Older adults show impaired modulation of attentional alpha oscillations: Evidence from the cocktail party


Volen National Center for Complex Systems, Brandeis University

Abstract
Auditory attention is critical for selective listening to speech from a single talker in a multi-talker environment (e.g., the cocktail party problem, Cherry, 1953). This type of listening is notoriously more difficult and more poorly encoded to long-term memory in older adults than in young adults (Tun, O’Kane, & Wingfield, 2002). Recent work by Payne et al. (2016) in young adults demonstrated a neural correlate of auditory attention in the directed dichotic listening task (DDLT), where listeners are required to attend to one ear while ignoring the other. Measured using electroencephalography (EEG), differences in alpha band power (8-13 Hz) between left and right hemisphere parietal regions mark the direction to which auditory attention is focused. Little prior research has been conducted on alpha power modulations in older adults, particularly with regard to auditory attention directed toward speech stimuli. In the current study, a sample of 19 older adults were run through the DDLT and delayed recognition procedure used by Payne et al. (2016). Compared to the young adults, older adults showed reduced selective attention in the DDLT, evidenced by a higher rate of intrusions from the unattended channel. Moreover, older adults did not exhibit the attention-related alpha modulation evidenced by young adults, nor did their ERPs to recognition probes differentiate between attended or unattended probes. Lastly, older adults did not show any behavioral differentiation between attended and unattended probes during a delayed recognition test, which Payne et al. (2016) reported finding a carryover suppression effect for unattended words by the young adults. We interpret these results as evidence for an age-related decline in selective auditory attention, potentially mediated by an age-related decline in the ability to modulate alpha oscillations.
Author Note

Correspondence concerning this article should be addressed to Dr. Chad S. Rogers, who is now at Department of Otolaryngology, Washington University in St. Louis, 660 South Euclid, Box 8115, St. Louis, MO 63110, phone: 314-362-1929, email: chad.s.rogers@wustl.edu

This research was supported in part by CELEST, an NSF Science of Learning Center (NSF SBE-0354378 and SMA-0835976), NIH T32 NS07292, and NIH R01 AG019714. The authors declare no competing financial interests.

*Author Chad Rogers’ current affiliation is Washington University in St. Louis, St. Louis, MO.

**Author Lisa Payne’s current affiliation is Swarthmore College, Swarthmore, PA.
Introduction

Listening to a conversation in a crowded or noisy room is a common and challenging auditory situation. Not only is the intended speech difficult to hear, but ignoring irrelevant speech is effortful and attention-demanding (Bregman, 1990). In addition to their age-related declines in hearing ability (e.g., CHABA, 1988; Humes, 1996), older adults tend to have general difficulty inhibiting irrelevant stimuli (Hasher & Zacks, 1988; Tun et al., 2002). This combination makes crowded and noisy environments more frustrating for older adults than for young adults (Shinn-Cunningham, 2009). Studies in a laboratory setting show that such difficult listening situations older adults have to expend greater effort than young adults to achieve similar levels of performance as the young, as evidenced by pupil dilation (Zekveld, et al., 2011), reaction time (Hornsby, 2014), and on performance on a secondary task (Tun et al, 2009).

Cherry (1953) famously coined the term “cocktail party problem” to describe the situation in which a listener tries to segregate speech from a single talker amid a group of other, irrelevant, speakers. In order to study the cocktail party problem, he devised the directed dichotic listening task (DDLT), during which listeners were presented with two talkers over earphones, one to each ear, and given cues to indicate to which speaker the participant should attend (e.g., a left-facing arrow cued attention to the speaker in the left ear). Using the DDLT, Craik (1965) found that older adults were more likely to confuse the attended and unattended channels, and attributed this pattern to declines in attentional control. Attentional control deficits have since been well documented in the field of cognitive aging (e.g., Balota et al., 2000). Such deficits have been found to exist even after taking into account age-related declines in basic perceptual processing (Mattys, et al., 2012).

