



OPEN ACCESS

EDITED BY

Sukhbinder Kumar,
The University of Iowa, United States

REVIEWED BY

Tobias Reichenbach,
University of Erlangen-Nuremberg, Germany
Sebastian Puschmann,
University of Oldenburg, Germany

Yi Du,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE

Gavin M. Bidelman
✉ gbidel@indiana.edu

SPECIALTY SECTION

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

RECEIVED 20 October 2022

ACCEPTED 17 January 2023

PUBLISHED 02 February 2023

CITATION

Lai J, Alain C and Bidelman GM (2023)
Cortical-brainstem interplay during speech
perception in older adults with and without
hearing loss.
Front. Neurosci. 17:1075368.
doi: 10.3389/fnins.2023.1075368

COPYRIGHT

© 2023 Lai, Alain and Bidelman. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Cortical-brainstem interplay during speech perception in older adults with and without hearing loss

Jesyin Lai^{1,2,3}, Claude Alain^{4,5} and Gavin M. Bidelman^{1,2,6,7*}

¹Institute for Intelligent Systems, University of Memphis, Memphis, TN, United States, ²School of Communication Sciences and Disorders, University of Memphis, Memphis, TN, United States, ³Department of Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, TN, United States, ⁴Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, ON, Canada, ⁵Department of Psychology, University of Toronto, Toronto, ON, Canada, ⁶Department of Speech, Language, and Hearing Sciences, Indiana University, Bloomington, IN, United States, ⁷Program in Neuroscience, Indiana University, Bloomington, IN, United States

Introduction: Real time modulation of brainstem frequency-following responses (FFRs) by online changes in cortical arousal state *via* the corticofugal (top-down) pathway has been demonstrated previously in young adults and is more prominent in the presence of background noise. FFRs during high cortical arousal states also have a stronger relationship with speech perception. Aging is associated with increased auditory brain responses, which might reflect degraded inhibitory processing within the peripheral and ascending pathways, or changes in attentional control regulation *via* descending auditory pathways. Here, we tested the hypothesis that online corticofugal interplay is impacted by age-related hearing loss.

Methods: We measured EEG in older adults with normal-hearing (NH) and mild to moderate hearing-loss (HL) while they performed speech identification tasks in different noise backgrounds. We measured α power to index online cortical arousal states during task engagement. Subsequently, we split brainstem speech-FFRs, on a trial-by-trial basis, according to fluctuations in concomitant cortical α power into low or high α FFRs to index cortical-brainstem modulation.

Results: We found cortical α power was smaller in the HL than the NH group. In NH listeners, α -FFRs modulation for clear speech (i.e., without noise) also resembled that previously observed in younger adults for speech in noise. Cortical-brainstem modulation was further diminished in HL older adults in the clear condition and by noise in NH older adults. Machine learning classification showed low α FFR frequency spectra yielded higher accuracy for classifying listeners' perceptual performance in both NH and HL participants. Moreover, low α FFRs decreased with increased hearing thresholds at 0.5–2 kHz for clear speech but noise generally reduced low α FFRs in the HL group.

Discussion: Collectively, our study reveals cortical arousal state actively shapes brainstem speech representations and provides a potential new mechanism for older listeners' difficulties perceiving speech in cocktail party-like listening situations in the form of a miss-coordination between cortical and subcortical levels of auditory processing.

KEYWORDS

age-related hearing loss, brainstem, frequency-following responses, arousal, cortical α , EEG

1. Introduction

Declines in auditory processing (Poth et al., 2001; Parthasarathy et al., 2010; Kortlang et al., 2016; Lai and Bartlett, 2018) and speech comprehension (Schneider et al., 2005; Peelle et al., 2010; Anderson et al., 2012)—especially in the presence of background noise (Dubno, 1984; Takahashi and Bacon, 1992; Souza et al., 2007; Anderson et al., 2011; Song et al., 2011; Jin et al., 2014; Presacco et al., 2016; Vermeire et al., 2016)—are ubiquitous during aging and age-related hearing loss. Age-related declines in the sensory (auditory) system (Parthasarathy and Bartlett, 2011, 2012; Fostick et al., 2013; Parthasarathy et al., 2014, 2016; Lai and Bartlett, 2015, 2018; Lai et al., 2017) [e.g., age-related impairments in sound source segregation (Alain et al., 1996; Alain and McDonald, 2007; Gallun and Best, 2020)], changes in cognitive function (Park et al., 2003), or a combination of both (Pichora-Fuller and Singh, 2006; Wingfield et al., 2005; Wayne and Johnsrude, 2015) could lead to listening and comprehension difficulties in elderly listeners.

Evidence suggests that speech-in-noise (SiN) problems could be related to dysfunctional connections and changes in speech processing between cortical and subcortical levels of the auditory system that emerge with age and age-related hearing impairment. It is well-established that SiN processing can be affected by many factors, such as attention (Saiz-Alfá et al., 2019; Price and Bidelman, 2021) and arousal state (Mai et al., 2019; Saderi et al., 2021). Several studies have shown neural correlates of these phenomena. For example, findings from EEG studies on emotion suggest that power in the cortical α band (8–12 Hz) is a useful indicator of arousal state (Aftanas et al., 2002; Uusberg et al., 2013). Moreover, parieto-occipital α power was shown to index cognitive processing, effortful listening (Wöstmann et al., 2015; McMahan et al., 2016; Dimitrijevic et al., 2017), the state of wakefulness (Pfurtscheller et al., 1996) and top-down processing (Henry et al., 2017). Alpha oscillatory activity has also been associated with adaptive, intentional, and top-down suppression of task-irrelevant information (Rihs et al., 2007; Jensen and Mazaheri, 2010; Händel et al., 2011; Klatt et al., 2020). Increased α power has been proposed to index inhibitory processing across sensory modalities (Klimesch et al., 2007; Weisz et al., 2007, 2011; Strauß et al., 2014), while decreased α oscillations are thought to facilitate sensory processing or neural firing (Haegens et al., 2011; Klatt et al., 2020). There is, however, no consensus regarding the mechanisms underlying α oscillations reported in these studies. In most studies α is treated as a unitary measure rather than reflecting different underlying processes. Meanwhile, evidence suggests that cortical α oscillations changes with aging (Yordanova et al., 1998; Böttger et al., 2002), such as a decrease in α frequency (Chiang et al., 2011) and reduced spontaneous entrainment of resting-state α oscillations (Gaál et al., 2010). Studying α power during SiN perception in older adults may reveal the impacts of aging in top-down attentional control that help facilitate the processing of target vs. non-target sounds, thus providing insight concerning why cocktail party-like situations are more difficult in older listeners (Pichora-Fuller et al., 2017).

In addition to cortical changes, age-related declines in speech coding have been widely observed at subcortical levels of the auditory system, both in terms of local processing within the brainstem but also its functional signaling to and from the cortex (Bidelman et al., 2019). In young adults, we recently observed that speech-evoked brainstem frequency-following response (FFR) amplitude varied as

a function of α power (Lai et al., 2022). Low FFR amplitude coincided with low α power whereas high FFR amplitude was associated with high α states. Notably, low α FFRs correlated positively with response times (RTs) for speech discrimination and more accurately decoded the input speech stimuli revealed by neural classifiers. Extending this approach to address questions of auditory aging, we analyzed neuroelectric FFRs recorded during active speech perception in age-matched older adults with normal (NH) or mild hearing loss (HL). This allowed us to investigate the effects of age-related hearing loss on cortical α state and its modulation of brainstem speech processing in real time. We aimed to determine the nature of auditory cortical-brainstem interplay in older adults, and more critically, whether such online corticofugal engagement during SiN listening is altered due to hearing loss, as suggested in prior work (Bidelman et al., 2019). Our results reveal that brainstem speech-FFRs are dynamically modulated by fluctuations in cortical α state in normal-hearing listeners but this cortical-subcortical interplay declines in age-related hearing loss.

2. Materials and methods

2.1. Participants

Detailed information on participants, informed consent, and demographics are reported in our original report detailing age-related changes in the brainstem and cortical evoked potentials (Bidelman et al., 2019). New analyses herein examine online changes in FFRs as a function of the simultaneous cortical state. All participants had no reported history of neurological or psychiatric illness. Participants were aged between 52 and 75 (69 ± 5.8 years; 16/16 M/F). There were divided into normal (NH) and hearing-impaired (HL) groups based on their pure-tone audiometry hearing thresholds. We used 25 dB HL as the cutoff to define normal hearing per standard clinical conventions (Gatlin and Dhar, 2021). NH listeners ($n = 13$) had average thresholds (250–8,000 Hz) better than 25 dB HL across both ears whereas HL listeners ($n = 19$) had average thresholds poorer than 25 dB HL. The pure-tone averages (PTAs) (i.e., mean of 500, 1,000, 2,000 Hz) of NH listeners were ~ 10 dB better than in HL listeners (mean \pm SD; NH: 15.3 ± 3.27 dB HL, HL: 26.4 ± 7.1 dB HL; $t_{2,71} = -5.95$, $p < 0.0001$; NH range = 8.3–20.83 dB HL, HL range = 15.8–45 dB HL). Both NH ($t_{12} = 0.15$, $p = 0.89$) and HL ($t_{18} = -2.02$, $p = 0.06$) groups otherwise had symmetric PTA between ears. Both NH and HL groups had elevated hearing thresholds at very high frequencies ($\geq 8,000$ Hz), typical of age-related presbycusis in older adults. Besides hearing, the two groups were matched in age (NH: 66.2 ± 6.1 years, HL: 70.4 ± 4.9 years; $t_{2,22} = -2.05$, $p = 0.052$) and sex balance (NH: 5/8 M/F; HL: 11/8; Fisher's exact test, $p = 0.47$). Age and hearing loss were not correlated (Pearson's $r = 0.29$, $p = 0.10$), suggesting these aging factors were largely independent in our sample.

