INNOVATIVE METHODOLOGY | *Sensory Processing*

Mutual information analysis of neural representations of speech in noise in the aging midbrain

Peng Zan,[1] Alessandro Presacco,[2] Samira Anderson,[3] and Jonathan Z. Simon[1,2,4]

[1]*Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland;* [2]*Institute for Systems Research, University of Maryland, College Park, Maryland;* [3]*Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland; and* [4]*Department of Biology, University of Maryland, College Park, Maryland*

Zan P, Presacco A, Anderson S, Simon JZ. Mutual information analysis of neural representations of speech in noise in the aging midbrain. *J Neurophysiol* 122: 2372–2387, 2019. First published October 9, 2019; doi:10.1152/jn.00270.2019.—Younger adults with normal hearing can typically understand speech in the presence of a competing speaker without much effort, but this ability to understand speech in challenging conditions deteriorates with age. Older adults, even with clinically normal hearing, often have problems understanding speech in noise. Earlier auditory studies using the frequency-following response (FFR), primarily believed to be generated by the midbrain, demonstrated age-related neural deficits when analyzed with traditional measures. Here we use a mutual information paradigm to analyze the FFR to speech (masked by a competing speech signal) by estimating the amount of stimulus information contained in the FFR. Our results show, first, a broadband informational loss associated with aging for both FFR amplitude and phase. Second, this age-related loss of information is more severe in higher-frequency FFR bands (several hundred hertz). Third, the mutual information between the FFR and the stimulus decreases as noise level increases for both age groups. Fourth, older adults benefit neurally, i.e., show a reduction in loss of information, when the speech masker is changed from meaningful (talker speaking a language that they can comprehend, such as English) to meaningless (talker speaking a language that they cannot comprehend, such as Dutch). This benefit is not seen in younger listeners, which suggests that age-related informational loss may be more severe when the speech masker is meaningful than when it is meaningless. In summary, as a method, mutual information analysis can unveil new results that traditional measures may not have enough statistical power to assess.

NEW & NOTEWORTHY Older adults, even with clinically normal hearing, often have problems understanding speech in noise. Auditory studies using the frequency-following response (FFR) have demonstrated age-related neural deficits with traditional methods. Here we use a mutual information paradigm to analyze the FFR to speech masked by competing speech. Results confirm those from traditional analysis but additionally show that older adults benefit neurally when the masker changes from a language that they comprehend to a language they cannot.

electroencephalography; entropy; information theory

Address for reprint requests and other correspondence: J. Z. Simon, Dept. of Electrical and Computer Engineering, 8223 Paint Branch Dr., College Park, MD 20742 (e-mail: jzsimon@umd.edu).

INTRODUCTION

Understanding speech in the presence of background noise becomes more challenging as humans age. Older listeners often report problems with listening to speech in noise even with clinically normal hearing sensitivity (Burke and Shafto 2008; Helfer and Freyman 2008). Behavioral studies have revealed age-related temporal processing deficits in a number of auditory tasks, such as pitch discrimination (Fitzgibbons and Gordon-Salant 1996), gap-in-noise detection (Fitzgibbons and Gordon-Salant 2001), and recognition of speech in noise (Frisina and Frisina 1997; Gordon-Salant et al. 2006; He et al. 2008; Schneider and Hamstra 1999). These results suggest a temporal processing degradation in the auditory pathway, consistent with observed age-related changes in response latency and strength in midbrain (Anderson et al. 2012; Burkard and Sims 2002; Clinard and Tremblay 2013) and cortical evoked responses (Lister et al. 2011; Presacco et al. 2016a, 2016b).

The neural mechanism underlying age-related temporal auditory process deficits has also been investigated in animal studies: decreased release of inhibitory neurotransmitters, such as γ-aminobutyric acid (GABA), in dorsal cochlear nucleus (DCN) (Caspary et al. 2005; Parthasarathy and Bartlett 2011; Schatteman et al. 2008; Wang et al. 2009), inferior colliculus (IC) (Caspary et al. 1995), and auditory cortex (de Villers-Sidani et al. 2010; Juarez-Salinas et al. 2010) have been found in aging mammals. Because the spectrotemporal fine structure of speech is encoded by synchronous neural firing in midbrain and the accurate processing of rapid fluctuations depends partly on inhibitory mechanisms, the representation of speech in midbrain may also deteriorate as a result of greater variability of neural firing (Walton et al. 1998; Yang et al. 1992) or loss of inhibition (Caspary et al. 2005, 2006; Walton et al. 1998). The midbrain frequency-following response (FFR), which tracks periodic components of speech or other sounds, may be detrimentally affected by the resulting neural jitter. In older listeners, jitter may be more prevalent than in younger listeners, as reflected by a decreased intertrial response consistency (Anderson et al. 2012) or, as we hypothesize here, by increased entropy and decreased mutual information as defined in the context of information theory (Cover and Thomas 1991; Shannon 1948).

Mutual information, in particular, can be interpreted as a reduction in auditory response variability due to the presentation of a stimulus (Nelken and Chechik 2007). It has been used to estimate transmission rates in the low-frequency fibers of the auditory periphery in bullfrog (Rieke et al. 1995) and applied to magnetoencephalography auditory responses to continuous speech (Cogan and Poeppel 2011). Auditory information transmitted from midbrain to auditory cortex has been observed to show greater redundancy in older listeners compared with younger listeners (Bidelman et al. 2014). However, given that older listeners have a weaker midbrain response than younger listeners (Presacco et al. 2016a, 2016b), it remains an open question whether the aging midbrain itself processes more information or less information than younger listeners.

The present study is a mutual informational analysis of auditory midbrain FFR. A more traditional analysis (evoked response) of this data set has already been published (Presacco et al. 2016a, 2016b). The goals of this new analysis are *1*) to describe these new and innovative methods in detail, *2*) to demonstrate rich examples of their use, and *3*) to demonstrate that the results are quite often stronger in statistical power than the more traditional methods. First, it is shown that the new analysis replicates the most basic earlier findings, that older listeners' midbrain FFR responses contain less auditory signal information about speech stimuli than those of younger listeners', at the fundamental frequency (F_0) of the FFR. We then generalize the analysis to harmonic frequencies, showing that speech information contained in the harmonics is similarly degraded with age (and falls off more quickly in frequency), consistent with earlier findings (Anderson et al. 2012). Finally, we also show that when the speech stimuli are degraded by the addition of a competing talker, the stimulus information contained in the midbrain FFR is more sensitive to informational masking (competing speech in a familiar vs. unfamiliar language) in older listeners than in younger listeners.

MATERIALS AND METHODS

Subjects

The data set used in this study has previously been described (Presacco et al. 2016a, 2016b). Seventeen younger listeners (14 women, 3 men) between 18 and 27 yr old (mean $\pm$ SD: 22.23 $\pm$ 2.27 yr) and fifteen older listeners (10 women, 5 men) between 61 and 73 yr old (mean $\pm$ SD: 65.06 $\pm$ 2.30 yr), recruited from the Maryland, Washington, DC, and Virginia areas, participated in the experiment. All subjects had clinically normal hearing with air conduction thresholds no greater than 25 dB hearing level (HL) from 125 to 4,000 Hz bilaterally and no interaural asymmetry. All of them were native English speakers and were free of neurological or middle ear disorders, and none of them spoke or understood the Dutch language. All participants were paid for their participation, and each of them gave written informed consent before the experiment. The experimental protocol and all procedures were reviewed and approved by the Institutional Review Board of the University of Maryland.

Stimuli and EEG Recording

The stimulus was a single speech syllable, a 170-ms /da/ (Anderson et al. 2012), synthesized at a 20-kHz sampling rate with a Klatt-based synthesizer (Klatt 1980) with a 100-Hz F_0. The syllable was chosen because it comprises both transient and steady-state components, the stop consonant /d/ is rich in phonetic information, and its perception is sensitive to background noise (Miller and Nicely 1955). Its waveform and spectrum are shown in Fig. 1. The speech syllable was presented diotically at 75 dB SPL with a repetition rate of 4 Hz. Stimuli were presented with alternating polarities to allow cancellation of potential stimulus artifact by summing the responses to each pair (Aiken and Picton 2008). The stimulus was presented to subjects both in quiet and in noise. For the noise conditions, a story narrated by a female competing speaker in either English or Dutch was used as a masker (a 1-min duration segment, continuously looped). The English story was an excerpt from *A Christmas Carol* by Charles Dickens (http://www.audiobooktreasury.com/a-christmas-carol-by-charles-dickens-free-audio-book/), and the Dutch story was *Aljaska en de Canada-spoorweg* by Anonymous (http://www.loyalbooks.com/book/
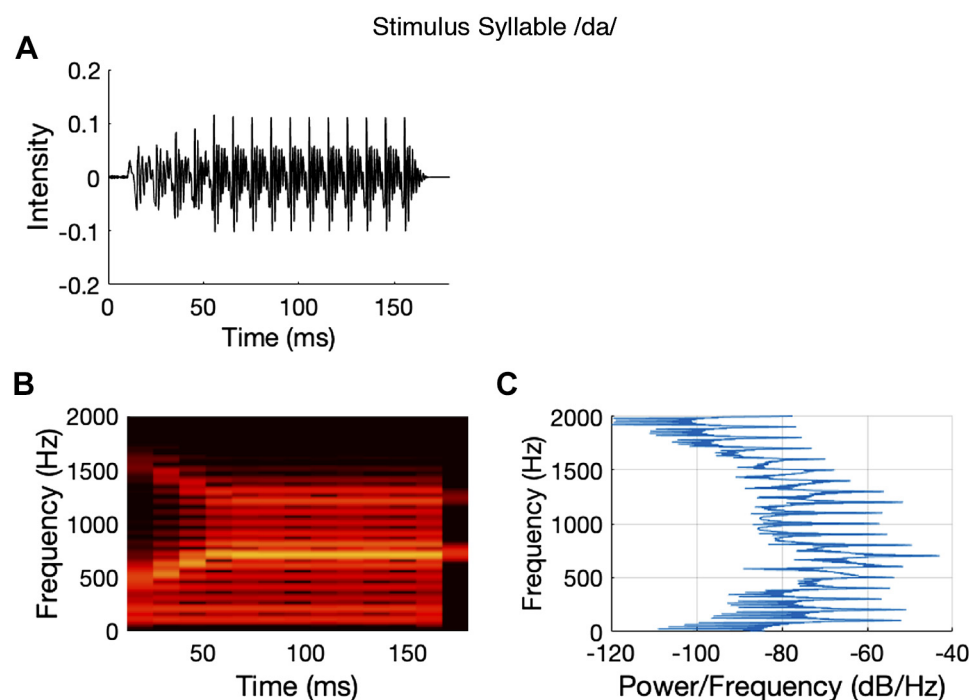
Stimulus Syllable /da/

A

B

C

Fig. 1. Stimulus waveform (*A*), spectrogram (*B*), and power spectral density (*C*) of 170-ms syllable /da/. The locations of the horizontal peaks in *C* indicate that the syllable has a fundamental frequency of 100 Hz with harmonic peaks at its multiples (Anderson et al. 2012, 2013).

Aljaska-en-de-Canada-spoorweg). For each of the two masker types, four signal-to-noise ratio (SNR) levels, $+3$, 0, -3, and -6 dB SNR, were created by using the logarithm of the ratio between root-mean-squared values of syllable /da/ and the long-duration masking speech. All stimuli were presented by insert earphones (ER1; Etymotic Research, Elk Grove Village, IL) via Xonar Essence One (ASUS, Taipei, Taiwan) with Presentation software (Neurobehavioral Systems, Berkeley, CA). FFRs were recorded at a sampling frequency of 16,384 Hz with the ActiABR-200 acquisition system (BioSemi B.V., Amsterdam, The Netherlands) with a standard vertical montage of five electrodes (Cz active, forehead ground common mode sense/driven right leg electrodes, earlobe references), and the recorded signal was filtered online by a band-pass filter with a cutoff band of 100 Hz to 3,000 Hz. During the 2-h recording session, subjects sat in a recliner and watched a silent captioned movie of their choice to facilitate a relaxed but wakeful state. For each of the nine conditions (1 quiet + 2 masker languages × 4 SNRs), at least 2,300 trials of response (to repetitions of syllable /da/) were recorded.