Many studies have subsequently reported that older adults suffer from a deficit in the ability to inhibit irrelevant stimuli (for a review, see Hasher, 2016), which is especially pronounced in groups that experience age-related hearing loss (Dey & Sommers, 2015). In addition, older adults are deficient relative to young adults in terms of elaborating to-be-remembered information (Naveh-Benjman et al., 2005), which could be conceived as a form of memory trace amplification. In cases of both suppression of irrelevant stimuli and elaboration of relevant stimuli, older adults report exerting more effort to achieve a similar level of listening performance than young adults (Tun & Lachman, 2008; Tun & Wingfield, 1995). The exertion of effort has downstream consequences, as effort exerted while processing perceptual stimuli leaves fewer resources available to encode that information into long-term memory (McCoy et al., 2005).
technique for indirectly measuring the downstream effects of inhibitory and elaborative processing is delayed recognition (e.g., Jacoby et al., 2005). Evidence of suppression during a listening task could be evidenced by impaired delayed memory for those target items.

Using a form of the DDLT task, Payne and colleagues (2016) presented young adults with recognition probes and asked them to indicate whether the probe was presented in the to-be-attended ear. Perceptual dampening of irrelevant perceptual stimuli is known to be correlated with alpha-band activity measured via electroencephalography (EEG; Payne & Sekuler, 2014). Increased power in the alpha band (8-14 Hz) is associated with increased ignoring of an irrelevant stimulus (e.g., Wöstmann, et al., 2015), increased working memory demands (e.g., Obleser et al., 2012), and attentional control (e.g., Shinn-Cunningham & Best, 2008). Spatial lateralization of auditory attentional control in the DDLT is marked by inter-hemispheric differences in EEG alpha power (Kerlin et al., 2010; Frey et al., 2014), and is believed to reflect suppression of the unattended auditory speech stream (Choi, et al., 2013; Golumbic et al., 2013; Chait, et al., 2010).

Payne et al. (2016) found that young adults’ successful filtering of unattended speech was associated with modulation of power in the alpha-band during encoding. Alpha modulation was found at two different locations, a lateral frontotemporal and a parietal site. Young adults performed well on the task, but were not perfect—they produced false alarms to words from the unattended ear at a higher rate than to new, extralist words. Also, on delayed recognition, young adults had good memory for words from the attended side, but poorer than chance memory for words from the unattended ear, indicating a kind of carryover suppression similar to that reported in the visual modality by Thomas and Hasher (2012).

The question of whether older adults modulate alpha power to the same extent as young adults remains relatively open unsettled. Prior work in the visual modality has suggested that, as compared to young adults, older adults have diminished alpha modulation (Vaden et al., 2012). However, also with visual stimuli, Schmiedt-Fehr et al. (2009) reported that older adults have intact modulation of alpha relative to young adults. Unfortunately, in both these studies, behavioral performance was at ceiling levels for young and older adults, which made it impossible to discern if modulation of alpha activity is a necessary correlate of task success. After all, it is known that young and older adults could have equivalent performance on a behavioral task, but may engage different cognitive processes and recruit different neurobiological mechanisms to achieve
similar levels of performance (e.g., Della-Maggiore et al., 2000). In addition, these studies were conducted with visual stimuli, in which is easier for researchers to correct for age-related perceptual degradation (e.g., glasses, contact lenses). As a result, work in vision may not generalize to an auditory modality, particularly speech (Weisz, et al., 2011).

In the current study, we invited a group of 19 older adults to the same cocktail party situation that was used by Payne et al. (2016). We then compared their performance to what had been observed in the young adult sample by Payne et al. (2016). Using the same stimuli and procedures as used by Payne and colleagues in their study of young adults, older adult participants listened to speech streams simultaneously presented dichotically (i.e., one stream to each ear), preceded by a cue to attend to the left or right ear. Short-term memory for the content of the to-be-attended ear was assessed immediately following the speech stream with a recognition probe. Following the dichotic task, participants completed a delayed recognition test for all words that had been presented in the study. Performance was assessed and compared to young adults in four ways: 1) alpha modulation during dichotic listening, 2) behavioral responses and response latencies to the short-term recognition probe, 3) whole brain ERPs to the short-term recognition probe, and 4) behavioral responses and latencies during delayed recognition. First, we sought to examine whether older adults’ performance, like young adults’, would also be associated with a pattern of alpha modulation. Second, we wished to test whether the indirect effect of attentional suppression on delayed recognition would generalize to older adults, or whether older adults would fail to show this carryover effect because of an inhibitory deficit.
Method

Participants.

Nineteen older adults gave written informed consent and were paid for participation. Of these, four participants were excluded from our analysis because of excessive EEG artifacts (epoch rejection rate > 50%). The age range of the remaining 15 participants was from 67 - 83 years (mean = 75.40, SD = 4.34) and seven were female. All were right-handed as determined by the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal visual acuity as determined by a minimum angle of resolution chart (ETDRS 2000 Series 2, www.good-lite.com).