2.2. QuickSiN test

The Quick Speech-in-Noise (QuickSiN) test was measured listeners' speech reception thresholds in noise (Killion et al., 2004). A list of six sentences with five keywords per sentence spoken by a female talker in a background of four-talker babble noise was heard by listeners during the test. Target sentences were presented at

70 dB sound pressure level (SPL) (binaurally) at signal-to-noise ratios (SNRs) decreasing from 25 dB (relatively easy) to 0 dB (somewhat difficult) in 5 dB steps. The number of keywords correctly recalled was logged, and a score was computed for each listener. The SNR-loss score indexes the difference between a listener's SNR-50 (i.e., the SNR required to identify 50% of the keywords correctly) and the average SNR threshold for normal-hearing adults (i.e., 2 dB) (Killion et al., 2004). A higher score reflects poorer SiN perception. Each listener's SNR-loss score was averaged from four lists of sentence presentations. In this study, NH listeners' scores ranged from -0.25 to 2.5 dB of SNR-loss ($M = 1.1$ dB, $SD = 0.8$ dB) while HL listeners' scores were higher and more variable, ranging from -2.5 to 8.5 dB of SNR-loss ($M = 2$ dB, $SD = 2.5$ dB) [see Figure 1D in Bidelman et al. (2019)]. Indeed, although mean QuickSiN scores did not differ between groups [$t_{2,35} = -1.43$, $p = 0.16$], HL listeners showed more inter-subject variability compared to NH listeners [Equal variance test (two-sample F -test): $F_{18,12} = 8.81$, $p = 0.0004$].

2.3. EEG stimuli and task

The stimuli and task are described fully in Bidelman et al. (2019) and illustrated here in Figure 1A. Three naturally produced English consonant-vowel phonemes (/ba/, /pa/, and /ta/), from the standardized UCLA version of the Nonsense Syllable Test (Dubno and Schaefer, 1992), were generated by a female talker. The duration of each phoneme was 100 ms and the average root mean square SPL of each phoneme matched. All three tokens had a common voice fundamental frequency (mean $F_0 = 150$ Hz), first and second formants ($F_1 = 885$, $F_2 = 1,389$ Hz). The resulting stimulus-evoked response (i.e., FFR) predominantly originates from the subcortex (Brugge et al., 2009; Bidelman, 2018b) since the stimulus F_0 is above the phase-locking limit of the cortical neurons and "cortical FFRs" (Coffey et al., 2016; Bidelman, 2018b; Bidelman and Momtaz, 2021). Indeed, we have shown that these types of speech tokens with an $F_0 = 150$ Hz elicit robust midbrain FFRs with no evidence of a cortical contribution [see Supplementary Figure 1 in Price and Bidelman (2021)]. Speech tokens were delivered binaurally to listeners in either clear (i.e., no noise) or noise-degraded conditions. A complete set of stimuli presented in each condition contained 3,000 /ba/, 3,000 /pa/, and 210 /ta/ tokens (spread evenly over three blocks to allow for breaks). The interstimulus interval between tokens was randomly jittered within 95–155 ms (5 ms steps, uniform distribution). The /ba/ and /pa/ tokens were presented more frequently than the /ta/ token in a pseudo-random manner such that at least two frequent tokens intervened between infrequent tokens. The rare /ta/ token was denoted as the target in which listeners were required to respond by pressing a button on the computer whenever they detected it. Both reaction time (RT) and detection accuracy (%) were recorded. For the noise-degraded condition, the same procedures as the clear condition were repeated, but the tokens were presented in an identical speech triplet mixed with eight talker noise babble (Killion et al., 2004) at a SNR of 10 dB. Six blocks (3 clear and 3 noise) were collected from each participant. Having the clear and noise conditions allowed us to compare behavioral performance in different backgrounds and evaluate the impact of noise on speech perception in NH vs. HL listeners, respectively. The task ensured listeners were actively engaged during speech perception and online EEG recording. Stimuli were controlled by a MATLAB program (The Mathworks, Inc., Natick, MA, USA) routed to a TDT RP2 interface (Tucker-Davis

Technologies; Alachua, FL, USA) and delivered binaurally through insert earphones (ER-3; Etymotic Research; Elk Grove Village, IL, USA). The speech stimuli were presented at 75 dB SPL (noise at 65 dB SPL) with alternating polarity.

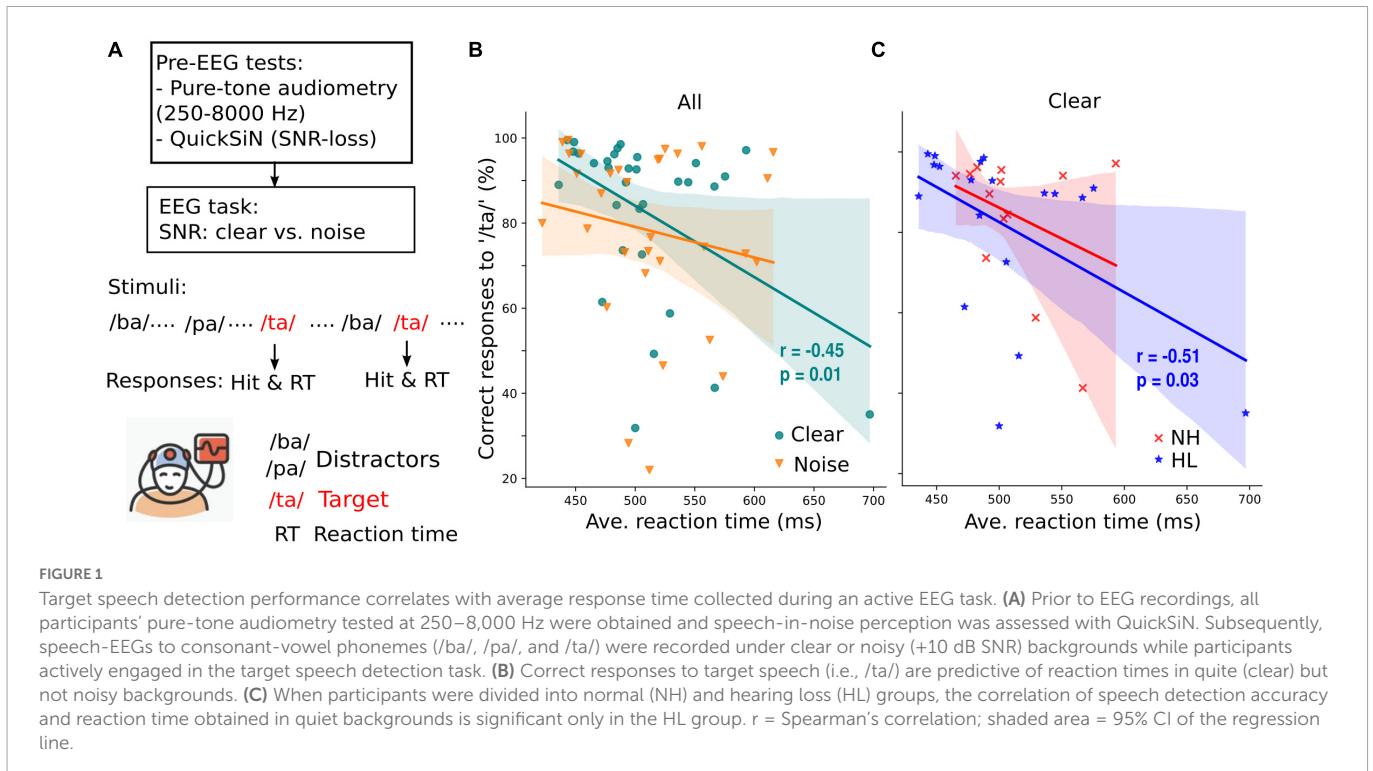
2.4. EEG recording and preprocessing

During the target speech detection task, neuroelectric activity was recorded from 32 channels at standard 10–20 electrode locations on the scalp (Oostenveld and Praamstra, 2001). Electrode impedances were ≤ 5 k Ω . EEGs were digitized at 20 kHz using SynAmps RT amplifiers (Compumedics Neuroscan; Charlotte, NC, USA). After EEG acquisition, the data were processed using the "mne" package in Python 3.9.7. EEG data were re-referenced offline to the mastoids (TP9/10) for sensor (channel-level) analyses. For source analysis of brainstem FFRs, we used a common average reference before source transformation (detailed below). Responses were then filtered 100–1,000 Hz [finite impulse response (FIR) filters; hamming window with 0.02 dB passband ripple, 53 dB stopband attenuation, -6 dB cutoff frequency, filter length = 661 samples/ 0.132 s] to isolate brainstem activity (Musacchia et al., 2008; Bidelman et al., 2013).

2.5. Derivation of source FFRs and cortical activities

The derivation of source FFR waveforms and isolation of cortical activities are similar to the methods described in Lai et al. (2022) for young adults. The 32-channel sensor data were transformed into source space using a virtual source montage. The source montage comprised of a single regional source (i.e., current flow in x , y , z planes) positioned in the brainstem and midbrain (i.e., inferior colliculus) [details refer to Bidelman (2018b), Bidelman and Momtaz (2021), Price and Bidelman (2021)]. Source current waveforms (SWF) from the brainstem source were obtained using the formula: $SWF = L^{-1} \times FFR$, where L is the brainstem source leadfield matrix (size 3×64) and FFR is the 32-ch sensor data ($64 \times$ NSamples). This essentially applied a spatial filter to all electrodes that calculated their weighted contribution to the scalp-recorded FFRs to estimate source activity within the midbrain in the x , y , and z directions (Scherg and Ebersole, 1994; Scherg et al., 2002). This model explains $\sim 90\%$ of the scalp-recorded FFR (Bidelman et al., 2019; Price and Bidelman, 2021). Only the z -oriented SWF was used for further analysis (x and y SWFs were not analyzed) given the predominantly vertical orientation of current flow in the auditory midbrain pathways relative to the scalp [x - and y -orientations contribute little to the FFR (Bidelman, 2018b)].

