Research Article

Time-Compressed Speech Identification Is Predicted by Auditory Neural Processing, Perceptuomotor Speed, and Executive Functioning in Younger and Older Listeners

JAMES W. DIAS,¹ CAROLYN M. MCCLASKEY,¹ AND KELLY C. HARRIS¹

¹*Department of Otolaryngology, Medical University of South Carolina, 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, USA*

Received: 10 January 2018; Accepted: 8 October 2018; Online publication: 19 November 2018

ABSTRACT

Older adults typically have difficulty identifying speech that is temporally distorted, such as reverberant, accented, time-compressed, or interrupted speech. These difficulties occur even when hearing thresholds fall within a normal range. Auditory neural processing speed, which we have previously found to predict auditory temporal processing (auditory gap detection), may interfere with the ability to recognize phonetic features as they rapidly unfold over time in spoken speech. Further, declines in perceptuomotor processing speed and executive functioning may interfere with the ability to track, access, and process information. The current investigation examined the extent to which age-related differences in time-compressed speech identification were predicted by auditory neural processing speed, perceptuomotor processing speed, and executive functioning. Groups of normal-hearing (up to 3000 Hz) younger and older adults identified 40, 50, and 60 % time-compressed sentences. Auditory neural processing speed was defined as the P1 and N1 latencies of click-induced auditory-evoked potentials. Perceptuomotor processing speed and executive functioning were measured behaviorally using the Connections Test. Compared to younger adults, older adults exhibited poorer time-compressed speech identification and slower perceptuomotor processing. Executive functioning, P1 latency, and N1 latency did not differ between age groups. Time-compressed speech identification was independently predicted by P1 latency,

perceptuomotor processing speed, and executive functioning in younger and older listeners. Results of model testing suggested that declines in perceptuomotor processing speed mediated age-group differences in time-compressed speech identification. The current investigation joins a growing body of literature suggesting that the processing of temporally distorted speech is impacted by lower-level auditory neural processing and higher-level perceptuomotor and executive processes.

Keywords: auditory temporal processing, time-compressed speech, P1, N1, connections test, perceptuomotor, executive functioning

INTRODUCTION

Compared to younger adults, older adults typically have more difficulty identifying speech in complex listening conditions, even when their audiometric thresholds fall within a range typically considered clinically normal for speech identification. These difficulties can manifest when the normal temporal structure of the speech signal is distorted by increasing speech rate (Gordon-Salant et al. 2014; Janse 2009; Koch and Janse 2016; Vaughan and Letowski 1997), reverberation (Adams et al. 2010; Fujihira et al. 2017; Gordon-Salant and Fitzgibbons 1993; Halling and Humes 2000; Neuman et al. 2010), and accentedness (Adank and Janse 2010; Gordon-Salant et al. 2010a, b, 2015), or when speech information is masked by competing noise (Dubno et al. 1984;

Correspondence to: James W. Dias · Department of Otolaryngology · Medical University of South Carolina · 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, USA. email: diasj@muscc.edu

Gelfand et al. 1986; Gordon-Salant and Fitzgibbons 1993; Humes 1996). The increased difficulties older adults have in processing distorted speech are thought to result in part from age-related declines in temporal processing (e.g., Anderson et al. 2011; Füllgrabe et al. 2015; Pichora-Fuller and Singh 2006).

In support of this hypothesis, older adults typically make more errors identifying time-compressed speech, compared to younger adults, even when presented in clear listening conditions (comfortable listening level, no noise) (Gordon-Salant and Fitzgibbons 1993, 2001). Time-compressed speech is created by taking normal speech utterances and removing periodic information across the length of the speech signal. As a result, the spectral characteristics of the original signal are largely preserved, but the phonetic information is presented across a shorter temporal interval. Time compression modulates the temporal speech envelope (e.g., Ahissar and Ahissar 2005; Nourski et al. 2009) and artificially increases the speech rate of a spoken stimulus while avoiding the articulatory distortions associated with speaking at faster rates (i.e., increased coarticulation and slurring of phonetic features) (Gordon-Salant et al. 2014; Janse 2009). Yet the ability to accurately identify speech spoken at faster rates is predictive of a listener's ability to accurately identify time-compressed speech (Gordon-Salant et al. 2014). Identification of time-compressed speech can be affected by the temporal resolution of phonetic information and by linguistic context, suggesting that both low-level and high-level processes can independently contribute to the processing of temporally distorted speech (Gordon-Salant and Fitzgibbons 2001; Pichora-Fuller 2003).

Age-related declines in time-compressed speech identification may result from declines in auditory temporal processing (e.g., Gordon-Salant and Fitzgibbons 2001; Pichora-Fuller 2003). Behavioral evidence suggests that perception of auditory temporal envelope and auditory temporal fine structure relate to speech identification in young normal-hearing adults (for a review, see Moore 2014). In older adults, declines in auditory temporal processing are related to declines in speech identification (Moore et al. 2012; Füllgrabe et al. 2015). Work from our lab has found that age-related loss of auditory neural synchrony can contribute to dysfunctional auditory temporal processing, such as reduced auditory gap detection, even after accounting for audiometric thresholds and global metrics of cortical neural processing speed (Harris and Dubno 2017). We have also found that slower auditory neural processing speed is associated with reduced auditory temporal processing, such that longer auditory-evoked response latencies predict poorer auditory gap detection (Harris et al. 2012). Importantly, gap detection, along

with other behavioral measures of auditory temporal processing (e.g., frequency discrimination and gap discrimination thresholds), has been found to predict age-related differences in time-compressed speech identification (e.g., Gordon-Salant and Fitzgibbons 1993, 1999). The results of these studies suggest that slower auditory neural processing may adversely affect time-compressed speech identification by negatively impacting the perceptual resolution of information contained within the auditory signal.

However, speech identification depends on more than auditory temporal processing. Processing speed measured behaviorally using the Connections Test and the Trail-Making Test has been found to predict auditory temporal processing (Füllgrabe et al. 2015; Harris et al. 2010, 2012) and the identification of accented speech (Adank and Janse 2010), interrupted speech (Bologna 2017), and speech-in-noise (Ellis and Munro 2013; Helfer and Freyman 2014; Woods et al. 2013). An important feature of the Connections and Trail-Making tests is their ability to measure both perceptuomotor processing speed and executive functioning. In the simple form of these tests, perceptuomotor processing speed is defined as the speed with which an individual is able to connect numbers or letters consecutively, requiring the individual to search for and identify targets, determine the next target, and search for, identify, and respond to the next target in a sequence. The complex form of these tests requires the individual to connect alternating numbers and letters consecutively, engaging the same perceptuomotor processing required for the simple test and executive functioning (selective attention and inhibition) to switch between numbers and letters (e.g., Reitan 1992; Salthouse 2000). Many studies investigating the relationships between speech perception and Connections or Trail-Making test scores do not distinguish between perceptuomotor processing speed and executive functioning, instead computing metrics of "processing speed." However, age-related differences in perceptuomotor processing speed, but not executive functioning, have been found using the Connections Test (Eckert et al. 2010; Salthouse 2011) and the Trail-Making Test (Füllgrabe et al. 2015).

Using the Trail-Making Test, there is some inconsistency regarding whether perceptuomotor processing speed or executive functioning is associated with age-related differences in speech identification. Füllgrabe et al. (2015) found that the complex form of the Trail-Making Test (which measures both perceptuomotor processing and executive functioning) was associated with speech-in-noise identification, but executive functioning computed as [(complex – simple) / simple] is not, suggesting that the relation-

ship between the complex test and speech-in-noise identification was due to perceptuomotor processing. However, Adank and Janse (2010) found that perceptuomotor processing speed (measured using the Digit-Symbol Substitution Task) did not predict accented speech identification, but executive functioning computed as [complex / simple] did. Cahana-Amitay et al. (2016) similarly found that executive functioning computed as [complex – simple] predicted speech-in-noise identification, but as they did not report the relationship between perceptuomotor processing speed and speech-in-noise identification, it is unclear whether associations were observed. Some of the inconsistency in the literature may have to do with how executive functioning was computed and its role in processing different types of speech stimuli (speech-in-noise vs. accented speech). A challenge when using the Trail-Making or Connections tests to measure executive functioning is isolating executive functioning from perceptuomotor processing in the complex form. As described above, different researchers have used different methods for computing executive functioning from simple and complex scores. A major problem with the popular methods for computing executive functioning, including difference scores ([complex – simple]) and ratios ([complex / simple] or [(complex – simple) / simple]), is that they fail to completely extract the perceptuomotor component from complex scores to isolate executive functioning (Salthouse 2011). However, using a regression approach, variability due to perceptuomotor processing can be almost completely extracted from complex scores to isolate executive functioning. This is accomplished by computing the residuals when predicting complex scores from simple scores (complex – [$a + b(\text{simple}) + \text{error}$], where [$a + b(\text{simple}) + \text{error}$] is the complex value predicted from simple scores). The residuals have no relationship to simple scores ($r < 0.001$) and serve as a more reliable metric of executive functioning (Salthouse 2011). We computed executive functioning using this residual metric for the current investigation.

