

# A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones

Li Zhu

Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, 100084, People's Republic of China

Hari Bharadwaj, Jing Xia, and Barbara Shinn-Cunningham<sup>a)</sup>

Center for Computational Neuroscience and Neural Technology, Boston University, 677 Beacon Street, Boston, Massachusetts 02215

(Received 25 July 2012; revised 7 April 2013; accepted 23 April 2013)

Two experiments, both presenting diotic, harmonic tone complexes (100 Hz fundamental), were conducted to explore the envelope-related component of the frequency-following response ( $\text{FFR}_{\text{ENV}}$ ), a measure of synchronous, subcortical neural activity evoked by a periodic acoustic input. Experiment 1 directly compared two common analysis methods, computing the magnitude spectrum and the phase-locking value (PLV). Bootstrapping identified which  $\text{FFR}_{\text{ENV}}$  frequency components were statistically above the noise floor for each metric and quantified the statistical power of the approaches. Across listeners and conditions, the two methods produced highly correlated results. However, PLV analysis required fewer processing stages to produce readily interpretable results. Moreover, at the fundamental frequency of the input, PLVs were farther above the metric's noise floor than spectral magnitudes. Having established the advantages of PLV analysis, the efficacy of the approach was further demonstrated by investigating how different acoustic frequencies contribute to  $\text{FFR}_{\text{ENV}}$ , analyzing responses to complex tones composed of different acoustic harmonics of 100 Hz (Experiment 2). Results show that the  $\text{FFR}_{\text{ENV}}$  response is dominated by peripheral auditory channels responding to unresolved harmonics, although low-frequency channels driven by resolved harmonics also contribute. These results demonstrate the utility of the PLV for quantifying the strength of  $\text{FFR}_{\text{ENV}}$  across conditions. © 2013 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4807498>]

PACS number(s): 43.64.Ri [TD]

Pages: 384–395

## I. INTRODUCTION

The frequency-following response (FFR), measured as a voltage on the human scalp, reflects how well subcortical portions of the auditory pathway encode periodic portions of an input acoustic stimulus (e.g., [Krishnan, 1999](#); [Galbraith et al., 2000](#); [Kraus and Nicol, 2005](#); [Akhoun et al., 2008](#); [Du et al., 2011](#)). Many studies of the FFR focus on those components that are phase locked to the envelope of the input stimulus ( $\text{FFR}_{\text{ENV}}$ ; the portion of the response that is the same for a stimulus and an inverted version of that stimulus), in part because many artifacts and non-neural signals (such as the cochlear microphonic) that can contaminate the measure are canceled when estimating  $\text{FFR}_{\text{ENV}}$  (e.g., see [Picton, 2011](#)). The strength of  $\text{FFR}_{\text{ENV}}$  is correlated with perceptual ability on a range of tasks, including at the level of individual subjects (e.g., [Krishnan et al., 2010](#); [Wile and Balaban, 2007](#); [Burman et al., 2008](#); [Carcagno and Plack, 2011](#); [Jerger and Hall, 1980](#); [Krizman et al., 2012](#); [Ruggles et al., 2011](#); [Ruggles et al., 2012](#)). This relationship with perceptual abilities in some tasks has led to recent surge of interest in  $\text{FFR}_{\text{ENV}}$ .

Many different methods have been proposed for extracting periodic neural signals like  $\text{FFR}_{\text{ENV}}$  amid the

significant noise present in the measures. For instance, the magnitude-squared coherence (MSC), which uses both phase and amplitude information to estimate how much of the power in raw measurements at a given frequency can be attributed to the signal, often performs equal to or better than other common methods for detecting a sinusoidal signal in auditory brainstem responses and similar noisy measurements (e.g., see [Dobie and Wilson, 1993, 1994, 1996](#); [Sturzebecher and Cebulla, 1997](#); [Cebulla et al., 2001](#)). However, the calculation of MSC depends on the noise being stationary; when this assumption is violated, methods based on across-trial consistency of phase (independent of amplitude), such as phase-locking value analysis (PLV, sometimes called the phase coherence or PC) can be equally powerful or even more sensitive than MSC (e.g., see [Dobie and Wilson, 1993, 1994](#); [Lachaux, et al., 1999](#)). Other approaches, including Hotelling's  $T^2$  test,  $T_{\text{circ}}^2$ , and various extensions of these approaches (and of MSC and PC) have also been proposed, but in head-to-head comparisons, most of these methods perform fairly similarly, depending on the exact characteristics of the signal and noise in the measurements (e.g., see [Victor and Mast, 1991](#); [Dobie and Wilson, 1993](#); [Cebulla et al., 2001](#)).

Perhaps the most intuitive approach to analyzing noisy neural data is simply to average many repetitions of responses to a fixed input and assess the spectral magnitude of the results. However, such measures contain many

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: [shinn@cns.bu.edu](mailto:shinn@cns.bu.edu)

different artifacts, ranging from line noise at multiples of the alternating electrical current to broadband neuro-electrical noise. This makes interpretation of the absolute spectral magnitude results challenging; to determine which frequencies in the response are driven by the acoustic input, the noise floor, which is not flat, must be estimated and/or the results normalized appropriately. In direct contrast, the unitless PLV metric (for instance) is directly interpretable, without need for estimation of the noise floor in the measures. Moreover, the simplest null-hypothesis model for the PLV, that the measurement phase at any given frequency is uniformly distributed between  $-\pi$  and  $\pi$  (i.e., unaffected by the stimulus presentation), has a distribution that depends only on the number of samples going into the calculation. In particular, the noise floor is independent of both frequency and the absolute power in the noise sources contributing to the measured responses, and hence is flat (see Bokil *et al.*, 2007). The PLV also has its own intuitive appeal to many auditory neuroscientists in that phase-locking measures such as vector strength are commonly used when assessing auditory coding in single-unit neurophysiology (e.g., see Joris *et al.*, 2004). Moreover, in some conditions, the PLV may be a more reliable metric than spectral magnitude approaches that estimate the signal-to-noise ratio of the power at the stimulus frequency (e.g., see Picton *et al.*, 1987; Stapells *et al.*, 1987), even though many FFR studies continue to employ spectral magnitude estimation when evaluating neural responses (e.g., see Krishnan, 1999; Galbraith *et al.*, 2000; Krishnan, 2002; Wile and Balaban, 2007; Gockel *et al.*, 2011; Hornickel *et al.*, 2012). Here, we directly compare spectral magnitude and PLV measures of the  $\text{FFR}_{\text{ENV}}$  evoked by complex tones to explore whether differences across subjects and conditions are better revealed by PLV analysis than by spectral magnitude analysis (Experiment 1; note that though many studies explore the temporal-fine-structure-related component of the FFR, the current study focuses exclusively on the envelope-related component).

Past experimental and theoretical studies suggest that  $\text{FFR}_{\text{ENV}}$  is generated primarily by subcortical structures including the cochlea, auditory nerve, cochlear nucleus, and inferior colliculus (Smith *et al.*, 1975; Sohmer and Pratt, 1977; Sohmer *et al.*, 1977; Gardi *et al.*, 1979; Davis and Britt, 1984; Dolphin and Mountain, 1992; Dau, 2003; Wile and Balaban, 2007; Harte *et al.*, 2010; Chandrasekaran and Kraus, 2010; Du *et al.*, 2011). Each peripheral auditory channel will respond to the portion of the acoustic signal falling within its critical band; the observable  $\text{FFR}_{\text{ENV}}$  is a sum of all of this activity, across frequency channels, and therefore depends on the phase locking within each channel as well as the phase and magnitude relationships among the phase-locked activity across all channels. For mid- to high-frequency peripheral auditory channels, multiple harmonics of the input stimulus are likely to fall within a critical band. In response to a periodic input, envelope-locked neural activity in these channels will contain energy at the fundamental frequency and low-order harmonics of the periodic input, even though they are tuned to higher acoustic frequencies (see, e.g., Gockel *et al.*, 2011; Shinn-Cunningham *et al.*, 2012). In addition, even phase-locked activity of a particular periodicity is not purely

sinusoidal, but will contain energy at different harmonics of that fundamental. As a result, it is difficult to predict exactly how different combinations of acoustic frequencies drive  $\text{FFR}_{\text{ENV}}$ . To begin to address this question, Experiment 2 uses the PLV to analyze  $\text{FFR}_{\text{ENV}}$  in response to different harmonic tone complexes: (1) low-frequency, resolved harmonics, (2) higher-frequency, partially resolved harmonics, (3) unresolved harmonics, and (4) both resolved and unresolved harmonics.

## II. METHODS

### A. Participants

Sixteen participants (ages 21 to 38 years old; 12 males) were recruited in Experiment 1. Twenty participants (ages 20 to 30 years old; 10 males) were recruited in Experiment 2. For all subjects, pure-tone audiometric thresholds were measured from 250 Hz to 8000 Hz at octave intervals. All participants had hearing thresholds within 15 dB of normal hearing level in each ear at all tested frequencies, and none had any history of central or peripheral auditory deficits. All gave written informed consent in accordance with procedures approved by the Boston University Charles River Campus Institutional Review Board and were paid for their participation.