The older adult participants in this study met a criterion of clinically normal hearing for speech defined as a best-ear pure tone threshold average (PTA) across 500, 1000 and 2000 Hz of less than 25 dB HL (mean = 13.92 dB HL, SD = 4.96; Mueller & Hall, 1997), and had no significant difference in acuity between the two ears (Left Ear mean = 15.60 dB HL, SD = 4.74; Right Ear mean = 17.50 dB HL, SD = 7.76; t[13] = 1.16, p>.26). The young adults reported by Payne et al. (2016) also met the above criterion for clinically normal speech (mean = 3.08 dB HL, SD = 4.61), and had no significant difference in acuity between the two ears (Left Ear mean = 6.77 dB HL, SD = 5.18; Right Ear mean = 7.5 dB HL, SD = 4.67; t<1, ns). Although the older adults had poorer hearing acuity than the young, we reiterate that they all fell within a range defined as clinically normal hearing for speech (PTA < 25 dB HL; Mueller & Hall, 1997).

Working memory was also measured for both young and older adults using the Reading Span task (Daneman & Carpenter, 1980), where older adults had lower average span scores than young adults (Young mean = 2.39, Older mean = 1.45, t[27] = 3.40, p<.01).

The young adults reported by Payne et al. (2016) had a mean pure tone threshold average of 3.08 dB HL (SD = 4.61) across 500, 1000 and 2000 Hz, with no significant difference in acuity between the two ears (Left Ear mean = 6.77 dB HL, SD = 5.18; Right Ear mean = 7.5 dB HL, SD = 4.67; t<1, ns). The older adult participants in this study had a mean better-ear PTA of 13.92 dB HL (SD = 4.96) and had no significant difference in acuity between the two ears (Left Ear mean = 15.60 dB HL, SD = 4.74; Right Ear mean = 17.50 dB HL, SD = 7.76; t[13] = 1.16, p>.26). Although the older adults had poorer hearing acuity than the young adults, all fell within a range defined as clinically normal hearing for speech (PTA < 25 dB HL; Mueller & Hall, 1997).

Working memory was also measured for both young and older adults using the Reading Span task (Daneman & Carpenter, 1980). As would be
expected (Salthouse, 1994), the older adults had lower average span scores
than young adults (Young mean = 2.39, Older mean = 1.45, t[27] = 3.40, p<.01).

All participants were native English speakers with no self-reported history
of stroke, Parkinson’s disease or other neurological disorder that might affect
task performance.

Stimuli and Apparatus.
A total of 689 monosyllabic words spoken by a male native speaker of
American English were recorded using an Audio Technica AT2045 microphone
and SoundEdit software (Macromedia, Inc., San Francisco, CA). Recordings
were digitized using a Presonus Audiobox USB audio interface at a sampling rate
of 44.1 kHz. A cosine-squared amplitude ramp of 5 ms rise time was applied to
the beginning and end of each word to avoid the popping noise from sudden
onsets or offsets.

The 689 words were divided into three groups that were balanced for
Kucera-Francis word frequency (Kucera & Francis, 1967) and assigned to one of
the three experimental conditions (Attended, Unattended, New). In each trial of
the DDLT, two streams of four words were presented simultaneously, one stream
to each ear, through Etymotic Research ER1 insert earphones. Durations for the
spoken words were between 422-758 ms (mean = 617 ms, SD = 72 ms). Stimuli
were concatenated into 2 streams of 4 words (i.e. left and right), and
synchronized so that each word pair began simultaneously at 0, 800, 1600, and
2400ms.

Electroencephalographic (EEG) signals were recorded from the scalp
using a high-density, 129-electrode array (Electrical Geodesics Inc., Eugene OR)
and high-impedance amplifiers. All channels were adjusted for scalp impedance
< 50kΩ. Sensor signals were sampled at 250 Hz with a 0–125 Hz analogue
bandpass filter, and stored for offline analysis. Bipolar periocular channels were
recorded from above and below each eye, and from a location near the outer
canthus of each eye.

Procedure.
Each participant completed one experimental session that included the
DDLT phase followed by a Delayed Recognition phase.