We isolated cortical α band activity from the EEG and used it as a running index of arousal state (high or low) during the target speech detection task. EEG at the Pz and Oz channels were filtered at 8–12 Hz (FIR filters, -6 dB cutoff frequency at 7 Hz and 13.5 Hz, filter length = 8,251 samples/1.65 s) and averaged (i.e., equivalent to POz) to obtain cortical α -band activity at a posterior scalp region. Filtered α activities were epoched with a time window of 195 ms (-50 to 145 ms in which 0 ms corresponded to the onset of a /ba/ or /pa/ token) to capture approximately 1–2 cycles of α band. This epoch window



encapsulated the entirety of the evoked FFR within the immediate trial with no spillover from the preceding or subsequent trial(s). Infrequent /ta/ tokens were excluded from the analysis due to their limited number. The root mean square (RMS) amplitude of single trial α activity was computed to quantify cortical arousal level over the duration of the target speech detection task. We then normalized RMS values to each run's median of RMS values, respectively. Next, the trial-wise normalized α RMS distribution was visualized using a histogram. We categorized trials of each participant per condition falling within the 0–35th percentile as “low α ” power and those falling within the 65–100th percentile as “high α ” power. This categorization was used because it provided $\sim 2,100$ trials for each low or high α power in each participant per condition, which is reasonable to obtain an average FFR with a robust SNR (i.e., ≥ 3 dB SNR) (Bidelman, 2018a). More detailed information on this methodology can be found in Lai et al. (2022) (see their Figure 2). We similarly measured cortical activity in another frequency band (e.g., β band; 18–22 Hz) from the same location (i.e., POz β) and α band from a different electrode site (i.e., Fz α). These control analyses allowed us to ensure that the observed changes in speech-evoked FFRs were specifically associated with cortical arousal level (indexed by α power) rather than general fluctuations in the EEG, *per se*.

2.6. Analysis of brainstem FFRs

We categorized source FFRs based on whether α amplitude in the same epoch was either high or low power, thus deriving FFRs according to the trial-by-trial cortical state. Source FFR waveforms (from the z-orientated dipole) were averaged for each α category and noise condition per participant. Subsequently, we analyzed the steady-state portion (10–100 ms) of FFR waveforms using the FFT (Blackman window; 11.1 Hz frequency resolution) to capture the spectral composition of the response. F0 amplitude was measured

as the peak spectral maximum within an 11 Hz bin centered around 150 Hz (i.e., F0 of the stimuli). To compare FFR F0 amplitudes during low vs. high α power, a normalized (within-subject) F0 ratio was calculated as follows:

$$F0 \text{ ratio} = \frac{F0amp_{high\alpha}}{F0amp_{low\alpha}} \quad (1)$$

where F0 ratios > 1 indicate stronger brainstem FFRs during states of high cortical α power and F0 ratios < 1 indicate stronger FFRs during states of low cortical α power (Lai et al., 2022).

2.7. Statistical analysis

We used mixed-model ANOVAs to compare brainstem F0 ratios among the clear vs. noise condition, and NH vs. HL group. Multiple pairwise comparisons (Mann-Whitney U test with Bonferroni corrections) between the NH and HL groups were performed using the “pingouin” package in Python. One sample t -tests (“scipy” package in Python) were also used to evaluate whether FFR F0 ratios were differed significantly from 1 (and thereby indicated the significance of α modulation). Wilcoxon signed-rank test was used when comparing raw F0 amplitudes at low vs. high α power within participants for each SNR condition and hearing group. To compare differences in α RMS values of all participants across (clear vs. noise) conditions, a non-parametric test was required because α RMS values were not normally distributed. We performed posthocConover's test (“scikit_posthocs” package in Python), which is a non-parametric pairwise test, with Bonferroni adjustment. To assess differences in raw F0 amplitudes (log-transformed) across factors for NH and HL groups, we first performed a $2 \times 2 \times 2$ (SNR \times α power \times hearing group) mixed model (participants = random factor) ANOVA (“lme4” package in Rstudio). Following a significant

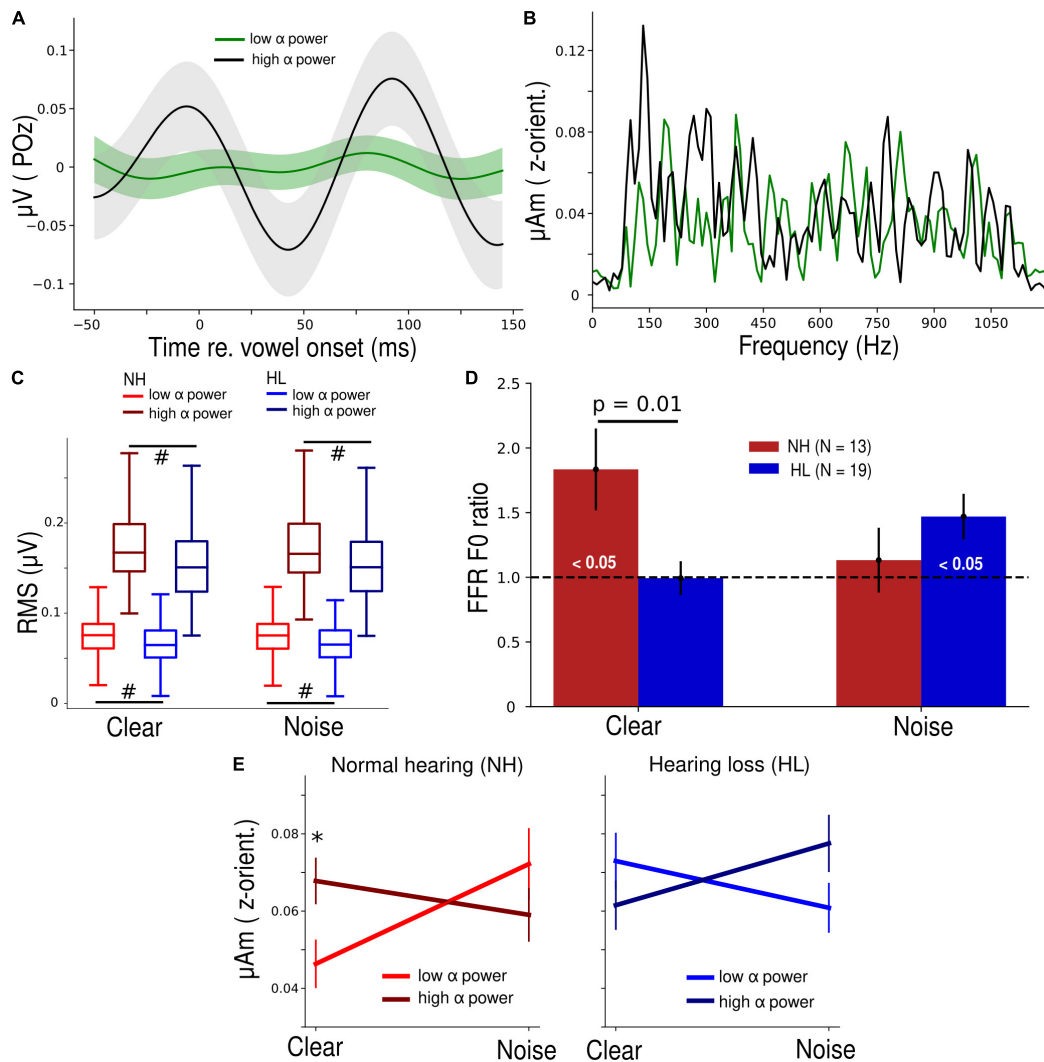


FIGURE 2

Normal hearing (NH) participants have a higher parieto-occipital α power and their brainstem speech processing is more strongly modulated by α band in clear backgrounds. (A) Average α waveform of low- and high-power trials plotted from a representative NH participant. (B) Frequency spectra of the steady state (10–100 ms) portion of low and high α brainstem FFRs from the same participant panel (A). Note the distinct response at ~ 150 Hz, corresponding to the voice pitch. (C) Root-mean-square (RMS) values of both low and high α of the NH group were significantly higher than the HL group. # $p < 0.01$ (Conover's test, non-parametric pairwise test, with Bonferroni adjustment). (D) FFR F0 ratios during low and high α trials. FFR F0 ratios were higher in the NH vs. HL group (Mann-Whitney U test) in the clear condition. Bars marked (< 0.05) are significantly larger than 1 (1-sample t -test) indicating enhancement of the FFR with changes in cortical α . (E) Grand average FFR F0 amplitudes as a function of SNR (clear vs. noise), α power (low vs. high), and group (NH vs. HL). Error bars = \pm s.e.m., * $p < 0.05$ (Wilcoxon signed-rank test).

interaction, we ran separate 2×2 (α power \times SNR) ANOVAs with random effects (subjects were considered randomly selected from a larger population) for the NH and HL groups, respectively. Initial diagnostics were performed using residual and Q-Q plots to assess the heteroscedasticity and normality of data. F0 amplitudes were log-transformed to improve normality and homogeneity of variance assumptions. Effect sizes are reported as η_p^2 . To check if low or high α FFR F0 amplitudes were associated with behavioral performance (i.e., QuickSiN, PTA, percent correct, and RTs) by pooling all NH and HL participants, we performed Spearman's correlations ("scipy" package in Python) to test their pairwise correlations of brain and behavior measures. Spearman's correlation was used because these measures were found to be not normally distributed ($p < 0.05$) from the test of normality using the Shapiro-Wilk test.