We have previously found that perceptuomotor processing speed measured using the Connections Test predicts auditory gap detection independent from auditory neural processing, suggesting that perceptuomotor processing plays an important role in auditory temporal processing (Harris et al. 2010). Unfortunately, this past investigation did not evaluate the extent to which executive functioning predicted auditory gap detection. Given the evidence for the role of perceptuomotor processing and executive functioning in speech perception, and for the role of perceptuomotor processing in auditory temporal processing, we measured perceptuomotor and executive functioning using the Connections Test to

evaluate the degree to which each factor independently accounts for variability in time-compressed speech identification.

Current Investigation

Previous work suggests that auditory neural response latencies and perceptuomotor processing speed independently predict auditory temporal processing (Harris et al. 2012) and that perceptuomotor processing and executive functioning predict auditory speech perception (Adank and Janse 2010; Cahana-Amitay et al. 2016; Füllgrabe et al. 2015). The current investigation expands upon these findings to determine the extent to which auditory neural processing, perceptuomotor processing speed, and executive functioning independently predict time-compressed speech identification in young and old normal-hearing listeners. Consistent with previous studies, the P1 and N1 components of click-induced auditory-evoked potentials were employed as a general (nonspeech-specific) neural measure of auditory neural processing (e.g., Narne and Vanaja 2008; Starr and Rance 2015). These components are obligatory cortical responses to auditory stimulation, the amplitude and latencies of which represent the degree of neural synchrony and time it takes for auditory stimulation to be transmitted from the ear to auditory cortex, respectively. Perceptuomotor processing speed and executive functioning were measured using the Connections Test, which we have previously found to predict auditory gap detection (Harris et al. 2012). Time-compressed speech identification was measured using the same methods and stimuli employed by Gordon-Salant and Fitzgibbons (1993).

We hypothesized that more accurate time-compressed speech identification would be predicted by shorter auditory response latencies, faster perceptuomotor processing, and more efficient executive functioning. Furthermore, consistent with theoretical accounts suggesting that low-level peripheral and high-level processes independently contribute to the processing of time-compressed speech (e.g., Gordon-Salant and Fitzgibbons 2001; Pichora-Fuller 2003), we hypothesized that auditory response latencies, perceptuomotor processing, and executive functioning would each account for *unique* variance in time-compressed speech identification. Given that perceptuomotor processing has been found to decline with age (Eckert et al. 2010; Füllgrabe et al. 2015), we also hypothesized that the variability in time-compressed speech identification accounted for by age would be mediated by perceptuomotor processing speed.

METHOD

All procedures were conducted in a sound-attenuated chamber shielded for electrophysiology recording. Time-compressed speech and click stimuli were presented to the right ear using an ER 3C insert earphone using a combination of MATLAB and a Tucker-Davis Technologies RZ6 Auditory Processor.

Participants

A group of 23 younger adults between the ages of 19 and 30 (15 female) and 33 older adults between the ages of 56 and 82 (23 female) participated in this study. Younger participants generally consisted of volunteer graduate students at the Medical University of South Carolina and volunteer undergraduate and graduate students from the College of Charleston who had previously participated in hearing studies. Older participants consisted of volunteers from an ongoing longitudinal study on presbycusis and volunteers from the greater Charleston, SC community. Many participants, both younger and older, were familiar with laboratory testing but were unfamiliar with the measures employed in the current investigation. All participants were native English-speaking and were found to have little to no cognitive impairment, having completed the Mini-Mental State Examination (Folstein et al. 1983) with three or fewer errors. Participants provided written informed consent before participating in this Medical University of South Carolina Institutional Review Board-approved study.

Pure-tone thresholds at conventional audiometric frequencies were measured using TDH-39 headphones connected to a Madsen OB922 clinical audiometer that was calibrated to appropriate American National Standards Institute (ANSI) specifications (ANSI 2010). Participants had audiometric thresholds ≤ 25 dB HL between 250 and 3000 Hz, with interaural asymmetries less than 10 dB HL at any frequency. Audiometric thresholds for the test ear (right ear) are reported in Fig. 1.

Time-Compressed Speech Identification

Time-compressed stimuli were a subset of the same stimuli used by Gordon-Salant and Fitzgibbons (1993, 2001), constructed from 180 low-probability sentences taken from the Revised Speech Perception in Noise Test (R-SPIN; Bilger et al. 1984). These low-probability sentences were selected because they provided enough acoustic material for the time compression distortion while not providing semantic or contextual cues to aid identification of the final target word in each sentence (e.g., Gordon-Salant and Fitzgibbons

Fig. 1. Right-ear pure-tone audiometric thresholds for older and younger listeners. Error bars represented 95 % confidence intervals

1993). For example, “They were interested in the strap” is grammatically correct, but the leading portion of the sentence provides little semantic or contextual information relevant to the final word. Sentences were digitized (10 kHz sampling rate), time-compressed (using the procedure described below), and low-pass-filtered (5000 Hz normal cutoff, 104 dB/octave attenuation rate) (Gordon-Salant and Fitzgibbons 1993, 2001).

The speech rate of the original untransformed sentences was approximately 200 words per minute (Gordon-Salant and Fitzgibbons 2004; Gordon-Salant et al. 2011). Time compression was accomplished by first dividing the sentences into four different lists: A 50-sentence 40 % time-compressed list, a 50-sentence 50 % time-compressed list, a 50-sentence 60 % time-compressed list, and a 30-sentence practice list comprised of 10 each 40, 50, and 60 % time-compressed sentences. The percentage given for compression describes the percentage that each sentence is compressed in length. For example, a 60 % time-compressed sentence is compressed by 60 % of its original duration, which makes the new duration 40 % of its original length. Time compression was achieved using the Global Duration option of WEDW software (Bunnell 2005). Periodic information was uniformly removed from the untransformed acoustic signals to reduce the duration of each sentence by 40, 50, or 60 %, depending on the stimulus condition. This transformation largely preserves the spectral aspects of the acoustic signal but reduces the duration in which the phonetic information is presented (Gordon-Salant and Fitzgibbons 1993, 1999, 2001; Gordon-Salant and Friedman 2011).

Gordon-Salant and Fitzgibbons (2001) found that the time compression of select consonant features had the strongest effect on time-compressed speech

identification. However, normal-hearing older and younger listeners still exhibited greater difficulty identifying time-compressed speech when periodic information was uniformly removed across speech utterances, suggesting that the time compression of phonetic features across an utterance can affect time-compressed speech identification. The current investigation is concerned with age-related differences in the general processing of time-compressed speech, without regard to specific phonetic features. Therefore, sentences were uniformly time-compressed, involving the uniform removal of periodic information across speech utterances, consistent with previous studies investigating age-related differences in time-compressed speech identification (Gordon-Salant and Fitzgibbons 1993, 1999; Gordon-Salant et al. 2014).

