Improving Older Adults' Understanding of Challenging Speech: Auditory training, rapid adaptation and perceptual learning

Rebecca E. Bieber, **Sandra Gordon-Salant**

Department of Hearing and Speech Sciences, 0100 LeFrak Hall, 7251 Preinkert Drive, University of Maryland, College Park, MD 20742

Abstract

The literature surrounding auditory perceptual learning and auditory training for challenging speech signals in older adult listeners is highly varied, in terms of both study methodology and reported outcomes. In this review, we discuss some of the pertinent features of listener, stimulus, and training protocol. Literature regarding the elicitation of auditory perceptual learning for time-compressed speech, non-native speech, and noise-vocoded speech is reviewed, as are auditory training protocols designed to improve speech-in-noise recognition. The literature is synthesized to establish some over-arching findings for the aging population, including an intact capacity for auditory perceptual learning, but a limited transfer of learning to untrained stimuli.

1. Introduction

Speech recognition in human adults is a remarkably robust, flexible process. During spoken communication, the same target lexical items are almost never produced with identical acoustic features, both within and across individual talkers. Additionally, this time-varying, somewhat idiosyncratic speech signal can be presented in challenging environments. Degradations to the speech signal can result from a variety of naturalistic (i.e. competing talkers, reverberation, non-native speech) and artificial (noise-vocoding, time-compression) sources, which can affect the temporal or spectral characteristics of speech, or both. Despite the variation in speech production and the distortion of the signal through the listening environment, most younger listeners with normal hearing are able to overcome these

---

multiple challenges in communication effortlessly. In contrast, older adults often report difficulty in communicating under these challenging environments, and this appears to be exacerbated in older adults with age-related hearing loss (ARHL). A growing body of research suggests that older listeners may overcome these communication difficulties through training programs that promote rapid adaptation and perceptual learning (Anderson & Kraus, 2013).

Perceptual learning is defined as 'long-lasting changes to [the] perceptual system that improve [the] ability to respond to [the] environment and are caused by this environment' (Goldstone, 1998). The early, fast stage of perceptual learning, which we will refer to as *rapid adaptation*, has been explored extensively for speech stimuli. Across the literature, the terms 'rapid adaptation' and 'perceptual learning' are used with varying degrees of consistency to distinguish two processes. In the context of auditory training, both adaptation and learning are typically understood to comprise changes to the perceptual system evident as changes in behavioral performance on a given task (i.e. improvements in speech recognition, faster reaction times, etc). The most apparent distinction between 'rapid adaptation' and 'perceptual learning' is the time period which is being described. 'Rapid adaptation' often implies a shorter term change, and refers to changes that occur within a single test session or condition. Many authors report on rapid adaptation as occurring within as few as 10–20 test items (Adank & Janse, 2009; Davis et al., 2005; Golumb et al., 2007; Peelle & Wingfield, 2005). 'Perceptual learning' however, implies a longer-term change and typically refers to a comparison across test sessions separated by at least a day. Longer-term learning may be evident both as retention over time of improved performance with training stimuli, as well as generalization to stimuli that differ from those used in the training protocol. This additional learning beyond initial rapid adaptation may be related to additional training time, or latent consolidation of learning that can occur in-between test sessions (Molloy et al., 2012; Ortiz & Wright, 2010).

The two concepts can be seen as different components of the overall learning process, where rapid adaptation represents the early phase of auditory perceptual learning. Much of the speech-based auditory training literature focuses on one component or the other, but a study by Manheim et al. (2018) is an example of an examination of both adaptation and learning within the same study and participant group. In this study, the two phases of learning are referred to as 'early learning' and 'later learning.' Participants completed two sessions on two consecutive days. In the first, a pre-test (3 blocks of 20 trials, duration not reported) was conducted. In the second session, listeners completed a training session (5 blocks of 60 trials, duration 40–55 minutes) and a post-test (3 blocks of 20 trials, duration not reported). Changes in performance *within the pre-test* were interpreted to reflect rapid adaptation/early learning, and changes in performance *between the pre-test and post-test* were taken to reflect longer-term learning. [Note: detailed findings of the Manheim et al., (2018) study will be discussed in Section 2.2.]

As Manheim et al. (2018) point out, the theoretical differences separating adaptation and perceptual learning are not clear in the literature. When examining behavior only, it can be challenging to identify whether the mechanisms underlying the two processes are distinct. However, there are findings from electrophysiology research that can provide some insight.

Event-related potentials (ERPs) have been used to distinguish true adaptation from simple procedural learning (i.e. task familiarity, Ben-David et al., 2011), but have also been used to describe and examine differences between rapid adaptation and longer-term learning (Atienza et al., 2002; Romero-Rivas et al., 2015; Tremblay et al., 2014). Consolidation of perceptual learning and long-lasting changes to the perceptual system would be the optimal outcome for auditory training programs targeting improvements in daily communication.

To date, there have been numerous investigations of auditory perceptual learning and the efficacy of auditory training programs directed at improving speech recognition for older adults in challenging environments. For many older listeners, audibility is not the only factor hindering successful communication. Targeted auditory rehabilitative intervention may hold promise for improving communication under challenging circumstances for older listeners, but the conditions that best facilitate rapid adaptation and perceptual learning for older listeners remain unclear. However, there is great variety in the design of these studies, including characteristics of participants, training stimuli/paradigms/duration, and methods of outcome assessment. Unsurprisingly, there is also variability in the reported success of various interventions. Given this range, it is challenging to distill the literature into a clear image of which factors (listener-related or protocol-related) contribute to a benefit of training. The objective of this review paper is to consider the factors that promote rapid adaptation and long-term perceptual learning by older listeners, and to examine recent studies targeting perceptual learning for various forms of speech degradation.

1.1 Perceptual learning and adaptation

Reverse Hierarchy Theory (RHT) is a model of perceptual learning proposed by Ahissar and Hochstein (2004), originally developed in relation to the visual system, and later applied to the auditory system (Ahissar et al., 2009). RHT suggests that learning begins at the higher levels and moves in a top-down fashion (i.e. 'backward search'), where the lower levels of perception are only accessed as needed. For example, RHT would posit that, under typical conditions, a listener perceives a word as a whole unit, rather than as the sum of its acoustic features. When challenging speech is encountered, listeners may not be able to perform this high-level whole-unit perception, and must rely on the low-level acoustic features for perception of words and phrases. Over the course of perceptual learning, there is an adjustment of the weights given to task-relevant and task-irrelevant information, such that listeners are able to 'fine-tune' processing by assigning less weight to the irrelevant properties of the input and more weight to those relevant features, improving their ability to make use of the low-level information. In this way, listeners are thought to decrease their use of the low-level detail that may be inconsistent or challenging, eventually adjusting their internal high-level representations of lexical items and allowing for an overall increase in efficiency of processing as a result of learning.

Overall, this re-weighting process allows for improvements in perception and recognition over time. This high-level adjustment of internal representations has been demonstrated in online processing (Dahan et al., 2008), suggesting that listeners perform flexible real-time adjustments when communicating. The results of these adjustments are also evident in off-line measures of speech recognition. Work by Pisoni and colleagues and others (Cai et al.,

2017; Nygaard et al. 1995, 1998; Palmeri et al. 1993; Pisoni, 1997) has documented that listeners retain not only lexical information, but indexical properties of the speakers they have heard (i.e., information about the talker, such as gender, age, language background, and identity), and can use this information to improve speech recognition. These findings, such as a benefit of talker familiarity (Campeanu et al., 2014, 2015; Johnsrude et al., 2013; Nygaard et al., 1994), have informed how the process of speech recognition is understood, with acoustic, lexical, and indexical properties of speech all contributing to the mapping of an incoming signal to flexible mental representations.

Though there have been no specific examinations of RHT in an older population, some predictions can be made about their performance in this framework. Given older adults' preserved lexical knowledge with age and increased reliance on lexico-semantic information in speech processing, the predictions of the RHT are expected to hold under listening conditions with minimal acoustic distortions. However, when the speech signal is distorted, RHT predicts an increased reliance on finer acoustic detail for perception. Older adults, particularly those with hearing impairment, are expected to be at a disadvantage in these circumstances, given age-related declines in auditory temporal processing (Fitzgibbons & Gordon-Salant, 1996) and spectral clarity (Alain et al., 2001; Florentine et al., 1980; Patterson et al., 1982). Therefore, under the predictions of RHT, age-related detriments in speech recognition performance and learning are not expected under optimal listening envirnments, but are expected to occur under degraded listening conditions.

1.2 Mechanisms underlying rapid adaptation and perceptual learning

The mechanisms underlying the cue re-weighting described by the RHT may result from processes including lexically guided adaptation and statistical learning. In addition to benefitting overall speech recognition performance, availability of lexico-semantic information is known to facilitate perceptual adaptation to unfamiliar speech signals. Lexical information drives perceptual adaptation to ambiguous phonemes in single word contexts (Eisner & McQueen, 2005; Norris et al., 2003). In these now classic studies, listeners heard ambiguous phonemes falling between /f/ and /s/ in the context of words that end in either /f/ or /s/. Following this exposure, listeners who had heard the ambiguous phoneme in an /-s/ final context were more likely to categorize tokens on an /f/-/s/ continuum as /s/, than listeners who had heard the ambiguous phoneme in /-f/ final words. These findings suggest that the listeners used the lexical information present in the exposure stimuli to adjust their internal boundaries of category representation to include the ambiguous phoneme. Lexically guided learning has been observed in older adults, although the time course of learning for older adults appears to be slower than in younger adults (Colby et al., 2018; Scharenborg & Janse, 2013).

Davis and colleagues (2005) conducted a series of experiments examining the influence of various forms of feedback and degrees of lexical information on perceptual adaptation to noise-vocoded speech. They found that providing feedback, including a clear undistorted iteration of the target stimulus followed by a repetition of the distorted stimulus, facilitated adaptation to noise-vocoded sentences, and that both auditory and visual feedback facilitated adaptation similarly. This suggests that listeners were using the higher-level, lexical

information rather than the low-level acoustics for learning. Davis et al. (2005) also explored the use of training stimuli with various levels of lexical information and found that any level of lexical information was beneficial; listeners who heard syntactic prose or syntactically correct sentences with content words replaced with non-words (i.e., Jabberwocky sentences) showed greater learning than those who heard non-word strings. Cooper and Bradlow (2016) extended these findings to naturalistic stimuli, documenting that multiple levels of lexical information facilitate perceptual learning of non-native speech.

Statistical learning is the process of learning of statistical regularities within sensory input, such as the probability with which patterns co-occur across inputs. This statistical learning requires continuous revision; as the individual is exposed to more input, the predictions and expectations about the probability of feature occurrence in relation to other features in the mental lexicon will be updated. The connections between statistical learning ability and both speech processing and language learning are clear. Statistical learning is classically understood to underlie infants' acquisition of phonetic categories in their native languages. In adulthood, statistical learning can contribute to adults' ability to adapt to unfamiliar input. This continuous updating of probabilistic expectations would be an important component of adaptation to an unfamiliar speech signal. Specifically, as listeners are exposed to signals with differing probability distributions from their usual input, they presumably update their expectations in order to achieve appropriate and successful predictive processing of speech.

Lexically guided and stastical learning processes have been documented in both younger and older adults. Given the role of lexical information and the potential for irregularities in the input for degraded speech, it is expected that both younger and older listeners would rely on some combination of lexically guided learning and statistical learning to aid in adaptation to challenging speech signals.

1.3 Methodological Variables Across Training Paradigms

The choice of experimental methods for auditory training paradigms varies widely across studies, and is undoubtedly the source of a wide range of findings in the literature. Here we review some of the critical issues in the conduct of experiments on rapid adaptation and long-term perceptual learning, which will provide a framework for evaluating specific training paradigms aimed at improving recognition of degraded speech by older listeners.

1.3.1. Type of training paradigm.—Most training regimens can be described as either passive exposure training paradigms or active training paradigms. Passive exposure paradigms are modeled from the common experience of improved perception of a talker who is difficult to understand, following a period of listening to that challenging talker. It is thought that over time, listeners modify their internal representation of acoustic information and alter the mapping of atypical sounds to learned phonetic features, based on the talker's systematic alterations in speech. This implicit learning, in turn, ultimately improves lexical access and semantic processing. The term "passive exposure training paradigm" is used in this article to refer to paradigms that do not involve explicit or external feedback, although it is recognized that internal or self-generated feedback may guide an individual's behavior. Passive exposure paradigms can include those where the listener simply hears the stimulus

without providing a response or where the listener provides a discrimination or recognition response without feedback. Improvements in speech recognition by both younger and older listeners have been demonstrated with this type of exposure paradigm using numerous types of degraded speech stimuli, including time-compressed speech (Golomb et al., 2007; Peelle & Wingfield, 2005), foreign-accented speech (Gordon-Salant et al., 2010), and dysarthric speech (Borrie et al., 2012; Lansford, et al., 2018).