Data Analysis

Encoding response amplitude. The EEG recordings were first converted into MATLAB format with the function *pop_biosig* from EEGLab (Delorme and Makeig 2004), and all remaining analyses were performed in MATLAB (version 2017b; MathWorks, Natick, MA). The EEG recordings were band-pass filtered off-line, to remove low-frequency neural oscillations, from 70 Hz to 2,000 Hz with a linear-phase finite impulse response (FIR) filter with low-pass transition band of 65–70 Hz and high-pass transition of 2,000–2,100 Hz. Filter delays were compensated by processing the data in both forward and backward directions with the MATLAB function *filtfilt* (MathWorks, Natick, MA). The response of each trial was analyzed in the time window -47 ms to 170 ms with respect to stimulus onset. Within this window, the response of each trial was band-pass filtered with linear-phase FIR filters of order 200, designed with least-square error minimization, into frequency bands centered at harmonics of 100 Hz, i.e., 100, 200, ..., 600 Hz, to investigate the midbrain representations of harmonics. Harmonics at or above 700 Hz, the first formant of the steady-state portion of the stimulus, were excluded from analysis. Sweeps with amplitudes larger than ± 30 μV were excluded, allowing 2,000 artifact-free sweeps to be used. To eliminate any possible electrical feedthrough artifacts, a 10-ms temporal response function centered at 0 ms with reference to the stimulus onset time was estimated per trial, and its contribution was subtracted from the response (Maddox and Lee 2018). Additionally, since two consecutive sweeps were always presented with opposite polarities, their responses were averaged into one effective sweep, leading to 1,000 such pair-averaged sweeps per subject and per condition that were then used for the analysis; the results for the same sweeps, with artifacts removed but not averaged (2,000 per condition) are presented in the APPENDIX. For each of the two analysis regions, the response waveforms were extracted from each sweep for every subject, for each of the nine conditions and six frequency bands.

Under each condition, for each subject and frequency band, a response matrix was obtained of size 1,000 trials × T samples. where T is the sample length of observation window. In addition to the entire response window 0–170 ms, the responses were also partitioned into two regions based on the acoustic properties of the syllable /da/, i.e., the transition (15–65 ms) and steady state (64–170 ms), for analysis of masker type influence on the response at 100 Hz. Here $T = 2,853$ samples for the entire response region, $T = 804$ samples for the transition region, and $T = 2,049$ samples for the steady-state region. The response amplitudes at each sample were subdivided into N bins, with the boundaries of the bins chosen so that approximately equal numbers of samples were assigned in each bin; each sample was then associated with its bin index (from 1 to N). The boundaries were chosen individually on the basis of each subject's response. Different

values of $N \in \{4,8,16,32,64,128\}$ were evaluated to verify a lack of any interaction with age ($F_{5,180} = 0.46$, $P = 0.809$ and $F_{5,180} = 0.18$, $P = 0.970$ by ANOVA test on interaction of age × bin number for amplitude and phase information, respectively). A final choice of $N = 32$ bins was selected as an optimal trade-off between increased resolution between bins and decreased samples per bin due to limited samples (too few bins or too few samples per bin both lead to estimation bias). The choice of 32 bins gave >30 samples/bin, on average, to estimate the conditional probability distribution.

Encoding response phase. For every sweep in each region, the phase for each frequency band was computed by first applying the Hilbert transform to the band-passed signal and then computing the phase of the resultant complex (analytic) signal, i.e.,

$$H\{x(t)\} = \mathrm{IFT}\left[-i \ sgn(f)\mathrm{FT}\{x(t)\}\right] \tag{1}$$

where FT is the Fourier transform, f is the frequency basis of the Fourier transform, $sgn(f)$ is the algebraic sign of f, and IFT is the inverse Fourier transform. Then

$$\theta(t) = \angle \left[x(t) + iH\{x(t)\}\right] \tag{2}$$

The phase-locking value (PLV) of the response in any single band can be computed as

$$\mathrm{PLV}(t) = \frac{1}{M}\left|\sum_{j=1}^{M} e^{i\theta_j(t)}\right| \tag{3}$$

where $\theta_j(t)$ is the phase of the jth trial at sample time t and M is the number of trials.

The set of phase responses $\theta_j(t)$ obtained for each frequency band were also subdivided into $N = 32$ bins, analogously to encoding the amplitude response; here the phase samples were divided into bins of width $\frac{2\pi}{32} = \frac{\pi}{16}$, with each sample encoded by its bin index (from 1 to N).

Mutual information. Under each condition, for each subject and each frequency band, the mutual information between stimulus and amplitude and the mutual information between stimulus and phase were estimated based on those integer-encoded responses. The response probability distribution was estimated as above (bin index for each of the T samples over 1,000 trials). The conditional distribution of $P(Y|X)$ was drawn from response samples at the same latency from 1,000 trials. The mutual information can then be estimated by the entropy of the response, whether amplitude or phase, minus the conditional entropy of the response given the (uniformly distributed) stimulus:

$$I(X;Y) = H(Y) - H(Y|X) \tag{4}$$

where X represents the stimulus distribution and Y is the response distribution, whether amplitude or phase. $H(Y)$ is the entropy of the response,

$$H(Y) = -\sum_y p(y)\log p(y) \tag{5}$$

where $p(y)$ is the probability of observing the response value y. $H(Y|X)$ is the entropy of the response conditioned by the stimulus X and is given by

$$H(Y|X) = \sum_x p(x)H(Y|X = x) \tag{6}$$

where

$$H(Y|X = x) = -\sum_y p(Y = y|x)\log p(Y = y|x) \tag{7}$$

The stimulus X is the amplitude or phase at each time point. The probability distribution of x is unknown but here assumed to be uniform $[p(x) = \frac{1}{T}$, a constant, so each bin contains roughly the same

number of stimulus value instances] for two reasons. First, when the actual stimulus distribution is unknown, this assumption minimizes estimation bias (Nelken and Chechik 2007). Second, while there is not yet evidence for any particular distribution (e.g., Gaussian or Laplacian), the assumption of uniform distribution was employed for stimulus amplitude by Cogan and Poeppel (2011) with encouraging results. Then, *Eq. 3* becomes

$$H(Y|X) = \frac{1}{T}\sum_{t=1}^{T} H(Y|X = x_t) \qquad (8)$$

where x_t is the amplitude or phase bin at sample t.

To illustrate, consider an analysis of the quiet condition over the steady-state region, which encompasses the time window from 64 ms to 189 ms with respect to stimulus onset, i.e., 2,049 samples, giving $T = 2{,}049$ and $p(x_t) = \dfrac{1}{2{,}049}$ for every value of t.

The distribution of the response, $P(Y)$, is estimated for each subject with all bin index-encoded samples in each of the 1,000 trials. The conditional distribution of Y given x_t, $P(Y|x_t)$, is estimated with 1,000 samples from trials at time point t. Then the conditional entropy is given by

$$\begin{aligned} H(Y|X) &= -\sum_{t=1}^{T}\sum_{i=1}^{N} p(X = x_t) p(Y = i|X = x_t) \log p(Y = i|X = x_t) \\ &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p(Y = i|X = x_t) \log p(Y = i|X = x_t) \end{aligned} \qquad (9)$$

where $i \in \{1,2,...,N\}$ is the bin number and N is the number of bins. The mutual information is therefore

$$\begin{aligned} I(X;Y) &= -\sum_{i=1}^{N} p(Y = i) \log p(Y = i) \\ &+ \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} p(Y = i|X = x_t) \log p(Y = i|X = x_t) \end{aligned} \qquad (10)$$

Statistics. To examine the effects of aging, frequency, masker type, and SNR level, multiple t tests with correction were performed, separately for both amplitude and phase information. To facilitate analysis of the information at fundamental frequency, linear models were constructed to test effects from interactions between aging and other factors, namely, masker type and SNR level, with the mathematical form $I \sim$ age × masker type + age × SNR. Tests were performed for both amplitude and phase, and for different temporal regions, separately. To test masker type influence within group, the mutual information difference between Dutch and English maskers for each subject was modeled as $I_{Dutch} - I_{English} \sim$ SNR, and the positivity of intercept was tested for both amplitude and phase, and for different temporal regions, separately. The results were justified by t tests on the intercept of linearly fitted regression lines for each subject and similar analysis for PLV.

Linear models with only fixed effects were analyzed in R (R Core Team 2017) with the function *lm*, which reports the model significance with an F test on the constructed model versus the null model with only the intercept and the significance of influence from fixed-effect factors with separate t tests on the slope of each factor. The assumption of homoscedasticity of the linear models was examined by global validation of linear model assumptions with toolbox *gvlma* (Peña and Slate 2006) in R. Responses at harmonic frequencies were analyzed with t tests. False discovery rate (FDR) correction (Benjamini and Hochberg 1995), to correct for multiple comparisons, was applied when appropriate.

Where appropriate, t tests for significance are supplemented with effect size (Cohen's d) and its 95% confidence interval (CI). When the CI excludes zero, this is alternate evidence that the result is statistically significant (i.e., the effect size is significantly greater than zero

at an α level of 0.05). Note, however, that the effect size analysis is not compensated for multiple comparisons even when the P value is.

The effective high-frequency cutoff for any frequency-decreasing statistical measure is defined to be the frequency at which the measure is not significantly higher than the noise floor (pure estimation bias). The noise floor is estimated using the same mutual information method as used elsewhere but instead using responses to quiet intervals between stimuli.

RESULTS

Here we report results from the mutual information analysis of pair-averaged polarity responses; the analogous analysis based on single sweeps is reported in the APPENDIX. Because our algorithm takes into account variations across trials, pair-averaging provides less variation and thus higher mutual information. Except for this overall scaling of mutual information, the results are typically comparable.

Information in FFR Amplitude

Amplitude information at 100 Hz. For the amplitude response at 100 Hz, to examine masker type and SNR interactions with both age groups, the linear model, $I \sim$ age × masker type + age × SNR, is tested. It is significant ($F_{5,250} = 4.99$, $P < 0.001$ for the entire region; $F_{5,248} = 2.93$, $P = 0.014$ for the transition region; $F_{5,249} = 6.11$, $P < 0.001$ for the steady-state region). Outliers that would otherwise cause the homoscedasticity requirement to be violated are excluded (2 samples from the transition region and 1 sample from the steady-state region). Results show no significant interactions between age and masker type ($t_{250} = 0.53$, $P = 0.587$ for the entire region; $t_{248} = 0.15$, $P = 0.884$ for the transition region; $t_{249} = 0.29$, $P = 0.773$ for the steady-state region) or between age and SNR ($t_{250} = 0.79$, $P = 0.428$ for the entire region; $t_{248} = 0.46$, $P = 0.645$ for the transition region; $t_{249} = 0.87$, $P = 0.386$ for the steady-state region). A linear model with no interactions was then constructed and tested, i.e., $I \sim$ age + masker type + SNR. The model itself is significant ($F_{3,252} = 8.05$, $P = 0.001$, $F_{3,250} = 4.84$, $P = 0.003$, and $F_{3,251} = 9.96$, $P < 0.001$ for the entire region and the transition and steady-state regions, respectively). Comparisons between the models show that younger listeners' responses contain significantly more information than older listeners' responses in the entire and steady-state regions ($t_{252} = 4.24$, $P < 0.001$ and $t_{251} = 4.99$, $P < 0.001$, respectively) and that information increases as SNR increases ($t_{252} = 2.37$, $P = 0.018$ for the entire region; $t_{250} = 2.86$, $P = 0.005$ for the transition region; $t_{251} = 2.15$, $P = 0.033$ for the steady-state region).

