1-1-2020

Neural Mechanisms Underlying Hierarchical Speech-in-Noise Processing

Caitlin Elise Nelms Price

NEURAL MECHANISMS UNDERLYING HIERARCHICAL SPEECH-IN-NOISE PROCESSING

by

Caitlin Elise Nelms Price

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Communication Sciences and Disorders

The University of Memphis

August 2020

Dedication

This manuscript is dedicated to my family, friends, professors, and mentors who have supported me throughout my academic journey, challenged me to never stop asking questions, encouraged me regardless of the circumstances, and ultimately inspired me to pursue dreams I never imagined possible.

Acknowledgements

This work was supported by grants from the Canadian Institutes of Health Research (MOP 106619) and the Natural Sciences and Engineering Research Council of Canada (NSERC, 194536) awarded to Claude Alain, and the National Institutes of Health (NIH/NIDCD R01DC016267) awarded to Gavin M. Bidelman, and the Institute for Intelligent Systems (IIS) Student Dissertation Grant awarded to Caitlin Nelms Price.

Preface

Chapter 2 was published as a manuscript in *Neuroscience*. Price, C. N., Alain, C., & Bidelman, G. M. (2019). Auditory-frontal channeling in α and β bands is altered by age-related hearing loss and relates to speech perception in noise. *Neuroscience, 423,* 18-28. doi:10.1016/j.neuroscience.2019.10.044

Abstract

One of the most commonly reported complaints related to hearing is difficulty understanding speech-in-noise (SIN). Numerous individuals struggle to effectively communicate in adverse listening conditions, even those with normal hearing. These difficulties are exacerbated due to age and hearing-related deficits such as hearing loss and auditory processing disorders. Despite the high prevalence of SIN deficits in individuals across the lifespan, the neural mechanisms underlying successful speech comprehension in noise are not well understood. Communication in noise is an incredibly complex process that requires efficient processing throughout the entire auditory pathway as well as contributions from higher-order cognitive processes including working memory, inhibition, and attention.

In a series of studies using electrophysiologic (EEG) and behavioral measures, this dissertation evaluated the neural correlates of SIN perception across subcortical and cortical levels of the auditory system to identify how top-down and bottom-up influences aid SIN understanding. The first study examined the effects of hearing loss on SIN processing in older adults at the cortical level using frequency-specific neural oscillations (i.e., brain rhythms) and functional connectivity (i.e., directed neural transmission). We found that low-frequency α and β oscillations within and between prefrontal and auditory cortices reflect the ability to flexibly allocate neural resources and recruit top-down predictions to compensate for hearing-related declines and facilitate efficient SIN perception. The second study, in younger adults, investigated the role of attention in SIN processing and how it interacts with early sensory encoding. Hierarchical processing in brainstem and cortex was assessed by simultaneously recording frequency-following responses (FFRs) and event-related potentials (ERPs) at the source level. We found that attention modulates SIN processing at both *subcortical* and *cortical* levels and

strengthens bidirectional neural signaling within the central auditory pathway. A relative disengagement of corticofugal transmission was observed in noise but only for passive listening suggesting attention aids SIN perception by maintaining top-down reinforcement of acoustic feature encoding within the primary auditory pathways. Taken together, these results indicate that the neural networks engaged during SIN perception depend on a complex interplay between bottom-up and top-down factors including signal clarity, listeners' hearing status, and attentional deployment.

Table of Contents

List of Figures

Chapter 1

General Introduction

Speech-in-noise (SIN) understanding is one of the most prevalent concerns voiced by adults seeking audiological care. Poor SIN perception influences an individual's ability to actively engage and effectively communicate in many everyday situations. In adults, difficulties hearing and communicating in noise have been associated with greater social isolation (Weinstein & Ventry, 1982), depression (Gopinath et al., 2009), poorer quality of life (Dalton et al., 2003), and accelerated cognitive decline (Lin et al., 2011; Lin et al., 2013). Deficits in SIN understanding are observed not only in older adults with hearing loss but also in individuals with normal hearing and those with auditory processing disorders (Gordon-Salant & Fitzgibbons, 1993; Guest, Munro, Prendergast, Millman, & Plack, 2018; Krishnamurti, 2001). While numerous studies have evaluated SIN perception behaviorally, the underlying neural mechanisms contributing to successful SIN comprehension remain unclear. Gaining a better understanding of the etiology and other factors contributing to perceptual SIN deficits can assist in the selection of clinical interventions, increase their effectiveness, and ultimately improve treatment outcomes.

The auditory system is comprised of ascending and descending pathways that work together to separate a target speech stream from competing background noise. Efficient processing and transmission within and between each level of the pathway is vital to facilitate accurate speech perception. When the efficiency and synchrony of bottom-up, perceptual processing of auditory stimuli is disrupted, listening difficulties are exacerbated due to an impaired ability to extract necessary acoustic features for comprehension (Anderson, Parbery-Clark, Yi, & Kraus, 2011; Parbery-Clark, Marmel, Bair, & Kraus, 2011; Song, Skoe, Banai, & Kraus, 2010). At the same time, higher-order cognitive processes, such as attention, inhibition,

and working memory, influence how auditory information is processed and perceived by a listener by prioritizing relevant sensory inputs and suppressing competing signals (Petersen & Posner, 2012; Starr & Golob, 2007) and by compensating for degraded acoustic signals due to poor hearing acuity or signal-to-noise ratio (SNR) (Bidelman, Price, Shen, Arnott, & Alain, 2019; Price, Alain, & Bidelman, 2019). Thus, SIN perception reflects the complex interplay between ascending peripheral encoding of acoustic features and descending cognitive influences. Because the system is so complexly interconnected, it is difficult to disentangle the respective contributions of individual processes to speech perception.

Due to their fine temporal resolution and presence throughout the auditory pathway, electrophysiologic responses enable these comparisons and are ideal for investigating the interaction of cognitive effects and early sensory encoding. Electrophysiologic studies provide unique insight into the underlying mechanisms of auditory processing across all levels of the auditory neuroaxis. The frequency-following response (FFR) is a sustained response optimally elicited by periodic stimuli such as tones or vowels and reflects the phase-locking ability of its neural generators, primarily within the rostral brainstem (Bidelman, 2018; Skoe & Kraus, 2010a; Smith, Marsh, & Brown, 1975; Sohmer, Pratt, & Kinarti, 1977; Tichko & Skoe, 2017). However, the FFR can include cortical activity if the evoking stimulus contains frequencies below the cortical phase-locking limit (~100 Hz) (Bidelman, 2018; Coffey, Herholz, Chepesiuk, Baillet, & Zatorre, 2016; Kuwada et al., 2002). The FFR mimics the spectrotemporal properties of the acoustic stimulus providing a measure of the fidelity and efficiency of early neural encoding. Cortical event-related potentials (ERPs), occurring approximately 50-250 ms following stimulus onset, consist of three primary components (P1, N1, P2) and are generated from primary and secondary auditory cortices and auditory association areas (Crowley & Colrain, 2004; Liegeois-

Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Naatanen & Picton, 1987). These obligatory ERPs reflect features of the acoustic stimulus as well as internal perceptual processes (N1 and P2 only; i.e., selective attention, auditory object formation, stimulus classification and categorization) (Bidelman, Moreno, & Alain, 2013; Crowley & Colrain, 2004; Luck, 2005). Neural oscillations, or rhythmic fluctuations in brain activity, provide insight into functional neural networks in which rhythms within different frequency bands are associated with unique processes underlying complex perceptual and cognitive functions such as those involved for SIN comprehension. Low frequency oscillations (i.e., θ, α, β) are associated with distributed cognitive processing across brain regions involving attention, inhibition, template matching, and tracking slower speech dynamics (Adrian & Matthews, 1934; Bidelman, 2015, 2017; Giraud & Poeppel, 2012; Klimesch, 2012; Pfurtscheller, 2001; Shahin, Picton, & Miller, 2009; Yellamsetty & Bidelman, 2018) while high frequency oscillations (i.e., γ) relate to localized sensory processing and extraction of acoustic features (Fontolan, Morillon, Liegeois-Chauvel, & Giraud, 2014; Giraud & Poeppel, 2012; von Stein & Sarnthein, 2000; Yellamsetty & Bidelman, 2018). Functional connectivity provides a measure of causal, directional information flow between 2 regions of interest (ROI) and enables insight into bottom-up and top-down signal transmission within the auditory system (Bidelman, Davis, & Pridgen, 2018; Bidelman, Price, et al., 2019; Lobier, Siebenhuhner, Palva, & Palva, 2014; Price et al., 2019). Using combinations of these measures within a study provide a more comprehensive understanding of underlying neural mechanisms contributing to SIN processing.

In this vein, our previous study used simultaneous electroencephalography (EEG) recordings of brainstem frequency-following responses (FFRs) and cortical event-related potentials (ERPs) to assess neural SIN processing within discrete levels of the auditory system

and, importantly, how the signal is transmitted between subcortical and cortical levels of the auditory pathway. We showed that age-related hearing loss alters the neural connectivity (signal transmission) between auditory brainstem and cortex while encoding within discrete levels and behavioral performance remain largely unaffected (Bidelman, Price, et al., 2019). While this study addressed the impact of peripheral hearing deficits on early SIN processing, questions regarding how and to what extent top-down cognitive influences engage in challenging listening conditions remained.

In a series of studies, this dissertation aimed to further evaluate the underlying neural mechanisms for SIN perception throughout the auditory pathway and prefrontal linguistic areas [(brainstem (BS)→primary auditory cortex (PAC)→inferior frontal gyrus (IFG)] and characterize the role of higher-order cognitive processes, particularly attention and other compensatory mechanisms, in SIN understanding. To investigate the effects of peripheral deficits on higher neural encoding (PAC→IFG) and identify compensatory top-down mechanisms (IFG→PAC), the first study evaluated patterns in neural oscillations and functional connectivity in older adults with and without hearing loss. Comparing neural responses between groups (i.e., normal hearing vs. hearing loss) and SNR (i.e. clean vs. noise) assessed how auditory and prefrontal linguistic processing centers contribute to speech understanding when acoustic inputs and feedforward signaling are degraded.

The second study focused more specifically on top-down, attentional modulation of SIN processing to determine whether attention influences subcortical speech encoding or alters connectivity between levels of the auditory pathway. We measured source-resolved FFRs and ERPs in normal hearing, young adults during active SIN perceptual tasks to evaluate speech activity generated at brainstem vs. cortical levels. Functional connectivity measures were used to

assess the strength and direction of neural signaling between these responses and identify "bottom-up" vs. "top-down" (corticofugal) communication within the auditory brainstem-cortical pathway. Comparisons between (i) active and passive SIN tasks and (ii) clean and noise responses evaluated attentional modulation of this circuit and whether added cognitive demands of noise altered hierarchical neural processing.

Chapter 2

Auditory-Frontal Channeling in α and β Bands is Altered by Age-Related Hearing Loss and Relates to Speech Perception in Noise

Introduction

Difficulty understanding speech-in-noise (SIN) is highly prevalent among the aging population including individuals both with and without hearing loss. Older adults exhibit greater listening effort (Anderson Gosselin & Gagné, 2011) and more significant performance deficits in adverse listening conditions than younger adults (Helfer & Wilber, 1990; Wong, Ettlinger, Sheppard, Gunasekera, & Dhar, 2010). Age-related hearing loss further exacerbates SIN difficulties (Helfer & Wilber, 1990). Previous studies characterizing the underlying mechanisms contributing to SIN difficulties reveal both peripheral and central brain mechanisms play a role in accurate and efficient SIN processing (Frisina & Frisina, 1997; Humes, 1996; Wong et al., 2010).

Indeed, neuroimaging studies reveal that structural and functional neural changes associated with aging (Bidelman, Mahmud, et al., 2019; Bidelman, Price, et al., 2019; Du, Buchsbaum, Grady, & Alain, 2016; Grady, 2012; Park & McDonough, 2013) contribute to older adults' SIN difficulties. Electrophysiological (EEG) studies often show exaggerated amplitudes and increased latencies of auditory cortical responses with aging, which has been taken as evidence for reduced inhibition (Alain & Woods, 1999; Bidelman, Villafuerte, Moreno, & Alain, 2014; Caspary, Ling, Turner, & Hughes, 2008; Chao & Knight, 1997) and decreased temporal fidelity in the aging auditory system (Tremblay, Piskosz, & Souza, 2003). The presence of hearing loss can amplify these changes due to the typical aging process (Lin et al., 2014; Pichora-Fuller & Levitt, 2012; Wayne & Johnsrude, 2015), resulting in even greater increases in

response amplitude and latency when those with hearing loss are compared to their normal hearing peers (Alain, Roye, & Salloum, 2014; Campbell & Sharma, 2013; Cardin, 2016).

To date, EEG studies have primarily relied on event-related potentials (ERPs) to infer the neural processes contributing to SIN perception. However, evaluating changes in gross activation within isolated brain regions can lead to misleading or ambiguous conclusions regarding the neurobiology of aging (Morcom & Henson, 2018; Wong et al., 2010). For instance, increases in evoked response amplitude commonly observed in older adults may be due either to the recruitment of additional neural resources (Bidelman et al., 2014; Wong et al., 2010), disinhibition (Bidelman et al., 2014; Caspary et al., 2008), or inefficient neural coding (Fabiani, Low, Wee, Sable, & Gratton, 2006). Evaluating ERPs alone prevents full understanding of the underlying mechanisms of aging, particularly how different brain regions might coordinate to orchestrate successful SIN perception. Alternative EEG analyses may better delineate the underlying neural mechanisms for speech processing that are not always apparent with traditional ERP approaches (Bidelman, 2015, 2017; Yellamsetty & Bidelman, 2018).

In this vein, neural oscillations have provided novel insight into functional neural networks underlying complex perceptual and cognitive functions. Therefore, evaluating oscillatory components of neural responses may provide a more sensitive measure and more thorough understanding of the neural correlates of speech processing. Different brain "rhythms" are thought to play unique roles in the hierarchy of speech processing. High frequency γ oscillations are thought to contribute to localized processing within sensory cortices (Fontolan et al., 2014; Giraud & Poeppel, 2012; von Stein & Sarnthein, 2000) and the extraction of acoustic features (Yellamsetty & Bidelman, 2018) while lower frequency α and β oscillations have been involved in global, distributed cognitive processing across brain regions (Fontolan et al., 2014;

von Stein & Sarnthein, 2000) including attention (Klimesch, 2012), inhibition of irrelevant cues (Adrian & Matthews, 1934; Klimesch, 2012; Pfurtscheller, 2001), working memory (Shahin et al., 2009; Zarahn, Rakitin, Abela, Flynn, & Stern, 2007), and template matching (Bidelman, 2015, 2017; Shahin et al., 2009; Yellamsetty & Bidelman, 2018). Evaluating how neural oscillations within different frequency bands of the EEG contribute to speech processing could provide further insight into the underlying processes supporting SIN perception in older adults.

In our ongoing studies on aging and the brain, we recently documented subtle neurophysiological changes in older adults with normal hearing (NH) and mild hearing loss (HL) that may reflect deficits in speech representations (Bidelman, Price, et al., 2019). Using source-resolved brainstem and cortical ERPs, we found somewhat spared region-specific responses to speech, at least in listeners with mild hearing impairment. More significant differences were identified in functional connectivity between the auditory brainstem and cortex, suggesting neural transmission within the early auditory pathway is critical for robust SIN processing in older adults. Additional full-brain, functional connectivity analysis revealed more widespread and less efficient connectivity patterns in HL compared to NH listeners suggesting more diffuse processing strategies are employed in those with hearing loss (Bidelman, Mahmud, et al., 2019). However, neither of these studies addressed the role of neural oscillations and how functionally distinct frequency channels of the EEG relate to senescent changes in SIN perception. Moreover, how the aging lemniscal hearing system (e.g., auditory cortex) interfaces with high-order brain regions that support linguistic decisions (e.g., prefrontal areas) is not well understood.

The current study aimed to examine contributions of neural oscillations and their role in neural signaling between auditory cortical and linguistic brain areas during SIN processing. In this reanalysis of our existing dataset (Bidelman, Price, et al., 2019), we measured frequency-

specific neural oscillations and functional connectivity via EEG in older adults with and without hearing loss during rapid SIN perception tasks. Based on previous studies on aging, the effects of hearing loss on SIN processing, and putative roles of neural oscillations, we hypothesized that differences in α activity would emerge in more difficult listening conditions and that HL listeners would demonstrate enhanced connectivity between auditory and prefrontal cortex to compensate for poorer signal transmission apparent in earlier stages of the speech hierarchy (e.g., diminished brainstem-cortical connectivity; Bidelman, Price, et al., 2019). Our findings reveal that (1) modulations in α phase coherence between clean and noise-degraded speech predicts accuracy in SIN tasks; (2) changes in functional brain connectivity precede measurable behavioral deficits in SIN processing; (3) "top-down" β connectivity increases in strength with increasing severity of hearing loss suggesting that the transfer of information between auditory-linguistic brain regions may be more sensitive to hearing-related changes than localized activity within regions.

Experimental Procedures

Analyses of the ERPs and behavioral responses associated with this dataset are reported in (Bidelman, Price, et al., 2019). New time-frequency analyses (applied here) were used to evaluate the correspondence between rhythmic brain oscillations and SIN perception in older adults.

Participants

Thirty-two older adults ranging in age from 52 to 75 years were divided into groups based on their average hearing thresholds (Fig. 1A). Listeners with average thresholds better than 25 dB HL comprised the normal hearing (NH; *n*=13) group while average thresholds worse than 25 dB HL classified participants with hearing loss (HL; *n*=19). The level of 25 dB HL reflects

the upper limit of the normal hearing range as specified by the clinical determination of hearing loss (Gelfand, 2009). The groups were otherwise matched for age (NH: 66.2±6.1 years, HL: 70.4±4.9 years; $t_{22.2}$=-2.05, p = 0.052) and gender (NH: 5/8 male/female; HL: 11/8; Fisher's exact test, p=0.47) (for complete demographic details, see Bidelman, Price, et al., 2019).