The 40, 50, and 60 % time compression rates were selected for the current investigation because their identification has previously been found to demonstrate robust age effects. Typically, younger and older adults similarly identify 40 % time-compressed speech near ceiling, but older adults have more difficulty identifying 50 and 60 % time-compressed speech (e.g., Gordon-Salant and Fitzgibbons 1993, 1999; Gordon-Salant and Friedman 2011). Employing multiple time compression rates allowed us to evaluate whether the amount of periodic information removed from a signal affects speech identification differently for older and younger adults (i.e., interaction of time compression rate and age group). Previous studies investigating time-compressed speech identification have included an untransformed, normal rate speech condition, which has proven important when considering the effects of hearing loss on the processing of time-compressed speech (Gordon-Salant and Fitzgibbons 1993, 1999, 2001; Gordon-Salant et al. 2007). However, younger and older normal-hearing adults, such as our participants, typically perform at ceiling and do not show significant differences in the identification of naturally spoken sentences presented in clear listening conditions (e.g., Gordon-Salant and Fitzgibbons 1993, 1999, 2001; Gordon-Salant et al. 2007; Gordon-Salant and Friedman 2011; Gordon-Salant et al. 2014). As such, we decided to forgo the use of an untransformed, normal rate speech condition for the current investigation, similar to other studies investigating age differences in time-compressed speech identification (e.g., Vaughan and Letowski 1997), interrupted speech identification (e.g., Benard et al. 2014; Kidd and Humes 2012), and speech-in-noise identification (e.g., Campbell et al. 2007; Dey and Sommers 2015; Helfer 1998; Presacco et al. 2016; Tye-Murray et al. 2007; Vaden et al. 2015).

For the speech identification task, each participant was fit with an insert earphone in their right ear and

presented time-compressed sentences at 90 dB SPL, an intensity that allows for comparison of the current results to past time-compressed speech studies (e.g., Gordon-Salant and Fitzgibbons 1993, 1999, 2001; Gordon-Salant et al. 2007) and below intensities found to induce rollover effects when identifying speech in quiet listening conditions (e.g., Molis and Summers 2003; Studebaker et al. 1999). The participants' task was to verbally identify as many discernable words as possible following the presentation of each sentence. Participants were instructed to make a guess if they were unsure of a word or phrase. First, they were presented the list of 30 practice sentences. Following the initial practice, participants were then presented the 40, 50, and 60 % lists of sentences in random order. A microphone equipped in the booth allowed a researcher to hear a participant's verbal responses from outside of the booth and score them for accuracy. Accuracy was coded as the proportion of final position words correctly identified across time-compressed sentences, consistent with how these time-compressed sentences have been scored before (Gordon-Salant and Fitzgibbons 1993) and how R-SPIN sentences are scored for clinical assessments (Bilger et al. 1984). Two older participants were excluded from further analysis for having accuracy scores more than two standard deviations below the mean across participants. Testing was performed only in the right ear. Right-handed individuals often demonstrated a right-ear advantage for speech perception, typically associated with a contralateral (left hemisphere) specification for speech and language processing (for a review, see Tervaniemi and Hugdahl 2003). However, the performance of the two left-handed older adults and one left-handed younger adult included in our sample fell within the range of our right-handed participants and so were not excluded from analysis. Consistent with previous studies, scores were arcsine-transformed to account for the small amounts of variability due to high performance in the low-time-compressed condition, providing more statistical power when evaluating the relationships between performance in the low-time-compressed condition with other factors (e.g., Gordon-Salant and Fitzgibbons 1993; Gordon-Salant et al. 2007).

Auditory-Evoked Response

Electroencephalography (EEG) was recorded while participants passively listened to a 100- μ s 100 dB pSPL alternating polarity click with a 2000-ms interstimulus interval (20 ms jitter) through an insert earphone in their right ear. Four hundred clicks were presented across two sessions. EEG was recorded using a 64-channel Neuroscan QuickCap based on the interna-

tional 10–20 system connected to a SynAmpsRT 64-Channel Amplifier. Recordings were done at a sampling rate of 1000 Hz. Bipolar electrodes were placed above and below the left eye to record vertical electro-oculogram activity. Curry 7 software was used to record the EEG signal. Offline, the recorded EEG data was down-sampled to 500 Hz, bandpass filtered from 1 to 30 Hz, re-referenced to the average, and corrected for ocular artifacts by removing ocular source components using independent components analysis. Individual trials were then segmented into epochs around the click onset (–100 to +500 ms) and then baseline-corrected (–100 to 0 ms). Epoched data was then subjected to an automatic artifact rejection algorithm. Any epochs contaminated by peak-to-peak deflections in excess of 100 μ V were rejected.

For each participant, epoched data was averaged across trials to compute the average waveform (event-related potential—ERP). ERPs were analyzed in a cluster of frontal-central electrodes: F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, and C2 (e.g., Harris et al. 2012; Tremblay et al. 2001; Narne and Vanaja 2008). To do this, ERPs were first averaged across these channels. Then the P1 and N1 peaks, defined as the local maxima in the first prominent positive peak and first prominent negative peak, respectively, were found. The latencies of these peaks were defined as the times after stimulus onset when the local maxima occurred. Intervals were then computed around the P1 and N1 peak latencies (± 25 ms). These intervals were then used to automatically find the P1 and N1 peak amplitudes and latencies in each of our channels of interest. The amplitudes and latencies across channels were then averaged to compute single measures of each participant's P1 and N1 peak amplitude and peak latency. In cases where a peak was not found in a channel, that channel was not considered in the computed averages (e.g., Anderer et al. 1996). Three participants (one older, two younger) were removed from further analysis due to a lack of measurable P1 or N1 across channels.

Processing Speed

Processing speed was measured using the Connections Test (Salthouse et al. 2000). The Connections Test, a variant of the Trail-Making Test (Reitan 1958, 1992), requires participants to connect circled letters and/or numbers in alphanumeric order. The letters and/or numbers are pseudorandomly organized into a 7×7 array on a single sheet of 8.5 in. $\times$ 11 in. paper. For any single target, the next target is located in an adjacent location, which is designed to limit the motor requirements for the task (Salthouse et al. 2000). There are two forms of the test. One form, Connections Simple, requires participants to connect as many numbers or

letters as possible in 20 s, requiring participants to search for and identify targets, determine the next target, and search for, identify, and respond to the next target in a sequence. Scores are averaged across two trials connecting letters and two trials connecting numbers. The second form, Connections Complex, requires participants to connect alternating numbers and letters in alphanumeric order. Scores are averaged across two trials starting with a letter and two trials starting with a number. Connections Complex requires the perceptual and motor skills required for Connections Simple but also involves the added task of switching between numbers and letters, engaging executive functioning (selective attention and inhibition). The Connections Test is postulated to measure both perceptuomotor processing speed (Connections Simple) and executive functioning (Connections Complex). Metrics for perceptuomotor processing speed and executive functioning derived from the Connections Test are strongly associated with scores on other perceptual processing speed and executive functioning tasks, respectively (Salthouse et al. 2000; Salthouse 2005, 2011).

Because both Connections Simple and Connections Complex measure perceptuomotor processing speed, their correlation was quite high, $r = 0.668$, $p < 0.001$. To extract executive functioning from Connections Complex scores, Connections Simple scores were regressed out and the standardized residuals used to represent executive functioning (e.g., Eckert et al. 2010; Salthouse et al. 2000). Using a regression approach, residuals were computed by subtracting Connections Complex scores predicted by Connections Simple scores ($\text{complex}_{\text{predicted}} = a + b(\text{simple}) + \text{error}$) from our observed Connections Complex scores ($\text{complex}_{\text{observed}} - \text{complex}_{\text{predicted}}$). The standardized residuals represent a reliable metric of the executive functioning component of Connections Complex scores that are unrelated to perceptuomotor processing speed (Connections Simple scores), $r < 0.001$, $p > 0.999$. We used the perceptuomotor and executive functioning metrics derived from the Connections Test to determine the extent to which each type of processing accounts for age-related variability in time-compressed speech identification.

Analytical Approach

Statistical analyses were conducted using SPSS and the lme4 package for R ($\times 64$ v3.3.2). Two-tailed independent samples t tests were used to determine whether age-group differences in time-compressed speech identification and our critical predictor variables were consistent with the extant literature. Linear mixed-effects regression (LMER) models were used to determine the extent to which time-compressed

speech identification was predicted by our critical predictor variables. LMER is a useful statistical tool for testing hypothesis-driven relationships between predictor and outcome variables while accounting for cluster-level attributes (nested variables). The standardized and unstandardized coefficients for each predictor in a LMER model represent the relationship between that predictor and the outcome variable after accounting for all other predictors in the model. Using a stepwise approach, the amount of total variance in the outcome variable accounted for by a model can be compared to the total variance accounted for by another model with additional predictor variables. If the addition of more predictors fails to significantly improve model fit, or significantly change the relationships between predictors and the outcome variable, then those additional predictors are typically excluded and the best-fitting model is used (e.g., Gelman and Hill 2007).