In contrast, active training paradigms employ correct-answer feedback, adaptive adjustment of the signal /distortion, and/or stimulus modeling following the listener's response on each stimulus trial, to promote active engagement of learning and shape listener re-tuning of stimulus representations. Wright and her colleagues have shown that adaptive training results in significant gains in auditory learning by young listeners with normal hearing in tasks of frequency discrimination, signal detection, temporal discrimination, temporal order, intensity discrimination, and spatial judgements (Banai et al., 2011; Fitzgerald & Wright, 2005; Mossbridge et al., 2008; Ning et al., 2019; Wright et al., 1996; Wright & Fitzgerald, 2017; Zhang & Wright, 2007). Wright and colleagues (2010) showed that comparable gains in perceptual learning can be attained in training regimens that alternate periods of adaptive training with periods of simple exposure on an auditory frequency discrimination task, thereby reducing the attention demands on the listener. Active training paradigms have similarly been incorporated in training protocols to improve recognition of degraded speech signals, and appear to result in greater gains in auditory learning than simple exposure paradigms, at least for certain types of degraded speech signals (Borrie et al., 2013; Woods et al., 2015). The strategy of alternating task practice with passive exposure has been applied to auditory learning of speech signals, both for a new speech contrast in a voice-onset time categorization task and for recognition of Mandarin-accented speech (Wright et al., 2015). For both speech tasks, young normal-hearing listeners demonstrated greater learning with combined practice and exposure compared to practice-only and exposure-only training, suggesting that the interaction between the two experiences produced greater gains than either type of training method alone. Whether or not older listeners receive a benefit of alternating passive exposure with adaptive training is currently unknown.

1.3.2. Outcome Measures.—Perceptual learning as a result of passive exposure or active training can be assessed though metrics generally falling into three categories: improvement in training stimuli (including the time-course of improvement), generalization of improvement to untrained stimuli or tasks, and retention over time of the improvements in trained or untrained stimuli. For perceptual learning of speech, improvement for training stimuli refers to changes in recognition performance for the speech stimuli used in training, including the specific training talker(s) and the identical speech items. The general finding is that listeners who are exposed to a stimulus during training (either actively or passively) will show improvements on those trained stimuli in a pre- to post-training comparison. These improvements among trained listeners are ideally observed in comparison to control groups who are not exposed to the stimuli during a training period (e.g., Wright et al., 2015).

Generalization of perceptual learning can be assessed for stimuli that share characteristics with the training stimuli (referred to as *near generalization*) or for stimuli and tasks that are distinctly different from those used in the perceptual training paradigm (referred to as *far*

generalization). Near generalization for speech stimuli used in training includes: 1) acoustic generalization, which refers to improvement in recognition of speech items used in training as spoken by unfamiliar talkers; and 2) semantic generalization, which refers to improvement in recognition of new speech items as spoken by the familiar talkers used in training (Banai & Lavner, 2019). Perceptual learning is often observed with acoustic generalization (i.e., Baese-Berk et al., 2013) or semantic generalization (i.e., Manheim et al., 2018), and may be observed in combinations of both acoustic and semantic generalization (Baese-Berk et al., 2013; Bieber & Gordon-Salant, 2017), but not in all cases (Manheim et al., 2018). In the lexically guided perceptual learning paradigm described above, learning for ambiguous phonemes is generally evident for listeners' categorization of unfamiliar, ambiguous stimuli. Results overwhelmingly show generalization with this paradigm for both younger and older listeners (Colby et al., 2018; McAuliffe & Babel, 2016; Scharenborg & Janse, 2013).

Many perceptual learning studies assess speech recognition in noise as an outcome measure, regardless of the training paradigm, because the ultimate goal in many of these investigations is to address the principal complaint of older adults with ARHL: difficulty understanding speech in noise. Results are mixed: some studies do not observe improvements in speech recognition in noise following perceptual learning on specific tasks (e.g., Ferguson et al., 2014; Woods et al., 2015), whereas others do observe transfer to speech in noise tasks that represent different stimuli and noise conditions (Ferguson & Henshaw, 2015). The conclusion reached by Ferguson et al. (2019) is that training with cognitively taxing auditory-based stimuli is necessary to achieve transfer of learning and generalization to new, challenging stimuli and tasks. Another form of generalization of perceptual learning is through the use of self-report measures (Tye-Murray et al., 2012). The rationale for incorporating self-report measures as an index of generalization is that individuals are more likely to complete training requirements if they perceive that the training is providing substantial benefit (Tye-Murray et al., 2012). A final form of generalization is the transfer of perceptual learning for acoustic signals to improved performance on cognitive measures. Because cognitive skills are an integral part of any perceptual learning paradigm, success as a result of this form of training may implicitly result in strengthening cognitive abilities. Cognitive skills that presumably support perceptual learning are working memory, attention, inhibition, and speed of processing. Although few studies report pre- and post-training performance on cognitive measures, those that do generally show some improvements on complex cognitive measures of divided attention and working memory by participants engaged in active training but not by control participants (Ferguson et al., 2014; Ferguson & Henshaw, 2015; Sweetow & Sabes, 2006).

1.3.3   Type of speech stimulus.—The choice of speech stimulus for perceptual learning is guided by the overall objective of training. Some investigations use nonsense syllables or syllable constituents for training, with the assumption that re-learning of the perceptual details of syllable onsets, codas, and nuclei will result in accurate word recognition (Ferguson et al., 2014; Miller et al., 2015; Woods et al., 2015). However, RHT (described above) suggests that perceptual learning at the syllable level is constrained and that training with higher level, ecologically relevant stimuli results in rapid and efficient

learning. Within this framework, many investigations of perceptual learning incorporate sentence-length materials as the target stimuli for training (Manheim et al., 2018; Miller et al., 2015; Sweetow & Sabes, 2006). Sentences are viewed as ecologically valid and also capitalize on listeners' use of cognitive resources to a greater extent than training with nonsense syllables or words (Sweetow & Sabes, 2006; Miller et al., 2015). Yet other studies employ word-learning paradigms that train recognition of words with varying lexical properties, in order to improve lexical access and thereby derive meaning from sentences (Burk & Humes, 2007, 2008; Humes et al., 2009). Multi-level training, using both syllable-level training and higher-level sentences, has been described to both shape retuning of phonetic details and capitalize on the use of cognitive resources including the benefit of context (e.g., Miller et al., 2015; Tye-Murray et al., 2012). However, the relative training gains of multi-level training have not yet been reported.

Another parameter of the speech stimulus is the type of signal degradation. The majority of clinical training programs promoting perceptual learning use a single form of speech distortion, and it is unclear if the magnitude of perceptual learning is equivalent across different forms of degraded speech stimuli, or if training benefit with one type of speech distortion transfers to another type of speech distortion. Two studies compared the benefits of training with multiple forms of speech distortion, including time-compressed speech, speech with multi-talker babble, and speech with a single competing talker (Karawani et al., 2016; Sweetow & Sabes, 2006). Older listeners showed comparable improvements following perceptual training across the three forms of speech degradation.

1.3.4   Other methodological variables.—In addition to the global methodological variables described above, there are many other differences in the conduct of studies of perceptual learning that limit direct comparison of findings across studies. In perceptual learning paradigms that include noise, the type of noise (speech spectrum, multiple talkers, single talker) and method of setting the signal-to-noise ratio (adaptive, fixed, multiple fixed signal-to-noise ratios [SNRs]) vary widely. The dose and duration of training are related variables that may impact the magnitude of learning (Banai & Lavner, 2014; Ferguson & Henshaw, 2015; Humes et al., 2014) as well as possible saturation of learning (Humes et al., 2014; Wright & Sabin, 2007). Additionally, the spacing of training sessions has also been examined, but older listeners did not show differences in the magnitude of learning in a massed vs. a spaced training protocol (Tye-Murray et al., 2017). Finally, the overall study design is of critical importance, with the gold standard being a randomized controlled, double-blind clinical trial (RCT) including both passive and active control participants (Ferguson et al., 2019). Inclusion of both types of control participants allows for a comparison of the magnitude of procedural learning (active control participants vs. passive control participants) and stimulus learning (experimental participants vs. active control participants) (Woods et al., 2015).

1.4   Individual factors influencing adaptation and perceptual learning

The process of learning may differ between younger and older adults. While the magnitude of adaptation and perceptual learning is often comparable between younger and older adults, the rate and pattern of learning may differ by age. There are numerous reports of different

patterns of rapid adaptation between younger and older listeners, in studies examining adaptation to non-native speech (Adank & Janse, 2010; Bieber & Gordon-Salant, 2017), time-compressed speech (Peelle & Wingfield, 2005), ambiguous phonemes (Scharenborg & Janse, 2013), and speech in noise (Karawani et al., 2016). These pattern-wise differences do seem to vary across studies. Some authors report 'unlearning' in their younger but not older listeners (Karawani et al., 2016; Scharenborg & Janse, 2013), while others report a plateauing and slight unlearning in older listeners (Adank & Janse, 2010). Others report some unlearning over the course of perceptual learning, differing by stimulus type rather than listener age (Colby et al., 2018). Some find a steeper, more rapid, or more linear rate of adaptation in younger versus older listeners (Adank & Janse, 2010; Bieber & Gordon-Salant, 2017; Peelle & Wingfield, 2005). In contrast, Manheim et al. (2018) report a steeper rate of adaptation for older hearing-impaired listeners compared to either younger or older normal-hearing listeners, though these authors note that this may be driven by differences in starting level (i.e. the older hearing-impaired listeners had more room to improve). This report of greater learning for participants at lower starting levels is not uncommon (Banks et al., 2015b; Henshaw et al., 2014; Manheim et al., 2018; Saunders et al., 2016). However, other investigations report no differences in the rates and patterns of adaptation between younger and older listeners (Erb & Obleser, 2013; Gordon-Salant et al., 2010; Neger et al., 2014). In a study of adaptation to time-compressed speech, Peelle and Wingfield (2005) report no age differences when the two groups are matched on starting performance level, but do find an age effect on the pattern of adaptation when the two groups are listening to the same degree of time compression.

Consideration of hearing loss varies widely across studies. In some cases, listeners with ARHL are treated as a separate group (Bieber & Gordon-Salant, 2017; Gordon-Salant et al., 2010; Manheim et al., 2018), while in other studies, the "older" participant groups combine listeners with and without ARHL (Adank & Janse, 2010; Erb & Obleser, 2013; Janse & Adank, 2012; Neger et al., 2014; Scharenborg & Janse, 2013). For a true understanding of the mechanisms underlying perceptual learning in older listeners, it is critical to account for differences in hearing sensitivity among older listeners, as hearing impairment may have a differential or exacerbating effect on recognition of distorted or challenging speech (Dubno et al., 1984; Presacco et al., 2019; Sommers, 1997).

While perceptual learning appears to be preserved with age, there is evidence that an individual's capacity for perceptual learning may be moderated by individual characteristics including executive functions like attention and inhibition (Banks et al., 2015b; Janse & Adank, 2012), and statistical learning ability (Colby et al., 2018; Neger et al., 2014). Under the predictions of RHT, listeners must be able to attend to the relevant features of the stimulus and inhibit irrelevant information in order to successfully and flexibly re-organize their internal lexical representations to achieve long-term learning. These and associated executive functions (i.e. working memory, attention-switching control) have been shown to contribute to perceptual learning in both younger and older listeners (Banks et al., 2015b, Janse & Adank, 2012, Scharenborg et al., 2015). Because advanced age is associated with decline in cognitive capacity, especially inhibitory mechansms (Hasher et al. 1991), working memory (Hasher & Zacks, 1988), and speed of processing (Salthouse, 2004), it may be

expected that an individual's declines in these cognitive abilities may substantially impact the facility for perceptual learning.