Since the stimulus has a fundamental frequency of 100 Hz and the phase locking of FFR is more robust in low frequencies than in high frequencies (Zhu et al. 2013), the 100-Hz FFR may contain significantly more information than its harmonics. To rule out the possibility that significant contributions to mutual information derive from averaging the opposite polarities, the same mutual information analysis is performed on single trials, where similar results are observed (see APPENDIX). Figure 2*A* displays the mutual information as a function of SNR level. Older listeners not only have a noticeably lower amount of information than younger listeners but also extract more speech information when the masker is Dutch than for English. To eliminate within-subject variance, a linear regression line of information by SNR was fitted for each subject and

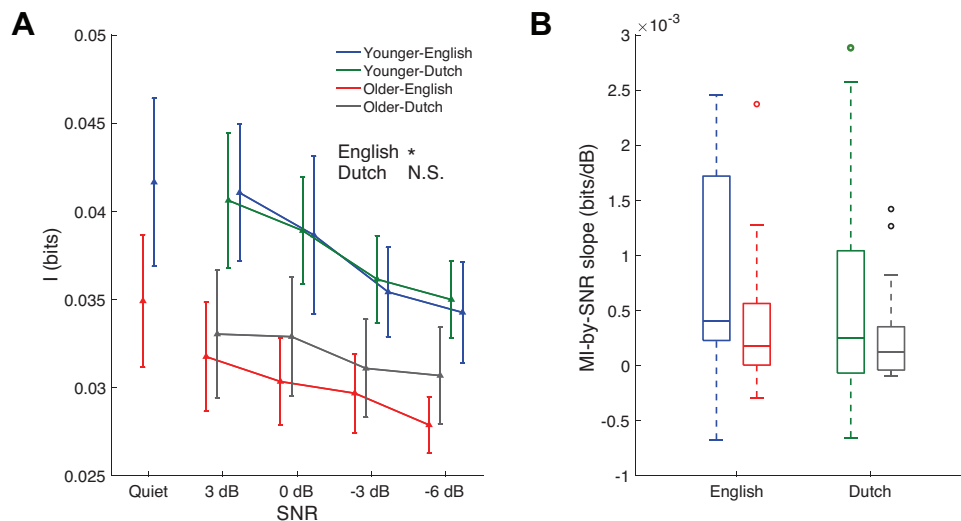## Amplitude Information at 100 Hz by Masker Type and Noise Level

Fig. 2. Mutual information between stimulus and response amplitude as a function of noise level for each age group and masker condition (masker language). *A*: mutual information (*I*) at the fundamental frequency as a function of noise level [quiet condition and 4 signal-to-noise ratio (SNR) levels] for younger listeners and older listeners with English and Dutch maskers. The response in younger listeners conveys noticeably more information than the response in older listeners for the English masker condition, but the difference for Dutch is not significant at 100 Hz. Older listeners show consistently higher mutual information for the Dutch masker than for the English (the younger listeners show no consistent difference), but the difference is not significant at 100 Hz. *B*: the mutual information (MI)-by-SNR slopes of the plots in *A* show decreasing trends as SNR worsens, regardless of masker type, for both age groups. Younger listeners show a steeper decrease than older listeners, but the difference is not significant at 100 Hz response. Error bars indicate SE. *$P < 0.05$. N.S., not significant.

its *y*-intercept and slope were analyzed, with results illustrated in Fig. 2. A one-tailed *t* test (younger > older) on the *y*-intercept shows a significantly larger amount of information in younger than older listeners for the English masker ($t_{30} = 1.71$, $P = 0.048$, $d = 0.75$, 95% CI = [0.032,1.469]). The difference is not significant for Dutch ($t_{30} = 1.41$, $P = 0.102$, $d = 0.51$, 95% CI = [−0.195,1.216]) (but, as seen below, it does become significant for higher harmonic frequencies). Both age groups demonstrate decreasing information with worsening SNR: a one-tailed *t* test on the negativity of the regression slope shows information loss for all cases except for older listeners with the Dutch masker ($t_{16} = 3.42$, $P = 0.002$ and $t_{16} = 2.54$, $P = 0.013$ for younger listeners for English and Dutch maskers, respectively, and $t_{14} = 2.32$, $P = 0.027$, $d = 0.60$, 95% CI = [2.55×10^{-5}, $+\infty$] and $t_{14} = 2.35$, $P = 0.059$, $d = 0.61$, 95% CI = [1.92×10^{-5}, $+\infty$] for older listeners). No significant difference is seen between the slopes across age groups (1-tailed *t* test: $t_{30} = 1.28$, $P = 0.106$, $d = 0.55$, 95% CI = [−0.155,1.260] for the English masker; $t_{30} = 1.20$, $P = 0.120$, $d = 0.73$, 95% CI = [0.018,1.452] for the Dutch masker, although the effect size CI is consistent with significance in the last case).

Amplitude information in harmonics of 100 Hz. To analyze aging-associated informational loss for the harmonics (200–600 Hz), similar tests are performed on mutual information in responses of these frequencies (analysis stops before 700 Hz, which represents the first formant of the steady-state portion of the stimulus). In each harmonic, a linear regression line of mutual information as a function of SNR is fitted for each subject under each masker type. First the *y*-intercept of the fitted line at 3 dB is analyzed for group differences (see Fig. 3).

One-tailed (younger > older) *t* tests (with FDR correction) and effect size analysis on the *y*-intercept (corresponding to 3 dB SNR) of the line fit across all SNR levels suggest that the aging

midbrain contains significantly less information than the younger midbrain in all frequencies from 100 to 600 Hz in the English masker condition. For *P* values near 0.05 (see Table 1), effect size analysis is further applied. For the English masker condition, the 100-Hz condition shows consistent significance from both tests ($t_{30} = 1.714$, $P = 0.048$, $d = 0.75$, 95% CI = [0.032,1.469]) and similarly for the Dutch masker condition at 300 Hz ($t_{30} = 2.05$, $P = 0.049$, $d = 1.236$, 95% CI = [0.478,1.993]), 500 Hz ($t_{30} = 2.27$, $P = 0.047$, $d = 0.787$, 95% CI = [0.0663,1.507]), and 600 Hz ($t_{30} = 2.26$, $P = 0.047$, $d = 1.053$, 95% CI = [0.312,1.794]) (see also Fig. 3*A*). In the English masker condition, one-tailed *t* tests on fitted regression line slopes of younger listeners compared with older listeners show significantly steeper slopes for younger listeners compared with older listeners at frequencies from 200 to 600 Hz (all *P* values are smaller than 0.05). All *P* values of multiple comparisons are corrected. Overall, higher harmonics contain significant information only for younger listeners, and the difference in information between the two age groups becomes more statistically significant as the observed frequency increases, which is consistent with the linear model analysis, where age × frequency interaction is significant.

Amplitude information frequency limits. As seen in Fig. 3*B*, the stimulus information contained in the response amplitude decreases with frequency for both age groups. The frequency-decreasing measure used here is the amplitude information's *y*-intercept at 3 dB of the fitted mutual information-by-SNR regression line. The frequency bands below 700 Hz are analyzed separately for different masker types. The measure at 600 Hz for older listeners is not statistically distinguishable from the noise floor ($t_{14} = 1.72$, $P = 0.107$ by 1-sample *t* test). For younger listeners, the measure is significantly higher than the noise floor at all frequencies ($t_{30} = 3.34$, $P = 0.002$ for English masker; $t_{30} = 2.26$, $P = 0.016$ for Dutch masker (younger >
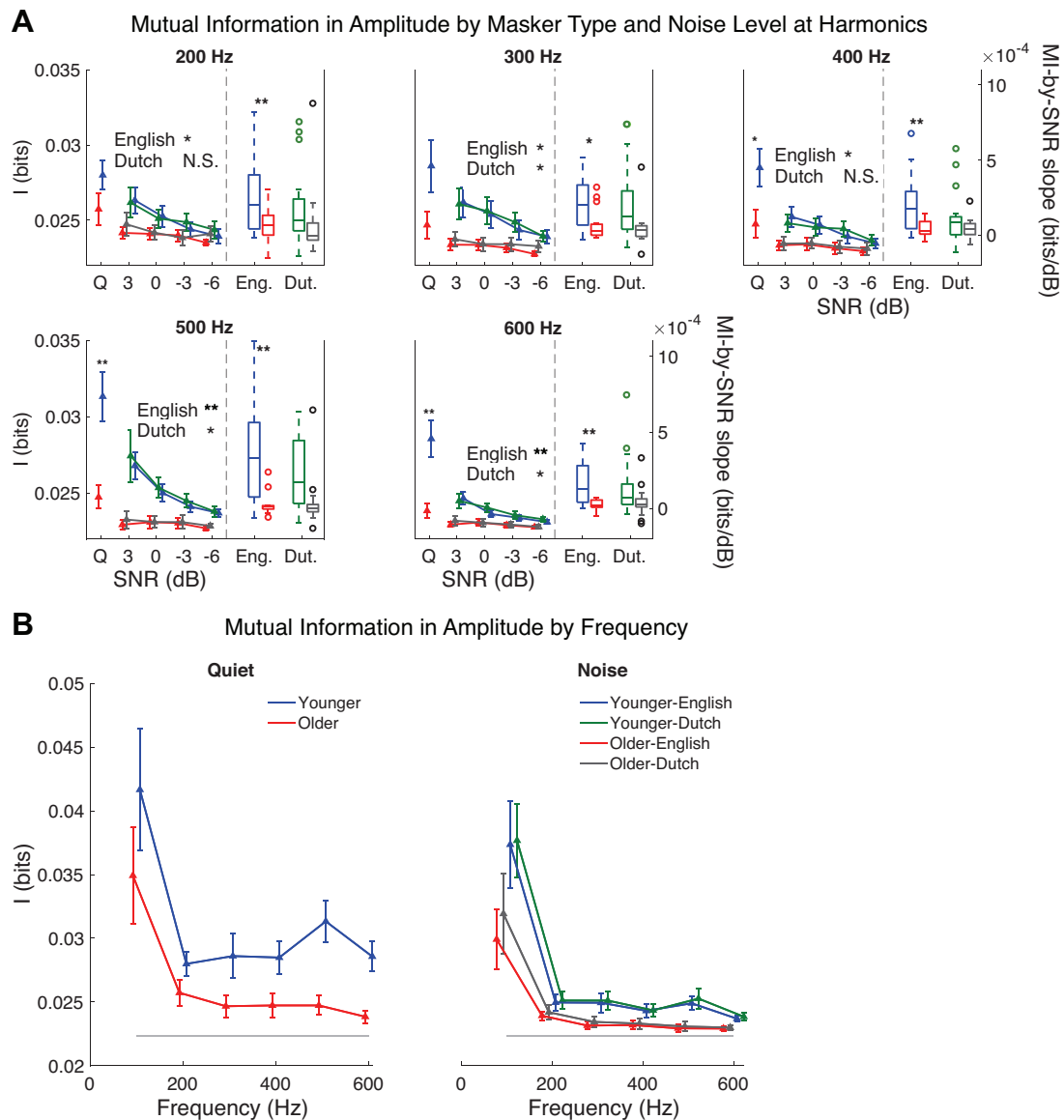
Fig. 3. *A*: mutual information (MI) for amplitude across frequency bands from 200 Hz to 600 Hz (separate subplot for each band). *Left*: the mutual information (*I*) as a function of signal-to-noise ratio (SNR), separately for age group and masker type. For the quiet condition (Q), asterisks above the error bars indicate the significance levels of group differences; text and asterisks above the plots demonstrate significance levels of group differences in the corresponding masker types. Only younger listeners convey a significant amount of information in the higher harmonics. *Right*: the bar plots depict the linearly fitted decreasing slopes (of plots shown on *left*) for the different age groups and masker types. In most bands, the mutual information decreases at a faster rate in younger listeners than in older. *B*: overall, both in quiet (*left*) and averaged over SNR levels (*right*), mutual information decreases with increasing frequency (except for a single increase at 500 Hz for younger listeners). For older listeners, the decreasing trend in mutual information levels off at 300 Hz, which is lower than the frequency (>600 Hz) at which amplitude information levels off in younger listeners. Error bars indicate SE. *$P < 0.05$, **$P < 0.01$. N.S., not significant.

older), both at 600 Hz where the lowest information is observed), i.e., the information for younger listeners has not yet reached the floor by 600 Hz. In contrast, the cutoff frequency for older listeners is 300 Hz: the information measure at 300 Hz is not significantly greater than that at 600 Hz ($t_{14} = 1.32$, $P = 0.130$ under the English masker; $t_{14} = 1.65$, $P = 0.095$ under the Dutch masker). Therefore, the results suggest a lower frequency limit in amplitude information of 300 Hz for older listeners than that of 600 Hz for younger listeners.