Insert Figure 1 about here

Figure 1. Schematic diagram illustrating a Directed Dichotic Listening (DDL)
trial’s event structure. Each trial began with a fixation cross that oriented the
participant to the region of the computer display at which the trial’s cues would be
centered. An arrow cued the subject to attend to either their right (») or left («) ear. Then, two simultaneous streams of four spoken words were presented, one stream to each ear. After a brief retention interval, a probe word was presented binaurally. Participants reported via keyboard key press whether or not the probe word had been a member of the to-be-attended word stream.

Directed Dichotic Listening Task. Figure 1 illustrates the DDLT procedure. Each participant completed 210 trials. The first 18 trials were practice trials that were not included in further analysis, resulting in 192 final trials. Participants were cued to attend to the left ear on half of these trials; on the remaining trials they were cued to attend to the right ear. The order of cues was randomly intermixed. For each trial, a visually displayed arrow presented for 0.5 s directed participants to attend to either their right or left ear. Then, 900-1100 ms (mean = 1000) later, simultaneous streams of four words were presented to the right and left ears through insert earphones. The duration of each four-word stream was 3200 ms. In order to generate the needed set of paired simultaneous, four-word streams, words were presented three times each across the entire DDLT. Half of these words were only presented in attended streams, and the other half were presented only in unattended streams. The order of attended and unattended word-pair presentations was randomly shuffled for each participant, with the constraint that all word pairs had to be presented in each condition before a given word could be repeated.

Following a brief, 1000 ms retention period, parentheses were displayed around the fixation cross and a single probe word was presented binaurally. For the 96 attend-left trials and 96 attend-right trials respectively: half of the probes (N=48) were target words from the attended stream (Target), 24 were from the unattended stream (Unattended Probe) and 24 were new words (New Probe) that were otherwise not presented during DDLT. The order of Target, Unattended Probe, and New Probe trials was randomized. If the probe was from the attended or an unattended stream, it was equally likely to have appeared at each of the four positions in that trial’s word set.

Participants rested two fingers from their right hand on a computer keyboard and responded via a key press to signal whether the probe had or had not been in the current trial’s attended stream. Participants had up to three seconds to make a response and were instructed to respond as quickly and as accurately as possible. During the 18 practice trials, feedback was provided after each trial in the form of a green fixation-cross for correct responses (attended hit,
unattended correct rejection, or new correct rejection) or red for incorrect responses (attended missed, unattended false alarm, or new false alarm). In each case the color was displayed for a 600 ms feedback interval. No feedback was given during the remaining trials.

Delayed Recognition. Delayed recognition instructions began after DDLT trials, and after the EEG cap had been removed. The time between the end of DDLT trials and the start of delayed recognition ranged from three to seven minutes (mean = 4 minutes). Participants completed the Delayed Recognition phase in less than ten minutes. During delayed recognition, words were presented binaurally through the insert earphones. The recognition test was comprised of 120 "old" words randomly drawn from the DDLT trials and 72 "new" words that were not previously heard in the experiment presented in random order. To distinguish the effects of attention during dichotic listening from the effects of attention to a single probe word, whether a word was used as a probe in the earlier DDLT was systematically varied. There were 24 old words from each of the following five categories: Words from the attended stream that had not served as probe words, words from the attended stream that had served as probe words, words from the unattended stream that had not served as probe words, words from the unattended stream that had served as probe words, and words that had been presented as new probes. Participants rested two fingers from their right hand on a computer keyboard. Following each word, the question, "Old or New?" was displayed on the computer screen. Participants were instructed to respond “Old” to any word that was presented in the DDLT phase of the experiment, regardless of whether it was attended or not. They had up to three seconds to make a response and were instructed to respond as quickly and as accurately as possible. The attended words and unattended words were counterbalanced for whether they had been presented to the left or right ear, and the position they had held within the stream.

Behavioral analysis.

Directed Dichotic Listening Task. For the DDLT, participants’ performance was assessed via hit rates to probes of attended items (Target Hit) and false alarm rates to probes that were new (FA New) and from the unattended stream (FA Unattended). Additionally we examined participants’ latencies to correct responses which included attended hits, new correct rejections (CR New), and unattended correct rejections (CR Unattended). To attenuate the undue influence of outliers on mean latencies, latencies outside 2.5 standard deviations of the mean for each condition and each participant were considered outliers and excluded from analysis. Further, to minimize the chance that overall response
latency differences (e.g., older adults being slower) would affect our interpretation of group \( \times \) condition interactions (Faust et al, 1999) we transformed latencies to z-scores relative to each participant’s overall mean latency.