### B. Stimuli

Stimuli were generated offline in MATLAB (Natick, MA) and stored for later playback using a sampling rate of 25 kHz. In Experiment 1, the stimulus was comprised of eight, equal intensity pure tones (100, 200, ..., 800 Hz), all in sine phase, creating a complex tone with fundamental frequency ( $F_0$ ) of 100 Hz. In the *quiet* condition, the Experiment 1 stimulus was presented alone; in the *noise* condition, a Gaussian noise (low-pass filtered at 3000 Hz) was added to the stimulus at an acoustic signal-to-noise ratio (root mean square) of +10 dB. The noise was continuous, generated online by special-purpose Tucker-Davis Technologies (Gainesville, FL) hardware.

In Experiment 2, we tested four different complex tones, each with  $F_0$  100 Hz, comprised of equal intensity pure tones presented in cosine phase, but containing different harmonics of 100 Hz. The first contained the first five harmonics of 100 Hz (LOW: 100, 200, 300, 400, and 500 Hz), which, for our normal-hearing listeners, should be easily resolved by the auditory periphery (e.g., see Moore, 2003); the second contained harmonics 6–10 (MID: 600, 700, 800, 900, and 1000 Hz), which should be partially resolved in the auditory peripheral response; the third contained harmonics 12–16 (HIGH: 1200, 1300, 1400, 1500, and 1600 Hz), which should be poorly resolved by the peripheral auditory system; the final stimulus contained the first 20 harmonics (BROAD: 100, 200, ..., 2000 Hz).

All of the complex tones in both experiments were 170 ms in duration, including 10 ms cosine-squared rise/fall times. All were presented diotically at 80 dB sound pressure level. The mean inter-stimulus interval was 770 ( $\pm$  100) ms, jittered from trial to trial to ensure that any signals unrelated to the acoustic stimulus were at a different, random phase at

the onset of each trial. The stimuli were presented with alternating starting polarities, so that half of all trials were presented in one polarity and the other half inverted to the opposite polarity.

### C. Equipment

A personal computer controlled all aspects of the experiment, including triggering sound delivery and storing data. Special-purpose sound-control hardware (System 3 real-time signal processing systems, including D/A conversion and amplification; Tucker Davis Technologies, Gainesville, FL) presented sound through insert phones (ER-1, Etymotic, Elk Grove Village, IL) coupled to foam ear tips.

FFR<sub>ENV</sub> responses were recorded using a BioSemi Active Two System (BioSemi, Amsterdam, Netherlands) at a sampling rate of 16.384 KHz. Although BioSemi makes special-purpose, high-impedance electrodes optimized for recording subcortical responses, here, we used the standard conductive Ag-AgCl scalp electrodes. The FFR<sub>ENV</sub> responses were recorded from active electrode Cz. Vertical eye movements (electrooculogram, EOG) were monitored with two external electrodes. Prior to data acquisition, we ensured that the offset voltage for each active electrode was stabilized at <20 mV.

### D. Procedures

After electrodes were placed on the scalp, participants were seated in an acoustically and electrically shielded booth (single-walled Eckel C-14 booth, Cambridge, MA). Throughout data collection, participants watched a silent, captioned movie of their choice, ignoring the acoustic stimuli.

Each participant in Experiment 1 performed one experimental session of ~2 h, including setup and data collection. In each session, the Experiment 1 stimulus was repeated 4000 times each in quiet and in noise. These data were collected in 8 blocks of 1000 trials each, alternating quiet and noise conditions from block to block (starting with a quiet block). Stimulus polarity alternated from trial to trial randomly within each block, resulting in 2000 repetitions of each stimulus (that is, 2000 repetitions each of 2 polarities and 2 conditions).

Each participant in Experiment 2 performed four sessions (one per day, maximum), each lasting ~1.5 h, including setup and data collection. The four sessions each presented one of the four different stimuli (LOW, MID, HIGH, BROAD, ordered identically for each subject) 2000 times in a single, long block, alternating stimulus polarity from trial to trial.

### E. Data pre-processing

All raw recordings were bandpass filtered from 80 Hz to 3000 Hz (zero-phase filtering, 800th-order finite impulse response filter, designed by the window method using a Hamming window in MATLAB, Natick, MA), to retain subcortical responses while minimizing lower-frequency noise and interference from cortical activity. The bandpass-filtered signals were then broken into epochs, from -50 to 250 ms

relative to the onset of each 170-ms-long complex tone. Any epochs containing EOG activity greater than 60  $\mu$ V (peak to peak) were removed to eliminate contamination from eye blinks. For all remaining epochs, the raw Cz responses were then referenced to the average of the two earlobe measures. For all participants, the number of artifact-free epochs was greater than 1800 in each of the four different stimuli (positive/negative polarities  $\times$  quiet/noise conditions) in Experiment 1, and greater than 900 for each polarity of each of the four stimulus conditions in Experiment 2.

In order to focus analysis on the quasi steady-state portion of the responses, the time-domain signals were time windowed with a 170-ms-long window (Slepian taper, achieving 5.9 Hz frequency resolution while minimizing spectral leakage; Thomson, 1982) that started 10 ms after the onset of the complex tone. The resulting time-domain waveforms from individual trials were then combined either by first averaging them and then analyzing the frequency content of this average (spectral analysis), or by analyzing the consistency of the phase of the individual trials' responses at each frequency (PLV analysis), as described further in Secs. II F and II G.

### F. Spectral analysis

Spectral analysis was undertaken separately for each subject, for each condition, in Experiment 1. The Appendix provides details of the processing stages undertaken; these steps are briefly summarized below.

Interpreting the raw magnitude spectrum of FFR<sub>ENV</sub> is challenging because the absolute strength of the noise floor varies with frequency roughly as  $1/f$  (e.g., see Buzsaki, 2006) and also contains some electrical line noise at multiples of 60 Hz (particularly at the odd harmonics, based on experience in our laboratory). Therefore, in order to determine which frequencies in the magnitude spectra were significantly driven by the stimuli, we employed a bootstrapping approach to estimate the noise floor as a function of frequency. We first computed a distribution of 100 raw FFR<sub>ENV</sub> magnitude spectrum estimates of the responses for each subject by averaging together 400 random samples (drawn with replacement) of the discrete Fourier transforms (DFTs) of the FFR waveform taken from individual positive and negative polarity trials. The average value of this distribution, which is roughly Gaussian distributed, was used as the estimate of the individual subject's raw magnitude spectrum response in a given condition. The standard deviation of these 100 estimates was used to estimate the variability in the Gaussian-distributed magnitude spectrum estimate of FFR<sub>ENV</sub>.

We then estimated the raw noise floor spectral magnitude separately for each subject and condition as a function of frequency, using an identical procedure except randomizing the phase of the 400 independent observations prior to averaging (i.e., assuming that the phase of the input signal is unrelated to the phase of any noise). This is a conservative null model, as it assumes that the magnitude of the response of each trial is dominated by noise, with a negligible increment due to the signal. However, the approach accounts for both  $1/f$  and line noise equally well.

By normalizing the raw magnitude responses by the estimated noise floor, we computed the spectral magnitude of  $\text{FFR}_{\text{ENV}}$  in units of dB SNR (signal-to-noise ratio), individualized for each subject. These magnitude responses in dB SNR can be understood readily and can be fairly combined across subject. We then could also compute the  $z$ -score separating the distribution of the data and the spectral magnitude expected by chance.

Figure 1 shows, for one typical subject, the estimated noise floor as well as the raw and normalized magnitude spectra. The raw magnitude response (thin red lines) varies dramatically, and similarly, with frequency for both quiet and noise conditions (left and right columns, respectively). The estimated noise floor, shown by the thin blue lines, also varies with frequency in a manner grossly like the raw magnitude responses. However, these large variations in magnitude are partially taken into account when the response is shown in dB SNR (bold black lines in Fig. 1); this analysis directly reveals the spectral components of  $\text{FFR}_{\text{ENV}}$  that arise at the stimulus  $F_0$  and its integer multiples.

To test whether the spectral magnitude of each harmonic was significantly greater than expected in the absence of a signal, we determined whether the observed spectral magnitude estimate was likely to be observed by chance, given the variability in the metric (see the Appendix). In order to summarize the results, we computed the across-subject means of the magnitude spectra in dB SNR (across 16 subjects, Experiment 1).

### G. PLV analysis

The same raw trial results were also analyzed by calculating PLVs as a function of frequency separately for each subject and condition. As with the spectral analysis, we computed the mean across 100 independent, random, identically distributed estimates of the PLV (each calculated from 400 randomly selected trials, with replacement). We also computed the standard deviation of these estimates, enabling us to compute the  $z$ -score between the data distribution and the expected mean under the null model. Bootstrapping was used to estimate the noise floor in the PLV values. However, because the distribution of PLVs depends only on the phase of individual trials and the number of trials averaged, independent of magnitude, the null model is identical for all

subjects, conditions, and stimuli used. Therefore, we computed only one distribution of mean PLV statistics under the null hypothesis, assuming a random phase distribution from trial to trial (see the Appendix for details of the processing), and computed the across-subject means (16 in Experiment 1; 20 in Experiment 2).