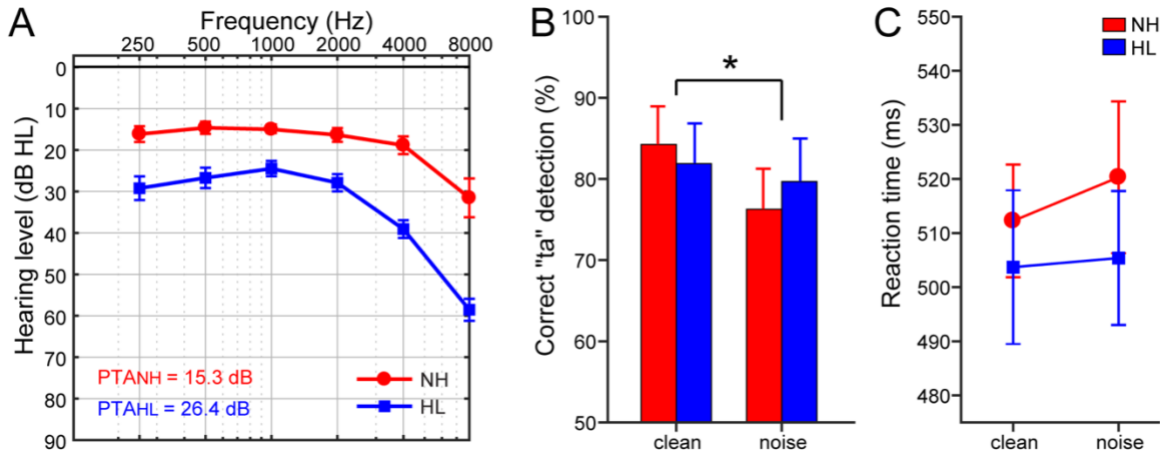

Figure 1: Audiometric and behavioral results. (A) Audiograms for listeners with normal hearing (NH) and hearing loss (HL). Hearing was ~10 dB better in NH vs. HL listeners. (B) Behavioral accuracy for detecting infrequent /ta/ tokens in clean and noise-degraded conditions. Noise-related declines in behavioral performance were prominent but no group differences were observed. (C) Reaction times (RTs) for speech detection were similar between groups and speech SNRs. errorbars = ± s.e.m., *p< 0.05

Stimuli and Task

Electrophysiologic responses were recorded while participants performed an active SIN perception task in which they were directed to identify an infrequent speech token (i.e., /ta/) via button press. The stimuli included three naturally produced English consonant-vowel phonemes (/ba/, /pa/, and /ta/) spoken by a female talker. The stimuli were presented binaurally in clean (i.e., no background noise) and noise-degraded conditions [10 dB signal-to-noise ratio (SNR) using 8-talker babble noise, cf. Killion et al., 2004]. In each condition, the frequent tokens /ba/ and /pa/ were each presented 3000 times while the infrequent, target token /ta/ was presented 210 times. Between presentations, the interstimulus interval was randomly jittered between 55-155

ms. Both speech detection accuracy (%) and reaction times (RTs) were logged. See Bidelman, Price, et al., 2019.

EEG Time-Frequency Analysis on Source Waveforms

The EEG recording protocol and data pre-processing is described in our original report (Bidelman, Price, et al., 2019). Briefly, cortical event-related potentials were recorded from 32 channels across the scalp. Ocular artifacts (saccades and blinks) were first corrected in the continuous EEG using a principal component analysis (PCA) (Picton et al., 2000). Cleaned EEGs were then epoched (-10-200 ms) and baseline corrected to the pre-stimulus period for each trial and stimulus condition per participant. The pre-stimulus interval was limited due to the pace of the perceptual task.[1]

To first reduce the dimensionality of the data and enable functional connectivity analysis between brain regions of interest (ROIs), full band (1-100 Hz), single trial scalp potentials were transformed to source space using the AEP virtual source montage in BESA (Scherg, Ille, Bornfleth, & Berg, 2002). This process applies a spatial filter to all electrodes and optimizes the relative weights of their contribution to the recorded scalp response to estimate the activity within each source while reducing overlapping activity from other brain regions (for details, see Scherg and Ebersole, 1994; Scherg et al., 2002). This allowed us to reduce each listener's EEG (32-channels) to 15 source channels with regional dipoles in bilateral primary auditory cortex (PAC), left/right frontal cortex near inferior frontal gyrus (IFG) (i.e., Broca's area), and left/right parietal cortex as well as sources along the mid-line (depicted in Fig. 1A of Zendel & Alain, 2014). From this model, we extracted the estimated neural current within single ROIs of the

[1] For the current study, the paradigm was designed to record frequency-following responses (FFRs) from the brainstem and cortical ERPs simultaneously. Because FFRs require many more trials (approximately 2000 per token) than traditional ERP measures, shortening the interstimulus interval was necessary to reduce the overall time required for data collection.

brain most relevant to our hypotheses including tangential and radial components of each auditory source as these orientations capture the majority of auditory cortical ERPs (Picton et al., 1999) and radial components of each frontal source (BESA default). Furthermore, the selection of these sources enabled us to assess the effects of hearing loss on band-specific connectivity between auditory (PAC) and linguistic (IFG) brain areas and potential recruitment of additional neural resources (e.g., compensatory processing) due to age-related hearing loss. Time-frequency analysis (TFA) was then performed on the single-trial epochs at the source level to improve spatial accuracy and reduce smearing due to volume conduction (Hoechstetter et al., 2004) using BESA® Research v7 (BESA, GmbH).

TFA assessed the frequency-specific contributions of time-locked neural oscillations to older adults' SIN processing. Prior to TFA analysis, additional artifact correction was performed using a threshold of ±120 μV. Initial analysis revealed negligible induced activity likely due to the restricted baseline (10 ms); therefore, subsequent analyses focused on phase-locked oscillatory activity. The time-frequency transformation was achieved using a sliding window complex demodulation (for detailed description, see Papp & Ktonas, 1977) using 10 ms/5 Hz resolution step sizes. These settings permitted analysis of frequencies ≥10 Hz (i.e., α band and higher) across the entire epoch window. The resulting time-frequency displays, akin to neural spectrograms (see Fig. 2), were then produced by computing inter-trial phase-locking (ITPL) at each time-frequency point across single trials (Hoechstetter et al., 2004).

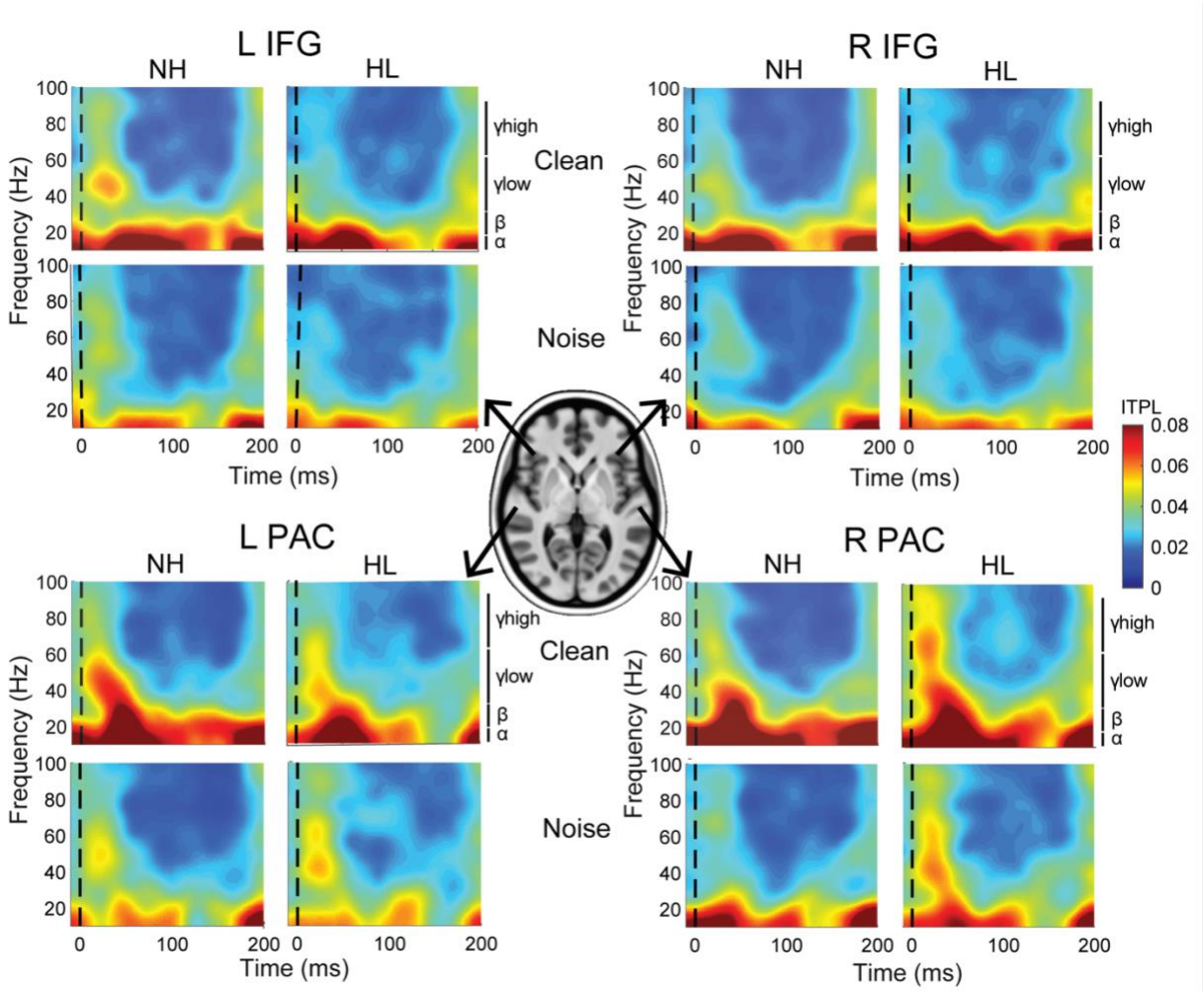

Figure 2: ITPL spectrograms for radial auditory and frontal sources by SNR and group. Time-frequency analysis demonstrates phase synchrony (ITPL) within each neural source across frequency and time. Trending differences are observed when comparing synchronicity across sources (PAC > IFG), SNR (clean > noise), and group. Hotter colors denote stronger neural phase synchrony across trials.

ITPL measures the phase consistency (i.e., trial-to-trial synchrony) of neural activity within each frequency band across time (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). Values range from 0 to 1 indicating the degree of phase synchronicity across trials (i.e., 0 – random noise; 1 – perfect trial-to-trial repeatability). For each ROI, we extracted band-specific time courses from the ITPL spectrograms in the α (10-12 Hz), β (15-29 Hz), low γ (30-59 Hz), and high γ (60-90 Hz) frequency bands (e.g., Bidelman, 2017) (see Fig. 3). We then measured the peak maximum ITPL and associated latency from each band waveform using MATLAB. Latency windows were guided

by visual inspection of the grand averaged traces [α: 25-100 ms; β: 25-75 ms; low/high γ:15-50 ms]. Peak responses were then used to assess the effects of SNR and hearing loss on neural oscillations involved in older adults' SIN perception.

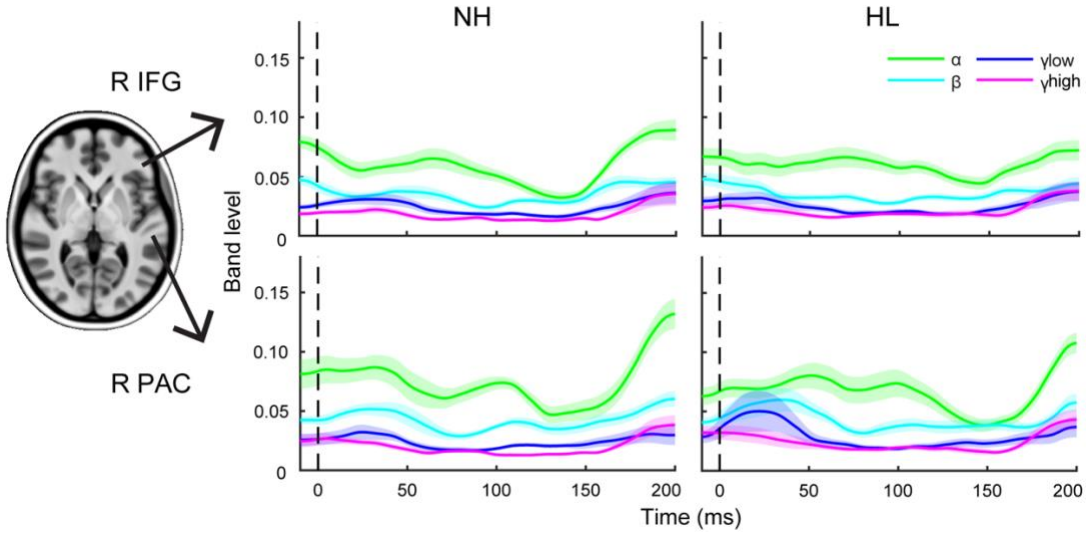

Figure 3: Band-specific time-course waveforms within auditory and frontal sources. The time-course waveforms illustrate the degree of phase synchronicity across trials over time for each frequency band. Bands were extracted from ITPL maps (see Fig. 2). Waveforms reflect grand averaged traces for each group from the right frontal and auditory sources in the noise condition. Clean and left hemisphere responses not shown.

Functional Connectivity

We measured band-specific functional connectivity between PAC and IFG sources (for each hemisphere) using phase transfer entropy (PTE) (e.g., Bidelman, Davis, et al., 2018; Bidelman, Price, et al., 2019; Lobier et al., 2014). PTE is a directional measure of signal dependence. Additionally, PTE can be implemented in a frequency-specific manner to assess connectivity in individual EEG bands (Lobier et al., 2014). We computed PTE between source signals in the PAC and IFG ROIs in both directions (i.e., X→Y and Y→X) to quantify differences in the strength of afferent/bottom-up (PAC→IFG) vs. efferent/top-down (IFG→PAC) connectivity within the auditory-linguistic pathway as a function of speech SNR and group.

Statistical Analysis

Mixed model ANOVAs were performed to assess all dependent variables of interest (GLIMMIX, SAS® 9.4, SAS Institute; Cary, NC) with participants serving as a random effect. Degrees of freedom were estimated using PROC GLIMMIX's containment option[2]. Unless otherwise specified, Bonferroni adjustments controlled for Type I error inflation. The significance level for all statistical analyses was set at $\alpha = 0.05$. Independent samples *t*-tests (un-pooled variance, two-tailed) were used to compare demographic variables between groups. Correlational analyses (Pearson's-*r*) and robust regression (bisquare weighting - achieved using the 'fitlm' function in MATLAB) were used to evaluate relationships between neural and behavioral measures. Specifically, to evaluate the relationship between neural oscillations and behavioral SIN perception, we used robust regression. We first collapsed clean and noise responses by computing their difference (clean - noise). We then conducted correlational analyses between neural responses (i.e., phase coherence peak amplitude/latency within each frequency channel and source) and the behavioral measures [i.e., pure-tone average (PTA), RT, %]. This allowed us to assess the degree to which modulations in neural oscillations between clean and noise-degraded speech were related to changes in hearing thresholds and behavioral performance. False discovery rate (FDR) was used to correct for multiple correlations (Benjamini & Hochberg, 1995). One RT data point was identified as an outlier and was excluded from correlation analyses. All analyses and results were collapsed across the frequent tokens (i.e., /ba/ and /pa/) to further reduce the dimensionality of the data. Responses to infrequent /ta/

[2]To satisfy model convergence and ensure estimable variance, it was necessary to remove the random term for the efferent (IFG-PAC) β connectivity variable. In this case, PROC GLIMMIX estimated degrees of freedom using the between-within approximation procedure (Schluchter & Elashoff, 1990), which divides the residual degrees of freedom into between-subject and within-subject portions.

tokens were not included in analysis due to the limited number of trials and to avoid mismatch negativities.

Results

Behavioral Data

Behavioral responses, reproduced from (Bidelman, Price, et al., 2019), are shown in Figure 1. Analyses of these results are reported in depth elsewhere (Bidelman, Price, et al., 2019). In short, we found no differences between groups in accuracy (Fig. 1B) nor RT speed (Fig. 1C) for target speech detection. However, noise had an expected detrimental effect on perceptual accuracy for both groups (Fig. 1B).

Electrophysiological Data

Time-frequency (ITPL) spectrograms for the PAC and IFG sources are shown for each SNR and group in Figure 2. Band time courses are shown in Figure 3. Diagnostics for amplitude analyses revealed a positive skew; thus, a cube-root transform was used. ANOVAs conducted on the transformed amplitude measures revealed significant effects of SNR for all frequency bands (all $p < 0.03$) but no main effect of group or SNR×group interaction. For latency, no significant group or SNR effects were observed for any frequency band. The lack of group effects might be anticipated given the relatively mild differences in hearing loss between groups and our previous study which did not observe differences in ERP responses (Bidelman, Price, et al., 2019). This further motivates the examination of band-specific oscillations in these data.

To determine whether neural activity within different frequency bands was associated with perceptual SIN measures, we used robust regression to assess brain-behavior relations. We found a significant negative correlation between α oscillations in the right frontal (IFG) source and speech detection accuracy [$r_{30} = 0.41$, $p_{FDR} = 0.007$; Fig. 4; left IFG: $r_{30} = 0.02$, $p_{FDR} =$

3.008; not shown]. This suggests that listeners who were more resistant to the detrimental effects of noise (i.e., performed equally as well or better in noise) also demonstrated less coherence of α activity while performing the SIN perception task. Little to no change in α phase coherence was observed in the listeners who performed more poorly in noise.