**Delayed Recognition.** The delayed recognition task was meant to assess indirect effects of auditory attentional control processes on words during the attention task. We hypothesized that words that served as probes during the DDLT might be remembered better because of their monaural presentation and additional processing at the time of short-term recognition. In order to control for this potential concern we limited our analysis only to words that were not probed during the DDLT. Participants’ delayed recognition memory sensitivity was measured using \( d' \) (Green & Swets, 1966), which was calculated as the difference between z-transformed false alarm rates to new items and either z-transformed hit rates to attended items or unattended items. This yielded two \( d' \) estimates: \( d' \)attended—an index of participants’ ability to distinguish attended words from new items, and \( d' \)unattended—an index of the extent to which participants distinguished between unattended words and new items. Both \( d' \) estimates were calculated using the same false alarm rate, but different hit rates. To avoid undefined values of \( d' \) when hit or false alarm rates equaled 1.00 or 0.00, we applied the log-linear rule recommended by Hautus (1995). Latencies to correct responses included attended hits, unattended hits, and correct rejections of new words and were transformed to z-scores and screened for outliers as described in the Directed Dichotic Listening Task section.

**EEG analysis.**

EEG signals were preprocessed using the EEGLAB toolbox (Delorme & Makeig, 2004) for MATLAB (The Mathworks, Inc., Natick, MA). The recorded signals were re-referenced to the grand average voltage. A 0.5 Hz Butterworth high-pass filter and a 60 Hz Parks-McClellan notch filter were applied. Trial epochs were segmented from – 2.2 to +8.2 seconds surrounding the onset of the dichotic stimuli. An initial visual inspection was performed to remove epochs containing gross artifact and to identify bad channels. Bad channels that extended more than 10% of trials were excluded from Independent Components Analysis (ICA). The ICA was performed with EEGLAB’s extended infomax ICA decomposition. Components clearly containing blinks, muscle artifacts, eye movements, electrical heartbeat or singular artifacts were manually identified and subtracted from the data. Bad channels were then interpolated using EEGLAB’s spherical spline interpolation. A final visual inspection was performed to reject any remaining epochs with artifacts.
Finally, incorrect trials were also excluded from analysis. The number of artifact-free, correct trials remaining totaled 2,121 trials for older adults (N=15) and 2,025 trials for young adults (N=14) in the Payne et al (2016) study. All subsequent analyses and plotting were performed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) for MATLAB.

**Insert Figure 2 about here**

Figure 2. Young and older adult’s alpha power during dichotic listening. Top row: Alpha power across 800 ms bins beginning pre-stimulus onset and continuing through the dichotic presentation. Bottom row: Topographical plots illustrating the average alpha power across the 3200 ms dichotic listening interval. Electrodes marked with an X showed a significant difference between young and older adults at p < .05. The time-frequency representations for young and older adults are averaged across the cluster of parietal electrodes marked by x’s in the young adult topographical plot. Stimulus onset at Time = 0 is marked by a solid white line. The dashed line at Time = 3200 marks the stimulus offset.

**Induced Alpha Power.** Induced (that is, non-phase locked) time-frequency power during DDLT was computed using the following method: Separately for every subject, electrode, and condition, the event-related potential (ERP) was calculated and this evoked (phase-locked) activity was subtracted from the EEG for each trial (Kalcher & Pfurtscheller, 1995). Following this single trial subtraction of the ERP, time-frequency representations were computed using Morlet wavelets with a width of 4 cycles per wavelet at center frequencies between 1 and 40 Hz, in 1 Hz steps. Topographic maps of alpha power (8 – 14 Hz) were generated for 800 ms epochs beginning pre-stimulus onset and extending through the duration of the dichotic presentation (Figure 2, Top Row). These maps illustrate the consistency of alpha power throughout the dichotic presentation in both young and older adults. Given the minimal spatiotemporal changes in alpha activity seen during dichotic listening, alpha power was averaged within each of the pre-stimulus (-800 to 0 ms) and post-stimulus (0 to 3200 ms) epochs for statistical comparison.