## III. RESULTS

### A. Experiment 1

The first experiment measured  $\text{FFR}_{\text{ENV}}$  in response to a 100-Hz  $F_0$  harmonic complex tone containing harmonics 1–8 to characterize the relative strength and robustness of the subcortical measure as a function of neural signal frequency. We present the basic results, analyzed first using the more traditional, spectral magnitude approach and then by using the PLV method. To test whether there are any advantages when undertaking a PLV analysis of  $\text{FFR}_{\text{ENV}}$  data rather than spectral analysis, we then perform statistical tests of the relationships across analysis methods and stimulus conditions.

#### 1. Spectral magnitudes in quiet and noise

$\text{FFR}_{\text{ENV}}$  spectral magnitude responses are summarized in Fig. 2, which plots the across-subject mean of  $\text{FFR}_{\text{ENV}}$  in both the quiet and noise conditions (left and right columns, respectively) in units of dB SNR.

$\text{FFR}_{\text{ENV}}$  is most robust at 100 Hz (the stimulus  $F_0$ ) and tends to decrease with increasing harmonic frequency. Adding external noise to the input stimulus reduces the  $\text{FFR}_{\text{ENV}}$  response, in general, as expected. However, the effect of the noise is not uniform. In particular,  $\text{FFR}_{\text{ENV}}$  at 100 Hz, which is the largest response to begin with, is hardly affected by the addition of the noise. In contrast, the other  $\text{FFR}_{\text{ENV}}$  components all decrease in average magnitude.

#### 2. Phase-locking values in quiet and noise

PLVs are shown in Fig. 3 in the same format as the magnitude responses in Fig. 2. Because the threshold for significance is independent of frequency and condition, the range of individual subject's PLVs that are not significantly greater than expected by chance is shown in light gray in each panel.

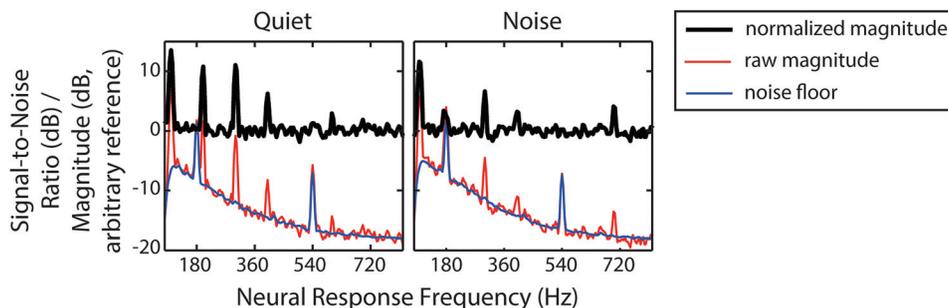


FIG. 1. Spectral magnitude results as a function of neural response frequency for one example subject (Experiment 1) in response to a complex tone consisting of harmonics 1–8 of a 100 Hz fundamental frequency. In order to calculate the response in units of dB SNR, the noise floor must be estimated and responses normalized by the frequency-dependent spectral content of the noise. The noise had both strong components at odd multiples of 60 Hz due to line noise and  $1/f$  magnitude noise from other sources.

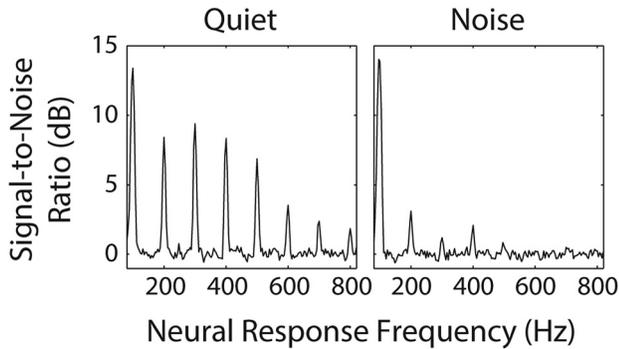


FIG. 2. Normalized spectral magnitude (in units of dB SNR) of  $FFR_{ENV}$  averaged across subjects as a function of neural response frequency (Experiment 1). The left panel shows results in quiet, while the right panel shows results in noise. Significant energy is present at multiples of the 100 Hz fundamental frequency of the input complex tone.

The only frequencies with PLVs significantly above the noise floor occur at harmonics of the input stimulus, as expected. The general trends described in the spectral magnitude analysis are also seen in the PLV results. In quiet,  $FFR_{ENV}$  is robust at  $F_0$ , and decreases in strength as harmonic number increases. The effect of the additive acoustic noise is negligible at 100 Hz, but is more evident for 200 Hz and above.

### 3. Comparison of spectral magnitude and PLV analyses

Theoretically, the spectral magnitude and PLV results should be closely related; both focus on extracting the component of the measured response that is consistent across trials (the spectral magnitude analysis, by averaging responses in the time domain; the PLV analysis, by examining the consistency of the phase response of individual trials across the distribution of trials). Here, we directly compare the two analyses methods to test whether this prediction holds up in practice. In particular, we explore whether (i) PLV analysis is more sensitive at revealing brainstem responses than spectral magnitude analysis by comparing the statistical power of the two analyses, and (ii) whether, aside from any differences in sensitivity, the two approaches produce results that are closely correlated.

To compare the relative sensitivity of the spectral magnitude analysis and the PLV analysis, we converted both raw

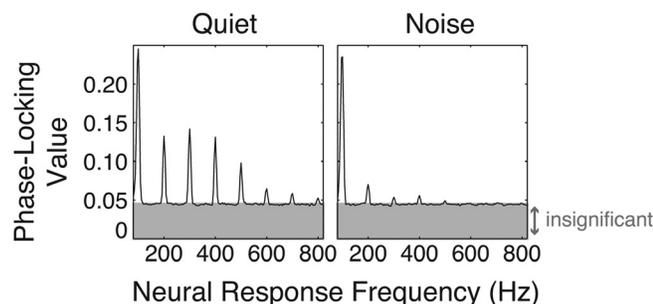


FIG. 3. Phase-locking value, averaged across subjects, as a function of neural response frequency (Experiment 1). PLVs are significant at multiples of the 100 Hz fundamental frequency of the input complex tone.

measures to  $z$ -scores (statistical separation between the obtained results and a null-model distribution). Figure 4(a) presents the  $z$ -scores (averaged across subjects) for the PLV analysis and the spectral magnitude analysis (squares and circles, respectively) both in quiet and in noise (left and right panels, respectively). In quiet, the mean PLV  $z$ -scores are slightly higher than are the spectral analysis  $z$ -scores of the same data at the low harmonics [squares are above circles in the left half of the left panel of Fig. 4(a)], but this modest difference disappears as frequency increases beyond 500 Hz and both metrics have low values. A similar pattern is seen in noise: at  $F_0$  and the lower harmonics, the mean  $z$ -scores for the PLV analysis are greater than the spectral magnitude  $z$ -scores; for harmonics from 600–800, the average  $z$ -scores are less than 1.0, near the floor.

Inter-subject variability in  $FFR_{ENV}$  is large compared to the small differences in the mean  $z$ -scores, which can visually obscure consistent pairwise differences. In order to illustrate the main effect of *method*, we therefore generated a scatter plot contrasting the  $z$ -scores obtained from the PLV and from the spectral magnitude, with each point representing one subject, *condition*, and *frequency* [Fig. 4(b)]. In this plot, only  $z$ -scores exceeding 4 in both methods were included to avoid floor effects. The identity line (diagonal where the  $z$ -scores calculated from the two methods are exactly equal) is plotted as a reference. Consistent with the repeated-methods analysis of variance (ANOVA) result showing a main effect of *method*, the majority of these points lie above the diagonal, showing that the  $z$ -scores obtained from the PLV method are greater than the  $z$ -scores obtained from the same raw data using spectral magnitude analysis.

The  $z$ -scores were subjected to repeated-measures ANOVA with main factors of method (spectral vs PLV), frequency (100–800 Hz), and condition (quiet vs noise), as well as their interactions. The residual error strata were partitioned. All main factors were statistically significant [method:  $F(1,15) = 16.3$ ,  $p < 0.01$ ; frequency:  $F(7,105) = 42.7$ ,  $p < 0.001$ ; condition:  $F(1,15) = 57.0$ ,  $p < 0.001$ ]. Respectively, these results support the observations that (1) PLV is a more sensitive measure than is spectral magnitude, (2)  $FFR_{ENV}$  strength decreases with frequency, and (3) additive acoustic noise degrades  $FFR_{ENV}$ . The interaction of method  $\times$  condition was not significant [ $F(1,15) = 1.2$ ,  $p = 0.285$ ]; however, the interaction of method  $\times$  frequency [ $F(7,105) = 3.1$ ,  $p < 0.01$ ] was significant, supporting the observation that the PLV  $z$ -scores are greater than the corresponding spectral magnitude  $z$ -scores as long as floor effects are not in play. The interaction between frequency  $\times$  condition was also significant [ $F(7, 105) = 5.8$ ,  $p < 0.001$ ], supporting the observation that the  $z$ -scores decrease more gradually with frequency in quiet than they do in noise.