RESULTS

Age-Group Differences Across Measures

Age-group differences for our critical variables are consistent with the extant literature. Table 1 reports the descriptive and inferential statistics for these age-group differences. Though the audiometric thresholds for all participants fell within a range that is typically considered clinically normal for speech identification, audiometric thresholds averaged across 250, 500, 1000, 2000, 3000, and 4000 Hz were significantly higher for older listeners' left and right ears. Older listeners had more difficulty accurately identifying 50 and 60 % time-compressed speech, though no age-group difference was found for 40 % time-compressed speech (Fig. 2), replicating previous studies using the same stimuli (e.g., Gordon-Salant and Fitzgibbons 1993, 1999, 2001). Average ERP waveforms for our older and younger listeners are shown in Fig. 3. Consistent with previous studies, older listeners exhibited larger P1 peak amplitudes (e.g., Woods and Clayworth 1986), but N1 amplitudes did not significantly differ between age groups. P1 and N1 peak latencies did not significantly differ between age groups, a finding that is not atypical of the broader literature (e.g., Basta et al. 2005; Gmehlin et al. 2011; Tremblay et al. 2004). Older adults were found to have lower Simple and Complex Connections scores, though the residual metric for executive functioning did not significantly differ between groups (Eckert et al. 2010; e.g., Salthouse 2000).

Age and Time Compression Rate

The first LMER models we tested were designed to confirm the effects of age and time compression rate observed in past studies (e.g., Gordon-Salant and

Fitzgibbons 1993, 1999; Gordon-Salant and Friedman 2011). Table 2 reports how an intercepts model—model 0—including time-compressed speech identification as the dependent variable and participant as a random factor, was significantly improved by adding age group, time compression rate, and their interaction—model 1a— $\chi^2(3) = 194.71$, $p < 0.001$. Replicating previous work (e.g., Gordon-Salant and Fitzgibbons 1993, 2001), model 1a indicates that time-compressed speech identification significantly decreased as age and time compression increased. The interaction of age group and time compression rate indicates that the relationship between compression rate and time-compressed speech identification is stronger (steeper slope) in older listeners, a relationship visible in Fig. 2. Adding audiometric averages for the test ear (right ear) to model 1a did not significantly improve model fit—model 1b— $\chi^2(1) = 0.130$, $p = 0.718$. Audiometric averages were not found to be a significant predictor of time-compressed speech identification, nor did their inclusion in the model significantly change the degree to which age group and time compression rate predicted time-compressed speech identification. This suggests that even though audiometric thresholds increase with age, as reported in Table 1, audiometric thresholds do not mediate the relationship between age and time-compressed speech identification. In other words, variability in time-compressed speech identification accounted for by age group is not explained by individual differences in audiometric thresholds.

Auditory Neural Processing

To test our prediction that the click-induced auditory-evoked response would predict time-compressed speech identification, LMER models independently assessed the role of the P1 and N1 components of the click-induced auditory-evoked response after accounting for age and time compression rate. As reported in Table 3, adding P1 peak latency to model 1a significantly improved model fit—model 2a— $\chi^2(1) = 10.330$, $p < 0.001$. After accounting for age group, time compression rate, and their interaction, P1 latency was found to be a significant predictor of time-compressed speech identification. Listeners with longer P1 peak latencies exhibited poorer time-compressed speech identification (Fig. 4, left). Adding N1 peak latency to model 2a failed to significantly improve model fit—model 2b— $\chi^2(1) = 3.095$, $p = 0.079$, nor did it prove a significant predictor of time-compressed speech identification (Fig. 4, right). Adding P1 and N1 peak amplitudes to model 2a also failed to improve model fit, $\chi^2(2) = 0.913$, $p = 0.634$.

TABLE 1

Summary statistics for age differences across measured variables

Variable	Younger (N = 21)		Older (N = 30)		t				
	M	SE	M	SE	MD	SE	t (df = 49)	p	r _{effect size}
Audiometric Thresholds									
Right ear	1.214	0.736	12.344	0.910	−11.130	1.251	−8.898	<0.001	0.780
Left ear	2.214	0.785	14.161	1.089	−11.947	1.460	−8.184	<0.001	0.753
Time-Compressed-Speech Identification									
40 %	0.955	0.010	0.927	0.013	0.028	0.018	1.634	0.109	0.223
50 %	0.868	0.021	0.795	0.018	0.073	0.028	2.725	0.009	0.357
60 %	0.740	0.038	0.537	0.025	0.203	0.043	4.859	<0.001	0.563
Auditory-Evoked Response									
P1 Peak Amplitude (μV)	0.454	0.076	0.812	0.094	−0.358	0.129	−2.772	0.008	0.362
P1 Peak Latency (ms)	49.722	3.164	46.916	1.885	2.806	3.472	0.808	0.423	0.112
N1 Peak Amplitude (μV)	−0.797	0.120	−0.922	0.133	0.125	0.189	0.664	0.510	0.093
N1 Peak Latency (ms)	93.977	1.620	91.567	1.896	2.410	2.644	0.912	0.366	0.127
Processing Speed									
Connections Simple Score	33.131	1.311	23.183	0.850	9.948	1.494	6.657	<0.001	0.682
Connections Complex Score	17.143	1.051	12.683	0.640	4.460	1.164	3.832	<0.001	0.473
Executive Functioning	0.032	0.256	−0.022	0.157	0.054	0.284	0.190	0.850	0.027

Audiometric Thresholds represent the average audiometric threshold across 250, 500, 1000, 2000, 3000, and 4000 Hz. Arcsine transforms of time-compressed speech accuracy were used for statistical analysis, though untransformed values are reported above for mean accuracy values. Values for Executive Functioning are standardized residuals after regressing Connections Simple scores from Connections Complex scores

Perceptuomotor Processing Speed and Executive Functioning

To test our prediction that perceptuomotor processing speed and executive functioning would predict time-compressed speech identification, LMER models independently assessed the relationships between perceptuomotor processing speed (Connections Simple scores) and executive functioning (the residuals after regressing Connections Simple scores from Connections Complex scores) after accounting for age and time compression rate. As reported in Table 4, adding perceptuomotor processing speed to model 1a improved model fit—model 3a— $\chi^2(1) = 6.768$, $p = 0.009$. Perceptuomotor processing speed significantly predicted time-compressed speech identification after accounting for age group, time compression rate, and their interaction. Listeners with faster perceptuomotor processing exhibited better time-compressed speech identification (Fig. 5, left). Importantly, the effect of age group, which was a significant predictor of time-compressed speech identification in model 1a, became a nonsignificant predictor in model 2a. Recall that perceptuomotor processing speed was slower in older listeners (lower Connections Simple scores—Table 1). The addition of perceptuomotor processing speed in model 2a mediated the effect of age group, suggesting that age-group differences in time-compressed speech identification may be accounted for by individual differences in perceptuomotor processing speed. Adding executive functioning to model 3a improved model fit—model 3b— $\chi^2(1) = 7.527$, $p = 0.006$. Executive

functioning was found to significantly predict time-compressed speech identification after accounting for age group, time compression rate, their interaction, and perceptuomotor processing speed. Listeners with more efficient executive functioning exhibited better time-compressed speech identification (Fig. 5, right).

Fig. 2. Time-compressed speech identification scores for younger and older listeners, based on the proportion of correctly identified final-position words from our time-compressed SPIN sentence stimuli. Older listeners are indicated by the dashed line. Younger listeners are indicated by the solid line. Error bars represent 95 % confidence intervals. Compared to younger adults, the identification scores of older adults were significantly lower at 50 and 60 % time compression, exhibiting greater decreases with increased time compression

FIG. 3. Average auditory-evoked responses for younger and older listeners. Older listeners are indicated by the dashed line. Younger listeners are indicated by the solid line. P1 amplitude was significantly higher for older listeners than for younger listeners, though N1 amplitude showed no significant difference between age groups. P1 and N1 latencies were not significantly different between age groups

Combined Model

To determine the extent to which auditory neural processing, perceptuomotor processing speed, and executive functioning independently predict time-compressed speech identification, models 2a and 3b were combined. Adding Connections Simple scores and executive functioning from model 3b to model 2a improved model fit—model 4— $\chi^2(2) = 13.328$, $p < 0.001$. As reported in Table 5, time compression rate, P1 peak latency, Connections Simple scores,

executive functioning, and the interaction between age group and time compression rate were all found to significantly predict time-compressed speech identification.