Effect of masker type on amplitude information. As seen in Fig. 2*B*, older listeners demonstrate a slower falloff in amplitude information as a function of SNR when the noise masker is Dutch than for English. To test for any potential amplitude information benefit from the Dutch masker over the English masker, the difference in information between the Dutch and English maskers is calculated for each subject in all SNR levels (for both transition and steady-state regions), and a linear model of $I_{Dutch} - I_{English} \sim$ SNR shows a significantly positive intercept for older listeners in the transition region ($t_{57} = 2.35$, $P < 0.001$ with 2 samples omitted) but not in the steady-state region ($t_{56} = 1.38$, $P = 0.173$ with 1 sample omitted). Younger listeners, however, do not show a significant positive intercept in either the transition ($t_{65} = 1.90$, $P = 0.061$ with 1 sample omitted) or steady-state ($t_{66} = -0.60$, $P = 0.549$) region. Samples were

Table 1. *Amplitude information*

| Harmonic, Hz | Quiet (Y > O) | | English Masker (Y > O) | | | | Dutch Masker (Y > O) | | | |
| | | | y-Intercept | | Slope | | y-Intercept | | Slope | |
	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P
100	1.056	0.150	**1.714**	**0.048**	1.275	0.106	1.405	0.102	1.199	0.120
200	1.542	0.080	**1.965**	**0.035**	**2.737**	**0.008**	1.223	0.115	1.262	0.120
300	1.871	0.053	**2.242**	**0.024**	**2.390**	**0.014**	**2.051**	**0.049**	2.019	0.108
400	**2.271**	**0.030**	**2.261**	**0.024**	**2.835**	**0.008**	1.767	0.066	1.502	0.108
500	**3.449**	**0.003**	**3.671**	**0.003**	**3.677**	**0.002**	**2.268**	**0.047**	1.830	0.108
600	**3.412**	**0.003**	**3.340**	**0.003**	**3.565**	**0.002**	**2.259**	**0.047**	1.629	0.108

One-tailed *t* test [younger (Y) > older (O)] results applied to the fitted *y*-intercepts (3 dB values) and slopes from the linear regression analysis of mutual information (for response amplitude) as a function of signal-to-noise ratio for each harmonic. *P* values are corrected for multiple comparisons by false discovery rate correction. Entries in bold indicate that the corresponding tests are statistically significant.

omitted from the tests to satisfy the homoscedasticity requirement. A regression line was fitted as a function of SNR to reduce within-subject variance. With a one-tailed *t* test on the y-intercept (effective mutual information benefit at 3 dB SNR) of the regression line against zero, the mutual information benefit from the Dutch masker over the English masker is significantly higher for older listeners in the transition region ($t_{14} = 2.35$, $P = 0.017$) but not the steady-state region ($t_{14} = 1.67$, $P = 0.058$). No significant benefit is found for younger listeners in either region ($t_{16} = 1.17$, $P = 0.130$ and $t_{16} = 0.51$, $P = 0.307$ for transition and steady-state regions, respectively). The regression slope is not significantly positive or negative for either group ($P > 0.05$ by 2-tailed *t* tests), as seen in the bar plots in Fig. 4, *C* and *D*, *right*.

Phase-Locking Value

Phase-locking value (PLV) is a traditional measure of intertrial coherence for a narrowband response. Figure 5 shows the grand average of PLV at 100 Hz by age and masker condition. Older listeners have lower PLVs than younger listeners ($t_{30} = 2.62$, $P = 0.007$ for 1-tailed *t* test) on the averaged PLVs across time and SNR levels. By one-tailed *t* tests ($PLV_{Dutch} - PLV_{English} > 0$), older listeners have significantly higher PLV under Dutch masking than English ($t_{14} = 2.74$, $P = 0.008$ for transition region; $t_{14} = 1.80$, $P = 0.047$ for steady-state region), whereas younger listeners' PLV is not significantly affected by informational masking ($t_{16} = 1.67$, $P = 0.058$ for transition region; $t_{16} = 0.05$, $P = 0.479$ for steady-state region).

Information in Phase of FFR

Phase information at 100 Hz. For the phase response at 100 Hz, the linear model, $I \sim \text{age} \times \text{masker type} + \text{age} \times \text{SNR}$, is significant ($F_{5,250} = 5.45$, $P < 0.001$ for the entire region; $F_{5,248} = 3.27$, $P < 0.007$ for the transition region; $F_{5,248} = 6.24$, $P < 0.001$ for the steady-state region). Outliers are excluded to satisfy homoscedasticity assumption (2 samples from transition region and 2 samples from steady-state region). The results show no significant interactions between age and masker type ($t_{250} = 0.56$, $P = 0.578$ for the entire region; $t_{248} = 0.22$, $P = 0.825$ for the transition region; $t_{248} = 0.06$, $P = 0.954$ for the steady-state region) and between age and SNR ($t_{250} = 0.86$, $P = 0.393$ for the entire region; $t_{248} = 1.05$, $P = 0.297$ for the transition region; $t_{248} = 0.66$, $P = 0.511$ for

the steady-state region). A linear model with no interactions was then constructed and tested, i.e., $I \sim \text{age} + \text{masker type} + \text{SNR}$. The model itself is significant ($F_{3,252} = 8.77$, $P < 0.001$, $F_{3,250} = 5.08$, $P = 0.002$, and $F_{3,250} = 10.32$, $P < 0.001$ for the entire region and the transition and steady-state regions, respectively). Comparisons show that younger listeners' responses contain significantly more information than older listeners' responses in the steady-state region ($t_{252} = 4.52$, $P < 0.001$ for the entire region; $t_{250} = 2.12$, $P = 0.035$ for the transition region; $t_{250} = 5.19$, $P < 0.001$ for the steady-state region) and that information increases as SNR increases ($t_{252} = 2.31$, $P = 0.022$ for the entire region; $t_{250} = 2.63$, $P = 0.009$ for the transition region).

Mutual information between stimulus and response phase is analyzed analogously to that of the response amplitude. Phase information at 100 Hz is examined separately from the higher harmonics. To examine the effect of age and noise level, a linear regression line is fitted for information by SNR for each subject in both noise contents. The fitted *y*-intercept is compared for group differences. A one-tailed *t* test (younger > older) effect size analysis on the *y*-intercept shows a significantly larger amount of information in younger than older listeners for the English masker ($t_{30} = 1.80$, $P = 0.041$, $d = 0.82$, 95% CI = [0.095,1.540]); the difference is not significant for Dutch ($t_{30} = 1.36$, $P = 0.092$, $d = 0.58$, 95% CI = [−0.133,1.284]) (Fig. 6*A*). Both age groups demonstrate decreasing information with worsening SNR: a one-tailed *t* test on the negativity of the regression slope shows information loss; however, the negativity is not significant for older listeners with the Dutch masker ($t_{16} = 3.31$, $P = 0.002$ and $t_{16} = 2.61$, $P = 0.013$ for younger listeners with English and Dutch maskers, respectively; $t_{14} = 2.17$, $P = 0.036$, $d = 0.56$, 95% CI = [$3.19 \times 10^{-5}, +\infty$], $t_{14} = 2.55$, $P = 0.061$, $d = 0.66$, 95% CI = [$3.84 \times 10^{-5}, +\infty$] for older listeners with English and Dutch maskers, respectively) (Fig. 6*B*). No significant difference is seen between the slopes across age groups ($t_{30} = 1.36$, $P = 0.091$ and $t_{30} = 1.34$, $P = 0.095$ for English and Dutch maskers, respectively). All tests have been corrected for multiple comparisons across the six frequency bands.

Phase information in harmonics of 100 Hz. To examine information in the harmonics of 100 Hz, a linear regression line is fitted for mutual information as a function of SNR for each subject under each masker type. One-tailed (younger > older) *t* tests on the *y*-intercept (with FDR correction) suggest that for

Amplitude Information at 100 Hz by Noise Level and Temporal Stage

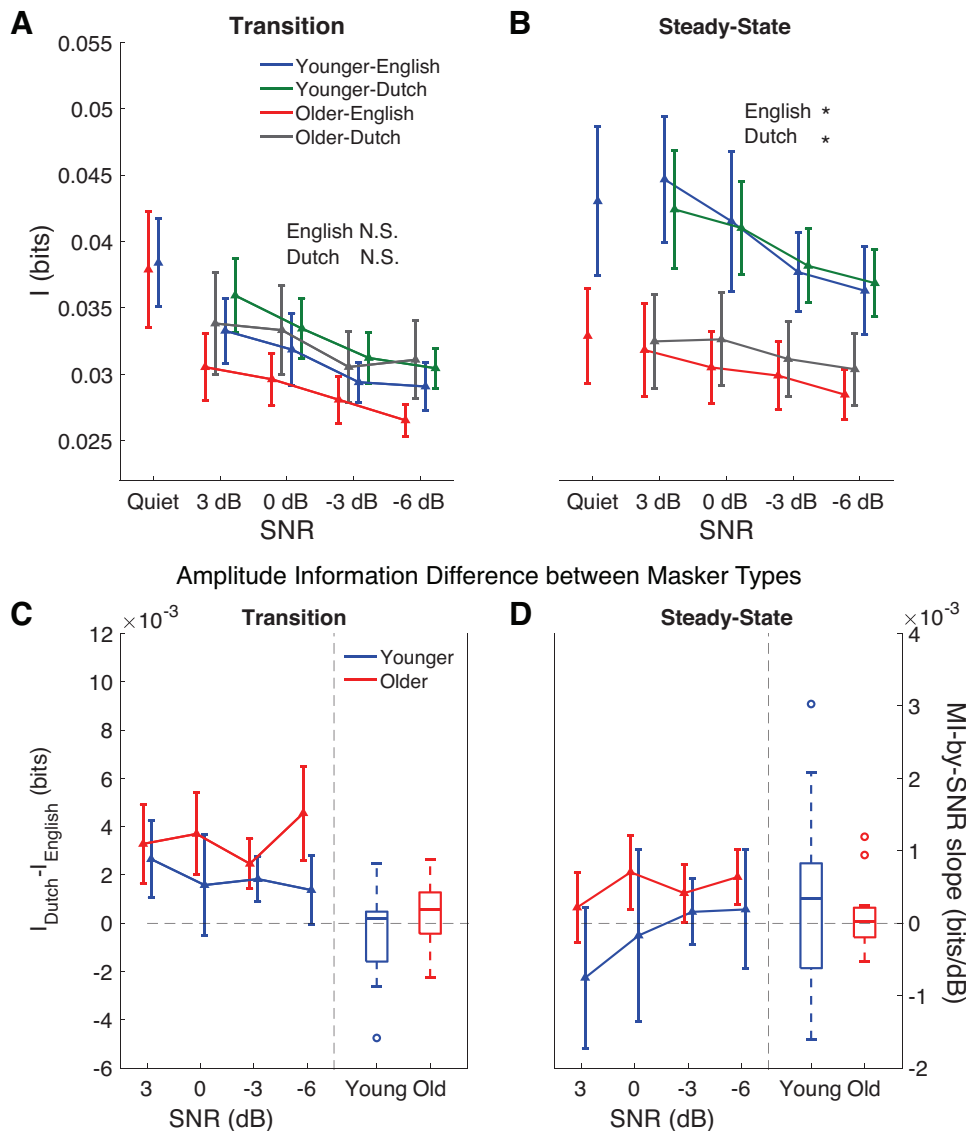## Amplitude Information Difference between Masker Types

Fig. 4. Mutual information of amplitude response by masker type and response region for younger listeners and older listeners with English and Dutch maskers. A and B: mutual information (I) as a function of signal-to-noise ratio (SNR) in the transition (A) and steady-state (B) regions. In the steady-state region, group differences are significant for both masker types, indicated by asterisks. C and D: mutual information (MI) difference between masker types (denoted $I_{Dutch} - I_{English}$) in the transition (C) and steady-state (D) regions. Left: information as a function of SNR. Right: a bar plot showing the slopes of the linear fits. The y-intercepts (corresponding to the fit at 3 dB SNR) are tested against 0 bits. Older listeners show significant benefit from the Dutch masker over English (denoted by asterisk) but only in the transition region. Error bars in all plots indicate SE. *$P < 0.05$. N.S., not significant.

all SNR levels the aging midbrain contains significantly less information than the younger midbrain in all frequencies from 100 to 600 Hz (Fig. 7A). For P values near 0.05 (see Table 2), effect size analysis is further applied. For the English masker condition the 100 and 200 Hz cases show consistent significance from both tests ($t_{30} = 1.80$, $P = 0.041$, $d = 0.82$, 95% CI = [0.095,1.541] and $t_{30} = 1.83$, $P = 0.041$, $d = 1.06$, 95% CI = [0.317,1.799]) and similarly for the Dutch masker condition at 300, 400, and 500 Hz, respectively ($t_{30} = 2.12$, $P = 0.042$, $d = 1.39$, 95% CI = [0.613,2.159], $t_{30} = 1.97$, $P = 0.044$, $d = 0.84$, 95% CI = [0.116,1.564], and $t_{30} = 2.28$, $P = 0.042$, $d = 1.64$, 95% CI = [0.838,2.443]) (see also Fig. 7A). The results show significant decreasing slope in both groups and show that the decrease with worsening SNR is faster for younger listeners than older listeners.