The literature regarding age differences in cognitive predictors of auditory perceptual learning is limited. Colby et al. (2018) examined perceptual learning for an ambiguous vowel in conditions with and without available lexical information, in older and younger listeners. They found that vocabulary was a significant predictor of learning-consistent behavior in both conditions, as was age, but that the two did not interact. Interestingly, aptitude for lexically guided learning was not a predictor for learning in the absence of lexical information, and vice versa. In a study of perceptual adaptation in non-native speech in older adults, Janse and Adank (2012) found that vocabulary and attention-switching control predicted recognition of non-native speech, but that these measures did not interact with block (i.e. were not predictive of learning behavior). Scharenborg et al. (2015) examined factors predicting perceptual learning of ambiguous phonemes in a group of older adults and found that attention-switching control significantly predicted the degree of learning-consistent behavior. Individual differences in lexical knowledge and vocabulary have been shown to predict adaptation and learning (Banks et al., 2015b; Colby et al. 2018; Janse & Adank 2012; Scharenborg & Janse 2013), which is logical given the role of lexical information in perceptual learning (Davis et al. 2005, Maye et al. 2008, Norris et al. 2003).

Learning aptitude may also play a role in a listener's capacity for perceptual learning of challenging speech and the ability to benefit from auditory training interventions. In younger adults, statistical learning ability (on a visual, non-linguistic task) positively predicts magnitude of learning for noise-vocoded speech (Neger et al., 2014). Interestingly, older adults did not demonstrate significant amounts of statistical learning, despite showing a similar degree of perceptual learning to the younger adults. Individual differences in learning aptitude also contribute to an individual's relative likelihood of success with different types of training. For example, Perrachione et al. (2011) found that listeners with greater learning aptitude were more likely to benefit from a high-variability training protocol than those with lower aptitude. This might hold for individual differences in cognitive capacity: listeners with strengths in certain cognitive domains may benefit more from specific types of training paradigms than those with relatively poorer cognitive abilities.

In the following sections, we will review a number of protocols designed to elicit rapid adaptation and longer-term perceptual learning in older adults for degraded speech. The studies are organized by type of signal degradation in order to illuminate any stimulus-related influences on age-related differences in learning. A summary of these studies and their design is included in Table 1. Finally, we will review studies that target improvements on speech-in-noise recognition. These studies are treated differently, as they often include multiple types of training and/or multiple stimulus types.

2.0   Adaptation and Perceptual Learning for Time-Compressed Speech

Accurate speech understanding requires listeners to rapidly and flexibly adjust to variations in speech rate, both within a talker and across different talkers. In a speech stream, listeners must segment word onsets and offsets correctly and use phrasal boundaries in order to derive

the intended meaning of a spoken utterance (Wingfield et al., 1992). These distinctions may become blurred when speech is produced at a rapid rate or with variations in rate. Older listeners have more difficulty than younger listeners in recognizing utterances spoken at a fast rate (Gordon-Salant et al., 2014) or with variations in speech rate (Sommers, 1997). Thus, listening to natural speech at a fast or variable rate is a form of challenging speech that older people encounter in daily communication situations. As a result, perceptual learning for fast speech is targeted in some auditory training programs.

In the laboratory, fast speech is simulated with time compression, a computer algorithm that removes silences and quasi-periodic segments of speech and concatenates the remaining signal to produce a signal free of spectral distortion (Gordon-Salant & Fitzgibbons, 1993; Schneider et al., 2005). With this simulation, the speech rate can be carefully controlled: the degree of time compression can be adjusted from minimal [e.g., 5% time compression ratio, where the time-compressed (TC) signal is 95% of the original signal duration] to maximal (e.g., 95% time compression ratio (TCR), where the TC signal is 5% of the original signal duration). Older listeners, with and without hearing loss, demonstrate significantly poorer recognition of TC sentences compared to younger listeners with comparable hearing sensitivity (e.g., Gordon-Salant & Fitzgibbons, 1993; Vaughan & Letowski, 1997; Wingfield et al., 1985; Wingfield et al., 1999), especially for fast speech rates in the range of 40–70% time compression. Older adults' difficulty in recognizing TC speech (relative to younger adults) appears to be associated with age-related deficits in auditory temporal processing (Dias et al., 2019; Schneider et al., 2005), coupled with senescent declines in cognitive abilities of speed of information processing (Dias et al., 2019; Wingfield et al., 1985; Wingfield, 1996), working memory (Vaughan et al., 2006), and executive function (Dias et al., 2019). These associations between deficits in recognition of TC speech and other perceptual and cognitive abilities suggest the intriguing notion that training paradigms to promote auditory perceptual learning of TC speech may have the potential to improve speed of neural processing of speech, in addition to strengthening cognitive abilities among older people.

2.1 Short-term exposure

Early work on rapid adaptation to TC speech generally indicates that older listeners benefit from short-term exposure to these stimuli. Peelle and Wingfield (2005) compared the magnitude and time course of auditory learning for TC speech by younger normal-hearing listeners and older listeners with normal to near-normal hearing. The investigators equated the starting level of performance between the younger and older listeners by individually adjusting the time-compression ratio to yield 30% correct performance and 70% correct performance. In the adaptation phase, listeners heard 20 sentences at a TCR that yielded 30% correct performance. Generalization of learning was assessed with new sentences presented at the TCR corresponding to 70% performance, both pre- and post-training. Listeners repeated the sentences heard, with no feedback provided. Both listener groups adapted to the trained stimuli rapidly and showed comparable rate and magnitude of learning over the course of 20 sentences. Transfer of learning was measured as the improvement in recognition performance pre- to post-training at the time compression ratio corresponding to

70% correct performance at pre-test (i.e., the slower speech rate). Significant improvement in performance was observed in the younger listeners, but not the older listeners.

A subsequent study (Golomb et al., 2007) examined the impact of disruption in the TC speech signal used for adaptation, by interspersing either slower TC sentences, natural-rate sentences, or silence during the adaptation phase with 20 sentences at 70% TCR for younger listeners and 60% TCR for older listeners. Younger and older listeners demonstrated comparable learning (approximately 20%) for trained stimuli, suggesting either that learning processes are resistant to age-related decline or that older listeners employ compensatory mechanisms that promote effective learning. Additionally, the effect of disruption condition was not significant for either listener group, indicating that adaptation was preserved with periods of less challenging signals (including silence), extending the work of Wright et al. (2010, 2015) to adaptation for TC speech. A similar paradigm to equate younger and older listeners for recognition accuracy prior to adaptation with TC speech was employed by Simhony et al. (2014), and comparable improvement in recognition of TC speech was observed after adaptation by both younger and older listeners. It appears that older listeners are capable of showing comparable gains from passive exposure to TC speech as younger listeners, when the two groups are performance-matched prior to adaptation. However, without performance matching, the magnitude of learning appears to be greater for younger listeners compared to older listeners (Peelle & Wingfield, 2005).

2.2 Active training

Active training paradigms also have incorporated TC speech for auditory perceptual learning with older listeners. Two studies (Karawani et al., 2016; Sweetow & Sabes, 2006) used TC speech as one of multiple forms of distorted speech targeted for training. Both auditory-based training programs included training with TC sentences over a 4-week program (described in detail in section 5.2), in which the TCR was modified adaptively based on the listener's response. Results of both studies showed significant learning on the TCR training task by older hearing-impaired listeners. The magnitude of improvement (pre- to post-training) was approximately a 6% change in TCR for speech recognition training (Sweetow & Sabes, 2006) and approximately a 25–30% change in TCR in a passage comprehension task (Fig. 3, Karawani et al., 2016). Importantly, a passive control group of older listeners did not show significant improvement in the TC speech task from pre- to post-training (Karawani et al., 2016).

Only one investigation to date has examined perceptual learning with an active training protocol using TC speech exclusively (Manheim et al., 2018). Three listener groups (young normal-hearing, older normal hearing, and older hearing-impaired) participated, with listener performance equated prior to training. Adaptive training with feedback was conducted over 300 trials in one session, using a 2-down, 1-up adaptive rule. The listeners judged the semantic plausibility of each sentence during training, but provided recognition judgments of both TC and naturally fast sentences during pre- and post-training tests. All listener groups showed learning for trained stimuli, but the pattern of learning varied between groups, with the two normal-hearing groups (but not the hearing-impaired group) exhibiting early learning during the pre-test. In addition, transfer of learning was measured

for naturally fast speech and recognition of untrained TC sentences. Older adults, both with and without ARHL, failed to show transfer to the untrained sentences, unlike the younger listeners with normal hearing.

Several conclusions may be drawn from these investigations. First, older listeners do show adaptation and perceptual learning for TC speech stimuli. Short-term exposure to TC sentences produces rapid gains in recognition of TC speech within the first 20 sentences presented. The magnitude of adaptation and perceptual learning appears to be equivalent to that observed for younger listeners, especially when younger and older listeners are presented with degrees of TC speech at which performance is equated at the outset. Interspersing the target TC stimuli with slower-rate stimuli or silent intervals does not diminish the magnitude of auditory learning, and may even improve learning while reducing attention demands. Second, active training paradigms that provide correct-answer feedback and alter the TCR based on the listener's response also show that older listeners derive training benefit. However, the magnitude of auditory perceptual learning varies widely across studies (6% reported by Sweetow & Sabes, 2006, vs. 25–30% reported by Karawani et al., 2016), and may depend on the dose of training (number of TC stimuli presented for training), the type of listener judgement required (recognition vs. comprehension), and whether or not training with TC speech is exclusive. Third, findings to date indicate that near generalization of active training with TC speech is minimal for older listeners, both with normal hearing and with hearing loss. Although the findings are somewhat disappointing, it should be noted that these studies had a relatively brief training regimen. Another study reported that older blind participants who regularly listened to rapid speech for leisure (i.e. audiobooks) showed superior perception of time-compressed speech under laboratory conditions (Gordon-Salant & Friedman, 2011). This suggests that frequent listening to rapid speech may result in significant perceptual learning, and may be feasible with today's ubiquitous smartphone and computer apps that can play back podcasts and audiobooks at fast rates. Future research needs for examining the benefits of passive exposure and active auditory training with TC speech include examining far generalization of training benefit to other forms of distorted speech, retention of training benefit, changes in neural processing for TC speech, and changes in cognitive skills as a result of training, particularly for cognitive skills shown to predict recognition of time-compressed speech.

3.0   Adaptation and learning for non-native speech

One naturalistic form of distortion to the auditory signal is the presence of a non-native accent. A non-native accent results from the combined influences of a non-native talkers' first language and the language in which they are speaking (Flege, 1988), which can change the segmental (i.e. sound substitutions or alterations), subsegmental (i.e. f0 range), and suprasegmental (i.e. altered timing and prosody) features of speech. Speech recognition is typically lower for non-native as compared to native speech (Floccia et al., 2006; Goslin et al., 2012). This 'accent effect' may be greater for older adults, especially those with ARHL (Burda et al., 2003; Gordon-Salant et al., 2010a, b), though this may depend on the particular features of the talker (Gordon-Salant et al., 2015) and some researchers do not find a clear difference between age groups (Ferguson et al., 2010). Age-related deficits in recognition of non-native speech may arise from age-related declines in auditory temporal

processing (Fitzgibbons & Gordon-Salant, 1994; Snell, 1997), or be related to changes in cognitive function and cortical mechanisms for speech recognition that accompany the aging process (Eckert et al., 2008; Erb & Obleser, 2013; Salthouse, 2004).

Rapid adaptation and perceptual learning for non-native speech has been documented for younger listeners with normal hearing in an extensive literature examining the both listener-related and protocol-related effects (Alexander & Nygaard, 2019; Banks et al., 2015b; Bradlow & Bent, 2008; Baese-Berk et al., 2013; Clarke & Garrett, 2004; Cooper & Bradlow 2016; Maye et al., 2008; Romero-Rivas et al., 2015; Sidaras et al., 2009; Tzeng et al., 2016; Wade et al., 2007; Xie et al., 2017). The literature regarding adaptation and learning of non-native speech in older adult listeners is relatively limited. These few studies vary in the type of stimulus used, including both artificially constructed accents (Adank & Janse, 2010; Janse & Adank, 2012) and naturalistic non-native accents (Gordon-Salant et al., 2010a; Bieber & Gordon-Salant 2017). Some examined rapid adaptation by looking strictly at improvement on the training stimuli (Adank & Janse, 2010; Janse & Adank, 2012; Gordon-Salant et al., 2010), while only one study measured learning by means of comparing post-test to pre-test performance for untrained stimuli (Bieber & Gordon-Salant, 2017).