DISCUSSION

Our previous work found that auditory temporal processing (gap detection) is predicted by auditory

TABLE 2

Linear mixed effects models - effects of age and time-compression rate

	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficients</i>		<i>df</i>	<i>t</i>	<i>p</i>
	<i>B</i>	<i>SE</i>	<i>β</i>	<i>SE</i>			
Model 0 - Intercepts Model							
Intercept	1.145	0.019	0.000	0.000	49	58.320	<0.001***
Model 1a - Age and Time Compression Rate							
Intercept	1.223	0.027	0.000	0.000	49	45.959	<0.001***
Age Group	−0.134	0.035	−0.272	0.071	49	−3.854	<0.001***
Time-Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 0 to Model 1a, $\chi^2_{change}(3) = 194.71, p < 0.001$							
Model 1b - Age, Time-Compression Rate, and Audiometric Thresholds							
Intercept	1.222	0.027	0.000	0.000	48	44.756	<0.001***
Age group	−0.149	0.057	−0.304	0.115	48	−2.637	0.011*
Time compression rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
Right-ear audiometric thresholds	0.001	0.004	0.040	0.115	48	0.350	0.728
Age group × time compression rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 1a to Model 1b, $\chi^2_{change}(1) = 0.130, p = 0.718$							

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

TABLE 3

Linear mixed effects models - effects of the click-induced auditory event-related potential

	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficients</i>		<i>df</i>	<i>t</i>	<i>p</i>
	<i>B</i>	<i>SE</i>	β	<i>SE</i>			
Model 2a - P1 Peak Latency							
Intercept	1.436	0.069	0.000	0.000	48	20.743	<0.001***
Age Group	−0.146	0.032	−0.296	0.065	48	−4.570	<0.001***
Time-Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
P1 Peak Latency	−0.004	0.001	−0.213	0.065	48	−3.283	0.002**
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 1a to Model 2a, $\chi^2_{change}(1) = 10.330, p = 0.001$							
Model 2b - P1 and N1 Peak Latencies							
Intercept	1.188	0.160	0.000	0.000	47	7.428	<0.001***
Age Group	−0.141	0.031	−0.287	0.064	47	−4.497	<0.001***
Time-Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
P1 Peak Latency	−0.005	0.001	−0.270	0.072	47	−3.762	<0.001***
N1 Peak Latency	0.003	0.002	0.123	0.072	47	1.715	0.093
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 2a to Model 2b, $\chi^2_{change}(1) = 3.095, p = 0.079$							

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

neural processing speed (Harris et al. 2012) and perceptuomotor processing speed (Harris et al. 2010). This past work also found that age differences in auditory temporal processing were mediated by individual differences in perceptuomotor processing speed (Harris et al. 2010). The results of the current investigation expand upon these findings, suggesting that auditory neural processing speed, perceptuomotor processing speed, and executive functioning predict time-compressed speech identifi-

cation in younger and older adults. The results also suggest that age-group differences in time-compressed speech identification are mediated by individual differences in perceptuomotor processing speed. These results support theoretical accounts suggesting that time-compressed speech identification is subject to both low-level peripheral processing and higher-level perceptuomotor processing speed and executive functioning ability (e.g., Gordon-Salant and Fitzgibbons 2001; Pichora-Fuller 2003).

FIG. 4. Time-compressed speech identification scores averaged across compression rate are plotted against click-induced P1 (left) and N1 (right) latencies. Older listeners are indicated by boxes. Younger listeners are indicated by dots. A line represents a significant

linear fit relationship. Though the graphs plot untransformed time-compressed speech scores, all inferential tests were conducted using arcsine transforms. Across age groups, listeners with shorter P1 latencies identified time-compressed speech more accurately

TABLE 4

Linear mixed effects models - effects of perceptuomotor processing and executive functioning

	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficients</i>		<i>df</i>	<i>t</i>	<i>p</i>
	<i>B</i>	<i>SE</i>	β	<i>SE</i>			
Model 3a - Perceptuomotor Processing Speed							
Intercept	0.952	0.107	0.000	0.000	48	8.902	<0.001***
Age Group	−0.052	0.045	−0.106	0.092	48	−1.155	0.254
Time-Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
Connections Simple	0.008	0.003	0.240	0.092	48	2.610	0.012*
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 1a to model 3a, $\chi^2_{change}(1) = 6.768, p = 0.009$							
Model 3b - Executive Functioning							
Intercept	0.944	0.100	0.000	0.000	47	9.398	<0.001***
Age Group	−0.048	0.043	−0.098	0.087	47	−1.128	0.265
Time-Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
Perceptuomotor Processing Speed	0.008	0.003	0.246	0.087	47	2.850	0.006**
Executive Function	0.042	0.015	0.171	0.063	47	2.734	0.009**
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 3a to model 3b, $\chi^2_{change}(1) = 7.527, p = .006$							

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Auditory Neural Processing

Shorter click-induced P1 latencies predicted better time-compressed speech identification, suggesting that individual differences in auditory neural processing speed can modulate the processing of time-compressed speech. Surprisingly, our results suggest a relationship between P1 and time-compressed speech, and not N1 and time-compressed speech. Previously, we reported that N1 latencies, when induced by the offset of a noise (onset of a silent gap), predicted auditory temporal processing (audito-

ry gap detection) (Harris et al. 2012). Similarly, the latency of N1 following a tone or syllable has been found to play a role in the discrimination of syllables differing in voice-onset time (Tremblay et al. 2004; Tremblay et al. 2002, 2003). Differences across these studies are likely due to the eliciting stimulus. In these past studies, the N1 response was elicited by the offset of noise or syllables. For example, in our previous gap detection study, the offset of noise did not generate a P1 response, and therefore, N1 represents the first electrophysiological indicator of stimulus change,

FIG. 5. Time-compressed speech identification scores averaged across compression rate are plotted against perceptuomotor processing speed (Connections Simple scores; left) and executive functioning (standardized residuals of Connections Complex scores after regressing out Connections Simple scores; right). Older listeners are indicated by boxes. Younger listeners are indicated by dots. Lines

represent significant linear fit relationships. Though the graphs plot untransformed time-compressed speech scores, all inferential tests were conducted using arcsine transforms. Across age groups, listeners with faster perceptuomotor processing and better executive functioning identified time-compressed speech more accurately

TABLE 5
Linear mixed-effects models - combined model

	<i>Unstandardized Coefficients</i>		<i>Standardized Coefficients</i>		<i>df</i>	<i>t</i>	<i>p</i>
	<i>B</i>	<i>SE</i>	β	<i>SE</i>			
Model 4 - Combined Model							
Intercept	1.131	0.111	0.000	0.000	46	10.174	<0.001***
Age Group	−0.060	0.039	−0.123	0.080	46	−1.531	0.133
Time Compression Rate	−0.169	0.016	−0.570	0.053	100	−10.760	<0.001***
P1 Peak Latency	−0.004	0.001	−0.181	0.060	46	−3.045	0.004**
Perceptuomotor Processing Speed	0.008	0.003	0.242	0.080	46	3.028	0.004**
Executive Function	0.033	0.015	0.132	0.059	46	2.226	0.031*
Age Group × Time Compression Rate	−0.085	0.020	−0.221	0.053	100	−4.168	<0.001***
Model 2a to model 4, $\chi^2_{change}(2) = 13.328, p = 0.001$							

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

comparable to P1 elicited by a click in the current study. Moreover, failing to find a significant relationship between the click-induced N1 and time-compressed speech identification is consistent with other studies failing to find a relationship between click-induced N1 latency and speech-in-noise identification (e.g., Narne and Vanaja 2008).

Individual differences in neural measures of auditory temporal processing have been theorized to depend on physiological factors associated with the number of neurons, their myelination, and the synchronicity of their firing (e.g., Rance 2005; Starr and Rance 2015). P1 is thought to be generated by primary auditory cortex and to represent one of the earliest cortical responses to auditory stimulation (e.g., Pratt 2012). Click-induced P1 has been thought to reflect a predominantly preattentive response (Lijffijt et al. 2009). Shorter click-induced P1 latencies, resulting from better myelination and more synchronous firing of neurons along the auditory path, may facilitate time-compressed speech identification by preserving the early neural representations of meaningful phonetic features within the acoustic signal.