The results in Fig. 4(b) suggest that although the PLV is more statistically sensitive than spectral analysis when  $FFR_{ENV}$  values are above the noise floor, the pattern of results may be quantitatively similar for the two methods. To test this prediction, we computed the across-subject Pearson correlation coefficients between the obtained spectral magnitude values and PLV values in quiet and in noise. Because

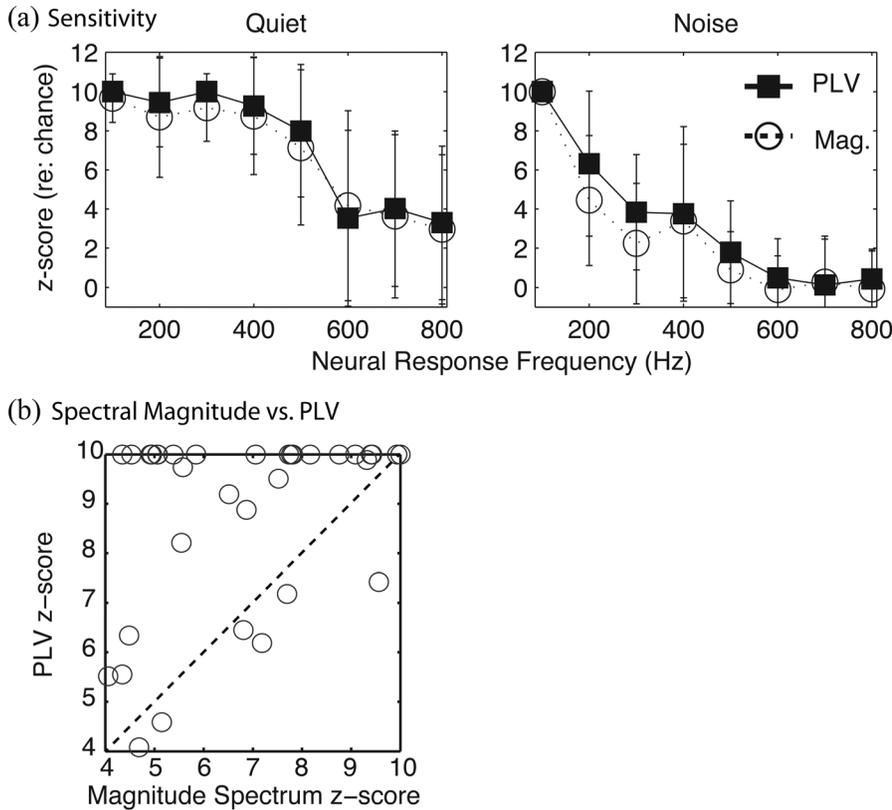


FIG. 4. Comparison of PLV and spectral magnitude responses in quiet and in noise as a function of neural response frequency (Experiment 1). (a) Across-subject average PLV (squares) and spectral magnitude (circles) as z-scores (relative to expected chance levels). Error bars show the across-subject standard deviation, which is large, obscuring consistent differences between the two analysis methods. (b) The relationship between PLV and spectral magnitude z-scores (relative to expected chance levels) for z-scores exceeding a value of 4, pooling over subjects, noise conditions, and frequencies. For a given subject and experimental condition, PLV z-scores are generally greater than spectral magnitude z-scores.

only results that are significantly above the noise floor of a given metric should have a value that reflects the signal, rather than the noise, we included in the analysis only those values that were statistically above the noise floor using both the spectral magnitude analysis and the PLV analysis for a given condition, harmonic, and subject. While this yielded 16 subjects for  $FFR_{ENV}$  at 100 Hz, for other harmonics, the number of points was lower; at some of the highest harmonics, there were not enough subjects with responses that were above the noise floor to allow a meaningful comparison. We found that for every harmonic and condition that contained enough data points to allow a direct assessment, spectral magnitude values and PLVs are very strongly correlated (see Table I).

Together, these results show that the PLV analysis is more sensitive than the spectral magnitude analysis at revealing  $FFR_{ENV}$  responses that are above the noise floor, but that the two approaches yield comparable patterns of results. Therefore, from this point on, we focus on PLV results.

#### 4. Statistical effects of additive acoustic noise

To assess the effect of additive noise, we directly compared the across-subject average PLVs. First we found those  $FFR_{ENV}$  components that were statistically above the noise floor in both conditions (based on a Tukey's *post hoc* test; harmonics 100, 200, 300, 400, 500 Hz), and computed the difference (the mean PLV in quiet – the mean PLV in noise). We then found the harmonics that were above the noise floor in the quiet condition, but below the noise floor in the noise condition (harmonics 600, 700, 800 Hz); for these harmonics,

we computed the difference between the measured value in quiet and the noise floor (the mean PLV in quiet – the PLV at the noise floor; the latter is a generous estimate of the phase locking evoked by the stimulus in noise). As seen in Table II, the noise caused a significant decrease in  $FFR_{ENV}$  at harmonic frequencies of 200,300,...,800. However, the noise did not cause a statistically significant change in the PLV at 100 Hz.

TABLE I. Correlations between spectral magnitude (dB SNR) and PLV measures across subjects in Experiment 1 for each harmonic of the stimulus fundamental frequency. To be included in the analysis, a subject had to have both a spectral magnitude and PLV measure that was statistically significant; a dash indicates that there were too few data points to conduct the particular analysis (most of the time, when a particular measure was in the noise floor using one method, it also was not significant using the other analysis approach). Here,  $L$  is the number of subjects for whom both spectral magnitude and PLV are significant at the given harmonic, and  $r$  is the linear correlation coefficient. For all comparisons with enough data points to address the question, the two metrics are highly correlated.

Harmonics (Hz)	Quiet			Noise		
	$r$	$p$	$L$	$r$	$p$	$L$
100	0.9069	$1.3 \times 10^{-6}$	16	0.9295	$1.9 \times 10^{-7}$	16
200	0.8709	$2.4 \times 10^{-5}$	15	0.9599	$7.6 \times 10^{-7}$	12
300	0.9145	$<1.0 \times 10^{-7}$	16	0.9669	$1.6 \times 10^{-3}$	8
400	0.9423	$<1.0 \times 10^{-7}$	15	0.9041	$1.3 \times 10^{-2}$	6
500	0.9305	$<1.0 \times 10^{-7}$	13	0.8329	$1.6 \times 10^{-1}$	4
600	0.9854	$<1.0 \times 10^{-7}$	7	—	—	0
700	0.9701	$1.0 \times 10^{-4}$	8	0.9840	$1.1 \times 10^{-1}$	3
800	0.8337	$3.9 \times 10^{-2}$	6	—	—	1

TABLE II. Effects of additive acoustic noise on PLV strength at harmonic frequencies of  $F_0$  in Experiment 1. Bold frequencies had PLV values that were significantly smaller in noise than in quiet. The noise had a statistically significant effect on  $\text{FFR}_{\text{ENV}}$  at harmonics 200–800 Hz, but not at the fundamental frequency.

Harmonic (Hz)	Change in PLV	Sign test $p$ value
100	+0.0103	0.7174
<b>200</b>	<b>+0.0623</b>	<b>0.0011</b>
<b>300</b>	<b>+0.0883</b>	<b>0.0004</b>
<b>400</b>	<b>+0.0756</b>	<b>0.0004</b>
<b>500</b>	<b>+0.0491</b>	<b>0.0009</b>
<b>600</b>	<b>+0.0191</b>	<b>0.0151</b>
<b>700</b>	<b>+0.0128</b>	<b>0.0097</b>
<b>800</b>	<b>+0.0070</b>	<b>0.0299</b>

## B. Experiment 2

Experiment 2 measured  $\text{FFR}_{\text{ENV}}$  responses to four different acoustic input signals, each consisting of different combinations of harmonics of 100 Hz. Figure 5 shows the across-subject average PLV of  $\text{FFR}_{\text{ENV}}$  in response to the four complex tones (rows).