Adank and Janse (2010) constructed an artificial accent involving modification of the vowel sounds in Dutch, to examine age-related differences in perceptual learning of an unfamiliar accent. Younger and older participants listened to and repeated back sentences in speech-shaped noise which varied adaptively in level; the signal-to-noise ratio (SNR) required for 50% correct performance was measured for 4 consecutive sentences lists. Perceptual learning was examined by measuring the change in SNR from list to list. Overall, older adults required significantly more favorable SNRs than younger adults for recognition of the accented speech. The results of the study also showed age-related differences in learning patterns. Younger listeners required progressively less favorable SNRs from list to list, indicating that they continued to adapt over the entirety of the relatively short experiment (approximately 30 minutes). Older listeners improved from List 1 to 2, but performance did not appear to improve for the final two blocks (see Fig. 1 from Adank & Janse, 2010). However, a comparison of the magnitude of learning between the two listener groups revealed no age differences. None of the measured individual characteristics (including age, hearing sensitivity, processing speed, and executive function) predicted patterns of learning.

In a subsequent study, this same constructed accent was used to examine perceptual learning in a cohort of older adults and investigated whether presentation modality [audiovisual (AV), or auditory only (A)] or individual characteristics influenced rate and magnitude of learning (Janse & Adank, 2012). This study utilized a passive paradigm; while participants were asked to respond to target stimuli, there was no adaptive signal adjustment or provision of feedback. The task was true/false categorization for sentences. Accuracy and reaction times (RTs) were both collected and measured over time to index learning. The two forms of outcome measures differed in finding an effect of modality on adaptation. Accuracy analyses showed that magnitude of learning across the two modalities (A and AV) was similar, though initial rate of learning was faster for those exposed to AV stimuli as compared to A-only stimuli. RT analyses showed that both groups adapted at a similar rate and magnitude to the accented speech. In the accuracy analyses, hearing thresholds were not predictive of the rate

of learning, but vocabulary knowledge and selective attention both were. In the RT analyses, none of the measured individual characteristics predicted rate of learning.

Gordon-Salant et al. (2010) explored rapid adaptation to Spanish-accented speech in three listener groups: younger adults with normal hearing, older adults with normal hearing, and older adults with ARHL. Adaptation was measured, both across lists and within an initial stimulus list, as improvement in word recognition over time. All three listener groups showed improvements between the first list and the final two lists, with no apparent age or hearing loss effects. A closer examination of the first list showed additional evidence of rapid adaptation; performance was significantly higher on the second half than the first half of the list for all listeners, regardless of age or hearing status. These findings indicate that listeners show adaptation to a non-native speech signal in as quickly as 20 trials, and that learning continues over the course of 140 additional trials.

While the studies detailed above investigated rapid adaptation to non-native speech, a subsequent study examined learning using post-training tests with both unfamiliar stimuli and unfamiliar talkers to examine the type and degree of transfer. Bieber and Gordon-Salant (2017) utilized a passive paradigm to examine the benefit of exposure to talkers from multiple language backgrounds for far generalization of learning to a talker with an unfamiliar non-native accent. The intent of this paradigm was to expose listeners to a high level of systematic variability in the realization of English lexical items by non-native talkers. Exposure to this variable input presumably allows listeners to adjust their internal representations in a more general manner and to be more flexible in mapping future input from non-native speakers, whose productions may vary from standard American English in similar ways. In each of two training sessions, listeners heard the same sentences repeated by talkers with different non-native accents. Post-testing was completed following both training sessions, and took the form of a far generalization task; listeners were asked to repeat unfamiliar sentences produced by unfamiliar talkers with accents that were not heard in the training. All three listener groups showed improvement from pre-test to post-test over the first day of training, but there was no additional improvement after a second day of training. This far generalization test was repeated at 1 week after training to assess retention of the training benefit; no groups showed retention. In fact, performance at the retention test did not differ significantly from pre-test performance. However, this study did not include any form of control group to distinguish true adaptation from practice effects.

In combination, these studies suggest that while ARHL contributes to overall poorer recognition of non-native speech (Bieber & Gordon-Salant, 2017; Gordon-Salant et al., 2010), it doesn't seem to have an additional detrimental effect on perceptual learning over aging (Adank & Janse, 2010; Bieber & Gordon-Salant, 2017; Gordon-Salant et al., 2010; Janse & Adank, 2012). In all of the studies described here, hearing sensitivity did not predict rate or magnitude of learning. Additionally, these studies of rapid adaptation and learning for artificially constructed and naturally occurring accents suggest that magnitude of learning and generalization are similar between older and younger listener groups (Adank & Janse, 2010; Bieber & Gordon-Salant 2017; Gordon-Salant et al., 2010). However, listeners do not retain these benefits at one week following training with a passive exposure paradigm (Bieber & Gordon-Salant, 2017).

The findings regarding age effects on patterns of adaptation are less clear. Adank and Janse (2010) found an age effect on the pattern of adaptation, while Gordon-Salant et al. (2010) did not. These contrasting findings could be due to the differences in study methodology. Adank and Janse (2010) constructed an accent using alterations to vowels only, whereas Gordon-Salant et al. (2010) used a naturalistic accent that involved multiple segmental and suprasegmental alterations. Adank and Janse (2010) also employed an adaptive paradigm to adjust the signal-to-noise ratio (SNR) to achieve 50% correct recognition perforance, while the listeners in the Gordon-Salant et al. (2010) study were tested at a fixed signal level in quiet. Additionally, Adank and Janse (2010) compared performance over blocks of 15 trials, while Gordon-Salant et al. (2010) presented four lists of 40 trials (though their within-list analysis provides additional support for a lack of age or ARHL-related differences). Older listeners do show different patterns of adaptation by stimulus modality type (Janse & Adank 2012), but it is unknown if these patterns differ for younger and older listeners.

4.0   Adaptation and Learning of Vocoded Speech

A commonly studied form of speech distortion is noise-vocoding, a signal adjustment which manipulates only the spectral components of the signal and leaves temporal information intact. Noise-vocoding is accomplished by dividing the speech signal into a number of frequency bands and extracting the amplitude envelope in those bands. Those envelopes are then used to modulate bands of noise. Generally speaking, the number of bands correlates positively with speech intelligibility; speech with 16 bands or greater is highly intelligible. This noise-vocoded speech is thought to approximate the experience of listening through a cochlear implant. Many researchers utilize experiments with noise-vocoded speech stimuli to make inferences about the performance of listeners who use cochlear implants. There is a large amount of variability in the performance of listeners who wear cochlear implants (Firszt et al., 2004), which makes it challenging to draw any specific conclusions about age effects on adaptation and learning for cochlear-implant processed speech in the older population. However, a handful of studies have examined adaptation and/or learning for noise-vocoded speech in older adults who do not use cochlear implants (Neger et al., 2014; Peelle & Wingfield 2005; Sheldon et al., 2008), which provide insight into the challenges of perceptual adaptation to spectrally degraded speech by older people.

Following their findings of absent age effects in rapid adaptation to time-compressed speech, Peelle and Wingfield (2005, see above) tested rapid adaptation to noise-vocoded speech in a group of older listeners. The goal was to confirm whether learning was differentially affected by age in the different domains of auditory distortion (temporal versus spectral for time-compressed and noise-vocoded, respectively). They found that, over the course of exposure to 40 noise-vocoded sentences (20 blocks of 2 sentences), both younger and older listeners showed robust adaptation, improving their sentence recognition by up to 20–30 percentage points without any active training. There were no differences between older and younger listeners in terms of overall performance or adaptation behavior, though it should be noted that the two listener groups were equated for starting performance level. Hearing thresholds were not reported for the older listeners, and the contribution of hearing to adaptation performance was not examined.

Similar findings are documented by Neger et al. (2014), who compared perceptual learning for noise-vocoded speech in younger and older adult listeners. Listeners heard noise-vocoded sentences and were asked to repeat them back; no feedback was given. Both age groups adapted well to the sentences, showing relative improvement of up to 30% by the final block of sentences. There were no differences in adaptation behavior between age groups, but within the older listener group, age in years was a significant predictor for learning. Hearing sensitivity did not contribute significantly to the variance in adaptation between listeners.

One study examined learning for noise-vocoded speech in younger and older normal-hearing adults with a more active training paradigm. Sheldon et al. (2008) conducted a series of experiments exploring recognition of noise-vocoded speech under different presentation conditions. In the first experiment (Experiment 1), the vocoded stimuli (monosyllabic words in a carrier phrase) were presented in a modified gating procedure, in which the number of bands was increased adaptively. Stimuli were presented first with 1-band vocoding applied. If the response was incorrect, the number of bands was increased by one until the word was correctly repeated. Feedback was provided following the responses. Each participant heard four lists of words; for each list per individual, the threshold number of bands required for 50% correct performance was calculated. Both the younger and older listener groups showed improvements (i.e. lower band threshold) across lists, and there was no effect of aging on this adaptation.

Two subsequent experiments followed. Experiment 2 used the same listeners as in Experiment 1 and occurred immediately after Experiment 1, while Experiment 3 had new participants. The goal was to examine performance with a different presentation method (i.e. blocking by bands, rather than gating), but the results of these experiments also can be used to assess generalization of learning from Experiment 1. The findings of Experiments 2 and 3 were identical: older listeners showed a poorer threshold for 50% performance than did younger listeners, despite similar psychometric function slopes across groups. The identical findings across experiments suggest that the listeners who participated in Experiment 1 did not generalize their learning to new stimuli in Experiment 2.

In sum, these studies all indicate that the aging process is not detrimental to the ability to adapt to spectrally distorted, noise-vocoded speech. Both older and younger adults show adaptation to noise-vocoded stimuli in as quickly as eight sentence presentations (Peelle & Wingfield, 2005), and continue to improve over the course of 200 words (Sheldon et al., 2008). This lack of age effects appears to be true whether listener groups are matched for starting performance (Peelle & Wingfield, 2005) or not (Neger et al., 2014). Additionally, age effects are absent when stimuli are presented at a single distortion level (Neger et al., 2014; Peelle and Wingfield, 2005), or when a gating-like procedure is used (Sheldon et al., 2008), and remain absent across a variety of distortion levels (four and five bands in Neger et al., 2014; 16 bands in Peelle and Wingfield, 2005). However, the limited evidence suggests that older adult listeners are not able to generalize this learning to new stimuli (Sheldon et al., 2008), though younger adults with normal hearing do show generalization (Huyck et al., 2017; Leobach & Pisoni, 2008).

Additional research is needed to determine the effects of ARHL on adaptation to spectrally degraded stimuli. The single study discussed here did not find that hearing sensitivity was a significant predictor of auditory perceptual learning to noise-vocoded speech (Neger et al., 2014). Considering the fact that hearing loss, per sé, among older people does not appear to exacerbate age-related auditory temporal processing deficits (Fitzgibbons & Gordon-Salant, 1996), it is logical to assume that ARHL would not be detrimental to perceptual learning for a stimulus in which temporal cues remain intact. However, these findings should be confirmed and extended to include an examination of the generalization of learning. This is especially critical in informing the development of auditory training protocols for older listeners with cochlear implants, who will likely have had a period of hearing loss prior to implantation and subsequent listening to a spectrally degraded speech signal.

5.0 Training to Improve Speech Recognition in Noise

Auditory training to improve speech recognition in noise is fundamentally different from protocols for improving recognition of distorted speech. In the case of distorted speech (accented speech, fast speech, vocoded speech), listeners learn to adjust their representations of the segmental and suprasegmental characteristics of speech, such that the altered speech matches a revised template to facilitate lexical access and semantic meaning. However, in most speech-in-noise situations, the speech signal itself is not distorted in any way. The listener must learn to retrieve an undistorted speech signal embedded in noise that either overlaps with the target signal in frequency and time (i.e., an energetic masker), or is composed of one or more other talkers' voices that distract the listener (i.e., an informational masker), or a combination of the two. Older listeners, with and without ARHL, typically exhibit much poorer speech recognition performance in these types of noises compared to younger listeners with comparable hearing sensitivity (e.g., Dubno et al., 1984; Stuart & Phillips, 1996), underscoring the importance of training aimed at this population to improve communication in typical listening scenarios that include noise. The task is to reduce the impact of the noise by increasing attention to the target speech signal, and/or inhibiting (suppressing) the deleterious effects of the noise. Different learning paradigms aim to accomplish these goals through varied auditory or cognitive regimens, as described below.

