NEUROSCIENCE
RESEARCH ARTICLE
Le Wang et al. / Neuroscience 407 (2019) 67–74

Assessing Cochlear-Place Specific Temporal Coding Using Multi-Band Complex Tones to Measure Envelope-Following Responses

Le Wang,[a] Hari Bharadwaj[b] and Barbara Shinn-Cunningham[ac]*

[a]*Biomedical Engineering, Boston University, 610 Commonwealth Ave, Boston, MA 02215, USA*

[b]*Speech, Language, and Hearing Sciences, Purdue University, Lyles-Porter Hall, 715 Clinic Dr., West Lafayette, IN 47907, USA*

[c]*Neuroscience Institute, Carnegie Mellon University, Room 115C, 4400 Fifth Ave., Pittsburgh, PA 15213, USA*

Abstract—Previous studies suggest that envelope-following responses (EFRs) reveal important differences in temporal coding fidelity amongst listeners who have normal hearing thresholds, consistent with these listeners differing in the degree to which they suffer from cochlear synaptopathy. Like conventional hearing loss, the severity of cochlear synaptopathy may vary along the cochlea. A number of earlier studies have suggested methods for estimating EFRs driven by specific frequency regions of the cochlea, which would allow synaptopathy to be estimated as a function of cochlear place. Here, we tested a method for measuring EFRs from multiple locations along the cochlea simultaneously, using narrowband stimuli. We compared responses to multiple simultaneous narrowband complex harmonic tones in three non-overlapping frequency bands, each having a unique fundamental frequency, to responses to the individual narrowband stimuli alone, and to responses when noise was added to different combinations of the frequency bands. Our results suggest that simultaneous presentation of multiple tone complexes with different fundamental frequencies leads to repeatable measures of temporal coding fidelity at the cochlear frequency regions corresponding to the narrowband carrier frequencies. Other results suggested that while off-frequency contributions to EFRs driven by narrowband signals (due to spread of excitation) can add destructively to the on frequency response, these interactions were small compared to EFR magnitude. Overall, our results point to the utility of using multi-band complex tone stimuli to estimate the profile of temporal coding fidelity, and thus the degree of synaptopathy, as a function of cochlear place.

This article is part of a Special Issue entitled: Hearing Loss, Tinnitus, Hyperacusis, Central Gain.

Key words: cochlear synaptopathy, hidden hearing loss, spread of excitation, coding fidelity.

INTRODUCTION

Responses phase locked to periodic sounds originating from either the subcortical or cortical portions of the auditory pathway often are collectively referred to as auditory steady-state responses or ASSRs (Galambos et al., 1981; Stapells et al., 1984). The envelope following response (EFR) is a specific form of ASSR measured by presenting a periodic input signal (typically with a periodicity in the 80–450 Hz range so that the subcortical portion of the response is emphasized) in opposite polarities and then averaging the responses (Goblick and Pfeiffer, 1969; Aiken and Picton, 2008).

The EFR reveals important differences in temporal coding fidelity in listeners with normal hearing thresholds, using either narrowband stimuli or broadband stimuli with a single envelope (Ruggles et al., 2012; Bharadwaj et al., 2015; Roberts et al., 2015). The finding that individual differences in auditory perceptual ability are related to differences in objective physiological measurements supports the idea that sensory coding fidelity differs amongst listeners with normal audiometric thresholds and that this affects hearing in everyday settings.

EFRs can be recorded to many different stimuli at the same time. Specifically, multiple carrier frequencies, each modulated by its own distinct, signature frequency, can be used to evoke distinct EFRs that are separable EFRs in the frequency domain, with each response reflected in the particular frequency at which the corresponding carrier is modulated (Lins and Picton, 1995; John et al., 1998). With this design, modulation sensitivity driven by the different locations along

*Corresponding author at: Biomedical Engineering, Boston University, 610 Commonwealth Ave, Boston, MA 02215, USA.
E-mail address: lwang@bu.edu (Le Wang). hbharadw@purdue.edu (Hari Bharadwaj). bgsc@cmu.edu (Barbara Shinn-Cunningham).

Abbreviations: AN, auditory nerve; ASSR, auditory steady-state response; CF, center frequency; cPCA, complex principal component analysis; EEG, electroencephalography; EFR, envelope-following response; IHC, inner hair cell; PCA, principal component analysis; PLV, phase-locking value; PSD, power spectral density; SNR, signal-to-noise ratio; SPL, sound pressure level.

https://doi.org/10.1016/j.neuroscience.2019.02.003
0306-4522/

the cochlea that correspond to the different carrier frequencies (i.e., different auditory frequency channels) can be probed simultaneously. Recording multiple responses simultaneously leads to a significant reduction in recording time. However, simultaneous presentation of multiple modulation bands may lead to significant interactions in the cochlea as well as the brain (e.g., see Ananthanarayan and Durrant, 1992; Gockel et al., 2015). These interactions could reduce EFR amplitudes evoked by a given carrier and interfere with the cochlear-place specificity of the responses, which could compromise interpretation of the results.