As in Experiment 1,  $\text{FFR}_{\text{ENV}}$  tends to be strongest at  $F_0$  and the lowest harmonics of the input stimulus—even for stimuli that contain no acoustic energy at those frequencies—both in terms of the magnitude of the PLV and the number of subjects with a significant PLV. The  $\text{FFR}_{\text{ENV}}$  components at low harmonics of  $F_0$  can be driven strongly by activity from many different peripheral channels, spanning a range of input acoustic frequencies; the envelope

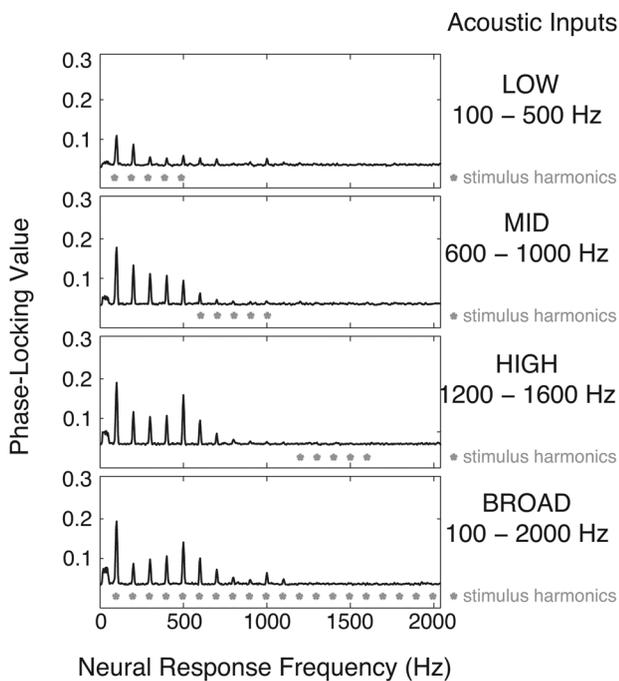


FIG. 5. Phase-locking value averaged across subjects as a function of neural response frequency for the four different acoustic input stimuli, all with fundamental frequency of 100 Hz (Experiment 2). For all stimuli, the envelope-related components of the FFR are greatest at  $F_0$  and low-order harmonics of the fundamental, even for stimuli that do not contain acoustic energy at  $F_0$ .

response in the lowest harmonics tends to be weakest for the LOW stimulus.

To directly compare the strength of  $\text{FFR}_{\text{ENV}}$  across the different stimuli, we examined the average PLV at 100 Hz in the different conditions (see Fig. 6). The LOW stimulus produced a much lower average PLV than the other three stimuli; the MID, HIGH, and BROAD stimuli all produced very similar values. To test these observations we performed paired  $t$ -tests with Bonferroni correction ( $p < 0.05/6 = 0.0083$ ) on all possible pairings of stimuli. These tests confirm that (i) the strength of  $\text{FFR}_{\text{ENV}}$  at  $F_0$  is significantly smaller for the LOW stimulus than for HIGH and BROAD stimuli [ $t(1, 19) = -4.0$ ,  $p = 0.001$  and  $t(1, 19) = -3.2$ ,  $p = 0.005$ , respectively], and trends toward being smaller for the LOW stimulus than for the MID stimulus [ $t(1, 19) = -2.9$ ,  $p = 0.009$ , which just misses reaching significance with the strict Bonferroni correction]; but that (ii), no pairs of PLV values are statistically significantly different from one another among the MID, HIGH, or BROAD stimuli [MID vs HIGH:  $t(1, 19) = 0.5$ ,  $p = 0.632$ ; MID vs BROAD:  $t(1, 19) = 0.6$ ,  $p = 0.544$ ; HIGH vs BROAD:  $t(1, 19) = 0.1$ ,  $p = 0.891$ ].

## IV. DISCUSSION

### A. Advantages of PLV analysis

In our study, PLVs estimated from  $\text{FFR}_{\text{ENV}}$  data were statistically more sensitive (had larger  $z$ -scores) than did the corresponding magnitude spectra, both in quiet and in noise, as long as the  $z$ -scores were moderately sized. To confirm that the PLV is more sensitive than spectral magnitude for detecting a periodic signal embedded in electroencephalography (EEG) noise and to determine how this might vary with the number of trials available for analysis, we did a

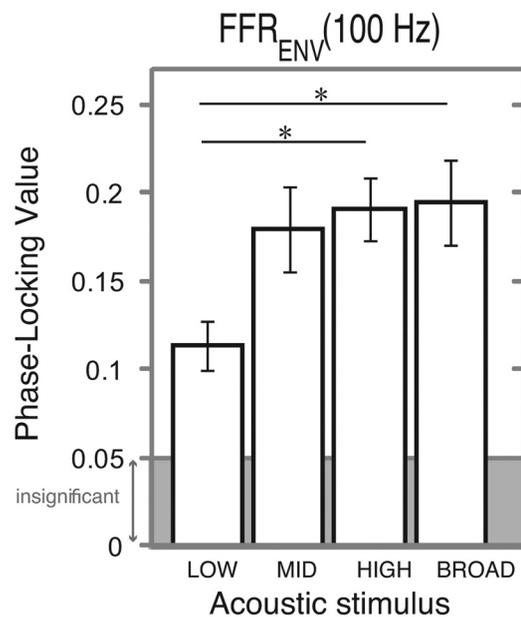


FIG. 6. Phase-locking value, averaged across subjects, at the fundamental frequency of 100 Hz for each of the four input stimuli. The LOW input stimulus produced a significantly lower FFR at 100 Hz than the other three stimuli (see the text).

simple simulation. We took resting-state EEG (measured in the absence of stimuli or task), extracted  $N$  300-ms-long epochs (trials), and used this to simulate the noise in our simulations. We used these  $N$  noise-only epochs to compute the PLV and spectral magnitude at the “target” frequency (250 Hz, which is near the middle of the range of  $\text{FFR}_{\text{ENV}}$  frequencies that were strong in the current experiments) using the bootstrapping method described in the Appendix. This produced the null-model distributions of the two metrics. We then added a 200 nV, 250-Hz sinusoid (a typical size for FFR components, using our system) to the  $N$  epochs and recomputed the PLV and spectral magnitude for these signal-plus-noise trials. We calculated  $z$ -scores for the signal-plus-noise vs noise-only distributions for both PLV and spectral magnitude. We repeated this whole process 20 times, so that we could evaluate the variability in the  $z$ -score estimates. We compared the  $z$ -scores for the PLV and spectral magnitude analyses for values of  $N = 50, 100, 150,$  and  $200$  (see Fig. 7). With only 50 trials, the  $z$ -scores have a value near one, and are equal for the two analyses. As more trials are analyzed, the  $z$ -scores increase using both analysis methods; however, the increase is more rapid with  $N$  for the PLV analysis than for the spectral magnitude analysis (in Fig. 7, triangles are above squares). This result confirms what we observed in the actual  $\text{FFR}_{\text{ENV}}$  data, that when trying to detect a weak sinusoidal signal embedded in EEG noise, PLV analysis is statistically more sensitive than is spectral magnitude analysis.

Nonetheless, across subjects and conditions, spectral magnitude and PLVs produce tightly correlated results. Together, these results suggest that as long as the number of trials is sufficient to reveal the signal, either spectral magnitude or PLV analysis can be applied successfully to  $\text{FFR}_{\text{ENV}}$  data to quantify differences in the brainstem response across

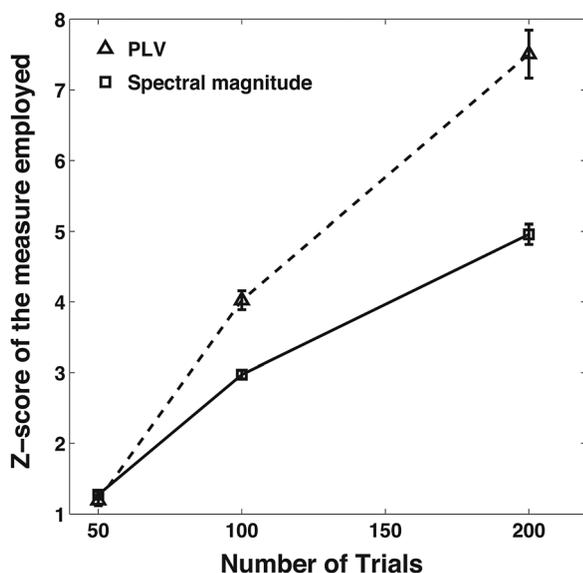


FIG. 7.  $z$ -scores for detection of a low-amplitude sinusoid embedded in EEG epochs taken during resting state for one volunteer subject, plotted as a function of the number of 300-ms-long epochs used. Error bars show the standard error of the mean over the 20 repetitions of the simulation. PLV analysis produces a larger  $z$ -score than the spectral magnitude analysis once the number of trials is sufficient to detect the signal.

subjects, conditions, or other factors. However, given that the PLV analysis is more sensitive, statistically, than the spectral magnitude analysis, it is likely to yield better results for conditions where information is just on the cusp of being statistically reliable using spectral magnitude analysis.