A data-driven, nonparametric clustering approach (Maris & Oostenveld, 2007) was used to compare alpha power between young and older adults. Separately for the pre-stimulus and post-stimulus epochs, an independent samples t-value for the mean amplitude of alpha power was calculated at each electrode. All comparisons for which the t-value exceeded the p < .05 significance level were clustered based on spatial adjacency. Then for each
cluster, the $t$-values of its member electrodes were summed, giving a cluster score that reflected both the extent of the cluster in space and the magnitude of the difference between the conditions. A reference distribution of test statistics was generated by randomly permuting the data across young and older adults, computing such scores for each resulting cluster, and taking the largest such cluster score on each of 1000 permutations. Where cluster-wise $p$-values are reported, they came from comparisons between the empirically obtained cluster score and the permuted reference distribution. A difference between young and older adults was deemed significant if less than 5% of the proportion of randomized values exceeded the observed test statistic.

Grand average time-frequency representations were created by averaging across participants and across the electrodes identified as part of the cluster.

**ERPs.** Event-related potentials evoked by the presentation of probe words were used to investigate possible differences in the evoked response between old and new words. The artifact-free, correct trials were low-pass filtered at 30 Hz and baseline corrected using the 200 ms interval preceding the onset of the probe word. ERPs time-locked to the onset of the probe word were generated for each participant by averaging across trials for each of the three probe types (attended, unattended, and new).

To determine an overall old/new effect, the ERP for attended words (Old) was first compared to the combination of new and unattended probe words (both considered New). The ERP to attended words was then compared to new and unattended probes words separately in order to investigate any differences in topography or magnitude between the two comparisons. The clustering approach (Maris & Oostenveld, 2007) described earlier, in the section on alpha power (above), was used for the three comparisons (attended vs. combination of unattended and new, attended vs. new, and attended vs. unattended) with the old/new epoch defined as 300 to 500 ms post probe onset (Rugg & Curran, 2007). Grand average ERPs were created by averaging across participants and across the electrodes that were significantly different between the Old (attended) and New (combined unattended and new) conditions.

**Results**

**Behavior.**

Unless otherwise indicated, we only report effects that were significant at the $p < .05$ significance level and were not involved in a higher-order of interaction. If an interaction provided to be significant, we used post-hoc $F$-tests that incorporated the Bonferroni correction for reduction of Type I error. Mauchly’s test of sphericity and Levene’s test for homogeneity of variance were also
calculated. If the assumption of sphericity was violated, the Greenhouse-Geisser correction for mean squared error (MSE) and degrees of freedom was applied. If homogeneity of variance was violated, degrees of freedom were adjusted.

**Insert Figure 3 about here**

Figure 3. Behavioral results from the older adults of the current study and young adult data reported by Payne et al. (2016). Panel A indicates hit and false alarm (FA) rates during the DDLT. Panel B indicates z-transformed latencies to correct responses during the DDLT. Panel C indicates sensitivity (d') during delayed recognition to attended and unattended words that were not probed during the DDLT. Error bars indicate 95% confidence intervals around the mean.

**Directed Dichotic Listening Task.** Participants’ performance is given in panel A of Figure 3, which depicts the hit rates to probes that had been presented in the attended channel, false alarms to new probes, and false alarms to probes that had been presented in the unattended channel. Older and young adults had equivalently high hit rates and low false alarm rates to new probes. However, older adults had elevated false alarm rates to probes from the unattended channel relative to young adults. A 3 (response: attended hit, new false alarm, unattended false alarm) × 2 (age: young, older) mixed-model repeated measures ANOVA revealed a significant response × age interaction, F(1.40, 37.87) = 5.01, partial η² = 0.16, p = .01. F-tests revealed that age group differences were not significant on attended hits, F<1, ns, or new false alarms, F(1, 27) = 3.48, partial η² = 0.11, p = .07, but were significant on unattended false alarms, F(1, 27) = 7.45, partial η² = 0.22, p = .01.

Z-transformed latencies to correct responses are given in panel B of Figure 4, which shows that young and adults were significantly slower to reject an unattended probe than they were to reject a new probe or to endorse an attended probe. A 3 (response: attended hit, new CR, unattended false alarm) × 2 (age: young, older) mixed-model repeated measures ANOVA did not reveal a significant response × age interaction, F(2, 54) = 1.75, η² = 0.06, p = .18., or main effect of age, F(1, 27) = 3.67, η² = 0.12, p = .07, but did reveal a significant effect of response, F(2, 54) = 44.01, η² = 0.62, p <.001. F-tests revealed that correct rejections to unattended probes were slower than attended hits, F(2, 54) = 8.68, p <.001, and correction rejections to new probes, F(2, 54) = 7.34, p <.001, which did not significantly differ from one another, F(2, 54) = 1.25, p=.65.