2.8. Classification of performance level from FFR frequency spectra via machine learning

All participants' performance in the clear condition was categorized into three levels (poor, average, and good) based on their percent correct of /ta/ detections. Participants with poor performance had percent correct scores \leq 30th-percentile while good-performance participants had scores \geq 70th-percentile. Subsequently, we classified poor- and good-performing participants using frequency spectra of their low or high α FFRs and a support vector machine (SVM) classifier (kernel = radius basis function, $C = 1,000$, gamma = scale) in the "scikit-learn" package in python. Due to the limitation in sample size (32 total observations), we were only able to perform a two-group rather than three-group

machine learning (ML) classification. There were ten participants in each poor- or good-performance category, which provided a total of twenty participants' frequency spectra to be used as data input for SVM. Frequency spectra were obtained from the FFT of average FFR waveforms (10 to 100 ms steady state portion) across ~2,100 trials of low or high α FFRs per participant. The absolute amplitudes of frequency spectra, which consisted of 226 amplitude-by-frequency points, were the input features; performance level (i.e., poor vs. good) served as the ground truth class labels. The ML classification procedures on FFRs were similar to those described by Xie et al. (2019) and Lai et al. (2022). During one iteration of training and testing, a four-fold cross-validation approach was used to train and evaluate the performance of the SVM classifier to obtain a mean classification accuracy [Figure 1 of Xie et al. (2019)]. In this process, a 4-fold stratification was performed to randomly and equally divide participants into 4 subgroups with 5 unique participants (almost similar number of poor- and good-performing participants) in each subgroup. Three of the 4 subgroups were selected as the training data while the remaining subgroup was used as the hold-out testing data. To mitigate the problem of imbalanced participant numbers in the two classes of the training set, we randomly over-sampled the minority class using the "imblearn" package in python. These steps were repeated within each iteration so that each subgroup was held-out as the test data whereas the other three subgroups were used to train the SVM classifier. Mean classification accuracy of poor- vs. good-performance was calculated across cross-validated iterations. We performed a total of $N = 5,000$ iterations to examine group classification for low vs. high α power FFRs.

To evaluate if the classifier accuracy (mean of $N = 5,000$ iterations) was statistically significant, we randomly shuffled the 226 data points of frequency spectra in each participant, and the same training and testing procedures described above were repeated to derive a null distribution of classification accuracies. We then calculated the p -value to determine the statistical significance of "true" classifier performance using the formula described in Phipson and Smyth (2010): $p = (a + 1)/(n + 1)$, where a is the number of classification accuracies from the null distribution that exceeds the median of the actual distribution of classification accuracies and n is the total number of classification accuracies from the null distribution.

2.9. Fitting linear regression models with brain and behavior measures

To compare the changes in low α FFR F0 amplitudes as behavior performance changed in the NH and HL groups, we fitted linear regression models for pooled SNR, clear and noise conditions, respectively, by using behavior performance and the group as main factors, their interaction factor, and without or with age as a covariate.

$$Y = \beta_0 + \beta_1 \cdot X + \beta_2 \cdot \text{group} + \beta_3 \cdot X \cdot \text{group} \quad (2)$$

$$Y = \beta_0 + \beta_1 \cdot X + \beta_2 \cdot \text{group} + \beta_3 \cdot X \cdot \text{group} + \beta_4 \cdot \text{age} \quad (3)$$

Y represents FFR F0 amplitudes during low α power, X is one of the behavior measures (QuickSiN, PTA, percent correct, or RTs), and the *group* is the dummy variable for NH ($\text{group} = 0$) and HL

($\text{group} = 1$). Meanwhile, β_0 is a constant and $\beta_1, \beta_2, \beta_3$ or β_4 is the coefficient or slope for the respective variable. Age was added as a covariate in Eq. (3) to confirm that the observed effects were not driven by participants' age, *per se*. After fitting regression models, we checked assumptions for the normality of residuals, homoscedasticity of residuals and linearity of the models. For residual normality, we tested if model residuals were normally distributed using the Shapiro-Wilk test and visualized the Q-Q plots of the residuals. For residual homoscedasticity, we used Breusch-Pagan and Goldfeld-Quandt tests. For model linearity, we visualized residual distributions by plotting residuals vs. predicted values. Result interpretation was then carried out after confirming the above three assumptions.

The fitted linear regression models allowed us to study if the slopes (i.e., the change of FFR amplitudes) of the NH group were significant as well as if the slopes of the HL group were significantly different from the NH group. In the NH group, i.e., $\text{group} = 0$, Eq. (2) can be written as $Y = \beta_0 + \beta_1 \cdot X$ and β_1 is the slope for the NH group. In contrast, in the HL group, i.e., $\text{group} = 1$, Eq. (2) can be written as $Y = \beta_0 + \beta_1 \cdot X + \beta_2 + \beta_3 \cdot X = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)X$, where $(\beta_1 + \beta_3)$ is the slope for the HL group and β_3 represents the difference in slope between the HL and NH group.

3. Results

3.1. Behavior performance of target speech detection

Behavioral responses during the EEG task (percent correct /ta/ detections vs. RTs) showed a negative correlation for the clear but not noise condition (Spearman's $r = -0.45, p = 0.01$, Figure 1B); participants with slower response speeds showed poorer speech detection accuracies. This is consistent with previous findings showing negative associations between hit responses and RTs in younger listeners (e.g., Lai et al., 2022). When separated into the NH and HL groups, we found a negative relationship between behavioral hit responses and decision speeds but only in the HL group (Figure 1C).

3.2. Cortical α band and brainstem speech-FFRs

Differences in cortical α -band amplitudes during low vs. high α states were prominent at the single participant (Figure 2A) as well as group level (Figure 2C). Spectral differences in the corresponding brainstem FFR for these same low vs. high cortical trials were also notable (Figure 2B). Cortical α (both low and high levels) was overall higher in the NH listeners ($p < 0.01$, non-parametric *post hoc* Conover's test with Bonferroni adjustment) but both groups showed clear separability of "low" vs. "high" α states during the speech detection task.

In response to clear speech (Figure 2D), brainstem F0 ratios (indexing cortical α -related FFR enhancement) in the NH group were significantly higher than 1 ($t_{12} = 2.64, p = 0.02$, 1-sample t -test) and higher than the F0 ratios in the HL group ($U = 192, p = 0.01$, Mann-Whitney U test). In response to the noise-degraded speech, this cortical-FFR enhancement was observed in the HL group ($t_{18} = 2.67, p = 0.02$, 1-sample t -test), probably due to compensation reasons, but

it was not significantly different from the FFR enhancement observed in the NH group. Repeating the same analysis for both controls (POz β and Fz α) revealed no difference ($p = 0.19$ and 0.053) in F0 ratios for the NH vs. HL group (Supplementary Figure 1), indicating cortical modulation of the FFR was restricted to POz α .

A 3-way mixed-model ANOVA performed on log F0 amplitudes revealed a significant SNR \times α power \times group interaction ($F_{1,96} = 9.5$, $p = 0.003$, $\eta_p^2 = 0.09$) (Figure 2E). To make sense of this complex interaction, we performed separate 2-way (SNR \times α power) mixed-model ANOVAs by hearing group. The SNR \times α power interaction was significant in both the NH ($F_{1,39} = 5.17$, $p = 0.03$, $\eta_p^2 = 0.12$) as well as the HL group ($F_{1,57} = 4.11$, $p = 0.05$, $\eta_p^2 = 0.07$). Though the comparison of effect sizes suggests this interaction was stronger in NH listeners, the interaction was distinct in direction compared to the HL group. In the NH group, FFR F0 amplitudes were significantly higher ($W = 17$, $p = 0.048$, Wilcoxon signed-rank test) during high α power for clear speech. This pattern was dampened and reversed in the HL group.

3.3. Brain-behavior relations in both the NH and HL groups

We next assessed associations between low or high α FFR F0 amplitudes and behavioral performance (QuickSiN, PTA, percent correct, and RTs) by performing Spearman's correlation analyses. The positive association between F0 amplitudes and percent correct scores was significant at low but not high α power (Spearman's $r = 0.3$, $p = 0.02$, Figure 3A), while the associations between F0 amplitudes (either during low or high α power) and other behavior measures were not significant (all p -values > 0.05). To further assess if FFRs during low or high α power in either clear or noise condition were more predictive of behavioral responses, we performed ML classification of percent correct scores (i.e., perceptual performance level) into poor (≤ 30 th of overall percent correct scores) or good (≥ 70 th of overall percent correct scores) using participants' FFR frequency spectra as input for the SVM classifiers. In the clear condition (Figure 3B), the classifier had a higher accuracy median (58%) in decoding participants' perceptual performance level using low α FFRs compared to the accuracy median (32%) when using high α FFRs. In contrast, in the noise condition (Figure 3C), classification accuracies did not differ between low or high α FFRs nor were they above the null distribution. These results provide evidence that adding background noise disrupted the relationship of low α FFRs with behavioral measures potentially as a result of compromising the SNR of the neural responses.