5.1. Word-level training

Burke and Humes (2007, 2008) developed an auditory word-learning training program to improve speech recognition in noise for hearing-impaired listeners who wear hearing aids. The underlying assumption was that strengthening a listener's access to isolated words in noise would enable the listener to better recognize those words in sentences or running discourse when presented in noise. The training paradigm presented target words in 2-talker noise-vocoded speech (ICRA noise; Dreschler et al., 2001) over 8 sessions (75–90 min/session) and included training with 94 selected frequent phrases, in addition to the words. The efficacy of the training program was evaluated for young normal-hearing listeners and older hearing-impaired listeners (Humes et al., 2009). Listeners in each group were tested at fixed SNRs intended to produce approximately 70% correct performance for each of the different types of speech stimuli (words, phrases, sentences). Significant and substantial improvement (20–30 RAUs) in trained words was observed, on average, across both

listening groups. More modest but significant improvements in recognition of untrained open-set sentences in noise were also observed, indicating that word-level training produced generalizable benefit. The authors attributed the observed generalization to the use of lexical items in training that overlapped considerably with the words featured in everyday sentences. Comparable improvements in recognition of trained and untrained stimuli with this training paradigm were reported by Kuchinsky et al. (2010), who also observed that older hearing-impaired listeners responded faster and exhibited a larger pupil size to speech stimuli after training, compared to stimuli presented at pre-test. The latter results indicated that listeners had more attentional focus on the speech stimuli following training.

5.2. Sentence-level training

Training with everyday sentences has obvious relevance, as spoken communication typically involves listening to running speech that contains prosodic, indexical, and suprasegmental information. In addition, sentence-length utterances incorporate lexical, syntactic, and semantic cues, enabling the listener to derive meaning from the spoken message based on perception of only some of the words in the sentence. Thus, the goal of auditory training paradigms that present sentences as training stimuli aim to strengthen the listener's attention to both the acoustic and contextual cues in running speech.

One auditory training program incorporating sentences in noise has been evaluated with clinical trials involving older listeners. The Listening and Communication Enhancement (LACE) Program (Sweetow & Sabes, 2006) involves several forms of training, including training with everyday sentences presented in either 6-talker babble or with one competing talker during the course of a 4-week home-based training program. Listeners are provided with correct-answer feedback following each trial, and the SNR is adapted based on response accuracy. Evaluation of LACE with an older group of hearing-impaired listeners showed significant improvement on the trained stimuli, with no improvement for a control group (Sweetow & Sabes, 2006). Although small but significant improvement in performance on the primary outcome measure, untrained sentence tests in noise, was reported in this clinical trial, these benefits can't be ascribed solely to training with speech in noise, because there are three other training modules in the program. These results were replicated with a group of young, normal-hearing listeners, who also exhibited significant enhancements in pitch-related neural encoding of a speech syllable presented in noise following training with LACE (Song et al., 2012). These findings indicate improvements in subcortical plasticity as a result of this auditory training program, at least for younger listeners. A subsequent assessment of the efficacy of LACE showed that older veterans with significant hearing loss who wore hearing aids gained no benefit from LACE training for recognizing untrained word stimuli or for using contextual cues in sentences presented in competing speech (Saunders et al., 2016). The findings suggest that while LACE may provide some benefit to older listeners with mild-to-moderate hearing loss who do not use amplification, the training paradigm (including sentence-level training in background speech) provides little, if any, benefit beyond that provided by a well-fit hearing aid.

A modification of the LACE training program was evaluated in a randomized controlled clinical trial with older normal-hearing and older hearing-impaired adults (Karawani et al.,

2016). Two of the three auditory training modules were speech in 4-talker babble and speech in the presence of a competing talker of the opposite sex. Passages of 1-min duration were presented and segmented into 1–2 sentence units. Based on listener responses to multiple choice questions following each unit, the SNR was varied adaptively to yield 70% correct recognition. Correct-answer feedback was provided following each response. Both trained listener groups showed significant learning over the course of training for the two speech tasks compared to a control group, with older normal-hearing listeners exhibiting greater gains than older hearing-impaired listeners for speech in 4-talker babble, but not speech in the single competing talker. However, transfer of learning, as measured with plausibility judgments to untrained sentences in 4-talker babble, was not observed for either training group. A small but significant transfer of learning to pseudowords in noise was reported for the older hearing-impaired listeners, consistent with the expectations of RHT. That is, RHT predicts that with a low quality signal (here associated with the adverse conditions of listening in noise coupled with ARHL), listeners must focus on lower-level acoustic information rather than higher-level contextual information to understand speech. Another interpretation is that the choice of speech tasks used in the generalization measures may not have adequately captured the auditory learning that ensued from this training paradigm.

In summary, older listeners with ARHL and with relatively normal hearing show auditory learning for sentence-level training in noise, but this learning appears to be limited to improvements in the training stimuli. Training benefit is also minimal for older listeners who wear well-fit hearing aids, at least for the multi-dimensional training program evaluated to date. Findings with younger listeners suggest that benefits of training result in neuroplastic changes in the brain that improve neural processing of key speech cues underlying accurate speech recognition performance in noise. Other specific mechanisms that may be strengthened with auditory learning realized as a result of speech-in-noise training are unknown at present. Mechanisms that could be targeted for evaluation in future research include improved use of acoustic cues for speech segregation (voice pitch, spatial segregation, dip listening), contextual information (syntactic, semantic), or suprasegmental, indexical, or prosodic information about the target speech.

5.3 Cognitive and Perceptuomotor Training

A substantial number of reports have now amassed demonstrating that deficits in understanding speech in noise by older people with and without hearing loss are linked to underlying limitations in cognitive abilities (i.e., Desjardins & Doherty, 2013; Gordon-Salant & Cole, 2015; Ronnberg et al, 2008, 2010, 2013; Schurman et al., 2014; Tun et al., 2002). As a result, some cognitive training programs have been developed with the overall objective of improving speech recognition in noise. Working memory is targeted in several of these training programs, because this cognitive domain is consistently linked to deficits in understanding speech in noise (i.e., Gordon-Salant & Cole, 2015; Rönnberg et al., 2008, 2010, 2013; Schurman et al., 2017).

One auditory-based cognitive training program teaches listeners to better perceive rapid transitions in consonant-vowel (CV) constituents by varying the duration of the formant transitions adaptively (Brain Fitness Cognitive Training Program; Posit Science). Different

training exercises increase the complexity of the listening task by embedding these rapidly changing formant transitions in sequences of syllables and words, commands, and stories of increasing duration to be recalled, thereby taxing working memory load. Older adults with ARHL who completed the 6-week training program showed significant improvements in measures of sentence recognition in multitalker babble, auditory short-term memory, and processing speed, while an active control group showed no improvements on these measures (Anderson et al., 2013). Additionally, the older adults who completed training showed better neural timing to the transition portion of a speech syllable presented in quiet and noise, relative to pre-training assessment, indicating that this type of auditory-based cognitive training improved neural processing of speech (Anderson et al., 2013).

The transfer of working memory training to speech recognition in noise has been evaluated for older people with normal hearing and with ARHL (Wayne et al., 2016). Participants completed a standardized 4-week cognitive training program that exclusively focused on working memory exercises (Cogmed Working Memory Training; Pearson; Klingberg et al., 2002), including adaptive training modules. Although the training group showed significant improvement on the training modules over the course of training, unlike the active control group, they did not show post-training improvement on two different measures of speech recognition in competing speech (Wayne et al., 2016). However, a significant correlation between working memory and recognition of low-context sentences in noise was observed, re-affirming that working memory is an important contributor to speech recognition with few contextual cues.

Engagement of attentional focus and learning circuits is also targeted in training programs aimed at boosting speech recognition performance in noise. Whitton et al. (2014) described a closed-loop audiomotor perceptual gaming paradigm that trained listeners and mice to better detect a tonal target in a background of continuous noise, using a foraging task in a visual scene. The closed-loop paradigm aims to focus attention continuously on target performance and suppress background noise as the gamer receives feedback during play. The benefit of a closed-loop gaming paradigm was evaluated for older listeners with ARHL who wore hearing aids, in which the listeners in the training group discriminated acoustic parameters of target tonal stimuli in a "tone cloud" of competing stimuli or in a background of speech babble (Whitton et al., 2017). The competing stimuli were adapted continuously based on gaming accuracy using a jigsaw puzzle task, thereby constantly focusing the listener's attention on subtle changes in the target acoustic stimulus as the listener received feedback and progressed through the task. An active control group engaged in training with a working memory task. Following 8 weeks of training, both groups showed improved performance on their respective training tasks, but only the audiomotor perceptual training group exhibited increased recognition of sentence stimuli presented at low SNRs. The investigators interpreted these results as reflecting training-related gains in focusing attention on low-level acoustic parameters in speech while inhibiting the distracting effects of background noise.

5.4. Summary of training paradigms to improve speech recognition in noise

As difficulty understanding speech in noise is a ubiquitous complaint of older people with ARHL, it behooves investigators to identify training paradigms that ameliorate this problem. Several conclusions can be reached from the foregoing review. First, auditory and cognitive training generally results in improved performance for trained stimuli by older listeners with and without ARHL. Second, few auditory training programs produce generalized benefit to speech understanding in noise for untrained stimuli. Third, benefits of general cognitive training programs that increase working memory capacity do not transfer to increases in speech recognition performance in noise. Finally, it appears that combined auditory-cognitive training programs that present acoustic stimuli while increasing the cognitive load, either through working memory or attention, hold promise for the transfer of training benefit to enhance speech understanding in competing backgrounds. As these types of training programs continue to be developed, assessment of their efficacy will require RCTs that include assessment of retention of learning.

6.   General Summary and Conclusions

The literature included in this review collectively shows that training can faciliate auditory learning for older people, but the benefits are varied. Protocol parameters vary widely, as do outcomes reported. There are a number of conclusions to be drawn from the present literature on facilitation of improvements in perception of challenging speech signals by older listeners. This synthesis of the literature also brings to light some significant areas in which additional research is warranted.

6.1.   The capacity for rapid adaptation and perceptual learning appears to be intact in older listeners.

In all of the protocols described in this review involving degraded speech signals (i.e. time-compressed speech, non-native speech, noise-vocoded speech), older listeners demonstrate improvement on trained stimuli, relative to initial performance prior to training. This holds both for protocols involving just a single training session and protocols involving multiple sessions. Significant improvements in recognition are evident in as quickly as 20 trials (Golomb et al., 2007; Gordon-Salant et al., 2010; Peelle & Wingfield, 2005; Simhony et al., 2014), and can persist in up to 300 trials (Manheim et al., 2018) in a single session or over multiple training sessions (Karawani et al., 2016). In general, the magnitude of both rapid adaptation and longer-term learning for trained stimuli appears to be similar between younger and older listeners, and there does not appear to be a detrimental effect of age-related hearing loss. The one exception to this conclusion is the study by Manheim and colleagues (2018), who observed a reduced magnitude in rapid adaptation among the older listeners compared to younger listeners, though longer-term learning was not reduced in the older listeners.

6.2.   Transfer of learning to untrained stimuli, when tested, is limited.

A small handful of the studies described in this review evaluated generalization to untrained stimuli. Of those, the single-session studies for learning of time-compressed speech found that younger adults showed near generalization of learning, while older adults did not

(Manheim et al., 2018; Peelle & Wingfield, 2005). All listener groups showed limited but comparable transfer of learning for non-native speech (Bieber & Gordon-Salant, 2017), while no listener groups showed transfer for noise-vocoded speech (Sheldon et al., 2008). Interestingly, many studies which included training on multiple forms of challenging speech and/or multiple training sessions did show transfer of learning in both younger and older adults, with and without hearing loss (Karawani et al., 2016; Sweetow and Sabes, 2006, Whitton et al., 2017, though see Humes et al., 2014). It is challenging to isolate which factors contribute to transfer of learning, but duration of training is one likely candidate. Prior studies of auditory perceptual learning and generalization indicate that generalization occurs at a significant delay relative to learning for trained stimuli (Wright et al., 2010b). Only a handful of studies included in this review reported results of retention testing. No listener groups retained a benefit of training at one week following short-term training with foreign-accented speech (Bieber & Gordon-Salant, 2017). Longer-term protocols showed mixed results, with some showing retention (Ferguson et al., 2014; Ferguson & Henshaw 2015), and others showing stimulus-dependent retention (Burk & Humes, 2008) or no retention (Whitton et al., 2017). Additional research is needed to assess whether training protocols involving longer-term training provide retention of learning, and which components of training best facilitate retention.

6.3. The time course of rapid adaptation appears to differ between younger and older listeners, with inconsistent findings regarding the differences across studies.

One factor which may relate to the inconsistent findings is the manner by which the time course data, when considered at all, are analyzed. The statistical analyses used to analyze rapid adaptation patterns in the prior literature often create training 'blocks' by averaging a number of trials, and use block as a categorical variable in analyses of variance or linear regressions. However, analyses which take advantage of trial-level data and account for non-linearities in the time course data, such as growth curve analysis or generalized additive modeling, may better capture group-wise differences in the time course data for perceptual learning (Hastie & Tibshirani, 1986; Mirman et al., 2008; Winter & Wieling, 2016; Wood, 2006).