PLVs have another distinct advantage over spectral analysis methods in that they are unit-less, with values ranging from 0–1. PLVs have the property that when making statistical inferences and testing hypotheses, the raw PLVs can be directly compared across conditions, as long as the number of trials being analyzed is the same in the different conditions (see Bokil *et al.*, 2007 for a consideration of how to compare conditions with different numbers of trials). This is in direct contrast with analysis of the average spectral magnitude. The magnitude of the measured signal depends not only on the power of the signal, but also on the frequency content and the power of noise sources unrelated to the stimulus that are present in the measures. For brainstem potentials measured on the scalp, the noise has a number of components, including electrical line noise, muscle artifacts, eye movement, and background EEG (unrelated to the stimulus) that, together, have power that is typically proportional to  $1/f$ . Because the noise measures can depend on the impedance of the electrodes on a given day and other factors that are difficult to control, rational analysis requires the noise floor to be estimated directly, and results interpreted relative to the noise in the measure. While this can be done in a number of ways (e.g., by estimating the noise from baseline measures that are analyzed in the same way as the signal-driven epochs), here, we adopted a bootstrapping method to normalize the spectral magnitude measures and transform them into units of dB SNR. Using this kind of metric, data can be compared across frequencies, or even across subjects. However, this kind of normalization requires extra stages of processing to quantify the noise as a function of frequency, a step that is unnecessary when considering the PLV. Thus, when analyzing  $\text{FFR}_{\text{ENV}}$ , PLVs not only provide a more sensitive measure, statistically, than spectral magnitude, but also are easily interpretable and provide meaningful data with fewer processing steps.

## B. Acoustic generators of FFR harmonics

Because the auditory periphery is nonlinear, the relationship between the acoustic frequencies present in an input stimulus and the frequency content of the resulting neural signal is complex. In Experiment 2, we considered PLVs in response to harmonic complex tones with the same  $F_0$  but different harmonics, to begin to determine how different acoustic frequencies contribute to  $\text{FFR}_{\text{ENV}}$ .

The LOW stimulus (harmonics 100–500 Hz) contained five individual frequency components that each should be well resolved by the auditory periphery (Moore, 2003). Because the acoustic harmonics are resolved into separate channels, the responses should be closely related to the raw stimulus components and thus very sensitive to stimulus polarity. When the responses to opposite polarity signals are averaged together, the negative and positive polarities will tend to cancel, resulting in a relatively weak  $\text{FFR}_{\text{ENV}}$  in

response to the LOW stimulus. Of course, the frequency content of the resulting electrical signal will not be purely sinusoidal; the auditory nerve response is half-wave rectified, and often includes a significant DC component (Pickles, 1982). Moreover, this activity is transformed as it progresses through the brainstem; our  $\text{FFR}_{\text{ENV}}$  measure is a sum of activity along this pathway. As a result, there may be a measurable  $\text{FFR}_{\text{ENV}}$  in response to the LOW stimulus, even though the dominant auditory nerve response is phase locked with opposite phase to positive and negative polarity trials. Consistent with these ideas, the LOW stimulus produced a significantly weaker response at  $F_0$  than all of the other stimuli, but one that was nonetheless statistically significant.

The MID stimulus (harmonics 600–1000 Hz) should be partially resolved in the cochlea (Moore, 2003). The peripheral channels excited by the MID stimulus will both phase-lock to the temporal fine structure of the input signal and to envelope fluctuations that occur at the acoustic stimulus fundamental frequency and its harmonics (e.g., Javel, 1980; Palmer, 1982; Joris and Yin, 1992). Therefore, the MID stimulus is likely to evoke stronger  $\text{FFR}_{\text{ENV}}$  responses at low-order harmonics of the acoustic input than those evoked by the LOW stimulus. The  $\text{FFR}_{\text{ENV}}$  response to the MID stimulus was consistent with this prediction, producing an average PLV at  $F_0$  that was significantly larger than for the LOW stimulus.

The acoustic components making up the HIGH stimulus (harmonics 1200–1600 Hz) should not be well resolved by the auditory periphery (Moore, 2003). Moreover, as acoustic carrier frequency increases, there should be an increase in the proportion of the total response of a peripheral channel that responds to the driving function envelope, rather than the carrier oscillations, both because the numbers of harmonics falling within the critical band will increase with increasing frequency and because the degree to which a channel phase-locks to the envelope, rather than only the carrier, of the driving signal increases with increasing frequency (Pickles, 1982). Indeed, the HIGH stimulus produced a strong  $\text{FFR}_{\text{ENV}}$  at the acoustic stimulus fundamental and its harmonics.

Finally, the BROAD stimulus (harmonics 100–2000 Hz) contains the sum of the LOW, MID, and HIGH stimuli (although it also contains a few harmonics not present in any of the other stimuli). Therefore, the peripheral channels tuned to the frequencies of the acoustic components in the LOW, MID, and HIGH frequency ranges should be excited by essentially the same inputs as they were for these individual narrowband inputs. To a first order, the  $\text{FFR}_{\text{ENV}}$  in response to the BROAD stimulus might equal the sum of the responses to the three narrowband signals plus the responses to those harmonics not present in any of the narrowband signals (1100 and 1700–2000 Hz). Of course, this approximation ignores nonlinearities in the auditory pathway, from cochlear mechanics and cross-frequency-channel neural processes to the far-field potentials produced by the population of neurons. Moreover, even if the system is quasi-linear and the response to the sum of these inputs is the sum of the elicited responses, at any given harmonic, the phases of the contributions from the different frequency regions may

differ, and can either add constructively or destructively. Indeed, the magnitude of  $\text{FFR}_{\text{ENV}}$  at  $F_0$  in response to the BROAD stimulus is much less than the sum of the magnitudes of the responses to the LOW, MID, and HIGH inputs, either because the phases of the responses elicited by the LOW, MID, and HIGH inputs differ, or due to system nonlinearities. This result illustrates how difficult it can be to understand what kind of  $\text{FFR}_{\text{ENV}}$  a particular input signal will evoke, and the importance of developing quantitative models of how the  $\text{FFR}_{\text{ENV}}$  is generated along the auditory pathway in order to predict and interpret these subcortical measures. Only with the advance of such models can the frequency following response be related to physiological mechanisms and perceptual function.

### C. Contrasting the neural response at $F_0$ to other neural components

Past studies have demonstrated that  $\text{FFR}_{\text{ENV}}$  is strongest at the fundamental frequency of the driving acoustic stimulus, even when the acoustic input does not contain energy at  $F_0$  (e.g., Greenberg *et al.*, 1987; Wile and Balaban, 2007; Aiken and Picton, 2006). Results of both Experiment 1 and Experiment 2 confirm that the neural response at the fundamental frequency of the input stimulus is both quantitatively and qualitatively different from the other frequency components of  $\text{FFR}_{\text{ENV}}$ . Quantitatively, at  $F_0$   $\text{FFR}_{\text{ENV}}$  is larger than at the other harmonics (Experiments 1 and 2). Qualitatively,  $\text{FFR}_{\text{ENV}}$  at  $F_0$  is more robust than the other components, as evidenced by the fact that additive noise has a greater impact on the responses at the second and higher-order harmonics than at  $F_0$  (Experiment 1).

In contrast to our results, some past studies have demonstrated that additive acoustic noise degrades the  $\text{FFR}_{\text{ENV}}$  response, including at  $F_0$  (e.g., Anderson *et al.*, 2011); however, the size of the effect of additive noise on  $\text{FFR}_{\text{ENV}}$  appears to depend not only on the acoustic properties of the signal and noise, but also on other factors, such as perceptual abilities that can vary across listeners (e.g., see Cunningham *et al.*, 2001; Song *et al.*, 2011). We do not know of any studies of the effects of additive acoustic noise on  $\text{FFR}_{\text{ENV}}$  that used exactly the same stimuli as we employed here (harmonic complexes consisting of the first eight harmonics of 100 Hz, at equal intensity; broadband Gaussian noise). Further work is necessary to determine if the robustness of  $\text{FFR}_{\text{ENV}}$  at the fundamental frequency is idiosyncratic to the stimuli we used here, or a result that will generalize to other stimuli.

## V. CONCLUSIONS

$\text{FFR}_{\text{ENV}}$  responses from the scalp can be effectively analyzed by computing either the spectral magnitude or the phase-locking value at  $F_0$  of the acoustic input stimulus and its integer multiples. While these two measures produce metrics that are very strongly correlated, for the same inputs the PLV can be more sensitive, statistically, at least when floor effects are not at issue. In addition, the PLV can be directly interpreted and compared across conditions as long as the number of trials is the same in the conditions; in contrast,

interpreting the spectral magnitude requires more sophisticated statistical analyses to estimate the frequency-specific noise in the measurements.

For harmonic complex tones, high-frequency auditory peripheral channels respond to multiple harmonics of the input, generating cochlear-induced envelopes that contribute to  $\text{FFR}_{\text{ENV}}$  components at the fundamental frequency of the input stimulus, as well as to low-order harmonics of this  $F_0$ . The  $\text{FFR}_{\text{ENV}}$  response at the fundamental frequency of a periodic input stimulus is less affected by additive noise than other higher-frequency components.