Phase information frequency limits. As seen in Fig. 7B, the stimulus information contained in the response phase decreases with frequency for both age groups. Similar to amplitude analysis, the frequency-decreasing measure used here is phase

information of y-intercept at 3 dB of the fitted mutual information-by-SNR regression line. The measure at 600 Hz for older listeners is not statistically distinguishable from the noise floor ($t_{14} = 0.11$, $P = 0.917$ by 1-sample t test). For younger listeners, the measure is significantly higher than the noise floor at all frequencies ($t_{30} = 3.74$, $P < 0.001$ for English masker; $t_{30} = 2.69$, $P = 0.007$ for Dutch masker (younger > older), both at 600 Hz where lowest information is observed), i.e., the information for younger listeners has not yet reached the floor by 600 Hz. In contrast, the cutoff frequency for older listeners is 500 Hz: the information measure at 500 Hz is not significantly greater than that at 600 Hz ($t_{14} = 0.74$, $P = 0.235$ under English masker; $t_{14} = 1.07$, $P = 0.152$ under Dutch masker). Therefore, the results suggest a lower frequency limit of 500 Hz for older listeners than beyond 600 Hz for younger listeners.

Effect of masker type on phase information. As seen in Fig. 6B, older listeners demonstrate a slower falloff in phase information as a function of SNR when the noise masker is Dutch

Fig. 5. The phase-locking value (PLV) of the 100 Hz frequency-following response (FFR) is shown for all signal-to-noise ratio (SNR) levels, averaged across subjects, with colors indicating age and masker language. *A–D*: the 4 SNR levels: 3, 0, −3, −6 dB. Younger listeners have visibly higher phase locking than older listeners. Older listeners have significantly better phase locking for the Dutch masker than for the English.
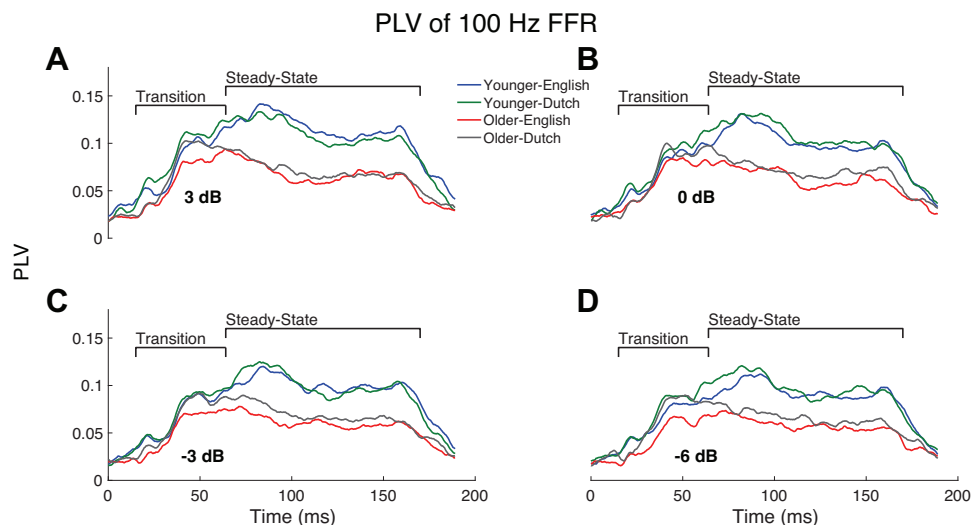
than for English. Analogous to amplitude analysis, the difference in mutual information between the Dutch and English maskers is calculated for each subject in all SNR levels (for both transition and steady-state regions) to examine phase information benefit from the Dutch masker over the English masker, and a linear model of $I_{\text{Dutch}} - I_{\text{English}} \sim$ SNR shows a significantly positive intercept for older listeners in the transition region ($t_{56} = 4.64$, $P < 0.001$ with 2 samples omitted) but not in the steady-state region ($t_{54} = 1.77$, $P = 0.083$ with 4 samples omitted). Younger listeners, however, do not show significant positive intercept in either the transition ($t_{64} = 1.75$, $P = 0.085$ with 2 samples omitted) or steady-state ($t_{66} = -0.64$, $P = 0.522$) region. Samples were omitted from the tests to satisfy the homoscedasticity requirement. For justification, a regression line was fitted as a function of SNR to reduce within-subject variance. With a one-tailed t test on the y-intercept (effective mutual information benefit at 3 dB SNR) of the regression line against zero, the mutual information benefit from the Dutch masker over the English masker is significantly higher for older listeners in the transition region ($t_{14} = 2.31$, $P = 0.018$) but not the

steady-state region ($t_{14} = 1.55$, $P = 0.072$). No significant benefit is found for younger listeners in either region ($t_{16} = 1.33$, $P = 0.102$ and $t_{16} = 0.44$, $P = 0.332$ for transition and steady-state regions, respectively). The regression slope is not significantly positive or negative for either group ($P > 0.05$ by 2-tailed t tests), as seen in the bar plots in Fig. 8, *C* and *D*, *right*.

DISCUSSION

On the basis of these results from the mutual information analysis of FFR amplitude and phase, we have provided supporting evidence that the neural response of the midbrain of older listeners is not merely less well synchronized than for younger listeners (Anderson et al. 2012; Presacco et al. 2016a, 2016b) but also actually contains less information, in both amplitude and phase. At the fundamental frequency, the informational loss for older listeners was seen only in the presence of a competing talker. In contrast, for higher frequencies the informational loss for older listeners was seen in both quiet and noisy conditions. Furthermore, the

Fig. 6. Mutual information between the stimulus and response phase as a function of noise level for each age group and masker condition (masker language). *A*: mutual information (*I*) at the fundamental frequency as a function of noise level. The response in younger listeners conveys noticeably more information than the response in older listeners for the English masker condition, but the difference for Dutch is not significant at 100 Hz. Older listeners show consistently higher mutual information for the Dutch masker than for the English (the younger listeners show no consistent difference), but the difference is not significant at 100 Hz. *B*: the mutual information (MI)-by-signal-to-noise ratio (SNR) slopes of the plots in *A* show decreasing trends as SNR worsens, regardless of masker type, for both age groups. Younger listeners show a steeper decrease than older listeners, but the difference is not significant at the 100 Hz response. Error bars indicate SE. *$P < 0.05$. N.S., not significant.
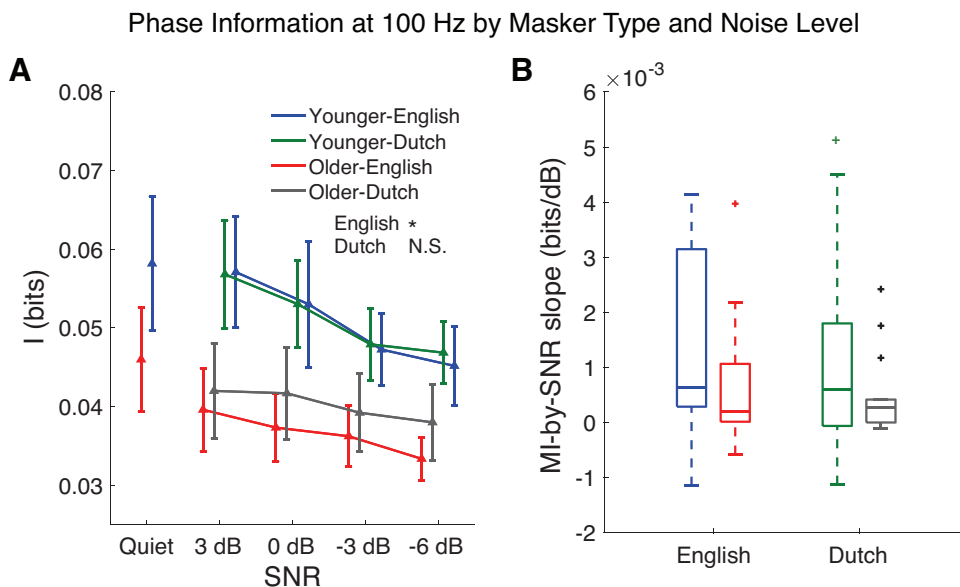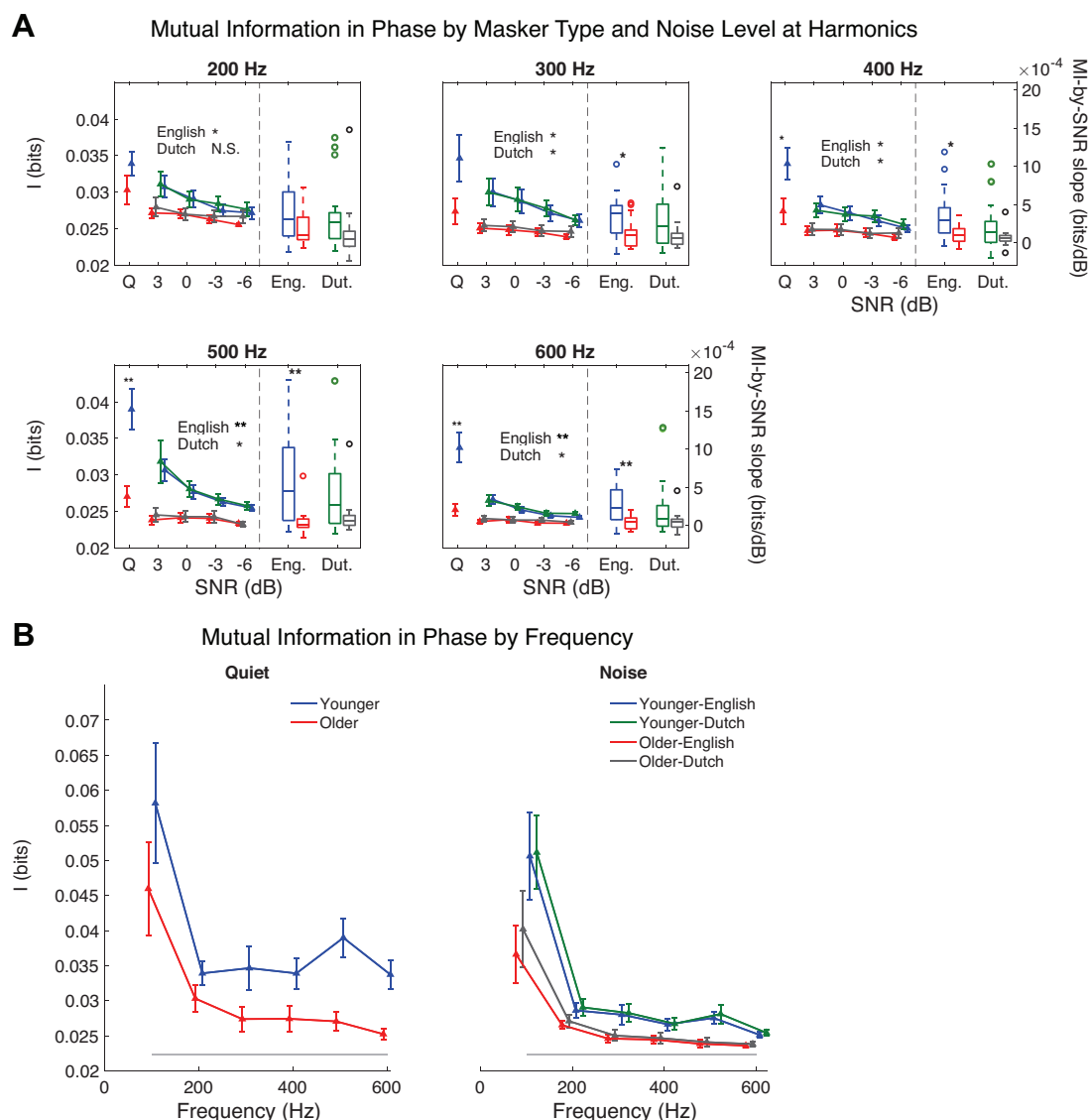
Fig. 7. *A*: mutual information (*I*) for phase across frequency bands from 200 Hz to 600 Hz (separate subplot for each band). Within each subplot, as in Fig. 3, the mutual information (MI) as a function of signal-to-noise ratio (SNR), separately for age group and masker type is shown at *left*; on *right*, the bar plots depict the linearly fitted decreasing slopes (of plots shown at *left*) for the different age groups and masker types. Q, quiet condition. *B*: overall, both in quiet (*left*) and averaged over SNR levels (*right*), mutual information decreases with increasing frequency (except for a single increase at 500 Hz for younger listeners). For older listeners, the decreasing trend in mutual information levels off at 500 Hz, which is lower than the frequency at which phase information levels off in younger listeners. Error bars indicate SE. *$P < 0.05$, **$P < 0.01$. N.S., not significant.