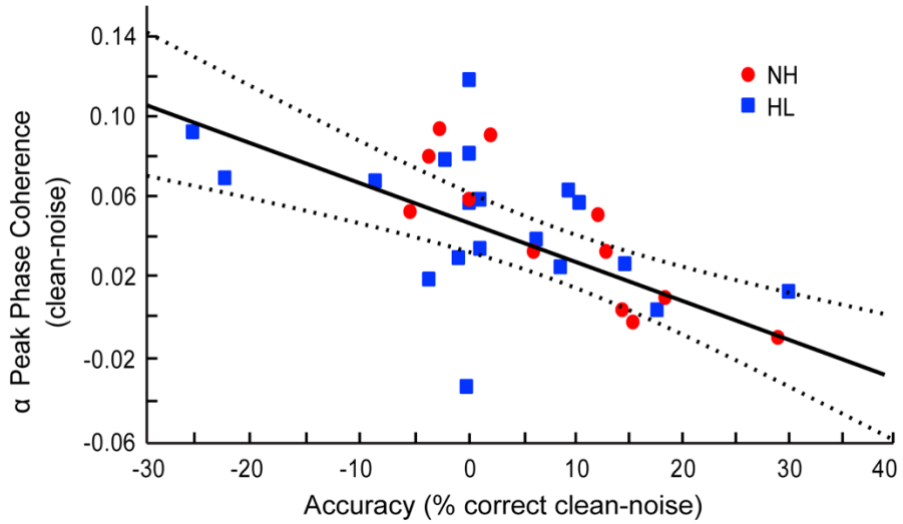

Figure 4: Phase coherence within right IFG α-band predicts accuracy of SIN performance. Difference scores between clean and noise conditions are plotted for α phase coherence within right IFG and /ta/ detection accuracy for each participant. Greater modulations in α band are observed in listeners whose behavioral performance was more resistant to the detrimental effects of noise. Dotted lines represent the 95% confidence interval.

Auditory-Frontal Functional Connectivity

We next asked whether differences in neural *transmission* (i.e., feedforward or feedback connectivity) between PAC and IFG is altered in individuals with mild hearing loss. Because initial inspection of the data revealed minimal connectivity within the low and high γ frequency bands (data not shown), subsequent analyses focused on connectivity within the α- and β-band channels. Mixed model ANOVAs (subjects=random effect) were performed for both afferent (PAC→IFG) and efferent (IFG→PAC) connectivity to evaluate the effects of group, hemisphere, and condition as well as potential interactions. These analyses were conducted separately by

frequency band and each dipole orientation (i.e., tangential and radial). These analyses revealed that HL listeners demonstrated stronger *efferent* β connectivity between IFG and the radial PAC component than NH listeners (*mean* ± *SE*; HL: 0.28 ± 0.02, NH: 0.20 ± 0.02; $F_{1,30} = 7.14$, $p = 0.0121$; Fig. 5A) regardless of SNR. In contrast, afferent (PAC→IFG) signaling did not differ between groups (HL: 0.27±0.03, NH: 0.29±0.03; $F_{1,30} = 0.08$, $p = 0.78$; Fig. 5A). None of the other comparisons or interaction effects remained statistically significant following correction for multiple comparisons.

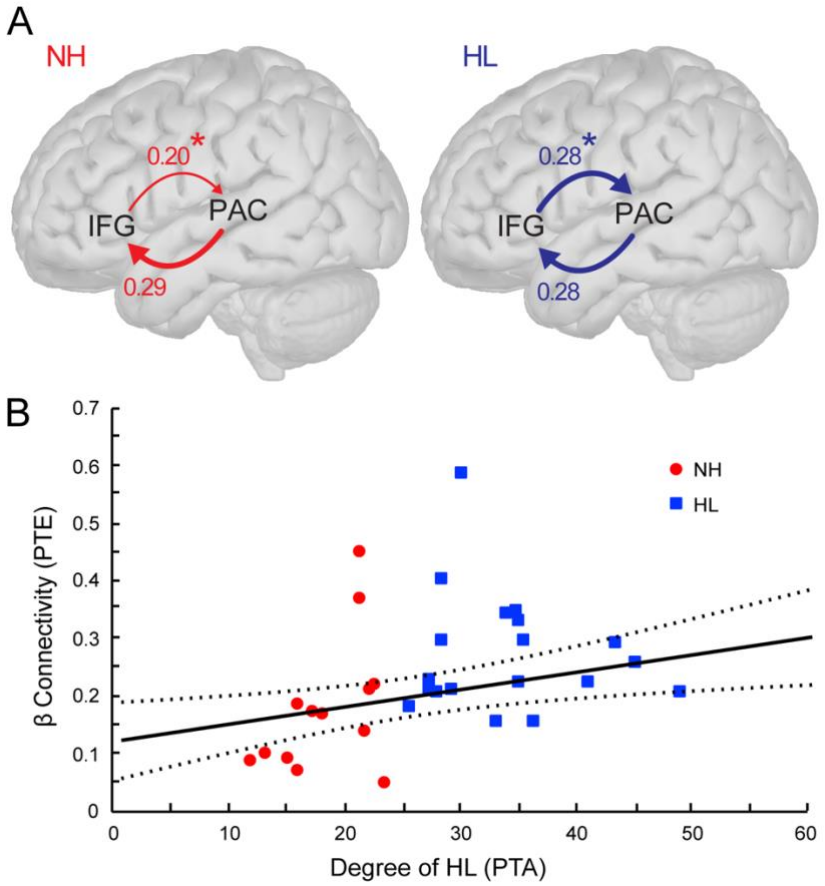

Figure 5: Efferent functional connectivity (IFG→PAC) within β band varies with degree of hearing loss. (A) Phase transfer entropy reflecting the directed (casual) afferent (PAC→IFG) and efferent (IFG→PAC) neural signaling between auditory and prefrontal cortex for the noise degraded speech condition. Efferent connectivity is stronger in listeners with hearing loss compared to those with normal hearing; afferent connectivity is similar between groups. (B) Efferent IFG→PAC connectivity increases in strength in listeners with poorer hearing (i.e., higher PTAs). Dotted lines represent the 95% confidence interval. *$p < 0.05$

Lastly, to relate neural connectivity effects to behavior, we conducted correlations between β connectivity (the only band showing group differences) and behavioral measures (i.e., PTA, RT, %). As in the previous correlation analyses, we used difference measures between clean and noise responses in these calculations. We found that efferent β connectivity (in noise) between IFG and radial PAC was positively correlated with PTA ($r_{30} = 0.24$, $p = 0.0044$; Fig. 5B) such that stronger efferent connectivity was associated with greater degrees of hearing loss. No other significant correlations were noted including those involving the clean speech responses.

Discussion

By measuring neural oscillations in older adults during SIN perception, our data reveal three primary findings: (1) modulations in α phase coherence between clean and noise-degraded speech predicts accuracy in SIN perception; (2) changes in functional brain connectivity precede measurable behavioral deficits in SIN processing; (3) "top-down" β connectivity from IFG to PAC increases in strength with increasing severity of hearing loss.

α Phase Coherence Predicts Accuracy in SIN Perception

We found that older adults who were more resistant to the detrimental effects of noise behaviorally demonstrated reduced α phase synchronicity in noise, particularly within right IFG. Previous studies suggest that α enhancement functions to inhibit task-irrelevant inputs (Adrian & Matthews, 1934; Pfurtscheller, 2001) while reductions in α facilitate task-relevant processing (Klimesch, 2012). Similar to our findings, greater event-related desynchronization (ERD) in α has been related to improved performance in semantic (Doppelmayr, Klimesch, Hodlmoser, Sauseng, & Gruber, 2005; Klimesch, Doppelmayr, Pachinger, & Ripper, 1997) and working memory (Bashivan, Bidelman, & Yeasin, 2014) tasks. Klimesch, Sauseng, and Hanslmayr

(2007) further suggest that ERD reflects "active information processing" related to excitatory rather than inhibitory processes in the brain and that this desynchronization is likely related to more generalized attentional demands required for the completion of a task. Furthermore, they posited that ERD may play a role in the release of inhibition related to spreading activation. Likewise, Proskovec, Heinrichs-Graham, and Wilson (2019) found greater decreases in α activation in high- compared to low-load conditions during a verbal working memory task. Attentional models further suggest that increasing task complexity, or cognitive load, leads to improved performance due to greater attentional focus (Kahneman, 1973) and requires higher levels of processing and attentional selection (Lavie, 1995; Lavie, Hirst, de Fockert, & Viding, 2004). Therefore, it is possible that our SIN detection task was less challenging, requiring less attentional and other neural resources, for listeners who showed greater α coherence during clean speech (those to left side of graph; Fig. 4) compared to listeners who were "low α modulators." However, when greater cognitive resources are required during more difficult noise conditions, less synchrony within α band, reflecting a release from inhibition, may enable the brain to deploy attention more flexibly to aid syllable detection accuracy. While these outcomes are limited to phase-locked neural oscillations, changes in induced activity may reveal different underlying mechanisms of SIN processing (Bidelman, 2015; Petersen, Wostmann, Obleser, Stenfelt, & Lunner, 2015). Future studies could incorporate analyses of induced activity to provide a more thorough representation of event-related neural processes contributing to SIN tasks.

Paralleling our data, previous studies have also shown that age-related changes in α activity are localized to frontal and sensorimotor regions (Dushanova & Christov, 2016; Nobukawa, Kikuchi, & Takahashi, 2019). Activity within frontal cortical areas may serve as a

compensatory mechanism for deficits in speech processing in older adults, particularly in more adverse listening conditions (Binder, Liebenthal, Possing, Medler, & Ward, 2004; Du et al., 2016; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006). Specifically, IFG and superior temporal gyrus (STG) activation within the right hemisphere is particularly salient for difficult sound contrasts (cf. our noise condition) (Doeller et al., 2003). Furthermore, increased α activity within the right hemisphere, particularly IFG, has been associated with inhibitory processes (Garavan, Ross, & Stein, 1999), which provides additional support to our conclusion that desynchronization in α activity within right IFG functions as a release from inhibition in older adults' speech-in-noise processing. Additional evidence of right lateralized compensation in SIN processing has been observed in passive listening tasks in normal hearing, young adults which reveals altered neural response laterality from being leftward dominant to include greater right hemispheric contribution within both PAC and IFG with decreasing SNR (Bidelman & Howell, 2016). It is possible that the compensatory rightward shifts in response laterality observed by Bidelman & Howell were exaggerated in our sample due to increased age and the presence of hearing loss in some of our listeners (Bidelman, Price, et al., 2019).

Changes in Functional Connectivity Precede Measurable Behavioral Deficits in SIN Processing

While no behavioral differences were observed between groups in our SIN detection task (Fig. 1), HL listeners demonstrated enhanced efferent β connectivity when processing SIN (Fig. 5). Overall, these data suggest that central compensation through the recruitment of additional, non-canonical auditory brain areas help overcome peripheral deficits to assist older adults' speech perception in noise (e.g., central gain compensation; Chambers et al., 2016). Numerous studies have described age-related changes in both brain structure and function, including inter-

regional connectivity (Betzel et al., 2014; Bidelman, Mahmud, et al., 2019; Bidelman, Price, et al., 2019; Grady, 2012; Sullivan & Pfefferbaum, 2006) and compensatory processing (Du et al., 2016; Grady, 2012; Park & McDonough, 2013). Hearing loss is thought to exacerbate the effects observed in typical aging (Lin et al., 2014; Pichora-Fuller & Levitt, 2012; Wayne & Johnsrude, 2015). In fact, studies have shown that increased recruitment of frontal cortical regions is associated with morphological changes particularly in auditory regions, and this additional recruitment has been further linked to behavioral performance (Tyler et al., 2010; Wong et al., 2009). Specifically, our results show enhanced connectivity directed from IFG to PAC in noise suggesting increased neural signaling from linguistic to auditory sensory areas in HL listeners. Previous studies have shown IFG contributes to "top-down" processing of speech in more adverse listening conditions (Binder et al., 2004; Zekveld et al., 2006), and it has also been associated with other cognitive functions like working memory (Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003; Specht, Shah, & Jancke, 2000) and template matching of the input stimulus to an internal representation within auditory memory (Zekveld et al., 2006). These processes are critical for SIN perception and may account for the hearing-related changes we find in IFG→PAC signaling.

Our data suggest that functional connectivity may perhaps provide a more sensitive measure of changes induced by hearing loss than behavioral measures. The recruitment of frontal sources in aging adults and those with hearing loss may reflect broader alterations within functional networks and compensatory cortical reorganization (Campbell & Sharma, 2013; Cardin, 2016). The stronger efferent (IFG→PAC) connectivity we observed in HL listeners suggests that even mild degrees of hearing loss can alter functional communication between cortical regions subserving speech-language functions. Such changes in functional connectivity

may provide a means by which older adults with HL could compensate for impoverished representations in auditory cortices.

Top-Down β Connectivity Increases in Strength with Poorer Hearing

Older adults with hearing loss demonstrated stronger β connectivity in noise between frontal and auditory regions (Fig. 5A) which also scaled with greater degrees of hearing impairment (Fig. 5B). Because no differences were observed in RTs between clean and noise-degraded conditions for either group (Fig. 1C), strengthened β connectivity is unlikely attributed to changes in general listening effort. Rather, we interpret these data to reflect alternative cognitive processing strategies that are utilized with impoverished auditory inputs. This notion aligns with previous studies that have related oscillatory β-band activity to cognitive processes associated with task demands including working memory (Shahin et al., 2009; Zarahn et al., 2007), encoding and integrating sensory information (Brovelli et al., 2004; von Stein & Sarnthein, 2000; Wang et al., 2017), speech template matching (Bidelman, 2015, 2017; Shahin et al., 2009; Yellamsetty & Bidelman, 2018), as well as predictive coding (Cope et al., 2017; Sedley et al., 2016).

Predictive coding utilizes prior knowledge and experience to form top-down predictions that assist in perception (Friston, 2005; Rao & Ballard, 1999) particularly when sensory inputs are degraded (Cope et al., 2017). Specifically, β-band activity has been related to the updating and precision of predictions (Cope et al., 2017; Sedley et al., 2016) and top-down signaling during speech processing (Fontolan et al., 2014; Wang, 2010). Under a predictive coding framework, fronto-temporal interactions would tend to increase in cases of degraded sensory information (Cope et al., 2017). Aging is associated with increased activation of frontal and motor cortex that helps compensate for impaired SIN perception in older adults (Bilodeau-

Mercure, Lortie, Sato, Guitton, & Tremblay, 2015; Du et al., 2016). Older listeners also show greater specificity of phoneme representations in frontal articulatory regions compared to auditory brain areas (Du et al., 2016). The increased IFG-PAC connectivity we find could reflect predictive coding that would naturally need to be stronger in listeners who have impoverished sensory encoding (i.e., HL listeners). The fact that this predictive inferencing is restricted to the β-band suggests the "top-down" mechanism observed here is not general attention or listening effort per se (which would be expected in α-band) but template matching and/or interactions between higher (IFG) and lower (PAC) order speech representations.

Additional studies have shown that older adults with hearing loss demonstrate reduced cognitive reserve which impacts higher order language processing as well as other complex processing and tasks that rely heavily on cognitive resources (i.e., SIN) (Cardin, 2016; Mishra, Lunner, Stenfelt, Ronnberg, & Rudner, 2013; Mishra, Stenfelt, Lunner, Ronnberg, & Rudner, 2014; Rudner, Foo, Ronnberg, & Lunner, 2009). Furthermore, lower cognitive reserve has been related to higher functional connectivity (Lopez et al., 2014), which in turn is related to perceptual SIN abilities (Bidelman, Mahmud, et al., 2019; Bidelman, Price, et al., 2019; Giordano et al., 2017). Because observed increases in functional connectivity only occurred in noise, it is likely that the increased task demands of more difficult listening conditions further reduced the spare capacity of available cognitive resources in HL listeners leading to the recruitment of frontal regions to overcome depleted sensory resources. The increased efferent connectivity within the β-band may reflect the online recruitment of these additional resources (e.g., IFG) to bolster the matching of sound to speech templates and facilitate SIN comprehension. Alternatively, increased prefrontal activity/connectivity in older, hearing impaired adults might instead reflect nonspecific neural responses (i.e., arousal, attention) rather

than compensation via recruitment of specific complementary neural regions to benefit task performance (Morcom & Henson, 2018).

In summary, our findings suggest that α desynchronization in challenging listening conditions reflects a release from inhibition contributing to better SIN performance. This finding supports the notion that a decrease in α functions to assist in active cognitive processing of task-relevant inputs. Additionally, even mild degrees of hearing loss in older adults result in neurophysiological changes in connectivity between cortical auditory and linguistic areas during SIN processing despite negligible behavioral deficits. That this top-down connectivity is restricted to β band suggests hearing loss increases the need to make high-order inferences on noisier sensory representations. Collectively, our findings suggest that functional connectivity is more sensitive to hearing-related changes than region-specific activation and that neural signaling is altered prior to observable behavioral changes. These results emphasize the importance of compensatory, top-down signal transmission in impaired systems to aid SIN perception.

Chapter 3

Attention Reinforces Hierarchical Speech-in-Noise Processing by Mitigating Noise Effects

Introduction

Speech-in-noise (SIN) perception is a complex process that requires the listener to isolate a talker's voice from competing background noise. Extraction of speech cues necessary for SIN perception relies on accurate and efficient bottom-up perceptual and top-down cognitive processes. Bottom-up encoding serves to provide rich neural representations of acoustic features while higher-order cognitive processes enable the prioritization and integration of inputs for comprehension. A breakdown at any point within this integrated system could lead to SIN deficits (Humes, 1996; Moore, 2015). Because SIN perception is highly variable even among normal hearing listeners (Gordon-Salant & Fitzgibbons, 1999; Guest et al., 2018), the issue cannot lie solely within the peripheral auditory system (i.e., the audibility of the signal). Rather, SIN difficulties likely arise from inefficient encoding within the central auditory system, general cognitive deficits, or an interaction between the two.