## ACKNOWLEDGMENTS

This work was supported by the grant from NIH (Grant No. DC009477 to B.S.C.), a NSSEFF fellowship to B.S.C., a China National Science Grant (Grant No. 30970756 to L.Z.), and Tsinghua-Yu-Yuan Medical Science Grant (Grant No. 20200521 to L.Z.). Scott Bressler, Siddarth Rajaram, and Dorea Ruggles assisted in many aspects of this work. Bertrand Delgutte (Massachusetts Eye and Ear Infirmary) provided helpful comments on aspects of this work.

## APPENDIX: ANALYSIS METHODS

As noted in Sec. II F, the noise floor in raw magnitude measures varies with frequency. We used each subjects' data to estimate the expected distribution of the noise floor. Using this distribution, we estimated the magnitude responses in dB SNR as a function of frequency. We also used the noise floor distribution to compute the probability of observing a given magnitude response by chance in order to test the statistical significance of the observed magnitude response. Finally, we analyzed the phase-locking values of the same data with similar methods.

### 1. Raw spectral magnitudes

To compute the magnitude responses, we took the following steps:

- (1) For each subject  $s$ , draw  $N=400$  trials at random from positive and negative polarities in condition  $x$  (quiet/noise in Experiment 1; the four different stimuli in Experiment 2), pooled together (on average, 200 each of positive and negative polarity). Of these trials, the number of positive polarity trials,  $N_{\text{pos}}$ , and negative polarity trials,  $N_{\text{neg}}$ , are random, but sum to 400. Define  $\text{POS}_i$  and  $\text{NEG}_i$  as the random sets of selected positive and negative trials, respectively, for this random draw ( $i$ ).
- (2) Calculate the DFT for each of the selected time-domain trials,  $S_{n,x}^s(f)$ , where  $n$  is the trial number,  $x$  is condition, and  $f$  is frequency.
- (3) For subject  $s$ , draw  $i$ , and condition  $x$ , calculate the mean raw spectral magnitude of  $\text{FFR}_{\text{ENV}}$  for the  $N$  sampled trials,  $M_{\text{ENV},x,i}^s(f)$  (in dB, relative to the raw magnitudes of the readings in voltage) by averaging the complex-valued spectra of the individual trials over positive and negative polarity trials

$$M_{\text{ENV},x,i}^s(f) = 20 \log \left| \frac{1}{400} \left( \sum_{n \in \text{POS}_i} S_{n,x}^s(f) + \sum_{m \in \text{NEG}_i} S_{m,x}^s(f) \right) \right|.$$

- (4) Repeat steps 1 through 3 each 100 times to generate distributions of the raw spectral magnitudes of  $\text{FFR}_{\text{ENV}}$  ( $i=1$  to 100).
- (5) Calculate the means of the raw spectral magnitudes for each subject and condition

$$\hat{M}_{\text{ENV},x}^{s,\text{raw}}(f) = \frac{1}{100} \sum_{i=1}^{100} M_{\text{ENV},x,i}^s(f).$$

Because these means are computed by averaging identically distributed random variables, they are well approximated as Gaussian, allowing us to use standard parametric statistical tests. The raw spectral magnitude as a function of frequency is shown by the dashed, thin lines in Fig. 1 for one example subject.

### 2. Spectral noise floor

To estimate the individualized estimate of the spectral magnitude noise floor, we repeated steps 1–5 from Sec. A (in Appendix) exactly 1000 times for each subject, but with the phases of the spectra set to random values in step 3,

$$\hat{N}_{\text{ENV},x}^{s,\text{raw},k}(f) = \frac{1}{100} \sum_{i=1}^{100} 20 \log \left| \frac{1}{N} \left( \sum_{n \in \text{POS}_i} |S_{n,x}^s(f)| e^{j\omega_n} + \sum_{m \in \text{NEG}_i} |S_{m,x}^s(f)| e^{j\omega_m} \right) \right|,$$

where  $\omega_n$  and  $\omega_m$  are random values, uniformly distributed between 0 and  $2\pi$ , selected independently for each positive polarity trial, negative polarity trial, repetition  $i$ , and distribution sample  $1 \leq k \leq 1000$ . The mean of this distribution was our estimate of the expected noise floor in the measures. This estimate is shown in Fig. 1 by the thin blue lines plotted in each panel for one example subject.

### 3. Normalized spectral magnitudes

The estimated noise floor can be used to calculate the mean magnitude spectra in units of dB SNR for each subject and condition. In particular, we normalized the mean raw magnitude spectra by the sample mean of the distribution of the raw noise magnitude spectra estimates to find the magnitude spectrum of the  $\text{FFR}$  in units of dB SNR, shown in Fig. 1 as the solid, thick lines,

$$\hat{M}_{\text{ENV},x}^s(f) = \hat{M}_{\text{ENV},x}^{s,\text{raw}}(f) - \frac{1}{1000} \sum_{k=1}^{1000} \hat{N}_{\text{ENV},x}^{s,\text{raw},k}(f).$$

As Fig. 1 illustrates, this normalization of the raw magnitude spectra, while computationally expensive, takes into account differences in absolute spectral magnitude of the responses that are not due to the neural response, but rather arise because of differences in the noise characteristics.

#### 4. Statistical significance of spectral magnitude results

For each subject, we then determined whether the observed spectral magnitudes,  $\hat{M}_{ENV,x}^s(f)$ , evaluated at each harmonic of 100 Hz ( $f = h \cdot 100$ ), were significantly greater than would be expected by chance, based on the estimated noise floor distribution generated from the raw measures. Specifically, we compared  $\hat{M}_{ENV,x}^{s,raw}(f)$  to the null hypothesis distribution  $\{\hat{N}_{ENV,x}^{s,raw,k}(f)\}_{1 \leq k \leq 1000}$ . For each relevant harmonic, we estimated the probability of observing the observed spectral magnitude by chance by finding the percentage of the samples of  $\{\hat{N}_{ENV,x}^{s,raw,k}(f)\}_{1 \leq k \leq 1000}$  that were equal to or larger than the observed value  $\hat{M}_{ENV,x}^{s,raw}(f)$ . Results were Bonferroni corrected for multiple comparisons; a spectral magnitude value was deemed significant if our estimate of the probability of observing it by chance was less than  $p < 0.05/M$ , where  $M$  is the number of harmonics examined ( $M = 8$  in Experiment 1 and  $M = 20$  in Experiment 2, yielding threshold probabilities of 0.00625 in Experiment 1 and 0.0025 in Experiment 2).

#### 5. Phase-lock value computations

The steps used to estimate the PLVs for each subject and condition were

- (1) For each subject  $s$ , randomly draw 400 trials from positive and negative polarities in condition  $\times$  (quiet/noise in Experiment 1; the four different stimuli in Experiment 2), pooled together. Of these trials, the number of positive polarity trials,  $N_{pos}$ , and negative polarity trials,  $N_{neg}$ , are random, but sum to 400. Define  $POS_i$  and  $NEG_i$  as the random sets of selected positive and negative trials, respectively, for this random draw ( $i$ ).
- (2) Calculate the phases of the DFTs of the selected time-domain trials,  $S_{n,x}^s(f)$ , defined above, where  $n$  is the trial number,  $x$  is condition, and  $f$  is frequency,

$$\phi_{n,x}^s(f) = \angle S_{n,x}^s(f).$$

- (3) For subject  $s$ , draw  $i$ , and condition  $x$ , calculate the PLV of  $FFR_{ENV}$ ,

$$P_{ENV,x,i}^s(f) = \frac{1}{400} \left| \left( \sum_{n \in POS_i} e^{j\phi_{n,x}^s(f)} + \sum_{m \in NEG_i} e^{j\phi_{m,x}^s(f)} \right) \right|.$$

- (4) Repeat steps 1 through 3 each 100 times to generate distributions of the PLVs ( $i = 1$  to 100).
- (5) Calculate the means of the PLVs for each subject and condition,

$$\hat{P}_{ENV,x}^s(f) = \frac{1}{100} \sum_{i=1}^{100} P_{ENV,x,i}^s(f).$$

Because these means are computed by averaging identically distributed random variables, they are well

approximated as Gaussian, allowing us to use standard parametric statistical tests.

#### 6. Statistical significance of PLV results

For each subject, we then determined whether the observed PLVs,  $\hat{P}_{ENV,x}^s(f)$ , evaluated at specific harmonics of 100 Hz ( $f = h \cdot 100$ ), were significantly larger than would be expected by chance. Since the null model is the same for all conditions, subjects, and frequencies, we generated only one random distribution of PLVs  $\{\hat{R}^k\}_{1 \leq k \leq 1000}$ , computed by repeating steps 1–5 1000 times, but with the phase in step 3 of Sec. E set to be random (uniformly distributed, from 0 to  $2\pi$ , selected independently for each trial and repetition  $i$ ). We then compared  $\hat{P}_{ENV,x}^s(f)$  to this distribution to estimate the probability of observing that PLV by chance as the percentage of the samples of  $\{\hat{R}^k\}_{1 \leq k \leq 1000}$  that were equal to or larger than the observed value  $\hat{P}_{ENV,x}^s(f)$ . As with the spectral magnitude analysis, results were Bonferroni corrected for multiple comparisons; a PLV was deemed significant if our estimate of the probability of observing it by chance was less than  $p < 0.05/M$  ( $M = 8$  in Experiment 1 and  $M = 20$  in Experiment 2).