masker type (Dutch vs. English) significantly affects the amount of stimulus information carried in the response at the fundamental frequency in the transition region for older listeners but not younger listeners. This last finding arises for the first time from this mutual information analysis and demonstrates that mutual information analysis provides access to response properties otherwise hidden by response variability.

Aging

Aging has different effects on subcortical and cortical auditory stages along the ascending pathway. Here we address its effect on midbrain representations of FFR from an information point of view. First we show a broadband (100–600 Hz) informational loss associated with aging in both quiet and noisy conditions, which is reflected in both

the amplitude and phase of the responses. The informational loss at the fundamental frequency can be attributed to the delayed and weakened responses in the aging midbrain (Anderson et al. 2012; Burkard and Sims 2002; Clinard and Tremblay 2013), which can be linked to age-related loss of inhibition. For example, DCN has been shown to represent signal and suppress background noise aided by glycinergic neurotransmitters, and aging rats display decreased glycinergic inhibition in DCN (Caspary et al. 2005, 2006). Another contribution may come from synaptopathy arising from a loss of inner hair cell ribbons and degeneration of ganglion cells (Sergeyenko et al. 2013) or from a decline in low-spontaneous-rate nerve fibers as has been seen in aging gerbils (Schmiedt et al. 1996). Together, synaptopathy and loss of inhibition in midbrain may both contribute to less information in midbrain FFR in older listeners.

Table 2. *Phase information*

Harmonic, Hz	Quiet (Y > O)		English Masker (Y > O)				Dutch Masker (Y > O)			
			y-Intercept		Slope		y-Intercept		Slope	
	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P
100	1.072	0.146	**1.798**	**0.041**	1.363	0.092	1.526	0.069	1.344	0.095
200	1.386	0.106	**1.833**	**0.041**	1.757	0.053	1.530	0.069	1.479	0.090
300	**1.898**	**0.050**	**2.219**	**0.026**	**2.089**	**0.034**	**2.122**	**0.042**	1.909	0.090
400	**2.170**	**0.038**	**2.407**	**0.022**	**2.694**	**0.011**	**1.967**	**0.044**	1.493	0.090
500	**3.609**	**0.002**	**3.740**	**0.001**	**3.352**	**0.003**	**2.280**	**0.042**	1.615	0.090
600	**3.579**	**0.002**	**3.738**	**0.001**	**3.446**	**0.003**	**2.690**	**0.035**	1.716	0.090

One-tailed *t* test [younger (Y) > older (O)] results applied to the fitted *y*-intercepts (3 dB values) and slopes from the linear regression analysis of mutual information (for response phase) as a function of signal-to-noise ratio for each harmonic. *P* values are corrected for multiple comparisons by false discovery rate correction. Entries in bold indicate that the corresponding tests are statistically significant.

Noise Level

In these results, the amount of information in FFR (both phase and amplitude) decreases as noise level increases (i.e., SNR decreases) for both younger and older listeners. This result is consistent with previous findings (Presacco et al. 2016a, 2016b) where the amplitude of FFR decreases with worsening noise level. Via linear regression, it is also seen that younger listeners have a more steeply decreasing slope (as a function of noise level) than older listeners, at both the fundamental frequency and its harmonics. This result may also be due to disrupted synchrony at auditory nerve fibers (Schmiedt et al. 1996) and the synapse (Sergeyenko et al. 2013). A loss of auditory nerve fibers in older listeners may lead to a reduced brain stem response, causing a decrease in information even in the quiet condition, leading to a slower rate of additional decrease with increasing noise level.

Masker Type

In this experiment background masker types included English (meaningful to all listeners) and Dutch (meaningless to all listeners). The results suggest that the informational content of the noise affects information in the midbrain FFR, in both amplitude and phase (in the transition region): older listeners benefit neurally from the masker being meaningless over meaningful. It is unexpected that a high-level feature such as language would affect midbrain neural responses, although this has been seen before for younger listeners (Presacco et al. 2016b). One explanation for the language-dependent response difference in the aging midbrain could be top-down modulation from cortical areas. Descending pathways from primary auditory cortex to inferior colliculus (IC) in the midbrain have been reported to mediate learning-induced auditory plasticity (Bajo et al. 2010), and IC neurons' sensitivity to sound frequency and intensity can be modified by cortical projections (Bajo and King 2013). Since older listeners benefit behaviorally from competing speech being nonmeaningful (Pichora-Fuller 2008; Tun et al. 2002), the cortical processing underlying this difference may also project back upstream to the midbrain.

Another explanation for this difference in FFR due to masking language is that the difference might be purely cortical, i.e., purely cortical FFR. Recent studies (Coffey et al. 2016, 2017) have shown that traditional EEG-measured FFR may not be purely subcortical at all. It would be substantially less surprising to see language-specific effects originating from cortex than midbrain, although, even so, these effects from the tran-

sition region (15–65 ms) are earlier than might be expected from a language-influenced cortical response.

High-Frequency Limit

We show that for both amplitude and phase information responses from older listeners in speech-in-noise conditions contain less information in the higher frequencies, and have lower high-frequency limits, than younger listeners. Such deficits might be also associated with lowered temporal precision arising from a loss of auditory nerve fibers and ganglion cells (Schmiedt et al. 1996; Sergeyenko et al. 2013), which affect all frequencies. The same analysis carried out on single sweeps (see APPENDIX) suggests that the decrease in information at high frequencies may not be due to the average of the two polarities.

Relation to Cortical Representation

Even though the stimulus representation at the level of auditory midbrain is weaker for older listeners, whether based on root mean square, correlation, or mutual information measures, it is paradoxically amplified at the level of auditory cortex (Brodbeck et al. 2018; Presacco et al. 2016a, 2016b). A negative association between subcortical FFR and cortical responses, as measured with mutual information, has been shown in older listeners in a task of categorical syllable perception (Bidelman et al. 2014). The analogous correlation between cortical speech representation and midbrain response amplitude was not seen, however, for temporal speech processing (Presacco et al. 2016b). Both attention and behavioral inhibition are used to enhance understanding of speech in noise, but the extent to which these high-level cortical processes are altered by auditory periphery deficits is not well known (Presacco et al. 2019). Furthermore, it is unclear where and how the neural representation of speech in older listeners shifts from degraded in midbrain to exaggerated in cortex, but mutual information is a promising tool to address these issues (Bidelman et al. 2014).

Summary

The approach employed here, using mutual information to analyze the relationship between a speech-in-noise stimulus and the FFR response, can be seen in at least two different lights. At one level it can be viewed as a mathematical measure derived from information theory (Cover and Thomas 1991; Shannon 1948). This places the present analysis on firm mathematical grounds, using concepts and measures from a well-

Phase Information at 100 Hz by Noise Level and Temporal Stage

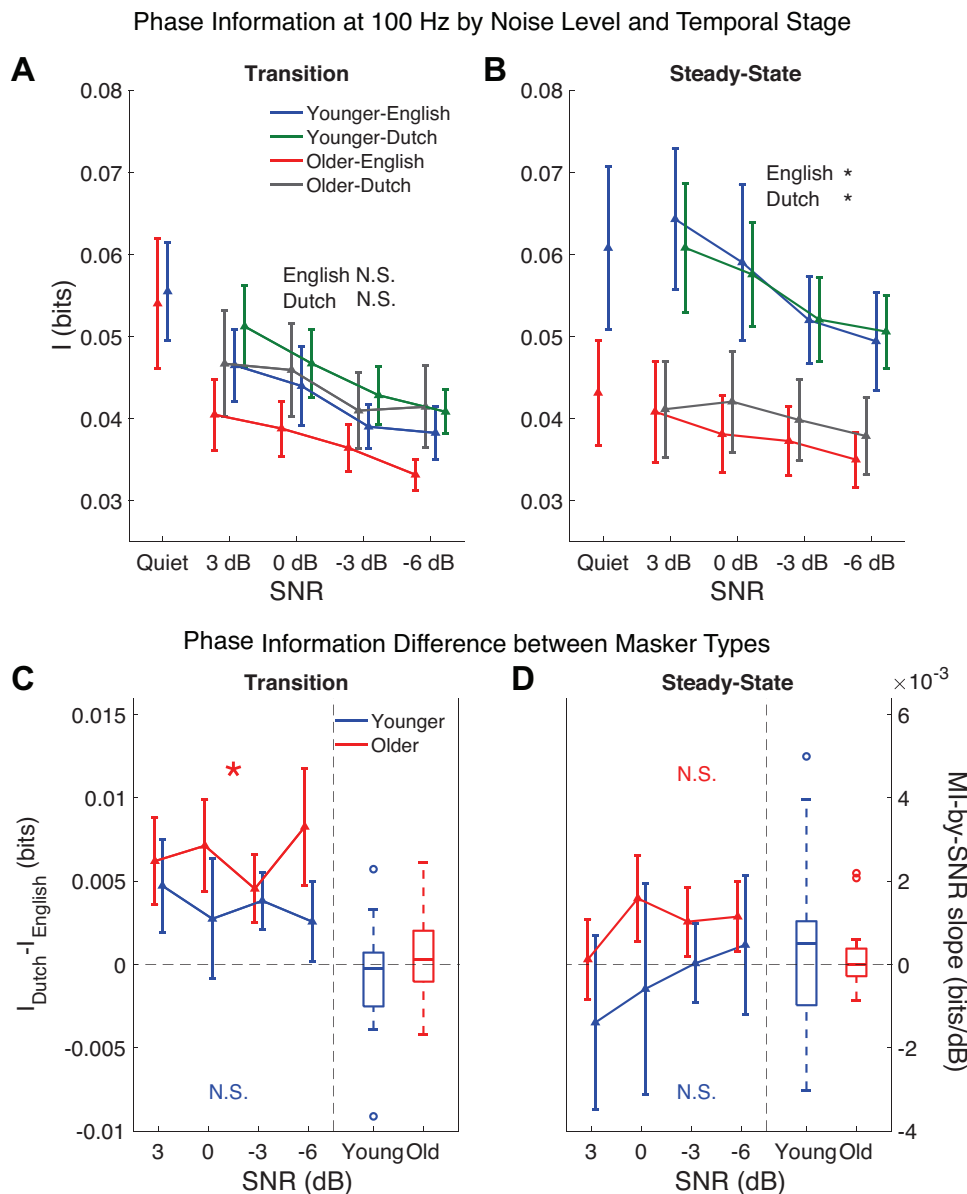Phase Information Difference between Masker Types

Fig. 8. Mutual information of phase response by masker type and response region for younger listeners and older listeners with English and Dutch maskers. *A* and *B*: mutual information (I) as a function of signal-to-noise ratio (SNR) in the transition (*A*) and steady-state (*B*) regions. In the steady-state region, group differences are significant for both masker types, indicated by asterisks. *C* and *D*: the mutual information (MI) difference between masker types (denoted $I_{Dutch} - I_{English}$) in the transition (*C*) and steady-state (*D*) regions. *Left*: information as a function of SNR. *Right*: a bar plot showing the slopes of the linear fits. The *y*-intercepts (corresponding to the fit at 3 dB SNR) are tested against 0 bits. Older listeners show significant benefit from the Dutch masker over English (denoted by asterisk) but only in the transition region. Error bars in all plots indicate SE. *$P < 0.05$. N.S, not significant.

established field of mathematical signal processing. At another level, the analysis can be viewed as an acknowledgment that the relationship between stimulus and response may have strongly nonlinear aspects, with mutual information being just one of several available nonlinear measures that allow us to move beyond conventional linear analysis methods (e.g., evoked response analysis) and conventional phase coherence methods.