3.4. Comparison of brain-behavior relations in the NH and HL groups

The aforementioned analyses showed that during high arousal states (i.e., with low α power), FFRs have a stronger relationship with behavior compared to low arousal states, especially in the clear condition, when pooling the NH and HL groups. Hence, we studied the changes in low α FFR amplitudes with behavioral performance (QuickSiN, PTA, percent correct, and RTs) systematically by fitting linear regression models [Eq. (2, 3)]. We observed significant coefficients or slopes (β) when PTA was used as X in Eq. (2). For both the clear and noise-degraded speech, we found a significant

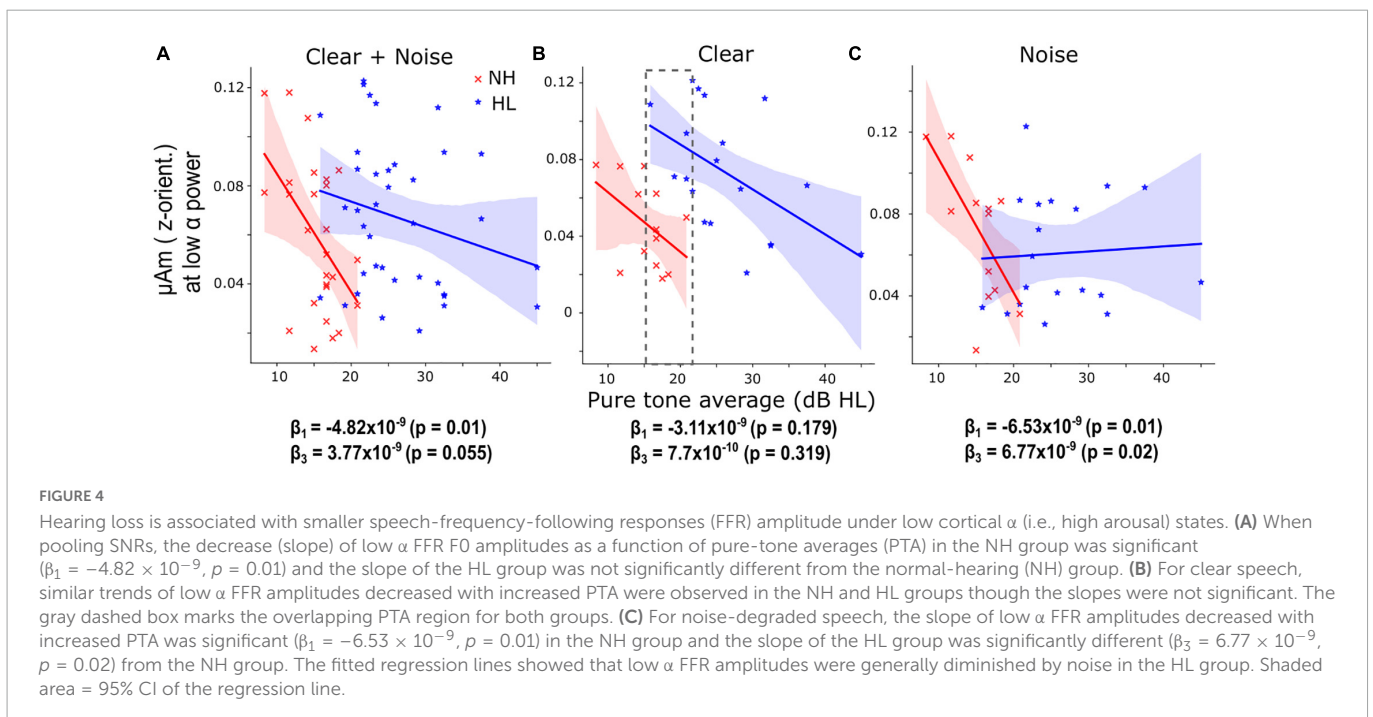
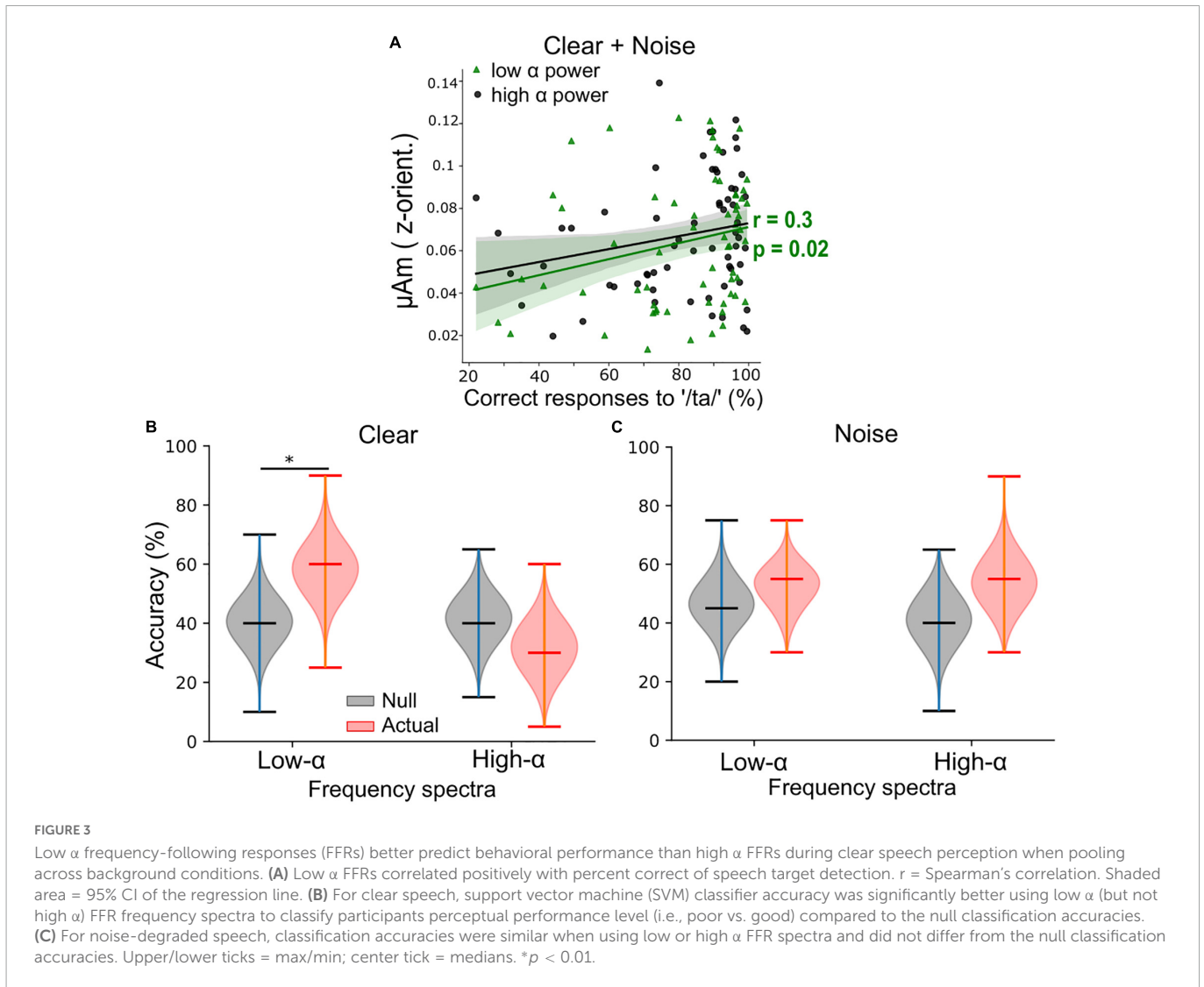
negative slope between low α F0 amplitudes and PTAs in the NH group ($\beta_1 = -4.82 \times 10^{-9}$, $p = 0.01$, Figure 4A) but the slope of the HL group was not significantly different from the NH group ($\beta_3 = 3.77 \times 10^{-9}$, $p = 0.055$). This indicates that even with clinically "normal" hearing, participants with slightly poorer thresholds have smaller FFRs during low α states. When separating the data by SNR, we observed similar trends of low α FFR amplitudes decreased with increased PTA in the NH and HL groups for clear speech (Figure 4B) though the slopes were not significant ($\beta_1 = -3.11 \times 10^{-9}$, $p = 0.179$; $\beta_3 = 7.7 \times 10^{-10}$, $p = 0.319$). For noisy speech (Figure 4C), we found that low α FFR amplitudes decreased significantly with increased PTA in the NH group ($\beta_1 = -6.53 \times 10^{-9}$, $p = 0.01$) and the slope of the HL group was also significantly different ($\beta_3 = 6.77 \times 10^{-9}$, $p = 0.02$) from the NH group. The fitted regression lines in Figure 4C showed that low α FFR amplitudes were generally diminished by noise in the HL group. Repeating the same analysis with age as covariate in Eq. (3) provided similar results. For both the clear and noise-degraded speech, β_1 was significant (-4.58×10^{-9} , $p = 0.014$) but not β_3 (3.45×10^{-9} , $p = 0.082$). For clear speech, none was significant ($\beta_1 = -2.83 \times 10^{-9}$, $p = 0.224$; $\beta_3 = 3.91 \times 10^{-10}$, $p = 0.874$). For noisy speech, both β_1 (-6.34×10^{-9} , $p = 0.019$) and β_3 (6.52×10^{-9} , $p = 0.024$) were significant. This helps rule out age as a confounding factor and suggests that hearing loss drove the observed group differences.