6.4. Individual predictors for learning have not been sufficiently examined, and the present findings are inconclusive.

In the literature reviewed here, only a small number of studies examined individual characteristics that may be predictive of learning for older adults, and each found different results. In their evaluation of an interactive, closed-loop, multi-modal training program, Whitton et al. (2017) found that inhibitory control (as measured by the Stroop task) was predictive of learning behavior in older adults with ARHL who used hearing aids. Janse and Adank (2012) found that vocabulary knowledge and selective attention predicted improvements in accuracy in older adults, while working memory and attention-switching control were not predictive of learning. Finally, Neger et al. (2014) found that, in older adults, learning was predicted by age, but not by any of the other measures evaluated (including hearing sensitivity, vocabulary knowledge, working memory, processing speed, attention-switching control, and statistical learning ability).

A number of other studies have examined the significance of individual characteristics that contribute to overall *recognition* of challenging speech signals. However, understanding the relationship between individual characteristics and *learning* behaviors is critical for providing optimal interventions to older adults who seek treatment. Researchers should not only consider these factors, but should be clear and consistent in reporting which measures were used, how the effects were calculated, and the potential for confounds related to measuring cognitive abilities through an auditory modality.

In addition to cognitive and linguistic characteristics of the listener, an important consideration is the listeners' hearing status. For the most part, the current literature provides little to no evidence that ARHL has a detrimental effect on rapid adaptation, perceptual learning, and generalization beyond that attributed to aging alone. However, the findings of Saunders et al. (2016) suggest that those with ARHL who already utilize amplification may not benefit from auditory training to improve speech-in-noise recognition (though see Whitton et al. 2017). Additional research is needed to determine the benefit of targeted aural rehabilitation over amplification alone.

6.5. There are a limited number of studies evaluating auditory training protocols that include RCTs, control groups, and/or independent investigation.

None of the single-session studies described in this review included any form of control group, while the multiple-session studies targeting speech-in-noise improvement did include some form of control group (Ferguson et al., 2014; Ferguson & Henshaw, 2015; Karawani et al., 2016; Sweetow & Sabes, 2006; Whitton et al., 2017). In the absence of true control groups, some studies have listeners serve as their own controls for pre-to-post-test comparisons or use cross-over designs, and in other cases listener groups were simply compared for the purposes of determining age effects or intervention-specific effects. While the lack of active control groups may have been logical for the research questions posed in these studies, their absence limits the utility of these findings for drawing conclusions about adaptation and learning abilities in older listeners. Although a considerable financial and timing burden is inherent in including control participants in studies of older listeners, it is nonetheless crucial that future studies involve at least a true active control group to determine if effects of training reflect perceptual adaptation rather than procedural learning. Other levels of control groups (i.e. passive controls, or no-intervention controls) may also be beneficial in determining the degree to which intervention facilitates procedural learning. It is further noted that in virtually all cases, the studies reviewed here include protocols that were evaluated by their own creators. There is a distinct lack of independent, objective evaluation of auditory training protocols; this is a significant area for future research.

6.6. Development of future successful protocols for speech-in-noise may benefit from research investigating the mechanisms underlying learning for this type of speech, as well as consideration of the domains (i.e. perceptual, cognitive, neural) and modalities (i.e. auditory, audio-visual) to be targeted by training.

For the most part, the studies examining a single form of signal degradation were successful in targeting improvement on trained stimuli, but were significantly limited in facilitating transfer of learning. As described above, auditory training protocols that include multiple

modalities, multiple forms of challenging speech, and/or target cognitive as well as perceptual processes seem to hold promise for improving speech-in-noise recognition for trained or untrained stimuli. Additional research is needed regarding the mechanisms by which transfer of learning is best facilitated.

Acknowledgements

References

Adank P, & Janse E (2010). Comprehension of a novel accent by young and older listeners. Psychology and Aging, 25(3), 736. [PubMed: 20853978]

Ahissar M, & Hochstein S (2004). The reverse hierarchy theory of visual perceptual learning. Trends in Cognitive Sciences, 8(10), 457–464. [PubMed: 15450510]

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

Alain C, McDonald KL, Ostroff JM, & Schneider B (2001). Age-related changes in detecting a mistuned harmonic. The Journal of the Acoustical Society of America, 109(5), 2211–2216. [PubMed: 11386572]

Alexander JE, & Nygaard LC (2019). Specificity and generalization in perceptual adaptation to accented speech. The Journal of the Acoustical Society of America, 145(6), 3382–3398. [PubMed: 31255164]

Anderson S, & Kraus N (2013). Auditory training: evidence for neural plasticity in older adults. Perspectives on Hearing and Hearing Disorders: Research and Diagnostics, 17(1), 37–57.

Anderson S, White-Schwoch R, Parbery-Clark A, & Kraus N (2013). Reversal of age-related neural timing delays with training. Proceedings of the National Academy of Sciences 110:11, 4357–4362.

Atienza M, Cantero JL, & Dominguez-Marin E (2002). The time course of neural changes underlying auditory perceptual learning. Learning & Memory, 9(3), 138–150. [PubMed: 12075002]

Baese-Berk MM, Bradlow AR, & Wright BA (2013). Accent-independent adaptation to foreign accented speech. The Journal of the Acoustical Society of America, 133(3), EL174–EL180. [PubMed: 23464125]

Banai K, & Lavner Y (2019). Effects of stimulus repetition and training schedule on the perceptual learning of time-compressed speech and its transfer. Attention, Perception, & Psycophysics 81(8), 2944–2955, doi: 10.3758/s13414-019-1714-7.

Banai K, Sabin A, & Wright BA (2011). Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. Hearing Research 280 (1–2), 219–227. [PubMed: 21664958]

Banks B, Gowen E, Munro KJ, & Adank P (2015a). Audiovisual cues benefit recognition of accented speech in noise but not perceptual adaptation. Frontiers in Human Neuroscience, 9, 422. [PubMed: 26283946]

Banks B, Gowen E, Munro KJ and Adank P (2015b). Cognitive predictors of perceptual adaptation to accented speech. Journal of the Acoustical Society of America. 137, 2015–2024.

Ben- David BM, Campeanu S, Tremblay KL, & Alain C (2011). Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. Psychophysiology, 48(6), 797–807. [PubMed: 21054432]

Bieber RE and Gordon-Salant S (2017). Adaptation to novel foreign-accented speech and retention of benefit following training: Influence of aging and hearing loss. Journal of the Acoustical Society of America. 141, 2800–2811.

Borrie SA, McAuliffe MJ, Liss JM, Kirk C, O'Beirne GA, & Anderson T (2012). Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthic speech. Lang. Cogn. Process 27, 1039–1055. [PubMed: 24009401]

Bradlow AR, & Bent T (2008). Perceptual adaptation to non-native speech. Cognition, 106(2), 707–729. [PubMed: 17532315]

Burda AN, Scherz JA, Hageman CF, and Edwards HT (2003). "Age and understanding speakers with Spanish or Taiwanese accents." Perc. Mot. Skills, 97, 11–20.

Burk MH, & Humes LE (2007). Effects of training on speech recognition performance in noise using lexically hard words. Journal of Speech, Language, and Hearing Research 50, 25–40.

Burk MH, & Humes LE (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. Journal of Speech Language Hearing Research 51(3), 759–771. Doi: 10.1044/1092-4388(2008/054).

Cai ZG, Gilbert RA, Davis MH, Gaskell MG, Farrar L, Adler S, & Rodd JM (2017). Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. Cognitive Psychology, 98, 73–101. [PubMed: 28881224]

Campeanu S, Craik FI, & Alain C (2015). Speaker's voice as a memory cue. International Journal of Psychophysiology, 95(2), 167–174. [PubMed: 25173195]

Campeanu S, Craik FI, Backer KC, & Alain C (2014). Voice reinstatement modulates neural indices of continuous word recognition. Neuropsychologia, 62, 233–244. [PubMed: 25080187]

Chisolm TH, Saunders GH, Frederick MT, McArdle RA, Smith SL, & Wilson RH (2013). Learning to listen again: the role of compliance in auditory training for adults with hearing loss. American Journal of Audiology, 22, 339–342. [PubMed: 24018575]

Clarke CM, & Garrett MF (2004). Rapid adaptation to foreign-accented English. Journal of the Acoustical Society of America 116, 3647–3658.

Colby S, Clayards M, & Baum S (2018). The role of lexical status and individual differences for perceptual learning in younger and older adults. Journal of Speech, Language, and Hearing Research 61, 1855–1874.

Cooper A, & Bradlow AR (2016). Linguistically guided adaptation to foreign-accented speech. The Journal of the Acoustical Society of America, 140(5), EL378–EL384. [PubMed: 27908062]

Committee on Hearing, Bioacoustics and Biomechanics (CHABA) (1988). "Speech understanding and aging." J. Acoust. Soc. Am 83, 859–895. [PubMed: 3281988]

Clarke CM, & Garrett MF (2004). Rapid adaptation to foreign-accented English. The Journal of the Acoustical Society of America, 116(6), 3647–3658. [PubMed: 15658715]

Dahan D, Drucker SJ, & Scarborough RA (2008). Talker adaptation in speech perception: Adjusting the signal or the representations?. Cognition, 108(3), 710–718.\ [PubMed: 18653175]

Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, & McGettigan C (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. Journal of Experimental Psychology: General, 134(2), 222. [PubMed: 15869347]

Desjardins JL, & Doherty KA (2013). Age-related changes in listening effort for various types of masker noises. Ear and Hearing, 34, 261–272. DOI: 10.1097/AUD.0b013e31826d0ba4. [PubMed: 23095723]

Dias JW, McClaskey CM, & Harris KC (2019). Time-compressed speech identification is predicted by auditory neural processing, perceptuomotor speech, and executive functioning in younger and older listeners. Journal of the Association for Research in Otolaryngology 20, 73–88. DOI: 10.1007/s10162-018-00703-1 [PubMed: 30456729]

Drescheler WA, Verschuure H, Ludvigsen C, & Westermann S (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology 40, 148–157. [PubMed: 11465297]

Dubno JR, Dirks DD, and Morgan DE (1984). "Effects of age and mild hearing loss on speech recognition," The Journal of the Acoustical Society of America, 76, 87–96. [PubMed: 6747116]

Eisner F, & McQueen JM (2005). The specificity of perceptual learning in speech processing. Perception & Psychophysics, 67(2), 224–238. [PubMed: 15971687]

Erb J, & Obleser J (2013). Upregulation of cognitive control networks in older adults' speech comprehension. Frontiers in Systems Neuroscience, 7, 116. [PubMed: 24399939]

Ferguson MA, & Henshaw H (2015). Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. Frontiers in Psychology, 6, 556 Doi: 10.3389/fpsyg.2015.00556 [PubMed: 26074826]

Ferguson MA, Henshaw H, Clark DPA, & Moore DR (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50- to 74-year-olds with mild hearing loss. Ear and Hearing, 35(4), e110–e121. [PubMed: 24752284]

Ferguson M, Maidment D, Henshaw H, & Heffernan E (2019, 2). Evidence-based interventions for adult aural rehabilitation: that was then, this is now. In Seminars in hearing (Vol. 40, No. 01, pp. 068–084). Thieme Medical Publishers.

Ferguson SH, Jongman A, Sereno JA, & Keum K (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. Journal of the American Academy of Audiology, 21(3), 153–162. [PubMed: 20211119]

Firszt JB, Holden LK, Skinner MW, Tobey EA, Peterson A, Gaggl W, … & Wackym PA (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. Ear and Hearing, 25(4), 375–387. [PubMed: 15292777]

Fitzgerald MB & Wright BA (2005). A perceptual learning investigation of the pitch elicited by amplitude-modulated noise. The Journal of the Acoustical Society of America, 118, 3794–3803. doi:10.1121/1.2074687 [PubMed: 16419824]

Fitzgerald MB, & Wright BA (2005). A perceptual learning investigation of the pitch elicited by amplitude-modulated noise. The Journal of the Acoustical Society of America, 118(6), 3794–3803. [PubMed: 16419824]

Fitzgibbons P, & Gordon-Salant S (1996). Auditory temporal processing in elderly listeners: Speech and non-speech signals. Journal of the American Academy of Audiology, 7, 183–189. [PubMed: 8780991]

Flege JE (1988). Factors affecting degree of perceived foreign accent in English sentences. The Journal of the Acoustical Society of America, 84(1), 70–79. [PubMed: 3411056]

Flege JE, Bohn OS, & Jang S (1997). Effects of experience on non-native speakers' production and perception of English vowels. Journal of phonetics, 25(4), 437–470.