APPENDIX: RESULTS WITHOUT AVERAGING POLARITIES

Analogously to the case of averaged polarities presented above, even without such polarity averaging older listeners still demonstrate a slower falloff in information as a function of SNR when the noise masker is Dutch than for English.

Information in Amplitude of FFR Without Averaging Polarities

For amplitude information, a regression line was fitted as a function of SNR to reduce within-subject variance. With a one-tailed *t* test on the

y-intercept (effective mutual information benefit at 3 dB SNR) of the regression line against zero, the mutual information in amplitude benefit from the Dutch masker over the English masker is significantly higher for older listeners in the transition region ($t_{14} = 1.80$, $P = 0.046$) but not the steady-state region ($t_{14} = 1.61$, $P = 0.065$); no significant benefit is found for younger listeners in either region ($t_{16} = 1.04$, $P = 0.156$ and $t_{16} = 0.16$, $P = 0.439$ for transition and steady-state regions, respectively) (see Table A1 for details and for harmonic frequency analysis). The regression slope is not significantly positive or negative for either group ($P > 0.05$ by 2-tailed *t* tests), as seen in the bar plots in Fig. A1, *C* and *D*, *right*.

Information in Phase of FFR Without Averaging Polarities

Similarly, for phase information, a regression line was fitted as a function of SNR to reduce within-subject variance. With a one-tailed *t* test on the *y*-intercept (effective mutual information benefit at 3 dB SNR) of the regression line against zero, the mutual information in phase benefit from the Dutch masker over the English masker is significantly higher for older listeners in the transition region ($t_{14} = 1.90$, $P = 0.039$) but not the steady-state region ($t_{14} = 1.45$, $P = 0.085$); no significant benefit is found for younger listeners in either region ($t_{16} = 1.04$, $P = 0.156$ and $t_{16} =$

Table A1. *Amplitude information*

Harmonic, Hz	Quiet (Y > O)		English Masker (Y > O)				Dutch Masker (Y > O)			
			y-Intercept		Slope		y-Intercept		Slope	
	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P
100	0.982	0.167	**1.700**	**0.050**	1.287	0.104	1.238	0.113	1.254	0.110
200	1.544	0.080	**1.918**	**0.039**	**2.583**	**0.011**	1.338	0.113	1.619	0.087
300	1.862	0.054	**2.161**	**0.029**	**2.060**	**0.029**	**2.138**	**0.041**	2.185	0.087
400	**2.441**	**0.021**	**2.380**	**0.024**	**2.699**	**0.011**	1.795	0.062	1.670	0.087
500	**3.466**	**0.002**	**3.640**	**0.003**	**3.612**	**0.002**	**2.247**	**0.041**	1.696	0.087
600	**3.536**	**0.002**	**3.370**	**0.003**	**3.546**	**0.002**	**2.168**	**0.041**	1.281	0.110

One-tailed *t* test [younger (Y) > older (O)] results applied to the fitted y-intercepts (3 dB values) and slopes from the linear regression analysis of mutual information (for response amplitude) as a function of signal-to-noise ratio for each harmonic. *P* values are corrected for multiple comparisons by false discovery rate correction. Entries in bold indicate that the corresponding tests are statistically significant.

0.25, *P* = 0.401 for transition and steady-state regions, respectively) (see Table A2 for details and for harmonic frequency analysis). The regression slope is not significantly positive or negative for either group (*P* > 0.05 by 2-tailed *t* tests), as seen in the bar plots in Fig. A2, *C* and *D*, *right*.

Fig. A1. Mutual information of amplitude response by masker type and response region for younger listeners and older listeners with English and Dutch maskers. *A* and *B*: the mutual information (*I*) as a function of signal-to-noise ratio (SNR) in the transition (*A*) and steady-state (*B*) regions. In the steady-state region, group differences are significant for only the English masker, indicated by asterisks. *C* and *D*: the mutual information (MI) difference between masker types (denoted $I_{Dutch} - I_{English}$) in the transition (*C*) and steady-state (*D*) regions. *Left*: information as a function of SNR. *Right*: a bar plot showing the slopes of the linear fits. The y-intercepts (corresponding to the fit at 3 dB SNR) are tested against 0 bits. Older listeners show significant benefit from the Dutch masker over English (denoted by asterisk) but only in the transition region. Error bars in all plots indicate SE. *$P < 0.05$. N.S., not significant.
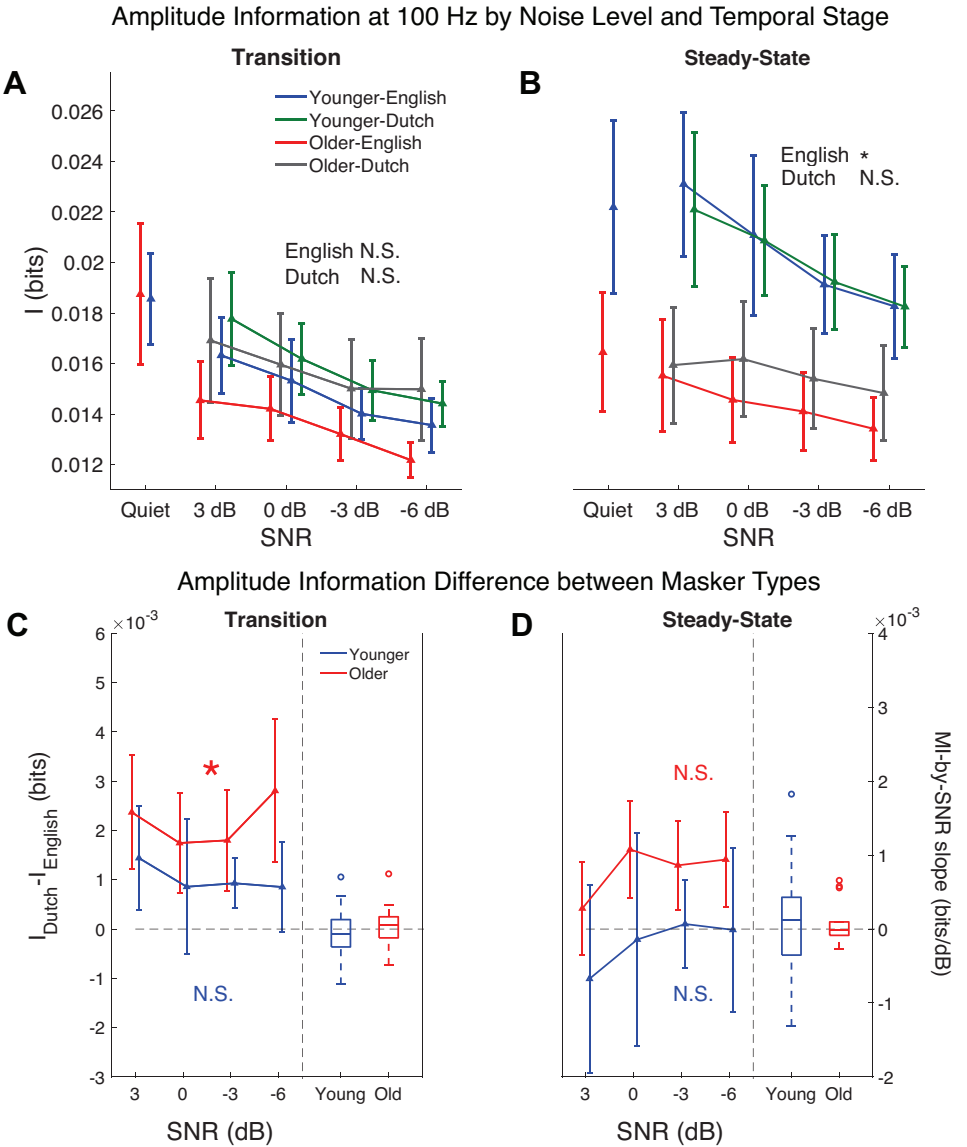
Table A2. *Phase information*

Harmonic, Hz	Quiet (Y > O)		English Masker (Y > O)				Dutch Masker (Y > O)			
			y-Intercept		Slope		y-Intercept		Slope	
	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P	t_{30}	P
100	1.005	0.161	**1.758**	**0.044**	1.334	0.096	1.313	0.100	1.302	0.101
200	1.514	0.084	**2.047**	**0.030**	**1.962**	**0.035**	1.782	0.051	1.947	0.061
300	1.822	0.059	**2.300**	**0.021**	**2.008**	**0.035**	**2.199**	**0.034**	2.167	0.061
400	**2.400**	**0.023**	**2.537**	**0.017**	**2.512**	**0.018**	**2.088**	**0.034**	2.054	0.061
500	**3.653**	**0.001**	**3.865**	**0.001**	**3.641**	**0.002**	**2.204**	**0.034**	1.556	0.081
600	**3.701**	**0.001**	**3.677**	**0.001**	**3.904**	**0.001**	**2.619**	**0.034**	1.533	0.081

One-tailed t test [younger (Y) > older (O)] results applied to the fitted y-intercepts (3 dB values) and slopes from the linear regression analysis of mutual information (for response phase) as a function of signal-to-noise ratio for each harmonic. P values are corrected for multiple comparisons by false discovery rate correction. Entries in bold indicate that the corresponding tests are statistically significant.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

P.Z., A.P., S.A., and J.Z.S. conceived and designed research; A.P. performed experiments; P.Z., A.P., and J.Z.S. analyzed data; P.Z., A.P., S.A., and J.Z.S. interpreted results of experiments; P.Z. and J.Z.S. prepared figures; P.Z. drafted manuscript; P.Z., A.P., S.A., and J.Z.S. edited and revised manuscript; P.Z., A.P., S.A., and J.Z.S. approved final version of manuscript.

REFERENCES

Aiken SJ, Picton TW. Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245: 35–47, 2008. doi:10.1016/j.heares.2008. 08.004.

Anderson S, Parbery-Clark A, White-Schwoch T, Drehobl S, Kraus N. Effects of hearing loss on the subcortical representation of speech cues. *J Acoust Soc Am* 133: 3030–3038, 2013. doi:10.1121/1.4799804.

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. *J Neurosci* 32: 14156–14164, 2012. doi:10.1523/JNEUROSCI.2176-12.2012.

Bajo VM, King AJ. Cortical modulation of auditory processing in the midbrain. *Front Neural Circuits* 6: 114, 2013. doi:10.3389/fncir.2012. 00114.

Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 13: 253–260, 2010. doi:10.1038/nn.2466.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B* 57: 289–300, 1995. doi:10.1111/j.2517-6161.1995.tb02031.x.