In complex listening environments, attention aids the selection and prioritization of behaviorally relevant inputs over irrelevant background noise. Top-down attentional enhancement of early speech processing may fine-tune auditory neural coding across the pathway (Atiani, Elhilali, David, Fritz, & Shamma, 2009; Gao & Suga, 2000; Suga & Ma, 2003) and improve SIN perception. Indeed, attention has been suggested to influence all stages of auditory processing from periphery to cortex (Galbraith, Olfman, & Huffman, 2003; Hernandez-Peon, 1966; Lukas, 1980, 1981; Picton & Hillyard, 1974). However, whether attention actively modulates early speech encoding (i.e., prior to cortex) and how it contributes to speech understanding remains unclear.

While attentional modulation of late cortical responses is well documented (Hillyard & Kutas, 1983; Picton & Hillyard, 1974), attentional effects upon early processing at the level of the auditory brainstem remain less clearly defined and highly debated. Most studies evaluating attentional effects on the click-evoked auditory brainstem response (ABR) reveal no difference in neural activity with attentional state of the listener (Picton & Hillyard, 1974; Picton, Hillyard, Galambos, & Schiff, 1971; Woods & Hillyard, 1978). However, more contradictory results emerge in frequency-following response (FFR) studies which employ more complex acoustic stimuli (i.e., speech). Some FFR studies suggest that attention enhances the robustness and efficiency of early brainstem encoding (Galbraith, Arbagey, Branski, Comerci, & Rector, 1995; Galbraith, Bhuta, Choate, Kitahara, & Mullen, 1998; Galbraith et al., 2003; Lehmann & Schonwiesner, 2014). Still, others demonstrate mixed (Holmes, Purcell, Carlyon, Gockel, & Johnsrude, 2018; Saiz-Alia, Forte, & Reichenbach, 2019) or even no attentional effects on the FFR (Galbraith & Kane, 1993; Varghese, Bharadwaj, & Shinn-Cunningham, 2015). The lack of attentional modulation found is further supported by the fact that brainstem responses can be reliably recorded even when a listener is asleep or sedated (Skoe & Kraus, 2010a) which leads to the conclusion and currently held assumption that subcortical processing is pre-attentive and automatic in nature. The mixed results obtained in previous studies likely result from differences in methodology (Varghese et al., 2015) and suggest that attentional influences on subcortical speech processing may be detected under certain conditions (e.g., rapid stimulus presentation rates, sufficiently challenging task demands) (Galbraith & Arroyo, 1993; Woldorff, Hansen, & Hillyard, 1987).

Neurobiologically, attention could influence early sensory encoding via top-down, corticofugal efferent projections from cortical layer V that relay information back to the inferior

colliculus (IC) of the midbrain (Schofield, 2010; Suga, 2008; Suga, Ma, Gao, Sakai, & Chowdhury, 2003). Animal studies reveal corticofugal stimulation sharpens auditory neural receptive fields for spectral, temporal, and spatial cues within brainstem IC neurons (Bajo & King, 2012; Suga, 2008; Suga et al., 2003). These acoustic cues are vital for SIN understanding and are precisely reflected within neural responses from IC (Bajo & King, 2012)—the primary generator of the FFR (Bidelman, 2018; Smith et al., 1975; Sohmer et al., 1977). Although the biological circuitry exists, human studies to date do not provide direct evidence of active corticofugal modulation of speech processing.

Corticofugal tuning of auditory inputs is thought to occur over time due to short-term training or life-long learning associated with particular auditory experiences (i.e., tonal language, musical training) (Bidelman, Gandour, & Krishnan, 2011; Carcagno & Plack, 2011; Krishnan, Xu, Gandour, & Cariani, 2005; Skoe, Krizman, Spitzer, & Kraus, 2013; Song, Skoe, Wong, & Kraus, 2008; Wong, Skoe, Russo, Dees, & Kraus, 2007). Yet, the extent of exposure required to establish a significant and measurable influence on auditory processing is debated. Some studies report that top-down influences are active and evolving during online speech processing (Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Skoe & Kraus, 2010b; Skoe et al., 2013) while others suggest top-down mechanisms only become engaged under challenging listening conditions (i.e., first-time exposure, early stages of learning or training, poorer signal-to-noise ratio) and function to reorganize local networks within a level of processing to maintain experience-dependent effects over time (Krishnan & Gandour, 2009; Krishnan, Gandour, & Bidelman, 2012; Song et al., 2010; Suga et al., 2003).

The goal of this study was to resolve ongoing debates in the literature regarding the role of attention in subcortical auditory processing and the engagement of the corticofugal system.

Identifying where within the auditory system attention influences neural encoding can provide further insight into how cognitive processes contribute to SIN performance and reveal important interactions between brainstem and cortical levels that support speech understanding. To this end, the current study aimed to delineate the role of attention in SIN processing across brainstem and cortical levels of the auditory system by (i) directly comparing the influence of attentional deployment on brainstem vs. cortical speech coding and (ii) evaluating corticofugal engagement during active and passive listening.

Materials & Methods

Participants

Twenty young adults ranging in age from 18 to 35 years (*M* = 24 years, *SD* = 3.4; 11 female) were recruited to participate in the study. An a priori power analysis (t-test estimates, 2-tailed, $\alpha = 0.05$, power = 0.95) revealed a sample size of 16-20 participants was sufficient to detect similar sized effects (*d* = 0.84, 1.0) as in our previous study using an analogous paradigm (Bidelman, Price, et al., 2019; GPower v3.1). All participants exhibited normal hearing thresholds (≤ 25 dB HL) for octave frequencies between 250 and 8000 Hz. Because prior musical and language experience influence FFRs/ERPs and SIN performance (Bidelman, Gandour, & Krishnan, 2010; Bidelman & Dexter, 2015; Bidelman et al., 2011; Yoo & Bidelman, 2019), participants had less than 3 years of formal musical training (*M* = 0.8 years, *SD* = 1.2) and were native English speakers. Participants were predominantly right-handed (*M* = 82.04%, *SD* = 21.04) (Oldfield, 1971) with no history of neuropsychiatric disorders. All participants provided written consent prior to their participation in accordance with dual protocols approved by the University of Memphis IRB (Protocol #2370 and PRO-FY2018-338).

Stimuli and Task

We recorded neurophonic FFRs and cortical ERPs simultaneously under active and passive listening conditions. Three synthesized vowel tokens (e.g., /a/, /i/, /u/) were presented during the recording of electrophysiologic responses since sustained periodic speech sounds, such as vowels, optimally evoke both the FFR and ERP (Picton, Woods, Baribeau-Braun, & Healey, 1976; Skoe & Kraus, 2010a). Each vowel had a duration of 100 ms and common voice fundamental frequency (F0=150 Hz). Using stimuli with F0s above cortical phase locking limits (~100 Hz) minimizes cortical contribution and ensures the brainstem is the primary generator of the recorded FFR (Bidelman, 2018; Coffey et al., 2016). The vowels were also matched in average root mean square amplitude. To vary the difficulty of the speech perception task, the vowels were presented in clean (i.e., no background noise) and noise-degraded conditions. For the noise condition, the stimuli were mixed with 8 talker noise babble (cf. Killion et al., 2004) at a signal-to-noise ratio (SNR) of 5 dB (speech at 75 dB_A SPL and noise at 70 dB_A SPL). In each block and condition, the frequent tokens /a/ and /i/ each were presented 4000 times while the infrequent token /u/ was presented 140 times. Tokens were presented in a random order back-to-back with a jittered interstimulus interval (95-155 ms, 5 ms steps, uniform distribution). Stimulus presentation was controlled by MATLAB (The Mathworks, Inc.; Natick, MA) routed to a TDT RP2 interface (Tucker-Davis Technologies; Alachua, FL) and delivered binaurally through shielded insert earphones (ER-3; Etymotic Research; Elk Grove Village, IL).

Attention was varied with the inclusion of active and passive listening blocks. During the active listening blocks, the participants were asked to rapidly detect the infrequent /u/ token via button press. This allowed us to measure reaction time (RT) and accuracy of speech detection during online EEG recording. A “hit” was defined as detection occurring within 5 tokens (~500

ms) of a target. For the passive listening blocks, participants watched a captioned movie to maintain arousal levels and were instructed to attend to the movie and ignore any sounds they heard. A control block was added (n = 8) to rule out potential visual confounds in the passive block. In these blocks, the transducers were unplugged while the participants watched the captioned movie as in the passive blocks. The presentation order of block (active, passive, control) and condition (clean vs. noise) were counterbalanced across participants to minimize order effects.

QuickSIN Test

The Quick Speech-in-Noise (QuickSIN) test was administered to each participant to provide a measure of speech reception thresholds in noise (Killion et al., 2004). The QuickSIN presents listeners with lists of 6 sentences spoken by a female talker embedded in four-talker babble noise. The sentences were presented at 70 dB SPL using pre-recorded SNRs decreasing in 5 dB steps from 25 dB (very easy) to 0 dB (very difficult). Recall of each sentence's 5 target words was used to calculate SNR-loss, or the SNR at which 50% of the words were repeated correctly (Killion et al., 2004). Higher scores are indicative of poorer SIN performance. Two lists were binaurally administered to each participant, and the average SNR-loss of these scores were calculated for each participant.

Electrophysiological Recording and Analysis

EEG Acquisition and Preprocessing. EEG recordings were obtained using 64-channels at standard 10-10 electrode locations across the scalp (Oostenveld & Praamstra, 2001). Additional electrodes were placed on the outer canthi of the eyes and superior and inferior orbit of the left eye to monitor eye movements and blink artifacts. Electrode impedances were kept below 5 kΩ. EEGs were digitized at 5 kHz (SynAmps RT amplifiers; Compumedics Neuroscan;

Charlotte, NC) using an online filter passband of DC—2000 Hz. This high sampling rate was needed to recover both the fast (FFR) and slower (ERP) frequency components of the compound speech-evoked potential.

Pre-processing was performed off-line in Curry 7 (Compumedics Neuroscan). Ocular artifacts (saccades and blinks) were first corrected in the continuous EEG using a principal component analysis (PCA) (Picton et al., 2000). Cleaned EEGs were then epoched (-10-200 ms), baseline corrected to the pre-stimulus period, and averaged in the time domain to obtain evoked responses containing both brainstem and cortical activity (Bidelman et al., 2013) for each stimulus condition per participant. Subsequent analyses were completed in BESA Research v7.0 (BESA, GmbH). Data were re-referenced using a common average reference. Neural responses to infrequent /u/ tokens were excluded from all analyses due to the limited number of trials and to avoid mismatch negativities.

Brainstem FFRs. To isolate the FFR from the compound response, the averaged responses were band-pass filtered (130-1500 Hz). This filtering limits the cortical contribution to the FFR and ensures the response originates solely from the midbrain (Bidelman, 2018; Coffey et al., 2016) while retaining the frequency content of the stimuli. The steady-state of the response (~10-100 ms) was analyzed using a Fast Fourier transform (FFT) which captured the spectral composition of the response. From each FFT, F0 amplitude was selected as the maximum value within a 10 Hz bin centered around 150 Hz (i.e., F0 of the stimuli). Analyzing the amplitude of F0 allowed us to evaluate whether the robustness of pitch encoding, which serves as an important cue for selecting and tracking a single voice within background noise (Assmann, 1996; Bidelman & Yellamsetty, 2017), was altered by attention or noise. FFRs were analyzed at both the

electrode (Fpz with mastoid reference) and source level to compare our findings to previous literature investigating attentional effects on the FFR.

Cortical ERPs. Averaged compound neural responses were filtered into a low frequency band (1-30 Hz) to isolate the slower cortical ERPs. ERP waves (i.e., P1, N1, P2) were quantified in amplitude and latency for each participant and condition using BESA's automated peak analysis. Latency windows were determined following visual inspection of grand average traces. P1 was identified as the maximum positive deflection occurring within 40-80 ms; N1 as the greatest negative deflection between 90-145 ms; and P2 as the maximum positive deflection within 145-175 ms (Hall, 1992). Our analysis focused primarily on N1 and P2 as these components are thought to be enhanced by attention (Naatanen, 1975; Picton & Hillyard, 1974) and are more reflective of SIN processing than earlier components (Bidelman, Price, et al., 2019; Billings, McMillan, Penman, & Gille, 2013).

Source Analysis. Scalp potentials were transformed to source space using BESA to reduce the dimensionality of the data and allow for connectivity analysis between our regions of interest [ROIs; auditory brainstem (BS) and primary auditory cortex (PAC)]. Full detail is described in Chapter 2. A source model (shown in Figure 6) comprised of putative neural generators of the FFR (Bidelman, 2018) was used to evaluate whether relative contributions from each source are modulated by attention or noise. The average goodness of fit (GoF) across conditions for our 3-dipole model (GoF = 90.6%) suggests that the model well explains the scalp recorded data [residual variance (RV) = 9.4 ± 1.6%]. The source model was applied to each individual's anatomy with fixed dipole locations and orientations. Rather than select a single dipole orientation from each source for subsequent analyses, the absolute magnitudes of each dipole orientation (e.g., x, y, z) were combined to provide a more accurate representation of the

activity generated within a given ROI (Coffey, Chepesiuk, Herholz, Baillet, & Zatorre, 2017). Because source analysis reduces the effects of volume conduction and estimates activity within specific ROIs, source analysis of SIN processing provided a more sensitive measure of the different underlying neural contributions to the speech FFR and how they change with noise and attention.

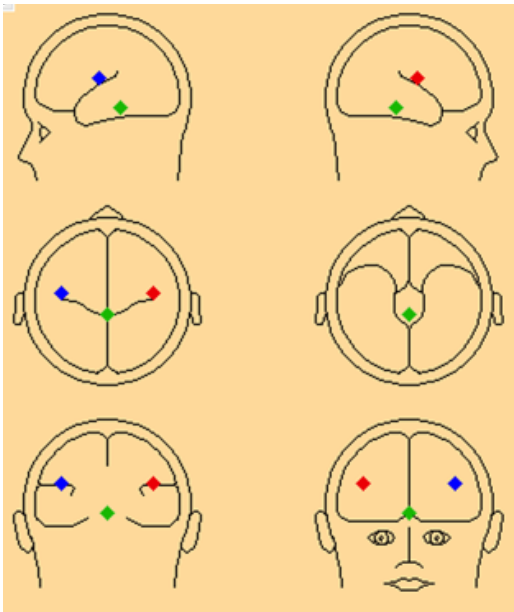

Figure 6: BESA source model. The dipole source model comprised of 3 sources; each seeded within an ROI (i.e., BS, right and left PAC).

MRI Scanning and EEG Co-Registration

A standard 3D T1-weighted anatomical volume was obtained on a Siemens 1.5T Symphony TIM scanner (tfl3d1 GR/IR sequence; TR=2000 ms, TE=3.26 ms, inversion time=900 ms, phase encoding steps=341, flip angle=8°, field of view=256x256 acquisition matrix, 1.0 mm thickness axial slices). Scanning was conducted at the Semmes Murphey Neurology Clinic in Memphis, TN. Electrode positions were mapped with a quad sensor Polhemus Fastrak digitizer (Polhemus, Colchester, VT). Scans were segmented in BESA MRI 2.0. Following inhomogeneity correction (Pham & Prince, 1999), images were automatically

partitioned into scalp, skull, CSF, and brain compartments (Chan & Vese, 2001), and the cortical surface was reconstructed to allow optional inflation of the brain volume (Fischl, Sereno, & Dale, 1999). MRI volumes were created in both ACPC and Talairach (Talairach & Tournoux, 1988) spaces using 3D spline interpolation. MRIs were not available for 4 participants. In these cases, we used a 4-shell spherical volume conductor head model (Berg & Scherg, 1994; Sarvas, 1987) projected onto the BESA adult MRI template brain (Richards, Sanchez, Phillips-Meek, & Xie, 2016).

Following segmentation, electrode locations were warped to the scalp surface (anchored to the nasion and preauricular fiducials) to co-register the sensor locations to each individual's anatomy. A 4-layer finite element head model (FEM) was then generated based on the MRI segmentation (Wolters, Anwander, Berti, & Hartmann, 2007) to construct each individuals' leadfield (forward volume conductor). The FEM leadfield describes the magnitude each source signal contributes at each electrode sensor (Scherg, 1990) and is less prone to spatial errors than other head models (e.g., concentric spherical conductor) (Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002). Collectively, this approach allowed us to source localize each individual's cortical and brainstem EEGs with high precision based on the anatomical constraints of their individual brain anatomy.

Functional Connectivity

Functional connectivity was measured between PAC and BS sources using phase transfer entropy (PTE) (e.g., Bidelman, Davis, et al., 2018; Bidelman, Price, et al., 2019; Lobier et al., 2014; Price et al., 2019). PTE is a measure of causal, directional information flow. We computed PTE between source signals in PAC and BS ROIs in both directions (i.e., X→Y and Y→X) to quantify differences in the strength of afferent/bottom-up (BS→PAC) vs. efferent/top-down

(PAC→BS) connectivity within the auditory pathway as a function of attention and speech SNR. This allowed us to evaluate how attending to auditory signals and listening in more adverse conditions influence bidirectional signaling within the central auditory pathway.

Statistical Analyses

We performed 2x2x2 (vowel x attention x SNR) mixed model (subjects=random factor) ANOVAs to assess all dependent variables of interest (GLIMMIX, SAS® 9.4, SAS Institute; Cary, NC) with Tukey-Kramer adjustments to control for Type I error inflation when necessary. Initial diagnostics revealed violation of normality and homogeneity of variance assumptions for all dependent variables except connectivity measures; therefore, statistical tests were conducted on log-transformed values. The significance level for all statistical tests was set to $\alpha = 0.05$. Paired samples *t*-tests (two-tailed) were used to compare behavioral performance between conditions. A generalized linear mixed effects model (GLME) was used to further evaluate relationships between neural measures and SIN performance. Neural measures from each ROI (FFR_{BS} – F0 amplitude, ERP_{PAC} – P2 magnitude), connectivity (aff – afferent, eff – efferent), and additional behavioral and conditional factors (QSIN – QuickSIN, PTA – pure-tone average of hearing thresholds, SNR - noise) were included as predictors of behavioral throughput, a measure of behavioral efficiency (i.e., thruPut ~ 1 + FFR + ERP + aff + eff + QSIN + PTA + SNR). Behavioral throughput reflects the time-accuracy tradeoff by dividing target detection accuracy by RT; therefore, slower response speeds result in poorer throughput and less overall perceptual efficiency (Bidelman et al., 2014; Salthouse & Hedden, 2002).