Floccia C, Goslin J, Girard F, & Konopczynski G (2006). Does a regional accent perturb speech processing?. Journal of Experimental Psychology: Human Perception and Performance, 32(5), 1276. [PubMed: 17002537]

Florentine M, Buus S, Scharf B, & Zwicker E (1980). Frequency selectivity in normally-hearing and hearing-impaired observers. Journal of Speech, Language, and Hearing Research, 23(3), 646–669.

Golomb JD, Peelle JE, & 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.

Gordon-Salant S, & Cole S (2016). "Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing." Ear and Hearing, 37, 593–602. [PubMed: 27232071]

Gordon-Salant S, & Fitzgibbons P (1993). Temporal factors and speech recognition performance in young and elderly listeners. Journal of Speech and Hearing Research 36, 1276–1285. [PubMed: 8114494]

Gordon-Salant S, & Friedman SH (2011). "Recognition of rapid speech by blind and sighted older adults." J. Speech Hear. Res 54, 622–631.

Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, & Schurman J (2010). Short-term adaptation to accented English by younger and older listeners. Journal of the Acoustical Society of America (Express Letters), 128, (EL200–EL204). DOI: 10.1121/1.3486199

Gordon-Salant S, Zion D, & Espy-Wilson C (2014). Recognition of time-compressed speech does not predict recognition of natural-fast speech sentences by older listeners. Journal of the Acoustical Society of America – Express Letters, 136, EL 268–274. doi: 10.1121/1.4895014.

Goslin J, Duffy H, & Floccia C (2012). An ERP investigation of regional and foreign accent processing. Brain and Language, 122(2), 92–102. [PubMed: 22694999]

Hasher L, Stoltzfus ER, Zacks RT, & Rypma B (1991). Age and inhibition. Journal of experimental psychology: Learning, memory, and cognition, 17(1), 163.

Hasher L, & Zacks RT (1988). Working memory, comprehension, and aging: A review and a new view. The psychology of learning and motivation, 22, 193–225.

Hastie T & Tibshirani R (1986). Generalized Additive Models. Statistical Science, 1, 297–310.

Hazan V, Sennema A, Iba M, & Faulkner A (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. Speech Communication, 47(3), 360–378.

Henshaw H, McCormack A, & Ferguson MA (2015). Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss. Frontiers in Psychology, 6, 1067. [PubMed: 26300796]

Humes LE, Burk MH, Strauser LE, & Kinney DL (2009). Development and efficacy of a frequent-word auditory training protocol for older adults with impaired hearing. Ear and Hearing 30(5), 613–627. Doi: 10.1097/AUD.0b013e3181b00d90. [PubMed: 19633564]

Humes LE, Kinney DL, Brown SE, Kiener AL, & Quigley TM (2014). The effects of dosage and duration of auditory training for older adults with hearing impairment. Journal of the Acoustical Society of America – Express Letters 136, EL224–EL230.

Huyck JJ, Smith RH, Hawkins S, & Johnsrude IS (2017). Generalization of perceptual learning of degraded speech across talkers. Journal of Speech, Language, and Hearing Research, 60(11), 3334–3341.

Ingvalson EM, Dhar S, Wong PC, & Liu H (2015). Working memory training to improve speech perception in noise across languages. The Journal of the Acoustical Society of America, 137(6), 3477–3486. [PubMed: 26093435]

Ingvalson EM, Lansford KL, Fedorova V, & Fernandez G (2017). Cognitive factors as predictors of accented speech perception for younger and older adults. The Journal of the Acoustical Society of America, 141(6), 4652–4659. [PubMed: 28679239]

Janse E, and Adank P (2012). "Predicting foreign-accented adaptation in older adults." Quarterly Journal of Experimental Psychology, 65, 1563–1585.

Johnsrude IS, Mackey A, Hakyemez H, Alexander E, Trang HP, & Carlyon RP (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice. Psychological Science, 24(10), 1995–2004. [PubMed: 23985575]

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. [PubMed: 26869944]

Klingberg T, Forssberg H, & Westerberg H (2002). Training of working memory in children with ADHD. Journal of Clinical and Experimental Neuropsychology 26, 781–791. Doi: 10.1076/jcen.24.6.781.8395.

Kuchinsky SE, Ahlstrom JB, Cute SL, Humes LE, Dubno JR, & Eckert MA (2014). Speech-perception training for older adults with hearing loss impacts word recognition and effort. Psychophysiology 51(10), 1046–1057. Doi: 10.1111/psyp.12242. [PubMed: 24909603]

Lansford KL, Luhrsen S, Ingvalson EM, & Borrie SA (2018). Effects of familiarization on intelligibility of dysarthric speech in older adults with and without hearing loss. American Journal of Speech-Language Pathology 27(1), 91–98. [PubMed: 29305612]

Loebach JL, & Pisoni DB (2008). Perceptual learning of spectrally degraded speech and environmental sounds. The Journal of the Acoustical Society of America, 123(2), 1126–1139. [PubMed: 18247913]

Luce P, & Pison D (1998). Recognizing spoken words: The neighborhood activation model. Ear and Hearing 19, 1–36. [PubMed: 9504270]

Manheim M, Lavie L, & Banai K (2018). Age, hearing, and the perceptual learning of rapid speech. Trends in Hearing 22, 1–18. DOI: 10.1177/23312/6518778651.

Maye J, Aslin RN, & Tanenhaus MK (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. Cognitive Science, 32(3), 543–562. [PubMed: 21635345]

McAuliffe M, & Babel M (2016). Stimulus-directed attention attenuates lexically-guided perceptual learning. Journal of the Acoustical Society of America 140(3), 1727–1738.

Miller JD, Watson CS, Dubno JR, & Leek MR (2015). Evaluation of speech-perception training for hearing aid users: A multisite study in progress. Seminars in Hearing 36(4), 273–283. DOI:10.1055/s-0035-1564453. [PubMed: 27587914]

Mirman D, Dixon JA, & Magnuson JS (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. Journal of Memory and Language, 59(4), 475–494. [PubMed: 19060958]

Molloy K, Moore DR, Sohoglu E, & Amitay S (2012). Less is more: latent learning is maximized by shorter training sessions in auditory perceptual learning. PloS one, 7(5), e36929. [PubMed: 22606309]

Mossbridge JA, Scissors BN, & Wright BA (2008). Learning and generalization on asynchrony and order tasks at sound offset: Implications for underlying neural circuitry. Learning and Memory 15, 13–20. [PubMed: 18174369]

Neger TM, 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. [PubMed: 25225475]

Ning R, Trosman SJ, Sabin AT, & Wright BA (2019). Perceptual-learning evidence for inter-onset-interval-and frequency-specific processing of fast rhythms. Attention, Perception, & Psychophysics, 81(2), 533–542.

Norris D, McQueen JM, & Cutler A (2003). Perceptual learning in speech. Cognitive Psychology, 47(2), 204–238. [PubMed: 12948518]

Nygaard LC, & Pisoni DB (1998). Talker-specific learning in speech perception. Perception & Psychophysics, 60(3), 355–376. [PubMed: 9599989]

Nygaard LC, Sommers MS, & Pisoni DB (1994). Speech perception as a talker-contingent process. Psychological Science, 5(1), 42–46. [PubMed: 21526138]

Nygaard LC, Sommers MS, & Pisoni DB (1995). Effects of stimulus variability on perception and representation of spoken words in memory. Perception & Psychophysics, 57(7), 989–1001. [PubMed: 8532502]

Ortiz JA, & Wright BA (2010). Differential rates of consolidation of conceptual and stimulus learning following training on an auditory skill. Experimental Brain Research, 201(3), 441–451. [PubMed: 19902196]

Palmeri TJ, Goldinger SD, & Pisoni DB (1993). Episodic encoding of voice attributes and recognition memory for spoken words. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(2), 309.

Patterson RD, Nimmo- Smith I, Weber DL, & Milroy R (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. The Journal of the Acoustical Society of America, 72(6), 1788–1803. [PubMed: 7153426]

Peelle JE, & 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. [PubMed: 16366792]

Perrachione TK, Lee J, Ha LY, & Wong PC (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. The Journal of the Acoustical Society of America, 130(1), 461–472. [PubMed: 21786912]

Pisoni DB (1997). Some thoughts on "normalization" in speech perception Mullenix (Ed.),Talker Variability in Speech Processing, 9–32.

Presacco A, Simon JZ, & Anderson S (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. PloS One, 14(3).

Reitan RM (1958). Validity of the Trail Making Test as an indicator of organic brain damage. Perceptual and motor skills, 8(3), 271–276.

Rönnberg J, Lunner T, Zekveld A, Sorqvist P, Danielsson H, Lyxell B, Dahlstrom O, Signoret C, Stenfelt S, Pichora-Fuller MK, & Rudner M (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. Frontiers in Systems Neuroscience, 7(31), 1–17. [PubMed: 23420631]

Rönnberg J, Rudner M, Foo C, & Lunner T (2008). Cognition counts: A working memory system for Ease of Language Understanding (ELU). International Journal of Audiology, 47 (Suppl. 2), S99–S105. [PubMed: 19012117]

Rönnberg J, Rudner M, Lunner T, & Zekveld AA (2010). When cognition kicks in: Working memory and speech understanding in noise. Noise & Health, 12, 263–9. [PubMed: 20871181]

Romero-Rivas C, Martin CD, & Costa A (2015). Processing changes when listening to foreign-accented speech. Frontiers in Human Neuroscience, 9, 167. [PubMed: 25859209]

Saunders GH, Smith SL, Chisolm TH, Frederick MT, McArdle RA, & Wilson RH (2016). A randomized control trial: Supplementing hearing aid use with listening and communication enhancement (LACE) auditory training. Ear and Hearing, 37(4), 381–396. [PubMed: 26901263]

Salthouse TA (2004). What and when of cognitive aging. Current Directions in Psychological Science, 13(4), 140–144.

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(2), 493–507.

Schneider BA, Daneman M, & Murphy DR (2005). Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing? Psychology and Aging 20, 261–271. [PubMed: 16029090]

Schurman J, Brungart DS, and Gordon-Salant S (2014). Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. Journal of the Acoustical Society of America, 136, 3337–3349.

Sheldon S, Pichora-Fuller MK, & Schneider BA (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. The Journal of the Acoustical Society of America, 123(1), 489–499. [PubMed: 18177176]

Sidaras SK, Alexander JE, & Nygaard LC (2009). Perceptual learning of systematic variation in Spanish-accented speech. The Journal of the Acoustical Society of America, 125(5), 3306–3316. [PubMed: 19425672]

Simhony M, Grinberg M, Lavie L, & Banai K (2014). Rapid adaptation to time-compressed speech in young and older adults. Journal of Basic Clinical Physiology and Pharmacology, 25(3), 285–288.

Snell KB (1997). Age-related changes in temporal gap detection. The Journal of the Acoustical Society of America, 101, 2214–2220. [PubMed: 9104023]

Sommers MS (1997). Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment. The Journal of the Acoustical Society of America 101(4), 2278–2287. [PubMed: 9104029]

Song JH, Skoe E, Banai K, & Kraus N (2012). Training to improve hearing speech in noise: Biological mechanisms. Cerebral Cortex, 22, 1180–1190. [PubMed: 21799207]

Stuart A and Phillips D (1996). "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners," Ear and Hearing, 17, 478–489. [PubMed: 8979036]

Sweetow R, & Sabes J (2006). The need for and development of an adaptive listening and communication enhancement (LACE) program. Journal of the American Academy of Audiology 17, 538–558. [PubMed: 16999250]

Sweetow RW, & Sabes JH (2007). Technologic advances in aural rehabilitation: applications and innovative methods of service delivery. Trends in Amplification, 11(2), 101–111. [PubMed: 17494876]

Tremblay K, Ross B, Inoue K, McClannahan K, & Collet G (2014). Is the auditory evoked P2 response a biomarker of learning?. Frontiers in Systems Seuroscience, 8, 28.

Tun PA, O'Kane G, & Wingfield A (2002). Distraction by competing speech in yung and older adult listeners. Psychology and Aging 17, 453–467. [PubMed: 12243387]

Tye-Murray N, Sommers MS, Mauze E, Schroy C Barcroft J, & Spehar B (2012). Using patient perceptions of relative benefit and enjoyment to assess auditory training. Journal of the American Academy of Audiology 23, 623–634. [PubMed: 22967737]

Tye-Murray N, Spehar B, Barcroft J, & Sommers MS (2017). Auditory training for adults who have hearing loss: A comparison of spaced vs. massed practice schedules. Journal of Speech, Language, and Hearing Research 60, 2337–2345.