**Delayed Recognition.** Delayed recognition to words that were not probed during the DDLT were assessed by examining d’ for words that were attended
and unattended, relative to new items. The results are shown in panel C of Figure 3. The young adults from Payne et al. (2016) showed greater memory discriminability for words that were attended rather than unattended. To examine this pattern in older and young adults, we performed a planned comparison on older adults’ \( d' \) and found no significant difference between hits that were attended and unattended \( t(14) = .14, p=.90 \). A similar comparison for young adults showed that the young did differ between attended and unattended items, \( t(13) = 3.10, p<.01 \). Young adults’ sensitivity for unattended items was likely so poor because of a low unattended hit rate (\( M = .41, SD = .12 \)). In fact, young adults’ unattended hit rates were significantly below chance, \( t(13) = 2.90, p=.01 \), indicative of a transfer of inhibition from the DDLT to delayed recognition effect. However, similar analyses on \( z \)-transformed response latencies did not reveal any significant effects.

**EEG and ERPs.**

*Alpha Power.* The cluster-based, permutation test revealed a significant difference in alpha power between young and older adults during the dichotic presentation. A posterior cluster of four electrodes (\( p < .05 \)) for which alpha power was greater for young adults than older adults was identified (Figure 1, Top Row). Time-frequency transforms averaged across the electrode cluster confirm the striking difference in posterior activity seen within the 8 – 14 Hz frequency band throughout the dichotic presentation (Figure 1, Bottom Row). There was no significant difference found between young and older adults for the pre-stimulus epoch.

**Insert Figure 4 about here**

*Figure 4.* Young adults’ ERP old/new effect. (A) - (C) Topographical display of ERP old/new effect from 300 to 500 ms post-probe onset. (A) Difference between Attended and both New and Unattended words. (B) Difference between Attended and New words. (C) Difference between Attended and Unattended words. Electrodes within the clusters marked with a circle showed a significant difference between the two test conditions at \( p < .01 \). Those marked with an \( x \) were significantly different at \( p < .05 \).

*ERPs.* For young adults, Figure 4 illustrates the midline cluster of 32 electrodes with a greater positive deflection for Old (attended) words than for New (unattended and new) words during the 300 to 500 ms following word onset (\( p < .01 \)), thus replicating the old/new recognition effect. As can be seen in b and c panels of Figure 4, this effect exists for both new words (\( p < .01 \)) as well as unattended words (\( p < .05 \)). For older adults, there were no significant
differences found in any of the three Old versus New comparisons. Figure 5 shows the ERPs time-locked to the onset of each of the three probe types for young and older adults respectively.

**Insert Figure 5 about here**

*Figure 5. Young and older adult's ERPs. Grand averaged ERPs time locked to the probe word onset for attended (red), unattended (blue), and new words (green). The topographical inset depicts the cluster of electrodes included in these averages as determined by the young adult comparison between Attended and both New and Unattended words (as seen in Figure 6).*

**Discussion**

In the DDLT, for two classes of probes, older adults’ recognition performance was not reliably different from what Payne et al. (2016) had found with young adults. In particular, older and young adults differed neither in recognizing probes that had been in the attended channel, nor extralist probes, that is, probes that had not been presented in either channel. This pattern of no reliable difference between age groups differs from classic results on dichotic listening (e.g., Craik, 1965), which found age-related deficits utilizing recall methods for assessing short-term memory. Consistent with Craik (1965)’s result however is the present finding that when recognition probes were words from the unattended stream, older adults were more likely to false alarm. Older adults’ good performance on identifying attended probes, rejecting new probes, but not for rejecting unattended probes is consistent with the notion that older adults were being less selective in their attention and that they were potentially attending to both channels. This “listen to everything approach” may reflect a form of compensation for dealing with an age-related decrease in signal loss (e.g., Getzman, 2012), and could be related to age-related changes in listening effort (Peelle & Wingfield, 2016).