4. Discussion

Previous neuroimaging work reveals weaker functional connectivity between the brainstem and cortex in older listeners with mild hearing loss compared to older adults with normal hearing for their age, and this interplay robustly predicts their SiN perceptual performance (Bidelman et al., 2019). Adding to these findings, we show the existence of active and dynamic modulation of brainstem speech processing in NH older listeners, which was dependent on online changes in listeners' cortical state. This active and dynamic cortical-brainstem modulation, however, is diminished when processing speech in noise and in older adults with HL. Compared to NH listeners, HL listeners showed weaker parieto-occipital α power but those with minimal hearing loss (i.e., smaller PTA) had unusually large FFRs during low α states (gray dashed box in Figure 4B). Although FFRs were smaller during low α power, they were predictive of perceptual speech measures (Figure 3A) when pooling NH and HL participants and especially for clear speech (Figure 3B). Collectively, our findings suggest that (i) increased F0 ratios were disrupted by noise in the NH group and diminished in the HL group for clear speech; (ii) FFRs during low α power (i.e., high cortical arousal states) have smaller F0 amplitudes but their spectra are more predictive of behavioral performance, (iii) decreased low α FFRs with increased PTA in both NH and HL participants, and (iv) increased low α FFRs in older adults with mild hearing loss for clear speech suggesting an increase in central gain.

4.1. Effects of age on cortical α power and cortical modulation of brainstem speech processing

Cortical α indexes states of wakefulness and arousal (Pfurtscheller et al., 1996; Aftanas et al., 2002; Uusberg et al., 2013). Still, there is also



evidence showing that α power may vary or index mind wandering during cognitive tasks (Compton et al., 2019; Maillet et al., 2020). Calculating the span length of low or high α trials for each listener showed averages of ~ 1.7 trials in both hearing groups, equating to several hundred milliseconds during our task. The relative speed of these fluctuations suggests that the α -modulations observed here are unlikely related to mind wandering *per se*, which presumably develops over longer time courses [tens of seconds (Pelagatti et al., 2020)]. Instead, we infer low α power tracks a high arousal state while high α power reflects task focus but in a state of wakeful relaxation. Induced α activity is crucial for SiN perception as it might suppresses irrelevant information like noise to aid target speech processing (Strauß et al., 2014). In our previous study conducted on younger listeners (18–35 years) using similar EEG tasks, we observed larger α power to noise-degraded compared to a clear speech during the active engagement [see Figure 2F in Lai et al. (2022)]. However, here in both NH and HL older adults, we do not find this same noise-related α effect. In general, α power was larger in NH than in HL older listeners (Figure 2C). We also observed that high α RMS values in young NH listeners (Figure 2F in Lai et al., 2022) had a larger range (0.1–0.9 μ V) than NH older adults (0.1–0.28 μ V). This observation advocates a reduction of α power with aging which has also been shown in other studies (Babiloni et al., 2006; Purdon et al., 2015). Decreased α activity is also related to declines in cognitive functions with increasing age (Klimesch, 1997, 1999).

More critically, we demonstrate the presence of dynamic and online modulation of brainstem speech encoding by fluctuations in cortical α state in older NH adults. When comparing the observations across ages at similar acoustic backgrounds, they are fundamentally different from those observed in younger, normal-hearing listeners (cf. Lai et al., 2022). In younger listeners, lower cortical α states positively correlate with smaller FFRs during SiN perception (Lai et al., 2022). Furthermore, low- α -indexed FFRs recorded in noisy backgrounds are predictive of behavioral RTs for rapid speech detection and have higher accuracies in token classification (Lai et al., 2022). Here, unlike younger listeners which require more difficult perceptual tasks (i.e., SiN perception) to tax the system and reveal effects of cortical arousal state on brainstem FFRs, we observed cortical modulation of FFRs in NH older adults during the perception of clear speech (Figures 2D, E). Moreover, low- α -indexed FFRs associated more with behavior [speech detection (Figure 3) and PTAs (Figure 4)] than high- α -indexed FFRs. Contrastively, in noise, low α and high α FFR amplitudes were not classifiable in terms of perceptual performance level. Taken together, the pattern of cortical-brainstem interactions in speech processing we found here in older NH listeners under clear backgrounds appears similar to what is found in younger listeners under challenging listening environments (cf. Lai et al., 2022). This indicates that aging might alter cross-talk between functional levels of the auditory system under challenging listening conditions as a means of compensatory processing. Similar maladaptive plasticity has been previously observed at higher cortical levels where frontal brain regions are more strongly engaged to aid auditory-sensory coding in the superior temporal gyrus (Price et al., 2019). This further suggests the presence of age-related deficits in top-down modulation of brainstem speech processing by the cortex and provides an explanation for why older listeners find it more exhausting to participate in cocktail party-like listening situations compared to younger listeners (Pichora-Fuller et al., 2017).

4.2. Effects of hearing loss on cortical α power and cortical modulation of brainstem speech processing

Compared to the NH group, we observed decreases in parieto-occipital α power in the HL group in both SNR conditions (Figure 2C). Acoustic signal detection especially in complex settings is found to increase working memory load (Shinn-Cunningham and Best, 2008) leading to enhanced power of α oscillations (Jensen et al., 2002). However, the neural mechanisms of working memory are affected by auditory degradation or hearing loss. Lower α power is reported in listeners with moderate hearing-loss across the age spectrum because they reach a ceiling level where no additional working memory resources can be recruited leading to decreased α power (Petersen et al., 2015). Moreover, PTA correlates negatively with pre-stimulus α power in older listeners (Alhanbali et al., 2021). These findings are partly concordant with our data since we found lower α power (during stimuli) in older listeners with mild to moderate hearing loss. Furthermore, in the HL group, we found no cortical-related enhancements of FFRs (i.e., F0 ratio ≈ 1) for clear speech, and responses were not different from the NH group in the noise condition though F0 ratio was > 1 (Figure 2D). The interaction effect of SNR \times α power was also distinct in direction between hearing groups. This finding implies that modulation of brainstem speech processing by cortical α state is altered in older listeners with mild hearing loss for both clear and noise-degraded speech processing.

In addition to parieto-occipital α power, trends of low- α -indexed FFRs to clear speech decreased with increased PTAs were observed in NH and HL listeners (Figure 4B) and the slope of decrease was significant for noise-degraded speech in NH listeners (Figure 4C). The reduction in low α FFRs with poorer PTAs is probably related to the decrease in peripheral hearing ability. However, when comparing across groups at a comparable hearing loss (PTA = 15–22 dB HL), we found enhanced FFR amplitudes in HL listeners (gray dashed box in Figure 4B). Speculatively, this could indicate an increase in central gain, probably related to high-frequency (> 4 kHz) hearing loss in the HL group [our NH listeners had normal audiometric thresholds up to 4 kHz, see Figure 1A in Bidelman et al. (2019)]. The average 4 kHz threshold of the NH group was ~ 20 dB HL but it was ~ 40 dB HL in the HL group. Similar central gain compensation secondary to peripheral hearing loss has been observed previously in both animal and human neuroimaging studies (Bidelman et al., 2014; Chambers et al., 2016). These phenomena were completely collapsed by noise in the HL group where low α FFRs were relatively smaller in most HL listeners (Figure 4C).

4.3. Association of brainstem speech processing during high arousal states with behaviors

Cortical α oscillations are used as a neural proxy of arousal (Pivik and Harman, 1995; Lai et al., 2022) and they were reported to play a significant role in functionally inhibiting the processing of task-irrelevant information (Jensen and Mazaheri, 2010; Foxe and Snyder, 2011). In our speech detection task, /ba/ and /pa/ tokens were the task-relevant distractors while /ta/ was the task-relevant target. Similar to the methods in Lai et al. (2022), we measured FFRs evoked

by /ba/ and /pa/ as there were insufficient counts (i.e., 210) of /ta/ token in each condition to provide reasonable FFRs. During low or decreased α power (i.e., high arousal state), NH listeners may have attended more to the target (/ta/) but less to the distractors (/ba/ and /pa/). Thus, FFRs to /ta/ were expected to increase while FFRs to /ba/ and /pa/ decrease. Whereas during high α power (i.e., low arousal state), responses to /ba/ and /pa/ were likely not inhibited because they were not task-irrelevant noise. Decreased α power was reported to associate with increased neural firing to attended stimuli (Haegens et al., 2011) and improved behavioral performance (Kelly et al., 2009; Gould et al., 2011; Haegens et al., 2011).

In younger listeners and under noisy backgrounds, we previously showed that neural decoding applied to low α FFRs offered higher accuracies in token classification as compared to high α FFRs [see Figure 5 in Lai et al. (2022)]. In this study, under no background noise, we observed that frequency spectra of low α FFR had better classification accuracies for perceptual performance level than high α FFR. These observations suggest that frequency spectra of low α FFRs were more robust in carrying information about the acoustic speech stimuli although their F0 amplitudes were smaller. For example, a better coding of harmonics or formants in the spectra renders the FFRs to be more representative of the acoustic speech waveforms and thus more predictive of perceptual performance. On the other hand, high α FFRs had larger F0 amplitudes, but the whole spectra of FFRs were noisier and contained less information about the acoustic speech. Taken together, better speech token discrimination is consistently observed in FFRs during high arousal states in both younger and older adults. Furthermore, low α FFRs were also observed to be decreased as PTA increased, especially in the NH group for noise-degraded speech. These observations suggested that brainstem FFRs during high arousal states have a strong association with behavior perception.