Results

Behavioral Data

Figure 7 shows behavioral accuracy and reaction time for detection of the infrequent /u/ tokens during the active listening blocks. T-tests conducted on behavioral measures revealed that noise reduced target detection accuracy [$t_{(19)} = 4.48$, $p < 0.001$] and slowed response speeds [$t_{(19)} = -4.78$, $p < 0.001$]. These findings suggest that poor SNRs detrimentally effect speech understanding.

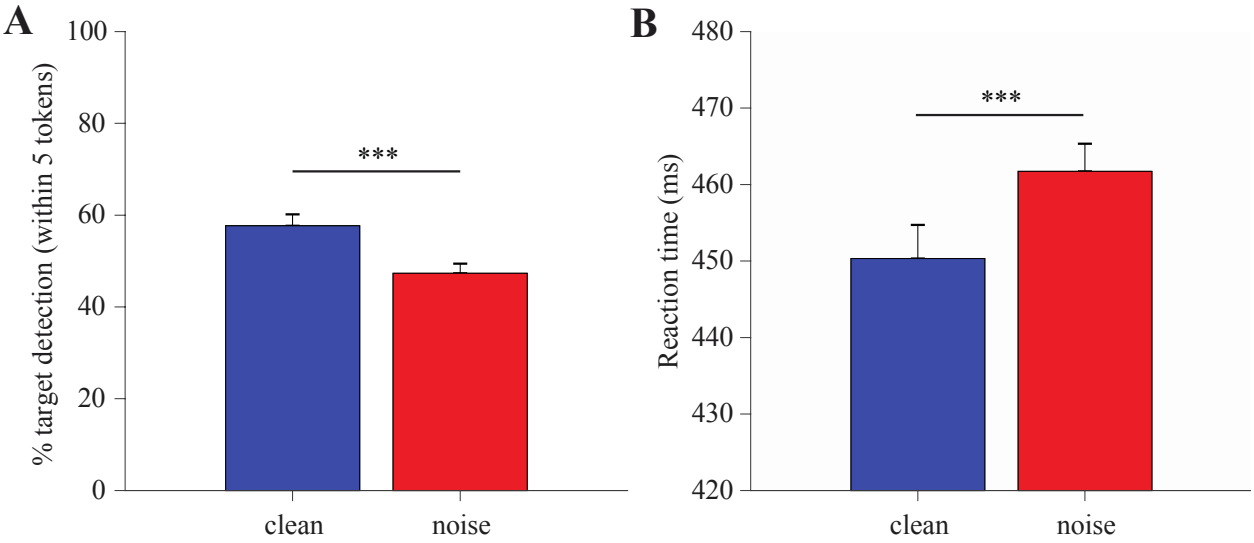

Figure 7: Target detection accuracy and reaction time. (A) Behavioral accuracy for detecting infrequent /u/ tokens in clean and noise-degraded conditions. Noise-related declines in accuracy were observed. (B) Reaction times (RTs) for speech detection increased for poor SNRs. errorbars = ± s.e.m., ***$p < 0.001$

Electrophysiological Data

Grand average source FFRs and ERPs are shown in Figure 8 for the pooled frequent tokens (/a/ and /i/) in active and passive listening (A, C) and clean and noise-degraded conditions (B, D).

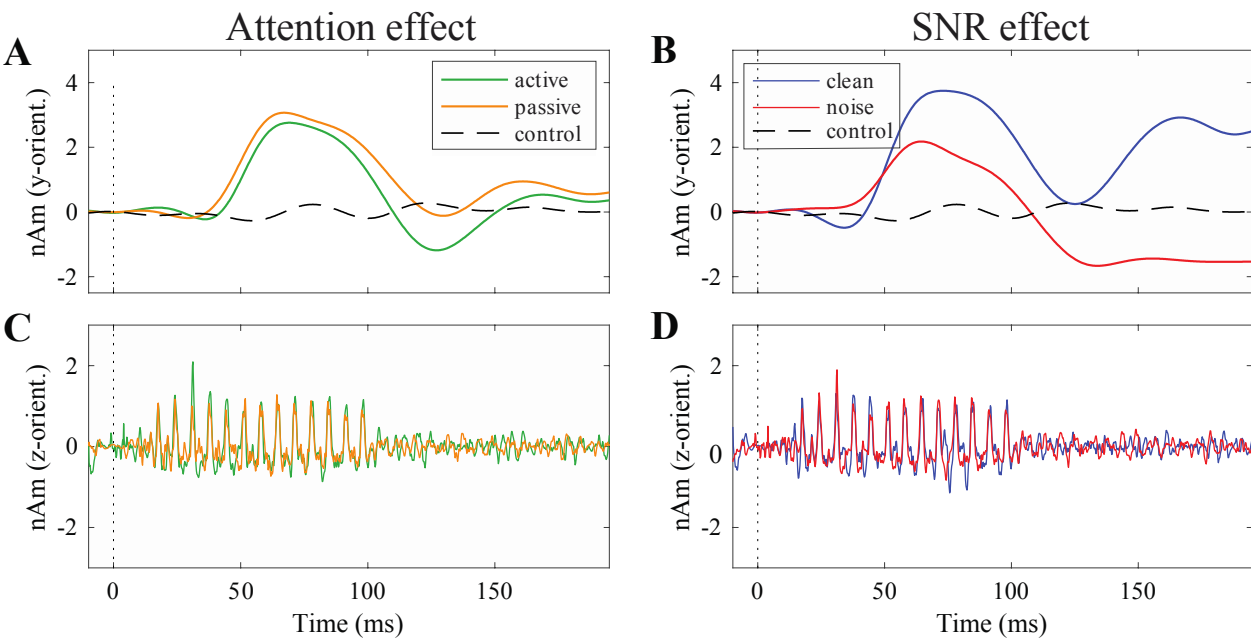

Figure 8: Source-level grand average neural responses. Average ERP responses are plotted in the top panels (A, B) with FFRs in the bottom panels (C, D). Averages reflect activity from a single dipole within each ROI (i.e., tangential component for ERPs and horizontal component for FFRs) and are pooled across frequent tokens. Attentional enhancements are observed in later ERP components, particularly within the range of N1 and P2 (A), as well as in the FFR (C). Large SNR effects are noted across essentially all ERP components (B) while little difference is seen at the brainstem level (D).

Brainstem FFRs. To identify whether attentional modulation of SIN processing occurs within the brainstem, F0 amplitudes of FFRs were assessed across SNR and attentional state. FFRs were analyzed at both the electrode and source levels to compare our results to previous literature. For scalp-recorded FFRs, noise weakened brainstem responses to speech [$F_{(1, 136)}$ = 8.15, p = 0.01], and similar to former studies (Picton & Hillyard, 1974; Picton et al., 1971; Woods & Hillyard, 1978), attention did not modulate FFRs [$F_{(1, 136)}$ = 1.68, p = 0.20]. Contrastively, at the source level, F0 amplitudes increased during active listening [$F_{(1, 136)}$ = 5.39, p = 0.02] but remained unaffected by noise [$F_{(1, 136)}$ = 2.22, p = 0.14]. Response spectra for channel and source FFRs are shown in Figure 9. F0 source amplitudes in the control condition

did not differ significantly from 0 [$t_{(83)} = 0.96$, $p = 0.34$] suggesting there were no visual confounds in the passive FFRs.

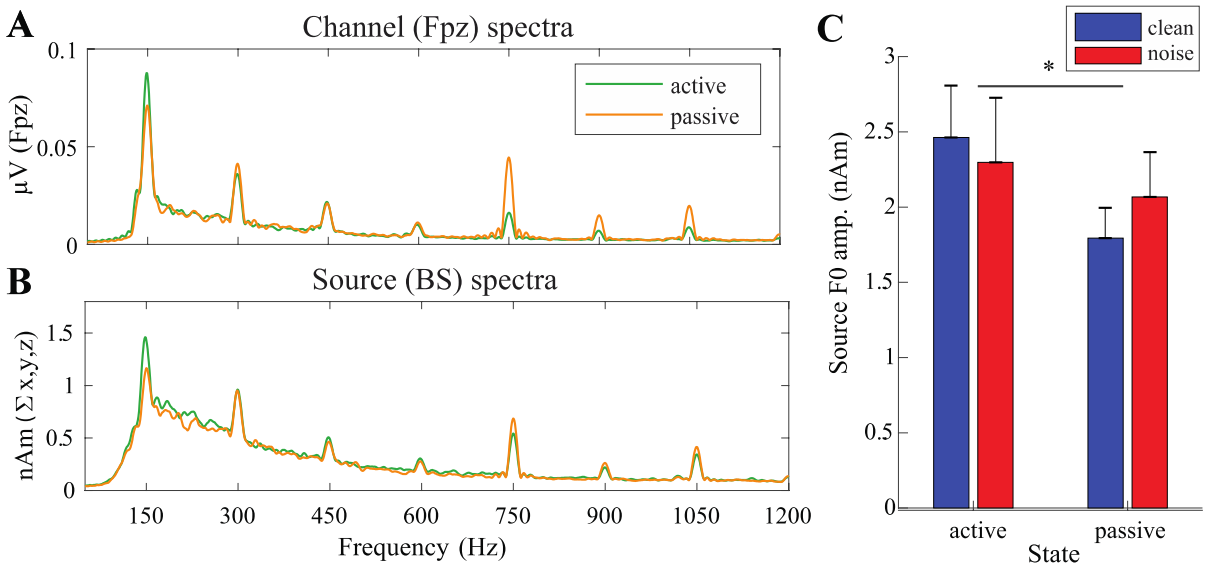

Figure 9: Response spectra for channel- and source-level FFRs. The plotted spectra are pooled across frequent tokens and SNR conditions. Attentional enhancement of F0 was only observed at the source level (B, C). Source FFRs showed no effect of SNR (C). errorbars = ± s.e.m, *p< 0.05

Cortical ERPs. Amplitudes and latencies of prominent ERP components (i.e., P1, N1, P2) were analyzed at the electrode and source levels. Our analyses focused on source level data to evaluate the neural generators of SIN processing more directly than electrode level analysis allows due to volume conduction and detection of activity from surrounding electrodes. ERP sensor data is reported in Appendix A. For ERP source analysis, peak magnitudes and latencies reflect PAC activity averaged across hemispheres. P1 and P2 magnitudes decreased with the addition of noise [P1: $F_{(1, 136)} = 77.33$, $p < 0.0001$; P2: $F_{(1, 136)} = 6.21$, $p = 0.01$] although differences in N1 magnitude were negligible [$F_{(1, 136)} = 0.04$, $p = 0.84$]. A vowel effect was also noted for P1 magnitude with /a/ eliciting a more robust response than /i/ [$F_{(1, 136)} = 4.18$, $p = 0.04$]. We attributed this to the differing spectral properties of the vowels since P1 magnitude is influenced by acoustic characteristics (Ceponiene, Torki, Alku, Koyama, & Townsend, 2008).

Attentional enhancement was noted for later component magnitudes including P2 and the N1-P2 complex [P2: $F_{(1, 136)} = 4.97$, $p = 0.03$; N1-P2: $F_{(1, 136)} = 4.14$, $p = 0.04$]. Attention and SNR effects on magnitudes are shown in Figure 10. For latencies, SNR affected all ERP components decreasing latency for P1 and P2 and increasing latency for N1 [P1: $F_{(1, 136)} = 6.39$, $p = 0.01$; N1: $F_{(1, 136)} = 16.24$, $p < 0.0001$; P2: $F_{(1, 136)} = 5.41$, $p = 0.02$]. No other significant effects or interactions were found.

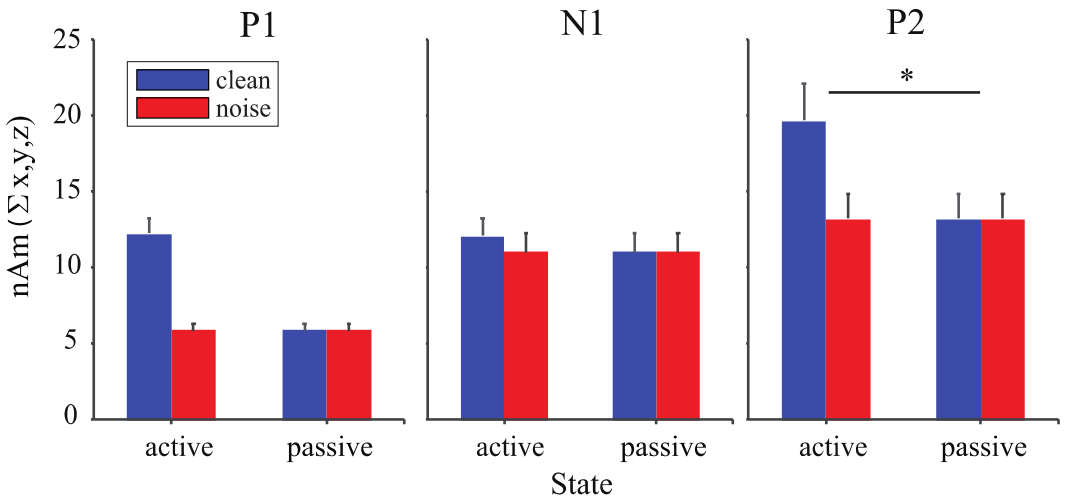

Figure 10: Attention and SNR effects on ERP source magnitudes. Attention increased P2 magnitude while SNR decreased P1 and P2 magnitudes; N1 magnitude was invariant to both attention and noise effects. errorbars = ± s.e.m, *$p < 0.05$

Brainstem-Cortical Functional Connectivity. Functional connectivity between BS and PAC is depicted in Figure 11. Connectivity strengthened in both bottom-up and top-down directions during active listening [$F_{(1, 295)} = 7.92$, $p = 0.01$] yet remained unaffected by SNR [$F_{(1, 295)} = 1.66$, $p = 0.20$]. Overall, efferent connectivity was much more robust than afferent regardless of SNR or attentional state [$F_{(1, 295)} = 1790.45$, $p < 0.0001$]. More critically, we found a direction x SNR interaction [$F_{(1, 295)} = 11.59$, $p < 0.001$]; whereas efferent (PAC→BS) connectivity weakened in noise while afferent (BS→PAC) connectivity was invariant [efferent: $t_{(295)} = 3.32$, $p = 0.01$; afferent: $t_{(295)} = -1.50$, $p = 0.44$]. Additional post-hoc testing revealed that

noise-related decreases in efferent connectivity occurred only in the passive condition [$t_{(292)}$ = 3.12, p = 0.04]; efferent feedback was unaffected by noise during active SIN perception [$t_{(292)}$ = 1.57, p = 0.77].

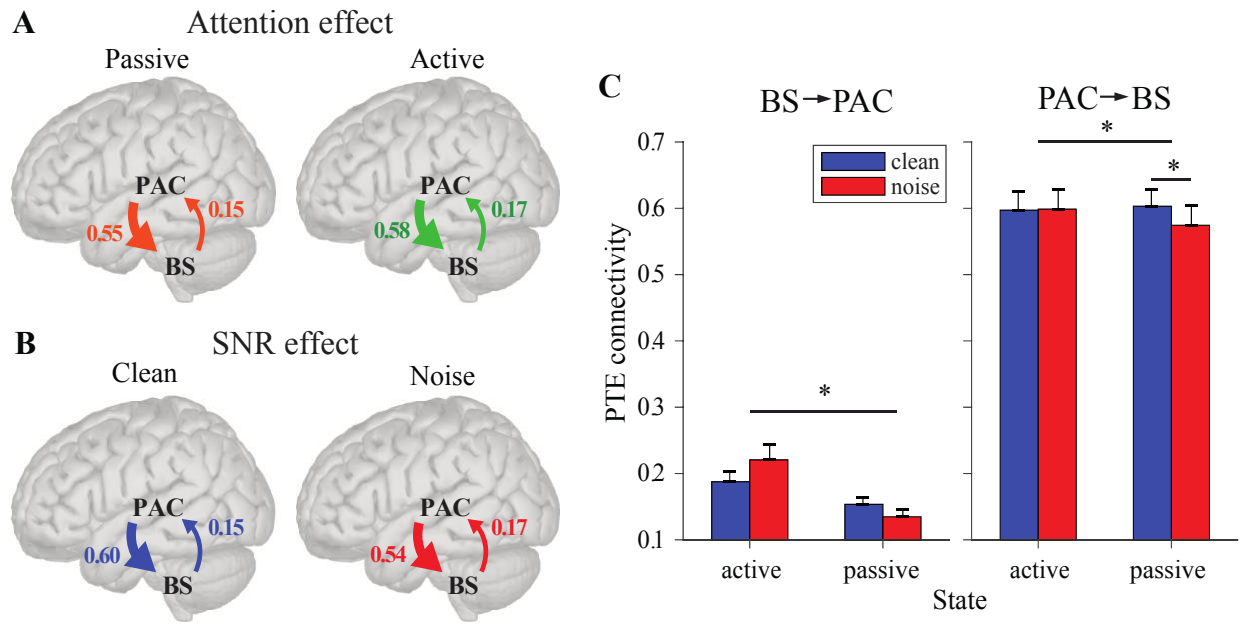

Figure 11: Attention and SNR effects on afferent and efferent connectivity. Attention increases bidirectional connectivity between BS and PAC (A, C). Efferent connectivity is decreased in noise (B), but this reduction only occurs during passive listening (C). Attention maintains efferent signaling strength in more challenging listening conditions (C). errorbars = ± s.e.m, *p < 0.05

Brain-Behavior Correlations. To evaluate how underlying neural mechanisms contribute to behavioral responses in noise, a GLME using neural responses (i.e., F0 amplitude, P2 magnitude, afferent and efferent connectivity), measures of hearing ability (i.e., QuickSIN, PTA), and condition (i.e., SNR) to predict perceptual efficiency was developed. The overall model was significant [$F_{(7,72)}$ = 5.83, p < 0.0001] explaining 27% of the variance (adjusted r^2 = 0.27) in behavioral throughput. Evaluating individual terms revealed significant predictors in F0 amplitude [$t_{(72)}$ = 2.22, p = 0.03), afferent connectivity [$t_{(72)}$ = -2.19, p = 0.03), QuickSIN scores [$t_{(72)}$ = -2.51, p = 0.01), and SNR [$t_{(72)}$ = -4.39, p < 0.0001]. Model fit parameters are reported in Table 1.