Tzeng CY, Alexander JE, Sidaras SK, & Nygaard LC (2016). The role of training structure in perceptual learning of accented speech. Journal of Experimental Psychology: Human Perception and Performance, 42(11), 1793. [PubMed: 27399829]

Vaughan NE & Letowski T (1997). Effects of age, speech rate, and type of test on temporal auditory processing. Journal of Speech, Language, and Hearing Research 40(5), 1192–1200. DOI:10.1044/jslhr5005.1192.

Vaughan NE, Storzbach D, & Furukawa I (2006). Sequencing versus nonsequencing working memory in understanding of rapid speech by older listeners. Journal of the American Academy of Audiology 17, 506–518. [PubMed: 16927515]

Wade T, Jongman A, & Sereno J (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. Phonetica, 64(2–3), 122–144. [PubMed: 17914280]

Wayne RV, Hamilton C, Huyck JJ, & Johnsrude IS (2016). Working memory training and speech in noise comprehension in older adults. Frontiers in Aging Neuroscience 8, 49 Doi: 10.3389/fnagi.2016.0049. [PubMed: 27047370]

Whitton JP, Hancock KE, & Polley DB (2014). Immersive audiomotor game play enhances neural and perceptual salience of weak signals in noise. Proceedings of the National Academy of Sciences 111(25), E2606–E2616. Doi:10.1073/pnas.1322184111.

Whitton JP, Hancock KE, Shannon JM, & Polley DB (2017). Audiomotor perceptual training enhances speech intelligibility in background noise. Current Biology 27, 1–11. DOI: 10.1016/j.cub.2017.09.014. [PubMed: 27916526]

Wingfield A (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. Journal of the American Academy of Audiology 7, 175–182. [PubMed: 8780990]

Wingfield A, Poon LW, Lombardi L, & Lowe D (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. Journal of Gerontology 50(5), 579–585.

Wingfield A, Wayland SC, & Stine EA (1992). Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences. Journal of Gerongology 47, P350–P356. DOI:10.1093/geronj/47.5.P350

Wingfield A, Tun PA, Koh CK, &Rosen MJ (1999). Regaining lost time: adult aging and the effect of time restoration on recall of time-compressed speech. Psychol. Aging 14, 380–389. doi: 10.1037/0882-7974.14.3.380 [PubMed: 10509694]

Winter B, & Wieling M (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. Journal of Language Evolution, 1(1), 7–18.

Woods DL, Doss Z, Herron TJ, Arbogast T, Younus M, Ettlinger M, & Yund EW (2015). Speech perception in older hearing impaired listeners: Benefits of perceptual training. PLoS ONE 10(3):e0113965 Doi: 10.1371/journal.pone.0113965 [PubMed: 25730330]

Wood SN (2017). Generalized additive models: an introduction with R. Chapman and Hall/CRC.

Wright BA, Baese-Berk M, Marrone N, & Bradlow AR (2015). Enhancing speech learning by combining task practice with periods of stimulus exposure without practice. Journal of the Acoustical Society of America 138, 928–237. DOI: 10.1121/1.4927411

Wright BA, Buonormano DV, Mahncke HW, & Merzenich MM (1996). Learning and generalization of auditory temporal-interval discrimination in humans. Journal of Neuroscience, 17, 3956–3963.

Wright BA, & Fitzgerald M (2017). Detection of tones of unexpected frequency in amplitude-modulated noise. The Journal of the Acoustical Society of America, 142, 2043–2046. [PubMed: 29092596]

Wright BA, & Sabin AT (2007). Perceptual learning: how much daily training is enough?. Experimental Brain Research, 180(4), 727–736. [PubMed: 17333009]

Wright BA, Sabin AT, Zhang Y, Marrone N, & Fitzgerald MB (2010). Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. The Journal of Neuroscience 30(38), 12868–12877. [PubMed: 20861390]

Wright BA, Wilson RM, & Sabin AT (2010b). Generalization lags behind learning on an auditory perceptual task. Journal of Neuroscience, 30(35), 11635–11639. [PubMed: 20810884]

Xie X, & Myers EB (2017). Learning a talker or learning an accent: Acoustic similarity constrains generalization of foreign accent adaptation to new talkers. Journal of Memory and Language, 97, 30–46. [PubMed: 28890602]

Zhang Y,. & Wright BA (2007). Similar patterns of learning and performance variability for human discrimination of interaural-time-differences at high and low frequency. The Journal of the Acoustical Society of America 121, 2207–2216. [PubMed: 17471734]

Highlights:

- Perceptual learning and adaptation appear intact in older listeners.

- Transfer of learning to untrained stimuli, when tested, is limited.

- Few studies which evaluate auditory training protocols include RCTs, control groups, and/or independent investigation.

- Future protocols may benefit from research into underlying mechanisms, and consideration of training domains/modalities.

Table 1.

Summary of the literature reviewed in sections 2.0, 3.0, 4.0. ONH = older adults with normal hearing; OHI = older adults with hearing impairment; PTA = Pure-tone average; TC speech = time-compressed speech

Authors,year	Subject groups? (n, age range)	Hearing status of older listeners	Control groups?	Training paradigm	Feedack provided?	# training sessions	# training trials/ session?	Training stimulus	Outcome measures	Improvement on trained stimuli?	Improvement on untrained stimuli?	Retention?
Golomb et al 2007	Younger (n=16, 18–22 years), Older (n=16, 65–85 years)	Mean PTA (.5, 1, 2 kHz) = 13.7 dB HL (sd = 8.2 dB)	None	Passive	No	1 per condition (4 total)	20 sentences	Time–compressed speech	% words repeated correctly	Yes. Similar magnitude of improvement between groups. Group differences not evaluated for within–session patterns.	Not tested	Not tested
Karawani et al., 2016	Older NH (n = 21, 60–71 yrs) Older HI (n = 35, 60–71 yrs	Older NH: thresholds 25 dB HL .25 – 6kHz Older HI: thresholds 60 dB HL through 8kHz (mild–mod loss)	passive control grp (n = 10), delayed training group served as control (n = 11)	Active	Yes	4 sessions w/TC speech +1/3 of final session (13 sessions altogether over 4 weeks, w/ 2 other training tasks)	3 blocks of training/ session; ea block asked questions based on a 3–6 min passage	time–compressed speech	speech–in–noise pseudoword discrim; sentence recognition in noise; duration discrim; frequency discrim	Yes. Significant training–based learning observed in ONH and OHI for TC speech, and improved more than untrained listeners. ONH and OHI showed similar patterns of learning over the training phase.	Yes, for pseudowords in noise, but observed as a result of 3 different training modules.	Not tested
Manheim, Lavie, & Banai 2018	Younger (n=57, 23–31 years), Older NH (n=56, 65–75 years), Older HI (n=36, 65–79 years)	Older NH: Mean RE PTA (.5, 1, 2, 4 kHz) = 21 dB HL. Older HI: Mean RE PTA (.5, 1, 2, 4 kHz) = 34 dB HL	passive control grp (n = 28 younger, n = 28 older NH, n = 18 OHI)	Active	Yes	1	300 sentences	Time–compressed speech	During training: TC rate for 71% correct performance. During test: plausability judgement	Yes. Greater magnitude of learning for normal–hearing listeners (YNH and ONH) re: control listeners. Rate differences may be confounded by differences in starting level.	Improvement seen for younger but not older	Not tested
Peelle and Wingfield 2005 *Experiment 1*	Younger (n=20, 18–22 years), Older (n=20, 65–78 years)	Mean PTA (.5, 1, 2 kHz) = 14.9 dB HL	None	Passive	No	1	20 sentences	Time–compressed speech	% words repeated correctly	Yes. Similar magnitude of improvement between groups. Possible group differences	Improvement seen for younger but not older	Not tested

Authors, year	Subject groups? (n, age range)	Hearing status of older listeners	Control groups?	Training paradigm	Feedack provided?	# training sessions	# training trials/ session?	Training stimulus	Outcome measures	Improvement on trained stimuli?	Improvement on untrained stimuli?	Retention ?
										in early time course.		
Simhoni et al., 2014	Younger (n=15, 20–30 years), Older (n=12, 60–75 years)	Not tested	None	Passive	Not reported	1 per condition (2 total)	20 sentences	Time–compressed speech	% correct repetition	Yes. Similar magnitude of improvement between groups. Rates of improvement not reported	Not tested	Not tested
Sweetow and Sabes 2006	Hearing impaired listeners (n = 65; ages 28–85 yrs; mean age = 64)	Mean PTA (trained = 37.7 dB; control = 40.0 dB).	passive control group (n = 27); experimental group (n = 38)	Active	Yes	20 sessions over 4 weeks (includes multiple types of training, not just TC speech)	Not reported	time–compressed speech; speech in 6–talker babble; speech w/ 1 competing talker	Q–SIN, HINT, LSPAN, Stroop	Yes. Significant improvement in TC speech for trained group, re: baseline performance	Improvement seen for experimental group on Q–SIN, HINT, LSPAN and Stroop, but observed as a result of a number of different training modules (not just TC speech).	Not reported
Adank and Janse 2010	Younger (n=20, 18–41 years), Older (n=30, 65–87 years)	Mean PTA: 25.5 dB HL (sd = 9.8 dB)	None	Active	No	1	60 sentences	Non–native speech	SRT for 50% correct performance	Yes. Similar magnitude of improvement between groups. Younger show steady improvement while older show plateau and possible subsequent declines	Not tested	Not tested
Bieber and Gordon–Salant 2017	Younger (n=15, 18–28 years), Older NH (n=13, 65–76 years), Older HI (n=15, 70–82 years)	NH: thresholds 25 dB HL from .25 – 4 kHz. HI: thresholds 26 dB HL from 2–8 kHz.	None	Passive	No	2 sessions, separated by 7–10 days	80 sentences	Non–native speech	% words repeated correctly	Yes. Rate and magnitude of improvement not analyzed.	Yes. Similar improvements across groups.	No groups retain improvement to trained or untrained stimuli.
Gordon-Salant et al., 2010	Younger (n=15, 18–30 years), Older	NH: thresholds 20 dB HL from .25 – 4	None	Passive	No	1 per condition (2 total, separated by	160 items (words, or sentences)	Non–native speech	% correct transcription	Yes. Similar rate and magnitude of improvement between groups.	Not tested	Not tested

Authors,year	Subject groups? (n, age range)	Hearing status of older listeners	Control groups?	Training paradigm	Feedack provided?	# training sessions	# training trials/ session?	Training stimulus	Outcome measures	Improvement on trained stimuli?	Improvement on untrained stimuli?	Retention?
	NH (n=15, 66–81 years), Older HI (n=15, 65–81 years)	kHz. HI: thresholds 26 dB HL from 3–8 kHz.				~2 weeks)						
Janse and Adank 2012	Older (n=66, 64–89 years)	Mean PTA: 27.2 dB HL (sd = 11.8 dB)	None	Passive	No	1	80 sentences	Non–native speech	Accuracy (T/F judgement) and RT	Listeners improve on trained stimuli. No age group compariso n.	Not tested	Not tested
Neger, Rietveld, and Janse 2014	YNH (n=60, 18–29 years), Older (n=73, 60–84 years)	Mean PTA (1, 2, 4 kHz): 23.31 dB HL (sd = 10.28)	None	Passive	Not reporte d	1	60 sentences	Noise–vocoded sentence s	% words repeated correctly	Yes. Similar rate and magnitud e of improvem ent between groups.	Not tested (confirm)	Not tested
Peelle and Wingfiel d 2005 *Experim ent 5*	YNH (n=30, 18–21 years), Older (n=30, 65–79 years)	Not reported	None	Passive	Not reporte d	1	40 sentences	Noise–vocoded and spectrall y shifted sentence s	% words repeated correctly	Yes. Similar magnitud e of improvem ent between groups. Possible group difference s in early time course.	Not tested	Not tested
Sheldon, Pichora–Fuller, Schneide r, 2008	YNH (n=12, 19–25), ONH (n=12, 66–74)	Thresho lds better than 25 dB HL from 0.25 – 3 kHz in test ear.	None	Active	Yes	1	200 words	Noise–vocoded words	Number of bands needed for 50% correct performa nce	Yes. Similar rate of improvem ent between groups.	No transfer for either age group.	Not tested