Perceptuomotor Processing and Executive Functioning

We found that faster perceptuomotor processing and more efficient executive functioning predicted better time-compressed speech identification across age groups. Importantly, we found that age-group differences in time-compressed speech identification were mediated by individual differences in perceptuomotor processing speed. Recall that perceptuomotor processing speed (Connections Simple) was significantly slower for older adults, though executive functioning and auditory neural processing speed were similar between age groups. When perceptuomotor processing speed was added to model 1a, which included age

group and time compression as significant predictors of time-compressed speech identification, the resulting model 3a found that time-compressed speech identification was significantly predicted by perceptuomotor processing speed and time compression, but not age group (Table 4). This mediation effect suggests that age-related declines in time-compressed speech identification may be partially accounted for by individual differences in perceptuomotor processing speed.

Faster perceptuomotor processing speed may affect time-compressed speech identification by facilitating the ability to track and extract meaningful phonetic features from an acoustic signal. The results are consistent with the findings of Adank and Janse (2010) and Füllgrabe et al. (2015). Recall from the introduction that Adank and Janse (2010) found that perceptuomotor processing measured using the Digit Symbol Substitution Task predicted identification of accented speech. Using the Trail-Making Task, Füllgrabe et al. (2015) found that better performance on Trail-Making Test B, the complex form that measures both perceptuomotor processing speed and executive function, predicted better processing of auditory temporal fine structures (tone discrimination) and better speech-in-noise identification. However, their metric for executive functioning derived from the simple and complex forms of the Trail-Making Test did not predict perception of temporal fine structure or speech-in-noise identification, suggesting the relationships with Trail-Making Test B scores were driven by perceptuomotor processing speed.

Our results contrast Füllgrabe et al. (2015) in that we find a significant relationship between executive functioning and time-compressed speech identification independent of perceptuomotor processing speed and auditory neural processing. The difference between these studies may be explained by the different types of test stimuli. Recall that our

sentences presented in quiet listening conditions were all low-probability SPIN sentences, where the final position key word could not be predicted from the preceding content. Füllgrabe et al. (2015) used the Adaptive Sentence Lists, which contain some contextual cues (Macleod and Summerfield 1990), presented with competing speech streams (two other talkers). In contrast, our results are consistent with the findings of Cahana-Amitay et al. (2016). Using the Trail-Making Test, they found that better executive functioning predicted better identification of low-probability, but not high-probability, SPIN sentences presented in low levels of noise. Their results suggest that in quieter listening environments, more similar to the quiet listening conditions our stimuli were presented in, executive functioning is more important for selectively attending to speech content and inhibiting irrelevant context in low-probability SPIN sentences. Similarly, our positive association between executive functioning and time-compressed speech identification is consistent with the findings of Adank and Janse (2010), who found that executive functioning measured using the Trail-Making Test predicted better identification of sentences spoken in an unfamiliar accent, but not sentences spoken in a familiar accent. Both time compression and accentedness change the temporal characteristics of speech and both appear to be affected by executive functioning. Taken together, the evidence is mounting in support of a role for executive functioning in processing temporally distorted speech.

The Connections Test is a visual-based test of perceptuomotor processing speed and executive functioning, measuring the speed at which an individual can connect letters and numbers. Yet scores still predict age-related declines in auditory temporal processing (Füllgrabe et al. 2015; Harris et al. 2012), speech-in-noise identification (Cahana-Amitay et al. 2016; Füllgrabe et al. 2015; Helfer and Freyman 2014), accented speech identification (Adank and Janse 2010), and now time-compressed speech identification. The relationships observed across these studies may suggest that speech identification reflects, in part, global modality-neutral perceptuomotor processing speed and executive functioning.

Individual differences and age-related declines in processing speed (including perceptuomotor processing and executive functioning) have been theorized to result from multiple contributing factors. For example, age-related slowing of processing speed has been associated with slower global cortical neural processing speed and degradation of white-matter microstructure (Valdés-Hernández et al. 2010; Vlahou et al. 2014). Work going forward should explore the degree to which the underlying neural mechanisms that predict processing speed, such as cortical oscillatory behavior and white-matter microstructure (Richard

Clark et al. 2004; Howard et al. 2017; Gierhan 2013), can predict age-related changes in perceptuomotor processing speed, executive functioning, and time-compressed speech identification.

Unique Roles of Auditory Neural Processing, Perceptuomotor Processing, and Executive Functioning

Across age groups, click-induced P1 latency, perceptuomotor processing speed, and executive functioning all account for unique variance in time-compressed speech identification, even after accounting for variability explained by age and time compression rate. These results support theories suggesting that time-compressed speech identification is subject to both low-level and high-level processes (e.g., Gordon-Salant and Fitzgibbons 2001; Pichora-Fuller and Levitt 2012). The results are also consistent with our previous work suggesting that variability in auditory temporal processing can be accounted for by individual differences in auditory neural and perceptuomotor processing speed (e.g., Harris and Dubno 2017; Harris et al. 2012).

The role of perceptuomotor processing speed in partially accounting for age-related variability in time-compressed speech identification has implications for theoretical accounts of age-related speech difficulties. Though we find that younger and older adults are similarly able to easily identify speech at low compression rates (40 %), older adults are more affected by increased compression, an effect observed across many studies investigating age differences in time-compressed speech identification (Gordon-Salant and Fitzgibbons 1993, 1999; Gordon-Salant and Friedman 2011; Gordon-Salant et al. 2007). Our results suggest that as perceptuomotor processing declines with age, the ability to process temporally distorted speech also declines (see also Adank and Janse 2010). Declines in perceptuomotor processing may account for the difficulties that normal-hearing older adults have processing many different types of temporally distorted speech (i.e., fast-rate, reverberant, and accented speech) (Adams et al. 2010; Adank and Janse 2010; Fujihira et al. 2017; Gordon-Salant and Fitzgibbons 1993; Gordon-Salant et al. 2010a, b; Gordon-Salant et al. 2015; Halling and Humes 2000; Neuman et al. 2010).

ACKNOWLEDGEMENTS

We would like to thank Sandra Gordon-Salant for providing the time-compressed stimuli for this study.

FUNDING INFORMATION

This work was supported (in part) by grants from the National Institute on Deafness and Other Communication Disorders (NIDCD) (R01 DC014467, P50 DC00422, and T32 DC014435). The project also received support from the South Carolina Clinical and Translational Research (SCTR) Institute with an academic home at the Medical University of South Carolina, National Institute of Health/National Center for Research Resources (NIH/NCRR) Grant number UL1RR029882. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program Grant Number C06 RR1 4516 from the National Center for Research Resources, NIH.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest The authors declare that they have no conflict of interest.

