, MN: IEEE. doi: 10.1109/ICASSP.1993.319366

Walden, B. E., Erdman, S. A., Montgomery, A. A., Schwartz, D. M., & Prosek, R. A. (1981). Some effects of training on speech recognition by hearing-impaired adults. *Journal of Speech and Hearing Research*, 24, 207–216. doi:10.1044/jshr.2402.207

Wingfield, A., Poon, L. W., Lombardi, L., & Lowe, D. (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology*, 40, 579–585. doi:10.1093/geronj/40.5.579

Wingfield, A., & Tun, P. A. (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Seminars in Hearing*, 22, 287–302. doi:10.1055/s-2001-15632

Wingfield, A., Tun, P. A., Koh, C. K., & Rosen, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology and Aging*, 14, 380–389. doi:10.1037/0882-7974.14.3.380