Table 1. GLME model fit parameters for prediction of behavioral throughput. Coefficients and significance tests for individual predictor variables including neural and hearing measures and SNR. Level of significance denoted *p< 0.05, ***p< 0.001

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
Intercept	0.1633	0.0197	8.29	72	4.41e-12	0.1240	0.2026
FFR*	0.0044	0.0020	2.22	72	0.03	0.0005	0.0084
ERP	-0.0004	0.0003	-1.55	72	0.13	-0.0009	0.0001
aff*	-0.0793	0.0363	-2.19	72	0.03	-0.1517	-0.0070
eff	0.0043	0.0241	0.18	72	0.86	-0.0437	0.0524
QSIN*	-0.0043	0.0017	-2.51	72	0.01	-0.0077	-0.0009
PTA	-0.0003	0.0007	-0.47	72	0.64	-0.0018	0.0011
SNR***	-0.0237	0.0054	-4.39	72	3.87e-05	-0.0345	-0.0130

Discussion

By simultaneously recording source-level brainstem FFRs, cortical ERPs, and connectivity between these ROIs, our study resolves that (i) attention actively modulates SIN processing throughout the auditory pathway as early as the brainstem and (ii) attention reinforces top-down neural signaling in adverse listening conditions. Thus, our study is one of the first to provide a more direct assessment and evidence of corticofugal involvement in the human auditory system during active speech processing.

SNR Differentially Affects Hierarchical Speech Processing

When analyzed at the source level, noise-related changes were evident only for cortical potentials with general reductions in magnitude and latency. Our findings of decreased magnitude for speech-evoked cortical responses are consistent with previous literature suggesting noise weakens neural representations within PAC (Bidelman & Howell, 2016;

Bidelman, Price, et al., 2019). While most studies report increased latency with poorer SNR (Bidelman & Howell, 2016; Billings et al., 2013; Billings, Tremblay, Stecker, & Tolin, 2009), latencies of cortical responses also decrease with increasing intensity (Picton, Hillyard, Krausz, & Galambos, 1974; Picton et al., 1976). It is likely that the observed decrease in latency reflects overall intensity changes (i.e., noise > clean) between SNR conditions. Contrasting cortical effects, F0 encoding within the brainstem is resistant to degrading noise effects because F0 is reinforced by upper harmonics in the complex acoustic stimulus (Smith, Marsh, Greenberg, & Brown, 1978). In fact, previous studies demonstrate the neural representation of F0 is often retained or even enhanced when the stimulus F0 is obscured by noise or is missing altogether (Bidelman, 2016; Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, & Kraus, 2009; Prevost, Laroche, Marcoux, & Dajani, 2013; Smalt, Krishnan, Bidelman, Ananthakrishnan, & Gandour, 2012; Smith et al., 1978).

Relatively few studies to date compare cross-level SIN differences enabled by simultaneous subcortical-cortical recordings. Using magnetoencephalography (MEG), Coffey et al. (2017) suggest cortical measures (i.e., cortical FFR F0 strength, P2 amplitude) are better predictors of SIN perception, but this study did not directly evaluate how noise influences responses across levels of processing. Our findings replicate results of Bidelman, Price, et al. (2019) who employed a similar rapid speech detection task in older adults. While our findings suggest that brainstem encoding of F0 is less susceptible to noise effects than cortical potentials, Bidelman, Davis, et al. (2018) conclude that noise modulates brainstem responses to a greater degree than cortical responses. However, this conclusion was based on their observation of noise strengthening F0 amplitude while having no effect on N1 amplitude. Similarly, our study found no noise-related changes in N1 magnitude but noted these effects in other ERP components.

Parbery-Clark et al. (2011) report noise effects in both brainstem and cortex with noise reducing subcortical response amplitudes and precision yet enhancing N1. A potential explanation for these cross-level differences is that ERP components reflect the encoding of a variety of acoustic cues, including pitch and timbre, whereas the FFR F0 reflects pitch encoding (Agung, Purdy, McMahon, & Newall, 2006; Alain, Roye, & Arnott, 2013; Bidelman & Krishnan, 2010; Skoe & Kraus, 2010a). Consequently, differential noise effects observed across levels of processing may reflect differences in acoustic features captured in each response. In fact, when the FFR analysis includes harmonics, or aspects of timbre, greater noise effects are observed (Bidelman, Davis, et al., 2018; Parbery-Clark et al., 2011). These findings demonstrate that certain acoustic characteristics may be less resistant to noise than others, but whether they are affected to the same degree at subcortical-cortical levels is not well understood. Although seemingly equivocal, these studies provide converging evidence that noise differentially influences speech coding at brainstem and cortical levels. Further study is needed to clarify the factors underlying differences across functional levels of the auditory pathway.

Attention Modulates SIN Processing as Early as the Rostral Brainstem

In comparing attentional state during SIN processing, we found active listening enhances neural encoding across the auditory system, including both subcortical and cortical levels. Attentional modulation of cortical potentials, particularly N1 and P2, is well-established in the literature. Moreover, several studies have linked these components to complex speech processing (Bidelman, Price, et al., 2019; Billings et al., 2013; Parbery-Clark et al., 2011). Studies evaluating attentional effects on N1-P2 amplitude demonstrate attention-driven enhancement of the N1 response with associated reductions in P2 amplitude (Crowley & Colrain, 2004; Naatanen, 1975), yet other studies demonstrate overall attentional enhancements of later cortical

components (Gross, Begleiter, Tobin, & Kissin, 1965; Picton & Hillyard, 1974; Picton et al., 1971). In the present study, we found increased P2 magnitudes for active listening, whereas N1 was invariant across attentional state. These findings are consistent with previous studies that posit P2 reflects attentional modulation of nontarget stimuli in auditory discrimination tasks and indexes stimulus classification (Garcia-Larrea, Lukaszewicz, & Mauguiere, 1992; Novak, Ritter, & Vaughan, 1992), which parallel the demands of our task. Although we did not observe attentional enhancements of N1 magnitude, it is likely that the rapid presentation rate resulted in adaptation of the N1 response (Naatanen & Picton, 1987; Picton et al., 1976). As seen in Figure 8, the grand average waveforms show minimal N1 response amplitude; therefore, any differences due to attention, if present, could not be adequately detected. Indeed, a recent study reports attentional effects on N1 within PAC but found that this effect interacts with the stimulus presentation rate (Neelon, Williams, & Garell, 2006).

Most notably, attentional enhancements were observed in "pre-attentive" subcortical levels, but these effects were only made apparent at the source level. Attentional modulation of brainstem responses has been a highly elusive effect in prior EEG studies (Dunlop, Webster, & Simons, 1965; Galbraith & Kane, 1993; Picton et al., 1971; Varghese et al., 2015). During an active speech perception task varying in attentional demand, Varghese et al. (2015) observed attentional influences only at the cortical level and concluded that if top-down modulation of attention occurs within the brainstem these changes are minimal. Such conclusions are supported by our channel-level data. While scalp-level recordings reflect summed activity of multiple generators (Luck, 2005), source analysis provides more accurate representations of localized contributions minimizing extraneous activity from adjacent regions or additional generators (Michel et al., 2004). By providing "purer" brainstem responses, source analysis proves to be

more sensitive to attentional changes. Additionally, typical channel-level FFRs are recorded using a mastoid reference which contains significant contributions from auditory nerve (Bidelman, 2018; Galbraith et al., 2000). With such a strong peripheral component, attentional effects would not be expected in brainstem responses (Picton & Hillyard, 1974; Picton et al., 1971; Woods & Hillyard, 1978).

Other contradictory conclusions in attentional studies of the FFR may be attributed to methodological differences across studies including varying task demands, SNR of electrophysiologic responses, and/or stimulus properties (Holmes et al., 2018; Varghese et al., 2015). Some studies demonstrating attentional effects use stimuli with F0s at or below the cortical phase-locking limit (Forte, Etard, & Reichenbach, 2017; Holmes et al., 2018; Saiz-Alia et al., 2019). In these instances, the FFRs contain phase-locked cortical activity and likely reflect attentional enhancement of cortical rather than subcortical encoding. Given the high F0 (150 Hz) of our stimuli, we ensure the subcortical origins of recorded FFRs. Hence, our results provide the most convincing evidence to date that attention enhances speech coding online as early as the brainstem.

Attention Reinforces Corticofugal Engagement in Difficult Listening Conditions

Beyond local enhancements to speech representations, our data also show that attention strengthens neural signaling in both feedforward (afferent) and feedback (efferent) directions. Similarly in visual studies, attention increases connectivity between regions involved in sensory processing (Buchel & Friston, 1997), and attentional selection is thought to be driven by the interaction of both feedforward and feedback mechanisms (Khorsand, Moore, & Soltani, 2015). Somewhat surprisingly, top-down connectivity was stronger throughout our task regardless of attentional state or SNR. The fact that the efferent system remained highly active regardless of

attention may explain why some studies have shown enhancements in FFR amplitude even in passive listening tasks (Chandrasekaran et al., 2009; Skoe & Kraus, 2010b).

Previous animal studies suggest that the corticofugal system can sharpen and enhance subcortical processing of auditory signals during short-term auditory learning (Bajo, Nodal, Moore, & King, 2010; Suga, 2008). Similarly, FFR enhancements in human listeners with long-term experience or training (Bidelman et al., 2011; Carcagno & Plack, 2011; Krishnan et al., 2005; Skoe et al., 2013; Wong et al., 2007) have been assumed to reflect similar top-down, corticofugal tuning of brainstem auditory processing (Chandrasekaran et al., 2009; Galbraith et al., 1998; Galbraith & Doan, 1995; Galbraith et al., 2003; Krishnan et al., 2012; Lukas, 1980, 1981; Oatman & Anderson, 1980; Skoe & Kraus, 2010b; Tzounopoulos & Kraus, 2009). To date, evidence of direct corticofugal involvement in human auditory processing has never been verifiable. Theoretically, increases in top-down contributions are expected in more challenging scenarios (e.g., during learning, degraded listening environments, increased attentional demands) to sharpen earlier processing and facilitate transmission of faithful neural representations of the acoustic input. Using direct measures of brainstem-cortical connectivity, we find that attention reinforces corticofugal signaling in more difficult conditions but does not enhance it, *per se*. Our findings suggest that attention maintains top-down signaling in noise to tune and enhance early pitch encoding (F0) which is associated with improved behavioral performance. Whether parametrically varying SNR or task difficulty in another manner would impact the degree of connectivity would be an interesting focus of future study.

Conclusions

Study Limitations

Just as with any conducted research, the current study is subject to limitations particularly related to analyses methods, their scope, and their generalizability. We used source analysis techniques to evaluate SNR- and attention-related differences in neural activity generated within auditory BS and PAC. Source analyses provide insight into both *when* and *where* activity is generated in the brain while traditional channel-level analyses reveal the time course of summed neural activity from multiple, unspecified generators (Luck, 2005; Scherg, Vajsar, & Picton, 1989). Yet, source analyses rely on a number of assumptions that may influence the resulting source waveforms. For instance, a discrete dipole model, as used in this study, requires the researcher to specify the number of generating sources (Michel et al., 2004). Although prior knowledge of neuroanatomical structures and their relation to sensory and perceptual processes may contribute to hypothesized source locations and/or number of sources, the true solution of underlying sources for the recorded scalp distributions cannot be known with complete certainty (Michel et al., 2004). However, the high average goodness of fit across conditions for our 3-dipole model (GoF = 90.6%) suggests that this somewhat simplistic model well-captures the scalp recorded data (RV = 9.4 ± 1.6%).

Additionally, our study evaluated attentional effects on SIN processing only in early auditory processing centers (i.e., BS and PAC). While neural responses from these regions reflect the encoding fidelity of acoustic properties and perceptual processes (i.e., auditory object formation and categorization) (Bidelman et al., 2013; Key, Dove, & Maguire, 2005; Skoe & Kraus, 2010a), these regions alone do not encompass the entire neural networks implicated in either attentional or perceptual speech processing. Frontal areas (i.e., anterior cingulate,

prefrontal, frontal cortices) and interhemispheric connections have been shown to contribute to auditory attentional processes (Bamiou, Sisodiya, Musiek, & Luxon, 2007; Benedict et al., 2002; Petersen & Posner, 2012), and activation within secondary auditory and prefrontal regions (i.e., posterior superior temporal gyrus, inferior frontal gyrus, prefrontal cortex) strengthens during speech perception tasks (Hickok & Poeppel, 2000; Wong et al., 2009). Because we did not analyze activity within these areas specifically, conclusions cannot be drawn as to how attention influences SIN encoding across the entire speech perception pathway or whether contributions of higher-order structures related to attention differ in more challenging listening conditions. Future study could investigate these remaining questions to provide a more comprehensive view of attentional influences on speech perception in noise.

Furthermore, as only speech stimuli were presented, we cannot determine from our findings whether the observed attentional enhancements are specific to speech processing. Attentional literature suggests that attentional enhancements may be elicited for behaviorally relevant inputs, including tonal stimuli in animal studies (Suga, 2008; Suga, Gao, Zhang, Ma, & Olsen, 2000). However, whether the observed enhancements in encoding and neural signaling in humans would be exhibited to the same degree across speech and non-speech stimuli remains unclear.

Summary

In sum, our results emphasize the complex interaction of auditory and cognitive processes, namely attention, during speech perception. Our results provide evidence to resolve ongoing debates regarding attentional influences on early brainstem encoding and corticofugal engagement during active listening in humans. Attention modulates online speech processing at both subcortical and cortical levels and strengthens neural signaling within the central auditory

pathway. Furthermore, attention serves as a mechanism to overcome detrimental noise effects and maintain efficient top-down signaling in challenging listening conditions. Overall, our findings suggest that attention serves to maintain feedback from descending pathways and enhance early encoding of acoustic features for subsequent processing.

Chapter 4

General Discussion

Using an array of electrophysiologic techniques (i.e., FFRs, ERPs, neural oscillations, and functional connectivity), these studies extend previous work by demonstrating cognitive influences on subcortical and cortical speech processing and providing a more thorough representation of neural activation and transmission during active SIN perception. Moreover, these studies capture the intricacies of the complex system underlying speech understanding by assessing how bottom-up perceptual aspects (i.e., hearing loss, noise) interact with top-down, cognitive processes (i.e., attention, inhibition, template matching). Together with our previous work (Bidelman, Price, et al., 2019), the results of the first study show that older adults with hearing impairment have weaker bottom-up (feed-forward) processing from auditory brainstem to cortex yet stronger top-down influences from linguistic brain regions (e.g., IFG) to auditory cortex. These findings reflect the use of alternative cognitive processing strategies to compensate for impoverished auditory inputs during SIN processing. The second study, investigating top-down attentional effects in young adults, provides novel evidence of active corticofugal engagement during challenging listening conditions. The results reveal that attention reinforces top-down signaling in poor SNRs and enhances early speech encoding at the level of the brainstem. Collectively, these studies demonstrate the importance of top-down contributions for accurate SIN perception.

Importantly, gaining a better understanding of the neurophysiological mechanisms, cognitive processes, and other factors contributing to SIN perception may reveal how deficits manifest in disordered systems and elucidate new avenues of exploration for clinical interventions. The objective, electrophysiologic measures used in these studies have the potential to identify the true area of deficit within the pathway (i.e. early sensory encoding within the

brainstem, early perceptual processing within auditory cortex, higher-order global cognitive processes, neural transmission within the pathway). Clarifying the area of deficit would disentangle the underlying mechanisms contributing to impaired perceptual processing and/or more global cognitive deficits. Additionally, isolating the processes that are most problematic for individuals with SIN deficits can inform the specific auditory or cognitive skill(s) to be targeted by auditory or cognitive training exercises or other rehabilitative interventions.

Ultimately, there is still much to be discovered regarding the study of cognitive processes, the auditory system, and SIN perception. Future studies should aim to further address how and under what circumstances top-down mechanisms, such as attention, interact with sensory inputs to influence speech understanding. Our study finds that noise differentially influences processing at subcortical and cortical levels, but the factors contributing to these differences remain unclear. Studies that parametrically vary the task difficulty (e.g., SNR, attentional engagement/arousal, cognitive load) or alter acoustic characteristics of the inputs may reveal the mechanisms responsible for such changes. These studies would also demonstrate whether activity within hierarchical levels of processing and neural signaling differs with increasing task demands. Applying similar paradigms to various patient populations (e.g., varying degrees of hearing loss, cognitive impairments) would identify whether similar compensatory processing patterns exist within a given population and how these may be optimized to improve interventions for these individuals. Lastly, new studies incorporating targeted auditory or cognitive training paradigms would evaluate focused training effects on SIN processing. This work would allow for the comparison of the effectiveness of targeting specific skills and also identify the most efficient intervention that provides the most robust, long-lasting benefits for behavioral SIN perception.

References

Adrian, E. D., & Matthews, B. H. (1934). The interpretation of potential waves in the cortex. *J Physiol, 81*(4), 440-471. doi:10.1113/jphysiol.1934.sp003147

Agung, K., Purdy, S. C., McMahon, C. M., & Newall, P. (2006). The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *Journal of the American Academy of Audiology, 17*(8), 559-572.

Alain, C., Roye, A., & Arnott, S. R. (2013). Middle-and long-latency auditory evoked potentials: what are they telling us on central auditory disorders. *Handbook of clinical neurophysiology: disorders of peripheral and central auditory processing, 10*, 177-199.