REFERENCES

- ADAMS EM, GORDON-HICKEY S, MOORE RE, MORLAS H (2010) Effects of reverberation on acceptable noise level measurements in younger and older adults. *Int J Audiol* 49(11):832–838. <https://doi.org/10.3109/14992027.2010.491096>
- ADANK P, JANSE E (2010) Comprehension of a novel accent by young and older listeners. *Psychol Aging* 25(3):736–740
- AHISSAR E, AHISSAR M (2005) Processing of the temporal envelope of speech. In: Heil P, Scheich H, Budinger E, König R (eds) *The auditory cortex: a synthesis of human and animal research*. Lawrence Erlbaum, Mahwah, NJ, pp 295–313
- ANDERER P, SEMLITSCH HV, SALETU B (1996) Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. *Electroencephalogr Clin Neurophysiol* 99(5):458–472. [https://doi.org/10.1016/S0013-4694\(96\)96518-9](https://doi.org/10.1016/S0013-4694(96)96518-9)
- ANDERSON S, PARBERY-CLARK A, YI H-G, KRAUS N (2011) A neural basis of speech-in-noise perception in older adults. *Ear Hear* 32(6):750–757. <https://doi.org/10.1097/AUD.0b013e31822229d3>
- ANSI (2010) Specification for audiometrics. American National Standards Institute, New York
- BASTA D, TODT I, ERNST A (2005) Normative data for P1/N1-latencies of vestibular evoked myogenic potentials induced by air- or bone-conducted tone bursts. *Clin Neurophysiol* 116(9):2216–2219. <https://doi.org/10.1016/j.clinph.2005.06.010>
- BENARD MR, MENSINK JS, BASKENT D (2014) Individual differences in top-down restoration of interrupted speech: links to linguistic and cognitive abilities. *The Journal of the Acoustical Society of America* 135(2):EL88–EL94. <https://doi.org/10.1121/1.4862879>
- BILGER RC, NUETZEL JM, RABINOWITZ WM, RZECZKOWSKI C (1984) Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research* 27(1):32–48. <https://doi.org/10.1044/jshr.2701.32>
- BOLOGNA WJ (2017) Age effects on perceptual organization of speech in realistic environments. (Doctor of Philosophy), University of Maryland, College Park, MD
- BUNNELL, T. (2005). SPEECH AND ASP SOFTWARE. RETRIEVED FROM WWW.ASEL.UDEL.EDU/SPEECH/SPCH_PROC/SOFTWARE.HTML
- CAHANA-AMITAY D, SPIRO A, SAYERS JT, OVEIS AC, HIGBY E, OJO EA ET AL (2016) How older adults use cognition in sentence-final word recognition. *Aging Neuropsychol Cognit* 23(4):418–444. <https://doi.org/10.1080/13825585.2015.1111291>
- CAMPBELL M, PREMINGER JE, ZIEGLER CH (2007) The effect of age on visual enhancement in adults with hearing loss. *The Journal of the Academy of Rehabilitative Audiology* 40:11–32
- DEY A, SOMMERS MS (2015) Age-related differences in inhibitory control predict audiovisual speech perception. *Psychol Aging* 30(3):634–646. <https://doi.org/10.1037/pag0000033>
- DUBNO JR, DIRKS DD, MORGAN DE (1984) Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America* 76(1):87–96. <https://doi.org/10.1121/1.391011>
- ECKERT M, KEREN N, ROBERTS D, CALHOUN V, HARRIS K (2010) Age-related changes in processing speed: unique contributions of cerebellar and prefrontal cortex. *Front Hum Neurosci* 4(10). <https://doi.org/10.3389/neuro.09.010.2010>
- ELLIS RJ, MUNRO KJ (2013) Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int J Audiol* 52(1):14–22. <https://doi.org/10.3109/14992027.2012.721013>
- FOLSTEIN MF, ROBINS LN, HELZER JE (1983) The mini-mental state examination. *Arch Gen Psychiatry* 40(7):812. <https://doi.org/10.1001/archpsyc.1983.01790060110016>
- FUJIIHARA H, SHIRASHI K, REMIJN GB (2017) Elderly listeners with low intelligibility scores under reverberation show degraded subcortical representation of reverberant speech. *Neurosci Lett* 637:102–107. <https://doi.org/10.1016/j.neulet.2016.11.042>
- FÜLLGRABE C, MOORE BCJ, STONE MA (2015) Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci* 6(347). <https://doi.org/10.3389/fnagi.2014.00347>
- GELFAND SA, PIPER N, SILMAN S (1986) Consonant recognition in quiet and in noise with aging among normal hearing listeners. *The Journal of the Acoustical Society of America* 80(6):1589–1598. <https://doi.org/10.1121/1.394323>
- GELMAN A, HILL J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, NY
- GIERHAN SME (2013) Connections for auditory language in the human brain. *Brain Lang* 127(2):205–221. <https://doi.org/10.1016/j.bandl.2012.11.002>
- GMEHLIN D, KREISEL SH, BACHMANN S, WEISBROD M, THOMAS C (2011) Age effects on preattentive and early attentive auditory processing of redundant stimuli: is sensory gating affected by physiological aging? *The Journals of Gerontology: Series A* 66A(10):1043–1053. <https://doi.org/10.1093/gerona/glr067>
- GORDON-SALANT S, FITZGIBBONS PJ (1993) Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research* 36(6):1276–1285. <https://doi.org/10.1044/jshr.3606.1276>
- GORDON-SALANT S, FITZGIBBONS PJ (1999) Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research* 42(2):300–311. <https://doi.org/10.1044/jshr.4202.300>
- GORDON-SALANT S, FITZGIBBONS PJ (2001) Sources of age-related recognition difficulty for time-compressed speech. *Journal of Speech, Language, and Hearing Research* 44(4):709–719. [https://doi.org/10.1044/1092-4388\(2001\)056](https://doi.org/10.1044/1092-4388(2001)056)
- GORDON-SALANT S, FITZGIBBONS PJ (2004) Effects of stimulus and noise rate variability on speech perception by younger and older adults. *The Journal of the Acoustical Society of America* 115(4):1808–1817. <https://doi.org/10.1121/1.1645249>