Bidelman GM, Villafuerte JW, Moreno S, Alain C. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging* 35: 2526–2540, 2014. doi:10.1016/j.neurobiolaging.2014. 05.006.

Brodbeck C, Presacco A, Anderson S, Simon JZ. Increased speech representation in older adults originates from early response in higher order auditory cortex (Preprint). *bioRxiv* 294017, 2018. doi:10.1101/ 294017.

Burkard RF, Sims D. A comparison of the effects of broadband masking noise on the auditory brainstem response in young and older adults. *Am J Audiol* 11: 13–22, 2002. doi:10.1044/1059-0889(2002/004).

Burke DM, Shafto MA. Language and aging. In: *The Handbook of Aging and Cognition*, edited by Craik FI, Salthouse TA. New York: Psychology Press, 2008, p. 373–443.

Caspary DM, Hughes LF, Schatteman TA, Turner JG. Age-related changes in the response properties of cartwheel cells in rat dorsal cochlear nucleus. *Hear Res* 216-217: 207–215, 2006. doi:10.1016/j.heares.2006.03. 005.

Caspary DM, Milbrandt JC, Helfert RH. Central auditory aging: GABA changes in the inferior colliculus. *Exp Gerontol* 30: 349–360, 1995. doi:10. 1016/0531-5565(94)00052-5.

Caspary DM, Schatteman TA, Hughes LF. Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: role of inhibitory inputs. *J Neurosci* 25: 10952–10959, 2005. doi:10.1523/ JNEUROSCI.2451-05.2005.

Clinard CG, Tremblay KL. Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *J Am Acad Audiol* 24: 590–599, 2013. doi:10.3766/jaaa.24.7.7.

Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7: 11070, 2016. doi:10.1038/ncomms11070.

Coffey EB, Musacchia G, Zatorre RJ. Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *J Neurosci* 37: 830–838, 2017. doi:10.1523/JNEUROSCI.1265-16.2016.

Cogan GB, Poeppel D. A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *J Neurophysiol* 106: 554–563, 2011. doi:10.1152/jn.00075.2011.

Cover TM, Thomas JA. *Elements of Information Theory*. New York: Wiley, 1991.

de Villers-Sidani E, Alzghoul L, Zhou X, Simpson KL, Lin RC, Merzenich MM. Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. *Proc Natl Acad Sci USA* 107: 13900–13905, 2010. doi:10.1073/pnas.1007885107.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9–21, 2004. doi:10.1016/j.jneumeth.2003.10. 009.

Fitzgibbons PJ, Gordon-Salant S. Auditory temporal processing in elderly listeners. *J Am Acad Audiol* 7: 183–189, 1996.

Fitzgibbons PJ, Gordon-Salant S. Aging and temporal discrimination in auditory sequences. *J Acoust Soc Am* 109: 2955–2963, 2001. doi:10.1121/ 1.1371760.

Frisina DR, Frisina RD. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res* 106: 95–104, 1997. doi:10.1016/S0378-5955(97)00006-3.

Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Barrett J. Age-related differences in identification and discrimination of temporal cues in speech segments. *J Acoust Soc Am* 119: 2455–2466, 2006. doi:10.1121/1. 2171527.

He NJ, Mills JH, Ahlstrom JB, Dubno JR. Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J Acoust Soc Am* 124: 3841–3849, 2008. doi:10.1121/1.2998779.

Helfer KS, Freyman RL. Aging and speech-on-speech masking. *Ear Hear* 29: 87–98, 2008. doi:10.1097/AUD.0b013e31815d638b.

Juarez-Salinas DL, Engle JR, Navarro XO, Recanzone GH. Hierarchical and serial processing in the spatial auditory cortical pathway is degraded by natural aging. *J Neurosci* 30: 14795–14804, 2010. doi:10.1523/JNEURO-SCI.3393-10.2010.

Klatt DH. Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 67: 971, 1980. doi:10.1121/1.383940.

Lister JJ, Maxfield ND, Pitt GJ, Gonzalez VB. Auditory evoked response to gaps in noise: older adults. *Int J Audiol* 50: 211–225, 2011. doi:10.3109/ 14992027.2010.526967.

Maddox RK, Lee AK. Auditory brainstem responses to continuous natural speech in human listeners. *eNeuro* 5: ENEURO.0441-17.2018, 2018. doi: 10.1523/ENEURO.0441-17.2018.
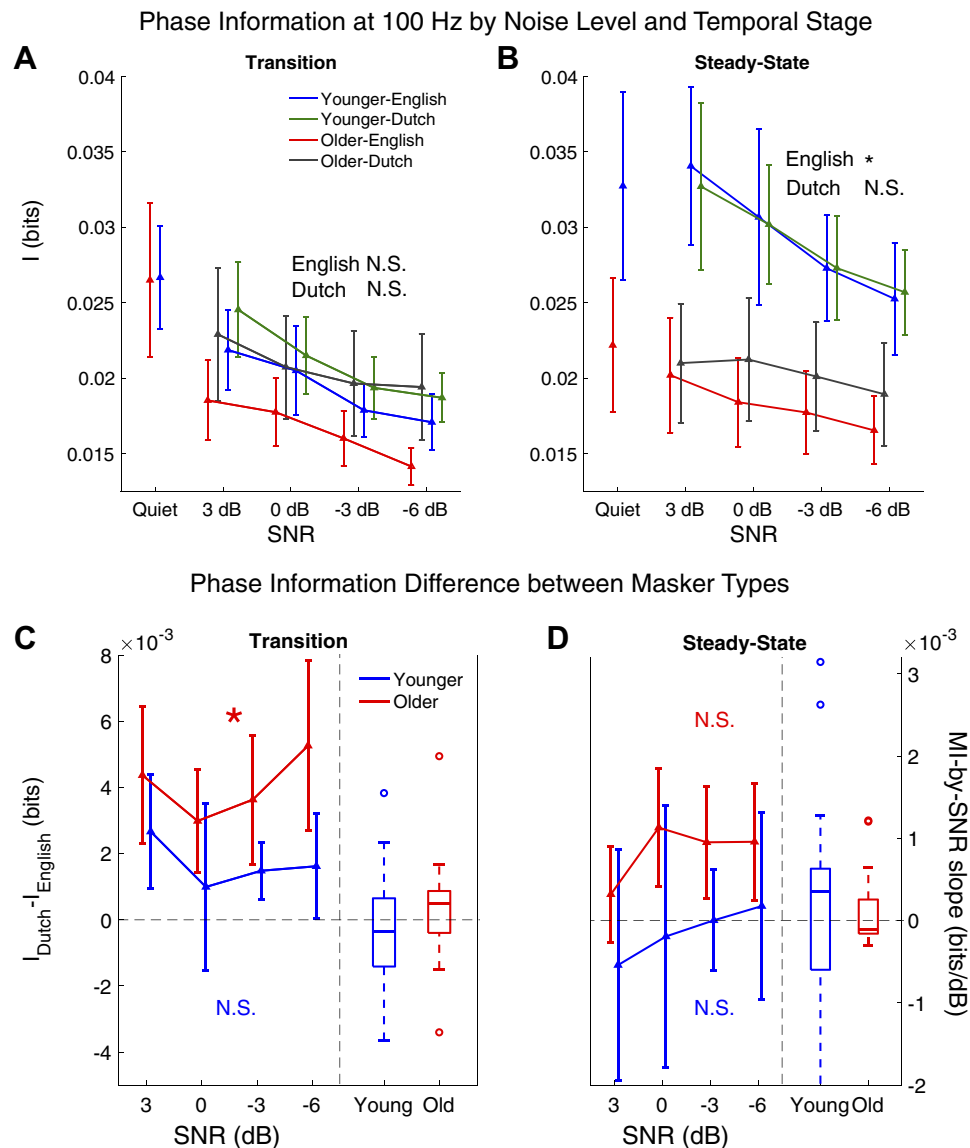
Fig. A2. Mutual information of phase response by masker type and response region for younger listeners and older listeners with English and Dutch maskers. *A* and *B*: the mutual information (*I*) as a function of signal-to-noise ratio (SNR) in the transition (*A*) and steady-stage (*B*) regions. In the steady-state region, group differences are significant for only the English masker, indicated by asterisks. *C* and *D*: the mutual information (MI) difference between masker types (denoted $I_{Dutch} - I_{English}$) in the transition (*C*) and steady-state (*D*) regions. *Left*: information as a function of SNR. *Right*: a bar plot showing the slopes of the linear fits. The *y*-intercepts (corresponding to the fit at 3 dB SNR) are tested against 0 bits. Older listeners show significant benefit from the Dutch masker over English (denoted by asterisk) but only in the transition region. Error bars in all plots indicate SE. *$P < 0.05$. N.S., not significant.

Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27: 338–352, 1955. [Erratum in *J Acoust Soc Am* 27: 617, 1955.] doi:10.1121/1.1907526.

Nelken I, Chechik G. Information theory in auditory research. *Hear Res* 229: 94–105, 2007. doi:10.1016/j.heares.2007.01.012.

Parthasarathy A, Bartlett EL. Age-related auditory deficits in temporal processing in F-344 rats. *Neuroscience* 192: 619–630, 2011. doi:10.1016/j.neuroscience.2011.06.042.

Peña EA, Slate EH. Global validation of linear model assumptions. *J Am Stat Assoc* 101: 341–354, 2006. doi:10.1198/016214505000000637.

Pichora-Fuller MK. Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int J Audiol* 47, Suppl 2: S72–S82, 2008. doi:10.1080/14992020802307404.

Presacco A, Simon JZ, Anderson S. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J Neurophysiol* 116: 2346–2355, 2016a. doi:10.1152/jn.00372.2016.

Presacco A, Simon JZ, Anderson S. Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol* 116: 2356–2367, 2016b. doi:10.1152/jn.00373.2016.

Presacco A, Simon JZ, Anderson S. Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS One* 14: e0213899, 2019. doi:10.1371/journal.pone.0213899.

R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing, 2017. https://www.R-project.org/.

Rieke F, Bodnar DA, Bialek W. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc Biol Sci* 262: 259–265, 1995. doi:10.1098/rspb.1995.0204.

Schatteman TA, Hughes LF, Caspary DM. Aged-related loss of temporal processing: altered responses to amplitude modulated tones in rat dorsal cochlear nucleus. *Neuroscience* 154: 329–337, 2008. doi:10.1016/j.neuroscience.2008.02.025.

Schmiedt RA, Mills JH, Boettcher FA. Age-related loss of activity of auditory-nerve fibers. *J Neurophysiol* 76: 2799–2803, 1996. doi:10.1152/jn.1996.76.4.2799.

Schneider BA, Hamstra SJ. Gap detection thresholds as a function of tonal duration for younger and older listeners. *J Acoust Soc Am* 106: 371–380, 1999. doi:10.1121/1.427062.

Sergeyenko Y, Lall K, Liberman MC, Kujawa SG. Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *J Neurosci* 33: 13686–13694, 2013. doi:10.1523/JNEUROSCI.1783-13.2013.

Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 27: 379–423 1948.

Tun PA, O'Kane G, Wingfield A. Distraction by competing speech in young and older adult listeners. *Psychol Aging* 17: 453–467, 2002. doi:10.1037/0882-7974.17.3.453.

Walton JP, Frisina RD, O'Neill WE. Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse.

J Neurosci 18: 2764–2776, 1998. doi:10.1523/JNEUROSCI.18-07-02764.1998.

Wang H, Turner JG, Ling L, Parrish JL, Hughes LF, Caspary DM. Age-related changes in glycine receptor subunit composition and binding in dorsal cochlear nucleus. *Neuroscience* 160: 227–239, 2009. doi:10.1016/j.neuroscience.2009.01.079.

Yang X, Wang K, Shamma SA. Auditory representations of acoustic signals. *IEEE Trans Inf Theory* 38: 824–839, 1992. doi:10.1109/18.119739.

Zhu L, Bharadwaj H, Xia J, Shinn-Cunningham B. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *J Acoust Soc Am* 134: 384–395, 2013. doi:10.1121/1.4807498.