With regard to long-term memory as measured by delayed recognition, older adults had poor discriminability both for items that were attended and unattended during the DDLT. This pattern suggests not only a lack of differentiation among items encoded to long-term memory but a lack of encoding into long-term memory more generally (see also Craik & Rose, 2012). This finding is consistent with the notion that effort exerted while decoding perceptual stimuli leaves fewer processing resources available to encode that information into long-term memory (Wingfield, 2016). In comparison, the young adults did
differentiate between items that were attended and unattended during the DDLT. In fact, young adults even showed below chance hit rates to unattended items, exhibiting a transfer of inhibition effect indicating a suppressive role of attention during dichotic listening (c.f., Thomas & Hasher, 2011).

The lack of a suppression effect may well be related to the lack of alpha power modulation that was observed in the older adults in this study. Changes in alpha power have been presumed to reflect auditory suppression of an unattended channel in a DDLT-type paradigm (Choi, et al., 2013; Zion Golumbic et al., 2013; Chait, et al., 2010). It is possible that this lack of alpha-based suppression is what led to the higher rate of unattended false alarms observed in older adults, as well as older adults’ “listen to everything approach.” While it appears that older adults do not engage alpha to the extent that young adults do, leading to observed age-related differences (e.g., Vaden et al, 2012), it is unclear if this is due to an inability or a strategic shift by older adults to not engage in processing that changes alpha. Further research on this topic will be needed to separate these potential causes.

These findings suggest a progressive top-down bias toward the representation of attended stimuli and degradation of unattended stimuli across the hierarchy of auditory processing (Lakatos et al., 2013; Mesgarani & Chang, 2012), which, novel to the current study, also declines with age. Unattended speech remains represented in low-level auditory areas in and near the superior temporal gyrus, but it is no longer represented in higher-order processing regions such as the inferior frontal cortex, and anterior and inferior temporal cortex (Zion Golumbic et al., 2013). Early representation of information arriving in the unattended ear is evidenced in faster response times to a target word in the attended ear if that word follows that same word presented to the unattended ear (Dupoux et al., 2003), or if a target word follows a semantically related word presented to the unattended ear (Bentin et al., 1995). More directly, the notion of early representation is corroborated by our finding of slower reaction times for correct rejection of unattended words than for correct rejection of new words. Bentin et al. (1995) indirectly demonstrated the lack of representation of the unattended information during higher-order processing. These authors found that the N400, an ERP component related to semantic processing (Kutas & Hillyard, 1980), for attended words, but not unattended words, was sensitive to semantic priming. In the current study, older adults’ greater rate of false alarms to the unattended channel indicate greater representation of information in the unattended ear than young adults, which suggests that the progressive top-down
bias toward the representation of attended stimuli and degradation of unattended stimuli across the hierarchy of auditory processing declines with age.

The ERPs from the DDLT in this study appeared to be very similar to old/new recognition effect previously reported with young adults in short-term recognition (Danker et al., 2008; Crites et al., 2000) and delayed recognition (Rugg & Curran, 2007). Our analysis revealed an "old/new" effect between attended targets and unattended lures. This result indicates that words from the attended stream were recognized better. Importantly, this result was true only for young adults; older adults' ERPs did not distinguish between words that were attended, unattended, or new. The timing of young adults' old/new effect, from 300 to 500 ms post stimulus onset supports the model of late selective inhibition (Lakatos et al., 2013; Zion Golumbic et al., 2013). Given the evidence for late suppression, it is likely that some false alarms to unattended words arose from early misallocation of attention to the wrong ear (Mesgarani & Chang, 2012). It was recently reported that the direction of auditory spatial attention can be predicted by multivoxel pattern analysis of fMRI activity in the auditory cortex (Kong et al., 2014). Future work could benefit from using such a technique to identify the trials on which attention was misallocated, and whether this is a model that can be used to understand hearing operations in older adults.

In summary, we invited a group of older adults to the cocktail party and found that they were less likely than young adults to be selective in their attention to dichotically presented speech. Older adults' behavioral pattern of listening did not co-occur with a pattern of EEG alpha modulation that was characteristic of the young adult listeners reported by Payne et al. (2016), nor did their ERPs to short-term recognition probes distinguish among probes that were attended, unattended, or new. These results suggest that older adults do not engage the same cortical correlates exhibited by young adults, which may in turn explain their age-related deficit to selective auditory attention. Lastly, the lack of a transfer of inhibition effect exhibited by older adults in delayed recognition suggests that the additional cognitive load caused by selectively attending to dichotic stimuli may interfere with older adults' ability to leave the cocktail party with lasting impressions of what was said to them.
References


Under review, *Psychophysiology*