Alain, C., Roye, A., & Salloum, C. (2014). Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. *Front Syst Neurosci, 8*, 8. doi:10.3389/fnsys.2014.00008

Alain, C., & Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. *Psychol Aging, 14*(3), 507-519. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10509703

Anderson Gosselin, P., & Gagné, J.-P. (2011). Older Adults Expend More Listening Effort Than Young Adults Recognizing Speech in Noise. *Journal of Speech, Language, and Hearing Research, 54*(3), 944-958. doi:10.1044/1092-4388(2010/10-0069)

Anderson, S., Parbery-Clark, A., Yi, H. G., & Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear & Hearing, 32*(6), 750-757. doi:10.1097/AUD.0b013e31822229d3

Assmann, P. F. (1996). Tracking and glimpsing speech in noise: Role of fundamental frequency. *The Journal of the Acoustical Society of America, 100*(4), 2680-2680. doi:10.1121/1.416961

Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron, 61*(3), 467-480. doi:10.1016/j.neuron.2008.12.027

Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Front Neural Circuits, 6*, 114. doi:10.3389/fncir.2012.00114

Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci, 13*(2), 253-260. doi:10.1038/nn.2466

Bamiou, D. E., Sisodiya, S., Musiek, F. E., & Luxon, L. M. (2007). The role of the interhemispheric pathway in hearing. *Brain Res Rev, 56*(1), 170-182. doi:10.1016/j.brainresrev.2007.07.003

Bashivan, P., Bidelman, G. M., & Yeasin, M. (2014). Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity. *Eur J Neurosci, 40*(12), 3774-3784. doi:10.1111/ejn.12749

Benedict, R. H. B., Shucard, D. W., Maria, M. P. S., Shucard, J. L., Abara, J. P., Coad, M. L., . . . Lockwood, A. (2002). Covert Auditory Attention Generates Activation in the Rostral/Dorsal Anterior Cingulate Cortex. *Journal of Cognitive Neuroscience, 14*(4), 637-645. doi:10.1162/08989290260045765

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289-300.

Berg, P., & Scherg, M. (1994). A fast method for forward computation of multiple-shell spherical head models. *Electroencephalography and Clinical Neurophysiology, 90*(1), 58–64.

Betzel, R. F., Byrge, L., He, Y., Goni, J., Zuo, X. N., & Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *Neuroimage, 102 Pt 2*, 345-357. doi:10.1016/j.neuroimage.2014.07.067

Bidelman, G. (2016). Relative contribution of envelope and fine structure to the subcortical encoding of noise-degraded speech. *Journal of the Acoustical Society of America, 140*(4), EL358-EL363. doi:10.1121/1.4965248]

Bidelman, G., Gandour, J., & Krishnan, A. (2010). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*(2), 425-434.

Bidelman, G. M. (2015). Induced neural beta oscillations predict categorical speech perception abilities. *Brain Lang, 141*, 62-69. doi:10.1016/j.bandl.2014.11.003

Bidelman, G. M. (2017). Amplified induced neural oscillatory activity predicts musicians' benefits in categorical speech perception. *Neuroscience, 348*, 107-113. doi:10.1016/j.neuroscience.2017.02.015

Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage, 175*, 56-69. doi:10.1016/j.neuroimage.2018.03.060

Bidelman, G. M., Davis, M. K., & Pridgen, M. H. (2018). Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. *Hearing Research, 367*, 149-160.

Bidelman, G. M., & Dexter, L. (2015). Bilinguals at the "cocktail party": dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners' speech-in-noise recognition deficits. *Brain Lang, 143*, 32-41. doi:10.1016/j.bandl.2015.02.002

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn, 77*(1), 1-10. doi:10.1016/j.bandc.2011.07.006

Bidelman, G. M., & Howell, M. (2016). Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. *Neuroimage, 124*(Pt A), 581-590. doi:10.1016/j.neuroimage.2015.09.020

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res, 1355*, 112-125. doi:10.1016/j.brainres.2010.07.100

Bidelman, G. M., Mahmud, M. S., Yeasin, M., Shen, D., Arnott, S. R., & Alain, C. (2019). Age-related hearing loss increases full-brain connectivity while reversing directed signaling within the dorsal-ventral pathway for speech. *Brain Struct Funct*. doi:10.1007/s00429-019-01922-9

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage, 79*, 201-212. doi:10.1016/j.neuroimage.2013.04.093

Bidelman, G. M., Pousson, M., Dugas, C., & Fehrenbach, A. (2018). Test-Retest Reliability of Dual-Recorded Brainstem versus Cortical Auditory-Evoked Potentials to Speech. *J Am Acad Audiol, 29*(2), 164-174. doi:10.3766/jaaa.16167

Bidelman, G. M., Price, C. N., Shen, D., Arnott, S. R., & Alain, C. (2019). Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. *Hear Res, 382*, 107795. doi:10.1016/j.heares.2019.107795

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging, 35*(11), 2526-2540. doi:10.1016/j.neurobiolaging.2014.05.006

Bidelman, G. M., & Yellamsetty, A. (2017). Noise and pitch interact during the cortical segregation of concurrent speech. *Hearing Research, 351*, 34-44. doi:10.1016/j.heares.2017.05.008

Billings, C. J., McMillan, G. P., Penman, T. M., & Gille, S. M. (2013). Predicting perception in noise using cortical auditory evoked potentials. *J Assoc Res Otolaryngol, 14*(6), 891-903. doi:10.1007/s10162-013-0415-y

Billings, C. J., Tremblay, K. L., Stecker, G. C., & Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hear Res, 254*(1-2), 15-24. doi:10.1016/j.heares.2009.04.002

Bilodeau-Mercure, M., Lortie, C. L., Sato, M., Guitton, M. J., & Tremblay, P. (2015). The neurobiology of speech perception decline in aging. *Brain Struct Funct, 220*(2), 979-997. doi:10.1007/s00429-013-0695-3

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci, 7*(3), 295-301. doi:10.1038/nn1198

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., & Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc Natl Acad Sci U S A, 101*(26), 9849-9854. doi:10.1073/pnas.0308538101

Buchel, C., & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex, 7*(8), 768-778. doi:10.1093/cercor/7.8.768

Campbell, J., & Sharma, A. (2013). Compensatory changes in cortical resource allocation in adults with hearing loss. *Front Syst Neurosci, 7*, 71. doi:10.3389/fnsys.2013.00071

Carcagno, S., & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol, 12*(1), 89-100. doi:10.1007/s10162-010-0236-1

Cardin, V. (2016). Effects of Aging and Adult-Onset Hearing Loss on Cortical Auditory Regions. *Front Neurosci, 10*, 199. doi:10.3389/fnins.2016.00199

Caspary, D. M., Ling, L., Turner, J. G., & Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *J Exp Biol, 211*(Pt 11), 1781-1791. doi:10.1242/jeb.013581

Ceponiene, R., Torki, M., Alku, P., Koyama, A., & Townsend, J. (2008). Event-related potentials reflect spectral differences in speech and non-speech stimuli in children and adults. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology, 119*(7), 1560-1577. doi:10.1016/j.clinph.2008.03.005

Chambers, A. R., Resnik, J., Yuan, Y., Whitton, J. P., Edge, A. S., Liberman, M. C., & Polley, D. B. (2016). Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation. *Neuron, 89*(4), 867-879. doi:10.1016/j.neuron.2015.12.041

Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *Trans. Img. Proc., 10*(2), 266–277. doi:10.1109/83.902291

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron, 64*(3), 311-319. doi:10.1016/j.neuron.2009.10.006

Chao, L. L., & Knight, R. T. (1997). Prefrontal deficits in attention and inhibitory control with aging. *Cereb Cortex, 7*(1), 63-69. doi:10.1093/cercor/7.1.63

Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun, 7*, 11070. doi:10.1038/ncomms11070

Coffey, E. B. J., Chepesiuk, A. M. P., Herholz, S. C., Baillet, S., & Zatorre, R. J. (2017). Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. *Front Neurosci, 11*, 479. doi:10.3389/fnins.2017.00479

Cope, T. E., Sohoglu, E., Sedley, W., Patterson, K., Jones, P. S., Wiggins, J., . . . Rowe, J. B. (2017). Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat Commun, 8*(1), 2154. doi:10.1038/s41467-017-01958-7

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126*(Pt 5), 1193-1201. doi:10.1093/brain/awg104

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol, 115*(4), 732-744. doi:10.1016/j.clinph.2003.11.021

Dalton, D. S., Cruickshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *Gerontologist, 43*(5), 661-668. doi:10.1093/geront/43.5.661

Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schroger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *Neuroimage, 20*(2), 1270-1282. doi:10.1016/S1053-8119(03)00389-6

Doppelmayr, M., Klimesch, W., Hodlmoser, K., Sauseng, P., & Gruber, W. (2005). Intelligence related upper alpha desynchronization in a semantic memory task. *Brain Res Bull, 66*(2), 171-177. doi:10.1016/j.brainresbull.2005.04.007

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

Dunlop, C. W., Webster, W. R., & Simons, L. A. (1965). Effect of attention on evoked responses in the classical auditory pathway. *Nature, 206*(988), 1048-1050. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/5839063

Dushanova, J., & Christov, M. (2016). Do the changes of event-related potentials and frequency band responses to sensory stimuli correlate to age cognitive decline? *International Journal of Neurology Research, 3*(1), 327-334.

Fabiani, M., Low, K. A., Wee, E., Sable, J. J., & Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *J Cogn Neurosci, 18*(4), 637-650. doi:10.1162/jocn.2006.18.4.637

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195-207. doi:10.1006/nimg.1998.0396

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A. L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat Commun, 5*, 4694. doi:10.1038/ncomms5694

Forte, A. E., Etard, O., & Reichenbach, T. (2017). The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *Elife, 6*. doi:10.7554/eLife.27203

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res, 106*(1-2), 95-104. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9112109

Friston, K. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 815-836. doi:10.1098/rstb.2005.1622

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., & Ebersole, J. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology, 113*(5), 702-712.

Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport, 6*, 2363-2367.

Galbraith, G. C., & Arroyo, C. (1993). Selective attention and brainstem frequency-following responses. *Biological psychology, 37*(1), 3-22.

Galbraith, G. C., Bhuta, S. M., Choate, A. K., Kitahara, J. M., & Mullen, T. A. (1998). Brainstem frequency-following response to dichotic vowels during attention. *Neuroreport, 9*, 1889-1893.

Galbraith, G. C., & Doan, B. Q. (1995). Brainstem frequency-following and behavioral responses during selective attention to pure tones and missing fundamental stimuli. *International Journal of Psychophysiology, 19*, 203-214.

Galbraith, G. C., & Kane, J. M. (1993). Brainstem frequency-following responses and cortical event-related potentials during attention. *Perceptual and Motor Skills, 76*, 1231-1241.

Galbraith, G. C., Olfman, D. M., & Huffman, T. M. (2003). Selective attention affects human brain stem frequency-following response. *Neuroreport, 14*(5), 735-738. doi:10.1097/01.wnr.0000064983.96259.49

Galbraith, G. C., Threadgill, M. R., Hemsley, J., Salour, K., Songdej, N., Ton, J., & Cheung, L. (2000). Putative measure of peripheral and brainstem frequency-following in humans. *Neurosci Lett, 292*(2), 123-127. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10998564

Gao, E., & Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc Natl Acad Sci U S A, 97*(14), 8081-8086. doi:10.1073/pnas.97.14.8081

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proc Natl Acad Sci U S A, 96*, 8301-8306.

Garcia-Larrea, L., Lukaszewicz, A. C., & Mauguiere, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia, 30*(8), 723-741. doi:10.1016/0028-3932(92)90042-k

Gelfand, S. A. (2009). *Essentials of Audiology* (3rd ed.). New York: Thieme.

Giordano, B. L., Ince, R. A. A., Gross, J., Schyns, P. G., Panzeri, S., & Kayser, C. (2017). Contributions of local speech encoding and functional connectivity to audio-visual speech perception. *eLife, 6*, e24763. doi:10.7554/eLife.24763

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci, 15*(4), 511-517. doi:10.1038/nn.3063

Gopinath, B., Wang, J., Schneider, J., Burlutsky, G., Snowdon, J., McMahon, C. M., . . . Mitchell, P. (2009). Depressive symptoms in older adults with hearing impairment: The Blue Mountains study. *Journal of the American Geriatrics Society, 57*(7), 1306-1308.

Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear Res, 36*(6), 1276-1285. doi:10.1044/jshr.3606.1276

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *J Speech Lang Hear Res, 42*(2), 300-311. doi:10.1044/jslhr.4202.300

Grady, C. (2012). The cognitive neuroscience of ageing. *Nat Rev Neurosci, 13*(7), 491-505. doi:10.1038/nrn3256

Gross, M. M., Begleiter, H., Tobin, M., & Kissin, B. (1965). Auditory Evoked Response Comparison during Counting Clicks and Reading. *Electroencephalogr Clin Neurophysiol, 18*, 451-454. doi:10.1016/0013-4694(65)90125-2

Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hear Res, 364*, 142-151. doi:10.1016/j.heares.2018.03.008

Hall, J. W. (1992). *Handbook of auditory evoked responses*. Needham Heights: Allyn and Bacon.

Helfer, K. S., & Wilber, L. A. (1990). Hearing loss, aging, and speech perception in reverberation and noise. *J Speech Hear Res, 33*(1), 149-155.

Hernandez-Peon, R. (1966). Physiological mechanisms in attention. In R. W. Russell (Ed.), *Frontiers in Physiological Psychology*. New York: Academic.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*(4), 131-138. doi:https://doi.org/10.1016/S1364-6613(00)01463-7

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annu Rev Psychol, 34*, 33-61. doi:10.1146/annurev.ps.34.020183.000341

Hoechstetter, K., Bornfleth, H., Weckesser, D., Ille, N., Berg, P., & Scherg, M. (2004). BESA source coherence: a new method to study cortical oscillatory coupling. *Brain Topogr, 16*(4), 233-238. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15379219

Holmes, E., Purcell, D. W., Carlyon, R. P., Gockel, H. E., & Johnsrude, I. S. (2018). Attentional Modulation of Envelope-Following Responses at Lower (93–109 Hz) but Not Higher (217–233 Hz) Modulation Rates. *Journal of the Association for Research in Otolaryngology, 19*(1), 83-97. doi:10.1007/s10162-017-0641-9

Humes, L. E. (1996). Speech understanding in the elderly. *J Am Acad Audiol, 7*(3), 161-167. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8780988

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, N.J.,: Prentice-Hall.

Key, A. P., Dove, G. O., & Maguire, M. J. (2005). Linking brainwaves to the brain: an ERP primer. *Dev Neuropsychol, 27*(2), 183-215. doi:10.1207/s15326942dn2702_1

Khorsand, P., Moore, T., & Soltani, A. (2015). Combined contributions of feedforward and feedback inputs to bottom-up attention. *Front Psychol, 6*, 155. doi:10.3389/fpsyg.2015.00155

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am, 116*(4 Pt 1), 2395-2405. doi:10.1121/1.1784440

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci, 16*(12), 606-617. doi:10.1016/j.tics.2012.10.007

Klimesch, W., Doppelmayr, M., Pachinger, T., & Ripper, B. (1997). Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. *Neurosci Lett, 238*(1-2), 9-12. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9464642

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev, 53*(1), 63-88. doi:10.1016/j.brainresrev.2006.06.003

Krishnamurti, S. (2001). P300 auditory event-related potentials in binaural and competing noise conditions in adults with central auditory processing disorders. *Contemporary Issues in Communication Sciences and Disorders, 28*, 40-47.

Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain Lang, 110*(3), 135-148. doi:10.1016/j.bandl.2009.03.005

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2012). Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport, 23*(8), 498-502. doi:10.1097/WNR.0b013e328353764d

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res, 25*(1), 161-168. doi:10.1016/j.cogbrainres.2005.05.004

Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., & D'Angelo, W. R. (2002). Sources of the scalp-recorded amplitude-modulation following response. *Journal of the American Academy of Audiology, 13*, 188-204.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J Exp Psychol Hum Percept Perform, 21*(3), 451-468. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7790827

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *J Exp Psychol Gen, 133*(3), 339-354. doi:10.1037/0096-3445.133.3.339

Lehmann, A., & Schonwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PLoS One, 9*(1), e85442. doi:10.1371/journal.pone.0085442

Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol, 92*(3), 204-214. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7514990

Lin, F. R., Ferrucci, L., An, Y., Goh, J. O., Doshi, J., Metter, E. J., . . . Resnick, S. M. (2014). Association of hearing impairment with brain volume changes in older adults. *Neuroimage, 90*, 84-92. doi:10.1016/j.neuroimage.2013.12.059

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25*(6), 763-770. doi:10.1037/a0024238

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., . . . Health, A. B. C. S. G. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern Med, 173*(4), 293-299. doi:10.1001/jamainternmed.2013.1868

Lobier, M., Siebenhuhner, F., Palva, S., & Palva, J. M. (2014). Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *Neuroimage, 85 Pt 2*, 853-872. doi:10.1016/j.neuroimage.2013.08.056

Lopez, M. E., Aurtenetxe, S., Pereda, E., Cuesta, P., Castellanos, N. P., Bruna, R., . . . Bajo, R. (2014). Cognitive reserve is associated with the functional organization of the brain in healthy aging: a MEG study. *Front Aging Neurosci, 6*, 125. doi:10.3389/fnagi.2014.00125

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique.*: MIT Press.