5. Conclusion

Collectively, our study reveals age-related hearing loss not only reduces cortical α power but differentially alters its dynamic relationship with subcortical speech processing. This cortical-brainstem modulation is especially prominent in the presence of noise and in listeners with age-related hearing loss. While brainstem speech processing is actively modulated by cortical arousal state in normal-hearing older adults for clear speech, this modulation is disrupted by the addition of background talker babble and diminished by hearing loss even in clear backgrounds. Speech-FFRs during low α states also offer a higher fidelity representation of the acoustic speech signature and are more predictive of perceptual performance than FFRs yoked to states of high cortical α . Enhanced FFRs in older adults with near-normal hearing (i.e., very mild hearing loss) suggest the presence of increased central gain compensation for reduced auditory input (Bidelman et al., 2014; Chambers et al., 2016).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Baycrest Hospital Human Subject Review Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JL contributed to study design, data and statistical analysis, data interpretation, and writing the manuscript. CA contributed to data curation, data interpretation, and editing the manuscript. GB contributed to conception, experimental design, data interpretation, and editing the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the National Institutes of Health (NIH/NIDCD R01DC016267) (GB), the Canadian Institutes of Health Research (MOP 106619) (CA), and the Natural Sciences and Engineering Research Council of Canada (NSERC, 194536) (CA).

Acknowledgments

We would like to thank Dawei Shen and Stephen R. Arnott for their assistance with data collection as well as Yimei Li and Tushar Patni for providing consultation in statistical linear regression analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2023.1075368/full#supplementary-material>

References

- Aftanas, L. I., Varlamov, A. A., Pavlov, S. V., Makhnev, V. P., and Reva, N. V. (2002). Time-dependent cortical asymmetries induced by emotional arousal: EEG analysis of event-related synchronization and desynchronization in individually defined frequency bands. *Int. J. Psychophysiol.* 44, 67–82. doi: 10.1016/S0167-8760(01)00194-5
- Alain, C., and McDonald, K. L. (2007). Age-related differences in neuromagnetic brain activity underlying normal sound perception. *J. Neurosci.* 27:1308. doi: 10.1523/JNEUROSCI.5433-06.2007
- Alain, C., Ogawa, K. H., and Woods, D. L. (1996). Aging and the segregation of auditory stimulus sequences. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 51, 91–93.
- Alhanbali, S., Munro, K. J., Dawes, P., Perugia, E., and Millman, R. E. (2021). Associations between pre-stimulus alpha power, hearing level and performance in a digits-in-noise task. *Int. J. Audiol.* 61, 197–204. doi: 10.1080/14992027.2021.1899314
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). Aging affects neural precision of speech encoding. *J. Neurosci.* 32, 14156–14164. doi: 10.1523/JNEUROSCI.2176-12.2012
- Anderson, S., Parbery-Clark, A., Yi, H. G., and Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear. Hear.* 32, 750–757. doi: 10.1097/AUD.0b013e31822229d3
- Babiloni, C., Binetti, G., Cassarino, A., Dal Forno, G., del Percio, C., Ferreri, F., et al. (2006). Sources of cortical rhythms in adults during physiological aging: A multicentric EEG study. *Hum. Brain Mapp.* 27, 162–172. doi: 10.1002/HBM.20175
- Bidelman, G. M. (2018a). Synchronization of scalp-recorded frequency-following responses (FFRs) offers improved response detection over conventional statistical metrics. *J. Neurosci. Methods* 293, 59–66. doi: 10.1016/j.jneumeth.2017.09.005
- Bidelman, G. M. (2018b). 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., and Momtaz, S. (2021). Subcortical rather than cortical sources of the frequency-following response (FFR) relate to speech-in-noise perception in normal-hearing listeners. *Neurosci. Lett.* 746:135664. doi: 10.1016/j.neulet.2021.135664
- Bidelman, G. M., Moreno, S., and 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., Price, C. N., Shen, D., Arnott, S. R., and 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., and Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol. Aging* 35, 2526–2540. doi: 10.1016/J.NEUROBIOLAGING.2014.05.006
- Böttger, D., Herrmann, C. S., and von Cramon, D. Y. (2002). Amplitude differences of evoked alpha and gamma oscillations in two different age groups. *Int. J. Psychophysiol.* 45, 245–251. doi: 10.1016/S0167-8760(02)00031-4
- Brugge, J. F., Nourski, K. v., Oya, H., Reale, R. A., Kawasaki, H., Steinschneider, M., et al. (2009). Coding of repetitive transients by auditory cortex on Heschl's Gyrus. *J. Neurophysiol.* 102:2358. doi: 10.1152/JN.91346.2008
- Chambers, A. R., Resnik, J., Yuan, Y., Whitton, J. P., Edge, A. S., Liberman, M. C., et al. (2016). Central gain restores auditory processing following near-complete cochlear denervation. *Neuron* 89, 867–879. doi: 10.1016/j.neuron.2015.12.041
- Chiang, A. K. L., Rennie, C. J., Robinson, P. A., van Albada, S. J., and Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clin. Neurophysiol.* 122, 1505–1517. doi: 10.1016/J.CLINPH.2011.01.040
- Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat. Commun.* 7:11070. doi: 10.1038/ncomms11070
- Compton, R. J., Gearing, D., and Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cogn. Affect. Behav. Neurosci.* 19, 1184–1191. doi: 10.3758/S13415-019-00745-9/FIGURE/2
- Dimitrijevic, A., Smith, M. L., Kadis, D. S., and Moore, D. R. (2017). Cortical alpha oscillations predict speech intelligibility. *Front. Hum. Neurosci.* 11:88. doi: 10.3389/FNHUM.2017.00088/BIBTEX
- Dubno, J. R. (1984). Effects of age and mild hearing loss on speech recognition in noise. *J. Acoust. Soc. Am.* 76, 87–96. doi: 10.1121/1.391011
- Dubno, J. R., and Schaefer, A. B. (1992). Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners. *J. Acoust. Soc. Am.* 91, 2110–2121. doi: 10.1121/1.403697
- Fostick, L., Ben-Artzi, E., and Babkoff, H. (2013). Aging and speech perception: Beyond hearing threshold and cognitive ability. *J. Basic Clin. Physiol. Pharmacol.* 24, 175–183.
- Foxe, J. J., and Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front. Psychol.* 2:154. doi: 10.3389/FPSYG.2011.00154/BIBTEX
- Gaál, Z. A., Boha, R., Stam, C. J., and Molnár, M. (2010). Age-dependent features of EEG-reactivity-spectral, complexity, and network characteristics. *Neurosci. Lett.* 479, 79–84. doi: 10.1016/J.NEULET.2010.05.037
- Gallun, F. J., and Best, V. (2020). Age-related changes in segregation of sound sources. *Audit. Res.* 72, 143–171. doi: 10.1007/978-3-030-49367-7_7
- Gatlin, A. E., and Dhar, S. (2021). History and lingering impact of the arbitrary 25-dB cutoff for normal hearing. *Am. J. Audiol.* 30, 231–234. doi: 10.1044/2020_AJA-20-00181
- Gould, I. C., Rushworth, M. F., and Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *J. Neurophysiol.* 105, 1318–1326. doi: 10.1152/JN.00653.2010
- Haegens, S., Nacher, V., Luna, R., Romo, R., and Jensen, O. (2011). α -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19377–19382. doi: 10.1073/PNAS.1117190108/-/DCSUPPLEMENTAL
- Händel, B. F., Haarmeier, T., and Jensen, O. (2011). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J. Cogn. Neurosci.* 23, 2494–2502. doi: 10.1162/JOCN.2010.12557
- Henry, M. J., Herrmann, B., Künke, D., and Obleser, J. (2017). Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nat. Commun.* 1, 1–11. doi: 10.1038/ncomms15801
- Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Front. Hum. Neurosci.* 4:186. doi: 10.3389/FNHUM.2010.00186/BIBTEX
- Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882. doi: 10.1093/CERCOR/12.8.877
- Jin, S.-H. H., Liu, C., and Sladen, D. P. (2014). The effects of aging on speech perception in noise: Comparison between normal-hearing and cochlear-implant listeners. *J. Am. Acad. Audiol.* 25, 656–665. doi: 10.3766/jaaa.25.7.4
- Kelly, S. P., Gomez-Ramirez, M., and Foxe, J. J. (2009). The strength of anticipatory spatial biasing predicts target discrimination at attended locations: A high-density EEG study. *Eur. J. Neurosci.* 30, 2224–2234. doi: 10.1111/J.1460-9568.2009.06980.X
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and 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.* 116:2395. doi: 10.1121/1.1784440
- Klatt, L. I., Getzmann, S., Begau, A., and Schneider, D. (2020). A dual mechanism underlying retroactive shifts of auditory spatial attention: Dissociating target- and distractor-related modulations of alpha lateralization. *Sci. Rep.* 1, 1–13. doi: 10.1038/s41598-020-70004-2
- Klimesch, W. (1997). EEG-alpha rhythms and memory processes. *Int. J. Psychophysiol.* 26, 319–340. doi: 10.1016/S0167-8760(97)00773-3
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/J.BRAINRESREV.2006.06.003
- Kortlang, S., Mauermann, M., and Ewert, S. D. (2016). Suprathreshold auditory processing deficits in noise: Effects of hearing loss and age. *Hear Res.* 331, 27–40. doi: 10.1016/j.heares.2015.10.004
- Lai, J., and Bartlett, E. L. (2015). Age-related shifts in distortion product otoacoustic emissions peak-ratios and amplitude modulation spectra. *Hear Res.* 327, 186–198. doi: 10.1016/j.heares.2015.07.017
- Lai, J., and Bartlett, E. L. (2018). Masking differentially affects envelope-following responses in young and aged animals. *Neuroscience* 386, 150–165. doi: 10.1016/J.NEUROSCIENCE.2018.06.004
- Lai, J., Price, C. N., and Bidelman, G. M. (2022). Brainstem speech encoding is dynamically shaped online by fluctuations in cortical α state. *Neuroimage* 263:119627. doi: 10.1016/J.NEUROIMAGE.2022.119627
- Lai, J., Sommer, A. L., and Bartlett, E. L. (2017). Age-related changes in envelope-following responses at equalized peripheral or central activation. *Neurobiol. Aging* 58:191. doi: 10.1016/J.NEUROBIOLAGING.2017.06.013
- Mai, G., Schoof, T., and Howell, P. (2019). Modulation of phase-locked neural responses to speech during different arousal states is age-dependent. *Neuroimage* 189, 734–744. doi: 10.1016/J.NEUROIMAGE.2019.01.049
- Maillet, D., Yu, L., Lau, B., Chow, R., Alain, C., and Grady, C. L. (2020). Differential effects of mind-wandering and visual distraction on age-related changes in neuroelectric brain activity and variability. *Neuropsychologia* 146:107565. doi: 10.1016/J.NEUROPSYCHOLOGIA.2020.107565
- McMahon, C. M., Boisvert, I., de Lissa, P., Granger, L., Ibrahim, R., Lo, C. Y., et al. (2016). Monitoring alpha oscillations and pupil dilation across a performance-intensity function. *Front. Psychol.* 7:745. doi: 10.3389/FPSYG.2016.00745

- Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res.* 241, 34–42. doi: 10.1016/j.heares.2008.04.013
- Oostenveld, R., and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112, 713–719. doi: 10.1016/S1388-2457(00)00527-7
- Park, H. L., O'Connell, J. E., and Thomson, R. G. (2003). A systematic review of cognitive decline in the general elderly population. *Int. J. Geriatr. Psychiatry* 18, 1121–1134. doi: 10.1002/GPS.1023
- Parthasarathy, A., and Bartlett, E. (2012). Two-channel recording of auditory-evoked potentials to detect age-related deficits in temporal processing. *Hear Res.* 289, 52–62. doi: 10.1016/j.heares.2012.04.014
- Parthasarathy, A., and Bartlett, E. L. (2011). Age-related auditory deficits in temporal processing in F-344 rats. *Neuroscience* 192, 619–630. doi: 10.1016/j.neuroscience.2011.06.042
- Parthasarathy, A., Cunningham, P. A., and Bartlett, E. L. (2010). Age-related differences in auditory processing as assessed by amplitude-modulation following responses in quiet and in noise. *Front. Aging Neurosci.* 2:152. doi: 10.3389/fnagi.2010.00152
- Parthasarathy, A., Datta, J., Torres, J. A., Hopkins, C., and Bartlett, E. L. (2014). Age-related changes in the relationship between auditory brainstem responses and envelope-following responses. *J. Assoc. Res. Otolaryngol.* 15, 649–661. doi: 10.1007/s10162-014-0460-1
- Parthasarathy, A., Lai, J., and Bartlett, E. L. (2016). Age-related changes in processing simultaneous amplitude modulated sounds assessed using envelope following responses. *J. Assoc. Res. Otolaryngol.* 17, 119–132. doi: 10.1007/s10162-016-0554-z
- Peelle, J. E., Troiani, V., Wingfield, A., and Grossman, M. (2010). Neural processing during older adults' comprehension of spoken sentences: Age differences in resource allocation and connectivity. *Cereb. Cortex (New York, NY)* 20:773. doi: 10.1093/CERCOR/BHP142
- Pelagatti, C., Binda, P., and Vannucci, M. (2020). A closer look at the timecourse of mind wandering: Pupillary responses and behaviour. *PLoS One* 15:e0226792. doi: 10.1371/JOURNAL.PONE.0226792
- Petersen, E. B., Wöstmann, M., Obleser, J., Stenfelt, S., and Lunner, T. (2015). Hearing loss impacts neural alpha oscillations under adverse listening conditions. *Front. Psychol.* 6:177. doi: 10.3389/fpsyg.2015.00177
- Pfurtscheller, G., Stancák, A., and Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. *Int. J. Psychophysiol.* 24, 39–46. doi: 10.1016/S0167-8760(96)00066-9
- Phipson, B., and Smyth, G. K. (2010). Permutation P-values should never be zero: Calculating exact P-values when permutations are randomly drawn. *Stat. Appl. Genet. Mol. Biol.* 9:Article39. doi: 10.2202/1544-6115.1585
- Pichora-Fuller, M. K., Alain, C., and Schneider, B. A. (2017). "Older adults at the cocktail party," in *The auditory system at the cocktail party*, eds J. C. Middlebrooks, J. Z. Simon, A. N. Poppe, and R. R. Fay (Cham: Springer), 227–259. doi: 10.1007/978-3-319-51662-2_9
- Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103
- Pivik, R. T., and Harman, K. (1995). A reconceptualization of EEG alpha activity as an index of arousal during sleep: All alpha activity is not equal. *J. Sleep Res.* 4, 131–137. doi: 10.1111/j.1365-2869.1995.tb00161.x
- Poth, E. A., Boettcher, F. A., Mills, J. H., and Dubno, J. R. (2001). Auditory brainstem responses in younger and older adults for broadband noises separated by a silent gap. *Hear Res.* 161, 81–86. doi: 10.1016/s0378-5955(01)00352-5
- Presacco, A., Simon, J. Z., and Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J. Neurophysiol.* 116, 2346–2355. doi: 10.1152/jn.00372.2016
- Price, C. N., Alain, C., and 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
- Price, C. N., and Bidelman, G. M. (2021). Attention reinforces human corticofugal system to aid speech perception in noise. *Neuroimage* 235:118014. doi: 10.1016/j.neuroimage.2021.118014
- Purdon, P. L., Pavone, K. J., Akeju, O., Smith, A. C., Sampson, A. L., Lee, J., et al. (2015). The ageing brain: Age-dependent changes in the electroencephalogram during propofol and sevoflurane general anaesthesia. *Br. J. Anaesth.* 115, i46–i57. doi: 10.1093/BJA/AEV213
- Rihs, T. A., Michel, C. M., and Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *Eur. J. Neurosci.* 25, 603–610. doi: 10.1111/j.1460-9568.2007.05278.x
- Saderi, D., Schwartz, Z. P., Heller, C. R., Pennington, J. R., and David, S. v (2021). Dissociation of task engagement and arousal effects in auditory cortex and midbrain. *Elife* 10, 1–25. doi: 10.7554/ELIFE.60153
- Saiz-Alia, M., Forte, A. E., and 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.* 1, 1–10. doi: 10.1038/s41598-019-50773-1
- Scherg, M., and Ebersole, J. S. (1994). Brain source imaging of focal and multifocal epileptiform EEG activity. *Neurophysiol. Clin.* 24, 51–60. doi: 10.1016/S0987-7053(05)80405-8
- Scherg, M., Ille, N., Bornfleth, H., and Berg, P. (2002). Advanced tools for digital EEG review: Virtual source montages, whole-head mapping, correlation, and phase analysis. *J. Clin. Neurophysiol.* 19, 91–112. doi: 10.1097/00004691-200203000-00001
- Schneider, B. A., Daneman, M., and Murphy, D. R. (2005). Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing? *Psychol. Aging* 20, 261–271. doi: 10.1037/0882-7974.20.2.261
- Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299. doi: 10.1177/1084713808325306
- Song, J. H., Skoe, E., Banai, K., and Kraus, N. (2011). Perception of speech in noise: Neural correlates. *J. Cogn. Neurosci.* 23:2268. doi: 10.1162/JOCN.2010.21556
- Souza, P. E., Boike, K. T., Witherell, K., and Tremblay, K. (2007). Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise. *J. Am. Acad. Audiol.* 18, 54–65. doi: 10.3766/jaaa.18.1.5
- Strauß, A., Wöstmann, M., and Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Front. Hum. Neurosci.* 8:350. doi: 10.3389/fnhum.2014.00350
- Takahashi, G. A., and Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech Hear Res.* 35, 1410–1421.
- Uusberg, A., Uibo, H., Kreegipuu, K., and Allik, J. (2013). EEG alpha and cortical inhibition in affective attention. *Int. J. Psychophysiol.* 89, 26–36. doi: 10.1016/j.ijpsycho.2013.04.020
- Vermeire, K., Knoop, A., Boel, C., Auwers, S., Schenus, L., Talaveron-Rodriguez, M., et al. (2016). Speech recognition in noise by younger and older adults: Effects of age, hearing loss, and temporal resolution. *Ann. Otol. Rhinol. Laryngol.* 125, 297–302. doi: 10.1177/0003489415611424
- Wayne, R. V., and Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Res. Rev.* 23, 154–166. doi: 10.1016/j.arr.2015.06.002
- Weisz, N., Dohrmann, K., and Elbert, T. (2007). The relevance of spontaneous activity for the coding of the tinnitus sensation. *Prog. Brain Res.* 166, 61–70. doi: 10.1016/S0079-6123(07)66006-3
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I., and Obleser, J. (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. *Front. Psychol.* 2:73. doi: 10.3389/fpsyg.2011.00073
- Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x
- Wöstmann, M., Herrmann, B., Wilsch, A., and Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J. Neurosci.* 35, 1458–1467. doi: 10.1523/JNEUROSCI.3250-14.2015
- Xie, Z., Reetzke, R., and Chandrasekaran, B. (2019). Machine learning approaches to analyze speech-evoked neurophysiological responses. *J. Speech Lang. Hear. Res.* 62, 587–601. doi: 10.1044/2018_JSLHR-S-ASTM-18-0244
- Yordanova, J. Y., Kolev, V. N., and Başar, E. (1998). EEG theta and frontal alpha oscillations during auditory processing change with aging. *Electroencephal. Clin. Neurophysiol.* 108, 497–505. doi: 10.1016/S0168-5597(98)00028-8