- GORDON-SALANT S, FITZGIBBONS PJ, FRIEDMAN SA (2007) Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners. *Journal of Speech, Language, and Hearing Research* 50(5):1181–1193. [https://doi.org/10.1044/1092-4388\(2007/082\)](https://doi.org/10.1044/1092-4388(2007/082))
- GORDON-SALANT S, FITZGIBBONS PJ, YENI-KOMSHIAN GH (2011) Auditory temporal processing and aging: implications for speech understanding of older people. *Audiology Research* 1(1):e4. <https://doi.org/10.4081/audiore.2011.e4>
- GORDON-SALANT S, FRIEDMAN SA (2011) Recognition of rapid speech by blind and sighted older adults. *Journal of Speech, Language, and Hearing Research* 54(2):622–631. [https://doi.org/10.1044/1092-4388\(2010/10-0052\)](https://doi.org/10.1044/1092-4388(2010/10-0052))
- GORDON-SALANT S, YENI-KOMSHIAN GH, FITZGIBBONS PJ (2010A) Recognition of accented English in quiet and noise by younger and older listeners. *The Journal of the Acoustical Society of America* 128(5):3152–3160. <https://doi.org/10.1121/1.3495940>
- GORDON-SALANT S, YENI-KOMSHIAN GH, FITZGIBBONS PJ (2010B) Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss. *The Journal of the Acoustical Society of America* 128(1):444–455. <https://doi.org/10.1121/1.3397409>
- GORDON-SALANT S, YENI-KOMSHIAN GH, FITZGIBBONS PJ, COHEN JI (2015) Effects of age and hearing loss on recognition of unaccented and accented multisyllabic words. *The Journal of the Acoustical Society of America* 137(2):884–897. <https://doi.org/10.1121/1.4906270>
- GORDON-SALANT S, ZION DJ, ESPY-WILSON C (2014) Recognition of time-compressed speech does not predict recognition of natural fast-rate speech by older listeners. *The Journal of the Acoustical Society of America* 136(4):EL268–EL274. <https://doi.org/10.1121/1.4895014>
- HALLING DC, HUMES LE (2000) Factors affecting the recognition of reverberant speech by elderly listeners. *Journal of Speech, Language, and Hearing Research* 43(2):414–431. <https://doi.org/10.1044/jslhr.4302.414>
- HARRIS KC, DUBNO JR (2017) Age-related deficits in auditory temporal processing: unique contribution of neural dyssynchrony and slowed neuronal processing. *Neurobiol Aging* 53:150–158
- HARRIS KC, ECKERT MA, AHLSTROM JB, DUBNO JR (2010) Age-related differences in gap detection: effects of task difficulty and cognitive ability. *Hear Res* 264(1–2):21–29. <https://doi.org/10.1016/j.heares.2009.09.017>
- HARRIS KC, WILSON S, ECKERT MA, DUBNO JR (2012) Human evoked cortical activity to silent gaps in noise: effects of age, attention, and cortical processing speed. *Ear Hear* 33(3):330–339. <https://doi.org/10.1097/AUD.0b013e31823fb585>
- HELPER KS (1998) Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology* 9:234–242
- HELPER KS, FREYMAN RL (2014) Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America* 136(2):748–759. <https://doi.org/10.1121/1.4887463>
- HOWARD CJ, ARNOLD CPA, BELMONTE MK (2017) Slower resting alpha frequency is associated with superior localisation of moving targets. *Brain Cogn* 117(Supplement C):97–107. <https://doi.org/10.1016/j.bandc.2017.06.008>
- HUMES LE (1996) Speech understanding in the elderly. *Journal-American Academy of Audiology* 7:161–167
- JANSE E (2009) Processing of fast speech by elderly listeners. *The Journal of the Acoustical Society of America* 125(4):2361–2373. <https://doi.org/10.1121/1.3082117>
- KIDD GR, HUMES LE (2012) Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences. *The Journal of the Acoustical Society of America* 131(2):1434–1448. <https://doi.org/10.1121/1.3675975>
- KOCH X, JANSE E (2016) Speech rate effects on the processing of conversational speech across the adult life span. *The Journal of the Acoustical Society of America* 139(4):1618–1636. <https://doi.org/10.1121/1.4944032>
- LIJFFIJT M, LANE SD, MEIER SL, BOUTROS NN, BURROUGHS S, STEINBERG JL, ... SWANN AC (2009). P50, N100, AND P200 SENSORY GATING: RELATIONSHIPS WITH BEHAVIORAL INHIBITION, ATTENTION, AND WORKING MEMORY. *PSYCHOPHYSIOLOGY*, 46(5), 1059–1068. <https://doi.org/10.1111/j.1469-8986.2009.00845.x>
- MACLEOD A, SUMMERFIELD Q (1990) A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br J Audiol* 24(1):29–43. <https://doi.org/10.3109/03005369009077840>
- MOLIS MR, SUMMERS V (2003) Effects of high presentation levels on recognition of low- and high-frequency speech. *Acoustics Research Letters Online* 4(4):124–128. <https://doi.org/10.1121/1.1605151>
- MOORE BCJ (2014) The role of TFS in speech perception. In: *Auditory processing of temporal fine structure*. World Scientific, pp 81–102
- MOORE BCJ, GLASBERG BR, STOEVE M, FÜLLGRABE C, HOPKINS K (2012) The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *The Journal of the Acoustical Society of America* 131(2):1003–1006. <https://doi.org/10.1121/1.3672808>
- NARNE V K, VANAJA C (2008) Speech identification and cortical potentials in individuals with auditory neuropathy. *Behav Brain Funct* 4(1):15. <https://doi.org/10.1186/1744-9081-4-15>
- NEUMAN AC, WROBLEWSKI M, HAJICEK J, RUBINSTEIN A (2010) Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear* 31(3):336–344. <https://doi.org/10.1097/AUD.0b013e3181d3d514>
- NOORSKI KV, REALE RA, OYA H, KAWASAKI H, KOVACH CK, CHEN H ET AL (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci* 29(49):15564–15574. <https://doi.org/10.1523/JNEUROSCI.3065-09.2009>
- PICHORA-FULLER MK (2003) Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. *International Journal of Audiology* 42:S59–S67
- PICHORA-FULLER MK, LEVITT H (2012) Speech comprehension training and auditory and cognitive processing in older adults. *Am J Audiol* 21(2):351–357. [https://doi.org/10.1044/1059-0889\(2012/12-0025\)](https://doi.org/10.1044/1059-0889(2012/12-0025))
- PICHORA-FULLER MK, SINGH G (2006) Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification* 10(1):29–59. <https://doi.org/10.1177/108471380601000103>
- PRATT H (2012) Sensory ERP components. In: Luck SJ, Kappenman ES (eds) *The Oxford handbook of event-related potential components*. Oxford University Press, Oxford, pp 96–114
- PRESACCO A, SIMON JZ, ANDERSON S (2016) Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol* 116(5):2356–2367. <https://doi.org/10.1152/jn.00373.2016>
- RANCE G (2005) Auditory neuropathy/dys-synchrony and its perceptual consequences. *Trends in Amplification* 9(1):1–43. <https://doi.org/10.1177/108471380500900102>
- REITAN RM (1958) Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills* 8(3):271–276. <https://doi.org/10.2466/pms.1958.8.3.271>
- REITAN, R. M. (1992). TRAIL MAKING TEST: MANUAL FOR ADMINISTRATION AND SCORING [ADULTS]: REITAN NEUROPSYCHOLOGY LABORATORY, TUCSON
- RICHARD CLARK C, VELTMAYER MD, HAMILTON RJ, SIMMS E, PAUL R, HERMENS D, GORDON E (2004) Spontaneous alpha peak frequency predicts working memory performance across the age span. *Int J Psychophysiol* 53(1):1–9. <https://doi.org/10.1016/j.jpsycho.2003.12.011>

- SALTHOUSE TA (2000) Aging and measures of processing speed. *Biol Psychol* 54(1–3):35–54. [https://doi.org/10.1016/S0301-0511\(00\)00052-1](https://doi.org/10.1016/S0301-0511(00)00052-1)
- SALTHOUSE TA (2005) Relations between cognitive abilities and measures of executive functioning. *Neuropsychology* 19(4):532–545. <https://doi.org/10.1037/0894-4105.19.4.532>
- SALTHOUSE TA (2011) What cognitive abilities are involved in trail-making performance? *Intelligence* 39(4):222–232. <https://doi.org/10.1016/j.intell.2011.03.001>
- SALTHOUSE TA, TOTH J, DANIELS K, PARKS C, PAK R, WOLBRETTE M, HOCKING KJ (2000) Effects of aging on efficiency of task switching in a variant of the Trail Making Test. *Neuropsychology* 14(1):102–111. <https://doi.org/10.1037/0894-4105.14.1.102>
- STARR A, RANCE G (2015) Auditory neuropathy. *Handb Clin Neurol* 129:495–508
- STUDEBAKER GA, SHERBECOE RL, MCDANIEL DM, GWALTNEY CA (1999) Monosyllabic word recognition at higher-than-normal speech and noise levels. *The Journal of the Acoustical Society of America* 105(4):2431–2444. <https://doi.org/10.1121/1.426848>
- TERVANIEMI M, HUGDAHL K (2003) Lateralization of auditory-cortex functions. *Brain Res Rev* 43:231–246
- TREMBLAY KL, BILLINGS C, ROHILA N (2004) Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 15(3):226–237. <https://doi.org/10.3766/jaaa.15.3.5>
- TREMBLAY KL, KRAUS N, MCGEE T, PONTON C, OTIS B (2001) Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear* 22(2):79–90
- TREMBLAY KL, PISKOSZ M, SOUZA P (2002) Aging alters the neural representation of speech cues. *NeuroReport* 13(15):1865–1870
- TREMBLAY KL, PISKOSZ M, SOUZA P (2003) Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol* 114(7):1332–1343. [https://doi.org/10.1016/S1388-2457\(03\)00114-7](https://doi.org/10.1016/S1388-2457(03)00114-7)
- TYE-MURRAY N, SOMMERS MS, SPEHAR B (2007) Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear Hear* 28(5):656–668. <https://doi.org/10.1097/AUD.0b013e31812f7185>
- VADEN KI, KUCHINSKY SE, AHLSTROM JB, DUBNO JR, ECKERT MA (2015) Cortical activity predicts which older adults recognize speech in noise and when. *J Neurosci* 35(9):3929–3937. <https://doi.org/10.1523/jneurosci.2908-14.2015>
- VALDÉS-HERNÁNDEZ PA, OJEDA-GONZÁLEZ A, MARTÍNEZ-MONTES E, LAGE-CASTELLANOS A, VIRUÉS-ALBA T, VALDÉS-URRUTIA L, VALDES-SOSA PA (2010) White matter architecture rather than cortical surface area correlates with the EEG alpha rhythm. *NeuroImage* 49(3):2328–2339. <https://doi.org/10.1016/j.neuroimage.2009.10.030>
- 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. <https://doi.org/10.1044/jslhr.4005.1192>
- VLAHOU EL, THURM F, KOLASSA I-T, SCHLEE W (2014) Resting-state slow wave power, healthy aging and cognitive performance. *Sci Rep* 4:5101. <https://doi.org/10.1038/srep05101>
- WOODS DL, CLAYWORTH CC (1986) Age-related changes in human middle latency auditory evoked potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 65(4):297–303. [https://doi.org/10.1016/0168-5597\(86\)90008-0](https://doi.org/10.1016/0168-5597(86)90008-0)
- WOODS WS, KALLURI S, PENTONY S, NOORAEI N (2013) Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *The Journal of the Acoustical Society of America* 133(6):4268–4278. <https://doi.org/10.1121/1.4803859>