Lukas, J. H. (1980). Human auditory attention: the olivocochlear bundle may function as a peripheral filter. *Psychophysiology, 17*(5), 444-452. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7465714

Lukas, J. H. (1981). The role of efferent inhibition in human auditory attention: an examination of the auditory brainstem potentials. *Int J Neurosci, 12*(2), 137-145. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7203823

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clin Neurophysiol, 115*(10), 2195-2222. doi:10.1016/j.clinph.2004.06.001

Mishra, S., Lunner, T., Stenfelt, S., Ronnberg, J., & Rudner, M. (2013). Visual information can hinder working memory processing of speech. *J Speech Lang Hear Res, 56*(4), 1120-1132. doi:10.1044/1092-4388(2012/12-0033)

Mishra, S., Stenfelt, S., Lunner, T., Ronnberg, J., & Rudner, M. (2014). Cognitive spare capacity in older adults with hearing loss. *Front Aging Neurosci, 6*, 96. doi:10.3389/fnagi.2014.00096

Moore, D. R. (2015). Sources of pathology underlying listening disorders in children. *Int J Psychophysiol, 95*(2), 125-134. doi:10.1016/j.ijpsycho.2014.07.006

Morcom, A. M., & Henson, R. N. A. (2018). Increased prefrontal activity with aging reflects nonspecific neural responses rather than compensation. *Journal of Neuroscience, 38*(33), 7303-7313. doi:10.1523/jneurosci.1701-17.2018

Naatanen, R. (1975). Selective attention and evoked potentials in humans--a critical review. *Biol Psychol, 2*(4), 237-307. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1156626

Naatanen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3615753

Neelon, M. F., Williams, J., & Garell, P. C. (2006). The effects of attentional load on auditory ERPs recorded from human cortex. *Brain Res, 1118*(1), 94-105. doi:10.1016/j.brainres.2006.08.006

Nobukawa, S., Kikuchi, M., & Takahashi, T. (2019). Changes in functional connectivity dynamics with aging: A dynamical phase synchronization approach. *Neuroimage, 188*, 357-368. doi:10.1016/j.neuroimage.2018.12.008

Novak, G., Ritter, W., & Vaughan, H. G., Jr. (1992). Mismatch detection and the latency of temporal judgements. *Psychophysiology, 29*(4), 398-411. doi:10.1111/j.1469-8986.1992.tb01713.x

Oatman, L. C., & Anderson, B. W. (1980). Suppression of the auditory frequency following response during visual attention. *Electroencephalography and Clinical Neurophysiology, 49*, 314-322.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. doi:10.1016/0028-3932(71)90067-4

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol, 112*(4), 713-719. doi:10.1016/s1388-2457(00)00527-7

Papp, N., & Ktonas, P. (1977). Critical evaluation of complex demodulation techniques for the quantification of bioelectrical activity. *Biomed Sci Instrum, 13*, 135-145. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/871500

Parbery-Clark, A., Marmel, F., Bair, J., & Kraus, N. (2011). What subcortical-cortical relationships tell us about processing speech in noise. *Eur J Neurosci, 33*(3), 549-557. doi:10.1111/j.1460-9568.2010.07546.x

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci, 29*(45), 14100-14107. doi:10.1523/JNEUROSCI.3256-09.2009

Park, D. C., & McDonough, I. M. (2013). The Dynamic Aging Mind: Revelations From Functional Neuroimaging Research. *Perspect Psychol Sci, 8*(1), 62-67. doi:10.1177/1745691612469034

Petersen, E. B., Wostmann, M., Obleser, J., Stenfelt, S., & Lunner, T. (2015). Hearing loss impacts neural alpha oscillations under adverse listening conditions. *Front Psychol, 6*, 177. doi:10.3389/fpsyg.2015.00177

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu Rev Neurosci, 35*, 73-89. doi:10.1146/annurev-neuro-062111-150525

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Res, 41*(10-11), 1257-1260. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11322970

Pham, D. L., & Prince, J. L. (1999). *An adaptive fuzzy segmentation algorithm for three-dimensional magnetic resonance images*. Paper presented at the Information Processing

in Medical Imaging. IPMI 1999. Lecture Notes in Computer Science, vol 1613, Berlin, Heidelberg.

Pichora-Fuller, M. K., & Levitt, H. (2012). Speech comprehension training and auditory and cognitive processing in older adults. *Am J Audiol, 21*(2), 351-357. doi:10.1044/1059-0889(2012/12-0025)

Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., . . . Trujillo, N. J. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiol Neurootol, 4*(2), 64-79. doi:10.1159/000013823

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr Clin Neurophysiol, 36*(2), 191-199. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/4129631

Picton, T. W., Hillyard, S. A., Galambos, R., & Schiff, M. (1971). Human auditory attention: a central or peripheral process? *Science, 173*(3994), 351-353. Retrieved from https://science.sciencemag.org/content/173/3994/351.long

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol, 36*(2), 179-190. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/4129630

Picton, T. W., van Roon, P., Armilio, M. L., Berg, P., Ille, N., & Scherg, M. (2000). The correction of ocular artifacts: a topographic perspective. *Clin Neurophysiol, 111*(1), 53-65. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10656511

Picton, T. W., Woods, D. L., Baribeau-Braun, J., & Healey, T. M. (1976). Evoked potential audiometry. *J Otolaryngol, 6*(2), 90-119. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1030745

Prevost, F., Laroche, M., Marcoux, A. M., & Dajani, H. R. (2013). Objective measurement of physiological signal-to-noise gain in the brainstem response to a synthetic vowel. *Clin Neurophysiol, 124*(1), 52-60. doi:10.1016/j.clinph.2012.05.009

Price, C. N., Alain, C., & Bidelman, G. M. (2019). Auditory-frontal channeling in alpha and beta bands is altered by age-related hearing loss and relates to speech perception in noise. *Neuroscience, 423*, 18-28. doi:10.1016/j.neuroscience.2019.10.044

Proskovec, A. L., Heinrichs-Graham, E., & Wilson, T. W. (2019). Load modulates the alpha and beta oscillatory dynamics serving verbal working memory. *Neuroimage, 184*, 256-265. doi:10.1016/j.neuroimage.2018.09.022

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci, 2*(1), 79-87. doi:10.1038/4580

Richards, J. E., Sanchez, C., Phillips-Meek, M., & Xie, W. (2016). A database of age-appropriate average MRI templates. *Neuroimage, 124*(Pt B), 1254-1259. doi:10.1016/j.neuroimage.2015.04.055

Rudner, M., Foo, C., Ronnberg, J., & Lunner, T. (2009). Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scand J Psychol, 50*(5), 405-418. doi:10.1111/j.1467-9450.2009.00745.x

Saiz-Alia, M., Forte, A. E., & Reichenbach, T. (2019). Individual differences in the attentional modulation of the human auditory brainstem response to speech inform on speech-in-noise deficits. *Sci Rep, 9*(1), 14131. doi:10.1038/s41598-019-50773-1

Salthouse, T. A., & Hedden, T. (2002). Interpreting reaction time measures in between-group comparisons. *J Clin Exp Neuropsychol, 24*(7), 858-872. doi:10.1076/jcen.24.7.858.8392

Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys.Med.Biol. , 32*, 11–22.

Scherg, M. (1990). Fundamentals of dipole source potential analysis. In: Auditory evoked magnetic fields and electric potentials. In F. Grandori, M. Hoke, & G. L. Romani (Eds.), *Advances in Audiology* (pp. 40-69). Basel: Karger.

Scherg, M., & Ebersole, J. S. (1994). Brain source imaging of focal and multifocal epileptiform EEG activity. *Neurophysiol Clin, 24*(1), 51-60. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8121338

Scherg, M., Ille, N., Bornfleth, H., & Berg, P. (2002). Advanced tools for digital EEG review: virtual source montages, whole-head mapping, correlation, and phase analysis. *J Clin Neurophysiol, 19*(2), 91-112. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11997721

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *J Cogn Neurosci, 1*(4), 336-355. doi:10.1162/jocn.1989.1.4.336

Schluchter, M. D., & Elashoff, J. D. (1990). Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures. *Journal of Statistical Computation and Simulation, 37*(69-87).

Schofield, B. (2010). Structural organization of the descending auditory pathway. In A. Rees & A. R. Palmer (Eds.), *Oxford Handbook of Auditory Science: The Auditory Brain* (pp. 43-64). New York: Oxford University Press.

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., . . . Griffiths, T. D. (2016). Neural signatures of perceptual inference. *Elife, 5*, e11476. doi:10.7554/eLife.11476

Shahin, A. J., Picton, T. W., & Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain Cogn, 70*(3), 259-266. doi:10.1016/j.bandc.2009.02.008

Skoe, E., & Kraus, N. (2010a). Auditory brain stem response to complex sounds: A tutorial. *Ear & Hearing, 31*(3).

Skoe, E., & Kraus, N. (2010b). Hearing it again and again: on-line subcortical plasticity in humans. *PLoS One, 5*(10), e13645. doi:10.1371/journal.pone.0013645

Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience, 243*, 104-114. doi:10.1016/j.neuroscience.2013.03.009

Smalt, C. J., Krishnan, A., Bidelman, G. M., Ananthakrishnan, S., & Gandour, J. T. (2012). Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hear Res, 292*(1-2), 26-34. doi:10.1016/j.heares.2012.08.001

Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology, 39*, 465-472.

Smith, J. C., Marsh, J. T., Greenberg, S., & Brown, W. S. (1978). Human auditory frequency-following responses to missing fundamental. *Science, 201*(4356), 639-641.

Sohmer, H., Pratt, H., & Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology, 42*, 656-664.

Song, J., Skoe, E., Banai, K., & Kraus, N. (2010). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*.

Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci, 20*(10), 1892-1902. doi:10.1162/jocn.2008.20131

Specht, K., Shah, N. J., & Jancke, L. (2000). Bilateral inferior frontal networks are involved in speech perception processes. *Neuroimage, 11*(5), S292.

Starr, A., & Golob, E. J. (2007). Cognitive factors modulating auditory cortical potentials. In R. F. Burkard, J. J. Eggermont, & M. Don (Eds.), *Auditory Evoked Potentials: Basic Principles and Clinical Application* (pp. 508-524). Baltimore, MD: Lippincott Williams & Wilkins.

Suga, N. (2008). Role of corticofugal feedback in hearing. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol, 194*(2), 169-183. doi:10.1007/s00359-007-0274-2

Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proc Natl Acad Sci U S A, 97*(22), 11807-11814. doi:10.1073/pnas.97.22.11807

Suga, N., & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat Rev Neurosci, 4*(10), 783-794. doi:10.1038/nrn1222

Suga, N., Ma, X., Gao, E., Sakai, M., & Chowdhury, S. A. (2003). Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. *Speech Communication, 41*(1), 189-200.

Sullivan, E. V., & Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neurosci Biobehav Rev, 30*(6), 749-761. doi:10.1016/j.neubiorev.2006.06.002

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system: an approach to cerebral imaging*. New York: Thieme Medical Publishers.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci, 16*(13), 4240-4249. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8753885

Tichko, P., & Skoe, E. (2017). Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. *Hear Res, 348*, 1-15. doi:10.1016/j.heares.2017.01.014

Tremblay, K. L., 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. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12842732

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb Cortex, 20*(2), 352-364. doi:10.1093/cercor/bhp105

Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron, 62*(4), 463-469. doi:10.1016/j.neuron.2009.05.002

Varghese, L., Bharadwaj, H. M., & Shinn-Cunningham, B. G. (2015). Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. *Brain Res, 1626*, 146-164. doi:10.1016/j.brainres.2015.06.038

von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int J Psychophysiol, 38*(3), 301-313. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11102669

Wang, L., Wang, W., Yan, T., Song, J., Yang, W., Wang, B., . . . Wu, J. (2017). Beta-Band Functional Connectivity Influences Audiovisual Integration in Older Age: An EEG Study. *Front Aging Neurosci, 9*, 239. doi:10.3389/fnagi.2017.00239

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev, 90*(3), 1195-1268. doi:10.1152/physrev.00035.2008

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res Rev, 23*(Pt B), 154-166. doi:10.1016/j.arr.2015.06.002

Weinstein, B. E., & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *J Speech Hear Res, 25*(4), 593-599. doi:10.1044/jshr.2504.593

Woldorff, M., Hansen, J. C., & Hillyard, S. A. (1987). Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. *Electroencephalogr Clin Neurophysiol Suppl, 40*, 146-154.

Wolters, C. H., Anwander, A., Berti, G., & Hartmann, U. (2007). Geometry-adapted hexahedral meshes improve accuracy of finite-element-method-based eeg source analysis. *IEEE Transactions on Biomedical Engineering, 54*(8), 1446-1453. doi:10.1109/TBME.2007.890736

Wong, P. C., Ettlinger, M., Sheppard, J. P., Gunasekera, G. M., & Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear Hear, 31*(4), 471-479. doi:10.1097/AUD.0b013e3181d709c2

Wong, P. C., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia, 47*(3), 693-703. doi:10.1016/j.neuropsychologia.2008.11.032

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci, 10*(4), 420-422. doi:10.1038/nn1872

Woods, D. L., & Hillyard, S. A. (1978). Attention at the cocktail party: Brainstem evoked responses reveal no peripheral gating. In D. Otto (Ed.), *Multidisciplinary Perspectives in Event-Related Brain Potential Research* Washington, D.C.: GPO.

Yellamsetty, A., & Bidelman, G. M. (2018). Low- and high-frequency cortical brain oscillations reflect dissociable mechanisms of concurrent speech segregation in noise. *Hear Res*. doi:10.1016/j.heares.2018.01.006

Yoo, J., & Bidelman, G. M. (2019). Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hear Res, 377*, 189-195. doi:10.1016/j.heares.2019.03.021

Zarahn, E., Rakitin, B., Abela, D., Flynn, J., & Stern, Y. (2007). Age-related changes in brain activation during a delayed item recognition task. *Neurobiol Aging, 28*(5), 784-798. doi:10.1016/j.neurobiolaging.2006.03.002

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage, 32*(4), 1826-1836. doi:10.1016/j.neuroimage.2006.04.199

Zendel, B. R., & Alain, C. (2014). Enhanced attention-dependent activity in the auditory cortex of older musicians. *Neurobiol Aging, 35*(1), 55-63. doi:10.1016/j.neurobiolaging.2013.06.022

Appendix A

Cortical ERP Electrode-Level Analysis

Sensor-level ERP peak response amplitudes and latencies were obtained from electrode Fz. Electrode selection was based on previous literature (Bidelman et al., 2013; Bidelman, Pousson, Dugas, & Fehrenbach, 2018; Picton et al., 1974) as well as topographic maps indicating maximum activation near and around Fz for essentially all experimental conditions. Noise affected the amplitudes of all ERP components decreasing the amplitudes of P1, P2, and the N1-P2 complex [P1: $F_{(1, 136)} = 63.23$, $p < 0.0001$; P2: $F_{(1, 136)} = 132.13$, $p < 0.0001$; N1-P2: $F_{(1, 136)} = 133.83$, $p < 0.0001$] but increasing N1 amplitude [$F_{(1, 136)} = 29.35$, $p < 0.0001$]. An interaction between SNR and attention [$F_{(1, 136)} = 6.45$, $p = 0.01$] revealed that P1 amplitude decreased less significantly in active listening conditions. Expectedly, attention increased N1 amplitude [$F_{(1, 136)} = 11.83$, $p < 0.001$] while decreasing P2 amplitude [$F_{(1, 136)} = 4.87$, $p = 0.03$] (Crowley & Colrain, 2004; Naatanen, 1975). For latency, SNR influenced only earlier components decreasing and increasing P1 and N1 latencies respectively [P1: $F_{(1, 136)} = 19.73$, $p < 0.0001$; N1: $F_{(1, 136)} = 41.83$, $p < 0.0001$]. Attention increased P2 latency [$F_{(1, 136)} = 5.26$, $p = 0.02$]. Vowel effects for N1 and P2 indicated increased latency for /i/ compared to /a/ [N1: $F_{(1, 136)} = 7.27$, $p = 0.01$; P2: $F_{(1, 136)} = 3.79$, $p = 0.05$]. No other significant effects or interactions were identified.

Appendix B
IRB Approval: Protocol #2370

Institutional Review Board
Division of Research and Innovation
Office of Research Compliance
University of Memphis
315 Admin Bldg
Memphis, TN 38152-3370

PI: Gavin Bidelman
Co-Investigator:
Advisor and/or Co-PI:
Department: IIS-Institute For Intelligent Systems
Study Title: Neural correlates of complex auditory perception
IRB ID: 2370
Submission Type: Renewal
Level of Review: Expedited

IRB Meeting Date:
Decision: Approved
Approval Date: April 15, 2020

Research Notes:
Findings:

The IRB has reviewed the renewal request. The University of Memphis Institutional Review Board, FWA00006815, has reviewed your submission in accordance with all applicable statuses and regulations as well as ethical principles.

Approval of this project is given with the following obligations:

1. If this IRB approval has an expiration date, an approved renewal must be in effect to continue the project prior to that date. If approval is not obtained, the human subjects consent form(s) and recruiting material(s) are no longer valid and any research activities involving human subjects must stop.
2. When the project is finished a completion form must be completed and sent to the board.
3. No change may be made in the approved protocol without prior board approval, whether the approved protocol was reviewed at the Exempt, Expedited or Full Board level.
4. Exempt approval are considered to have no expiration date and no further review is necessary unless the protocol needs modification.
5. Human subjects training is required every 2 years and is to be kept current at citiprogram.org.

Thank you,
James P. Whelan, Ph.D.
Institutional Review Board Chair

Appendix C
IRB Approval: PRO-FY2018-338

Institutional Review Board
Division of Research and Innovation
Office of Research Compliance
University of Memphis
315 Admin Bldg
Memphis, TN 38152-3370

September 27, 2019

PI Name: Gavin Bidelman
Co-Investigators:
Advisor and/or Co-PI:
Submission Type: Modification
Title: MRI studies of human auditory system
IRB ID : #PRO-FY2018-338
Level of Review: Full Board

Approval: September 27, 2019
Expiration: --*

The University of Memphis Institutional Review Board, FWA00006815, has reviewed your submission in accordance with all applicable statuses and regulations as well as ethical principles.

The modification is approved.

Approval of this project is given with the following obligations:

1. This IRB approval for modification has an expiration date, an approved renewal must be in effect to continue the project prior to that date. If approval is not obtained, the human subjects consent form(s) and recruiting material(s) are no longer valid and any research activities involving human subjects must stop.
2. When the project is finished a completion form must be submitted.
3. No change may be made in the approved protocol without prior board approval.
4. Human subjects training is required every 2 years and is to be kept current at citiprogram.org.

**Modifications do not extend the expiration of the original approval*

Thank you,
James P. Whelan, Ph.D.
Institutional Review Board Chair
The University of Memphis.