- Aiken, S. J., and Picton T. W. (2006). "Envelope following responses to natural vowels," *Audiol. Neuro-Otol.* **11**, 213–232.
- Akhoun, I., Gallego, S., Moulin, A., Menard, M., Veuillet, E., Berger-Vachon, C., Collet, L., and Thai-Van, H. (2008). "The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults," *Clin. Neurophysiol.* **119**, 922–933.
- 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.
- Bokil, H., Purpura, K., Schoffelen, J. M., Thomson, D., and Mitra, P. (2007). "Comparing spectra and coherences for groups of unequal size," *J. Neurosci. Methods* **159**(2), 337–345.
- Burman, D. D., Bitan, T., and Booth J. R. (2008). "Sex differences in neural processing of language among children," *Neuropsychologia* **46**, 1349–1362.
- Buzsaki, G. (2006). *Rhythms of the Brain* (Oxford University Press, New York), pp. 3–29.
- Carcagno, S., and Plack C. J. (2011). "Subcortical plasticity following perceptual learning in a pitch discrimination task," *J. Assoc. Res. Otolaryngol.* **12**, 89–100.
- Cebulla, M., Sturzebecher, E., and Wernecke K. D. (2001). "Objective detection of the amplitude modulation following response (AMFR)," *Audiology* **40**, 245252.
- Chandrasekaran, B., and Kraus, N. (2010). "The scalp-recorded brainstem response to speech: Neural origins and plasticity," *Psychophysiology* **47**, 236–246.
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., and Kraus, N. (2001). "Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement," *Clin. Neurophysiol.* **112**, 758–767.
- Dau, T. (2003). "The importance of cochlear processing for the formation of auditory brainstem and frequency following responses," *J. Acoust. Soc. Am.* **113**, 936–950.
- Davis, R. L., and Britt, R. H. (1984). "Analysis of the frequency following response in the cat," *Hear. Res.* **15**, 29–37.
- Dobie, R. A., and Wilson M. J. (1993). "Objective response detection in the frequency domain," *Electroencephalogr. Clin. Neurophysiol.* **88**, 516–524.
- Dobie R. A., and Wilson, M. J. (1994). "Objective detection of 40 Hz auditory evoked potentials: Phase coherence vs. magnitude-squared coherence," *Electroencephalogr. Clin. Neurophysiol.* **92**, 405–413.
- Dobie, R. A., and Wilson M. J. (1996). "A comparison of  $t$  test,  $F$  test, and coherence methods of detecting steady-state auditory-evoked potentials,

- distortion-product otoacoustic emissions, or other sinusoids," *J. Acoust. Soc. Am.* **100**, 2236–2246.
- Dolphin, W. F., and Mountain, D. C. (1992). "The envelope following response: Scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic signals," *Hear. Res.* **58**(1), 70–78.
- Du, Y., Kong, L., Wang, Q., Wu, X., and Li, L. (2011). "Auditory frequency-following response: A neurophysiological measure for studying the 'cocktail-party problem,'" *Neurosci. Biobehav. Rev.* **35**, 2046–2057.
- Galbraith, G. C., Threadgill, M. R., Hemsley, J., Salour, K., Songdej, N., Ton, J., and Cheung, L. (2000). "Putative measure of peripheral and brainstem frequency-following in humans," *Neurosci. Lett.* **292**, 123–127.
- Gardi, J., Merzenich, M., and Mckean, C. (1979). "Origin of the scalp recorded frequency-following response in the cat," *Audiology* **18**, 358–381.
- Gockel, H. E., Carlyon, R. P., Mehta, A., and Plack, C. J. (2011). "The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch," *J. Assoc. Res. Otolaryngol.* **12**, 767–782.
- Greenberg, S., Marsh, J. T., Brown, W. S., and Smith, J. C. (1987). "Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones," *Hear. Res.* **25**, 91–114.
- Harte, J. M., Ronne, F., and Dau, T. (2010). "Modeling human auditory evoked brainstem responses based on nonlinear cochlear processing," in *Proceedings of 20th International Congress on Acoustics*.
- Hornickel, J., Anderson, S., Skoe, E., Yi, H. G., and Kraus, N. (2012). "Subcortical representation of speech fine structure relates to reading ability," *NeuroReport* **23**, 6–9.
- Javel, E. (1980). "Coding of AM tones in the chinchilla auditory nerve: Implications for the pitch of complex tones," *J. Acoust. Soc. Am.* **68**, 133–146.
- Jerger, J., and Hall, J. (1980). "Effects of age and sex on auditory brainstem response," *Arch. Otolaryngol.* **106**, 387–391.
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). "Neural processing of amplitude-modulated sounds," *Physiol. Rev.* **84**, 541–577.
- Joris, P. X., and Yin, T. C. (1992). "Responses to amplitude-modulated tones in the auditory nerve of the cat," *J. Acoust. Soc. Am.* **91**, 215–232.
- Kraus, N., and Nicol, T. (2005). "Brainstem origins for cortical 'what' and 'where' pathways in the auditory system," *Trends. Neurosci.* **28**, 176–181.
- Krishnan, A. (1999). "Human frequency-following responses to two-tone approximations of steady-state vowels," *Audiol. Neuro-Otol.* **4**, 95–103.
- Krishnan, A. (2002). "Human frequency-following responses: Representation of steady-state synthetic vowels," *Hear. Res.* **166**, 192–201.
- Krishnan, A., Bidelman, G. M., and Gandour, J. T. (2010). "Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices," *Hear. Res.* **268**, 60–66.
- Krizman, J., Skoe, E., and Kraus, N. (2012). "Sex differences in auditory subcortical function," *Clin. Neurophysiol.* **123**, 590–597.
- Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). "Measuring phase synchrony in brain signals," *Hum. Brain Mapp.* **8**, 194–208.
- Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing*, 5th ed. (Academic, San Diego, CA), pp. 177–212.
- Palmer, A. R. (1982). "Encoding of rapid amplitude fluctuations by cochlear-nerve fibers in the guinea-pig," *Arch. Oto-Rhino-Laryngol.* **236**, 197–202.
- Pickles, J. O. (1982). *An Introduction to the Physiology of Hearing* (Academic, New York), pp. 26–77.
- Picton, T. W. (2011). *Human Auditory Evoked Potentials* (Plural, San Diego), pp. 213–245.
- Picton, T. W., Vajsar, J., Rodriguez, R., and Campbell, K. B. (1987). "Reliability estimates for steady-state evoked potentials," *Electroencephalogr. Clin. Neurophysiol.* **68**, 119–131.
- Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2011). "Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication," *Proc. Natl. Acad. Sci. U.S.A.* **37**, 15516–15521.
- Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2012). "Why middle-aged listeners have trouble hearing in everyday settings," *Curr. Biol.* **22**, 1417–1422.
- Shinn-Cunningham, B. G., Ruggles, D., and Bharadwaj, H. (2012). "How early aging and environment interact in everyday listening: From brainstem to behavior through modeling," *International Symposium on Hearing*, 23–27 July, Cambridge, UK.
- Smith, J. C., Marsh, J. T., and Brown, W. S. (1975). "Far-field recorded frequency-following response evidence for the locus of brainstem sources," *Electroencephalogr. Clin. Neurophysiol.* **39**, 465–472.
- Sohmer, H., and Pratt, H. (1977). "Identification and separation of acoustic frequency following responses (FFRs) in man," *Electroencephalogr. Clin. Neurophysiol.* **42**(4), 493–500.
- Sohmer, H., Pratt, H., and Kinarti, R. (1977). "Sources of frequency following responses (FFR) in man," *Electroencephalogr. Clin. Neurophysiol.* **42**(5), 656–664.
- Song, J. H., Skoe, E., Banai, K., and Kraus, N. (2011). "Perception of speech in noise: Neural correlates," *J. Cogn. Neurosci.* **23**, 2268–2279.
- Stapells, D. R., Makeig, S., and Galambos, R. (1987). "Auditory steady-state responses: Threshold prediction using phase coherence," *Electroencephalogr. Clin. Neurophysiol.* **67**, 260–270.
- Sturzebecher, E., and Cebulla, M. (1997). "Objective detection of auditory evoked potentials: Comparison of several statistical tests in the frequency domain on the basis of near-threshold ABR data," *Scand. Audiol.* **26**, 7–14.
- Thomson, D. (1982). "Spectrum estimation and harmonic analysis," *Proc. IEEE* **70**, 1055–1096.
- Victor, J. D., and Mast, J. (1991). "A new statistic for steady-state evoked potentials," *Electroencephalogr. Clin. Neurophysiol.* **78**, 378–388.
- Wile, D., and Balaban, E. (2007). "An auditory neural correlate suggests a mechanism underlying holistic of pitch perception," *PLoS ONE* **2**, e369.