This paper investigates the recording multiple concurrent EFRs at supra-threshold sound levels. The investigation involves two experiments that specifically focused on the limitations of simultaneous presentation of multiple, simultaneous modulation bands. In the first experiment, we compared the EFR measured when multiple independent stimulus bands were presented simultaneously to the EFRs measured when each of the component stimulus bands were presented individually. This experiment is useful in assessing the degree to which presentation of modulated frequency bands at other frequencies influences the EFR in response to the modulation in a particular carrier. In the second experiment, multiple stimuli were simultaneously presented along with Gaussian noise added to some or all frequency bands of the stimuli. This experiment evaluated the cochlear place specificity of the simultaneous, multi-band paradigm by comparing the EFR amplitudes in one frequency band with and without additive noise in other bands.

EXPERIMENTAL PROCEDURES

Participants

Forty participants (age 18 to 53 years old) were recruited in accordance with procedures approved by the Boston University Charles River Campus Institutional Review Board and were paid for their participation. Eight of these participants performed Experiment 1. Six who participated in Experiment 1 also performed Experiment 2, along with thirty-two additional participants (total of 34 participants in Experiment 2). For all subjects, pure-tone audiometric thresholds were measured from 125 Hz to 8000 Hz at octave intervals. All participants had hearing thresholds within 20 dB HL in each ear at all tested frequencies, and none reported any history of central or peripheral auditory deficits.

Stimuli and Procedures

Throughout data collection, participants watched a muted, captioned movie of their choice, ignoring the acoustic stimuli. Each subject in each experiment participated in a single experimental session lasting up to about one hour. Both experiments presented different trial types within each session, which were intermingled in random order during the session. Each trial type was presented 1000 times; for a randomly selected set of 500 out of the 1000 trials of each stimulus type, the polarity of the complex tones was inverted. Averaging the responses to both the original and inverted stimuli allows the responses phase-locked to the cochlear-induced envelope to be separated from the responses phase-locked to the temporal fine-structure of the acoustic inputs. To reduce spectral splatter, the onset and offset of each stimulus were tapered using the rising and falling half of a 20-ms-long Slepian sequence, respectively (the first Slepian sequence of time half bandwidth product = 1). The inter-trial interval was randomly chosen from trial to trial from a uniform distribution between 300 and 400 ms. Stimuli for both Experiment 1 and Experiment 2 were generated in MATLAB (Natick, MA) using a sampling rate of 48,828 Hz.

In Experiment 1, each trial was 300-ms-long and consisted of three harmonic complex tones with different fundamental frequencies (F0s), which were resolvable in a frequency-domain decomposition for the chosen duration of the tokens. The harmonic complexes were chosen to have non-overlapping frequency components. The low band complex tone has a frequency range of 500–1500 Hz and a F0 of 114 Hz; the mid band complex tone has a frequency range of 2000–3500 Hz and a F0 of 170 Hz; and the high band complex tone has a frequency range of 4000–6000 Hz and a F0 of 236 Hz (Fig. 1). The three complex tones were either presented individually to get single-band EFRs or were presented concurrently to obtain multiple EFRs simultaneously (multi-band EFR), for a total of four different trial types. The sound level of each complex tone was 70 dB Sound Pressure Level (SPL), so the overall stimulus intensity was ~75 dB SPL for trials with all three complex tones presented simultaneously. For each subject, a total of 4000 trials was presented in Experiment 1.

In Experiment 2, the same multiband mixture of three complex tones used in Experiment 1 was presented with different masking noises. For each trial, band-limited noise was added to either a single band (Low, Mid or High), or combinations of frequency bands (Low + Mid, Low + High, Mid + High, and Low + Mid + High). The condition where no noise was added to any band was also included as a control condition, leading to 8 distinct types of trials in total (Fig. 1). The intensity of the noise in each band, when present, was 70 dB SPL, resulting in a maximal stimulus intensity of ~78 dB SPL when noise was added to all three bands. The spectral level of the noise is different across bands because the 70 dB SPL noise energy was distributed over different bandwidths. For each subject, a total of 8000 trials were presented in Experiment 2.

Equipment

A personal computer controlled all aspects of the experiment, including triggering sound delivery and storing data.

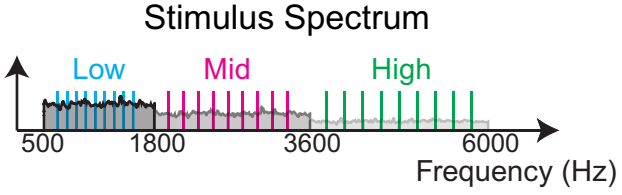

Fig. 1. Diagram for the spectrum of the multi-band stimulus with noises in all bands. Vertical lines represent harmonic components of the complex tone in low band (cyan), mid band (magenta), and high band (green). Shaded area represents the additive noise in each band. Both the complex tone and the additive noise in each band, if included, were presented at 70 dB SPL.

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. Scalp responses were recorded in 32 channels using a BioSemi Active Two System (BioSemi, Amsterdam, Netherlands) at a sampling rate of 4096 kHz in a sound-shielded booth (single-walled Eckel C-14 booth, Cambridge, MA). Prior to data acquisition, we ensured that the offset voltage for each active electrode was stabilized at <30 mV.

Data Analysis

For both experiments, scalp recordings in each channel were re-referenced offline to the average potential recorded at the two mastoids using additional surface electrodes. The measurements were then high-pass filtered in Matlab at 70Hz using an FIR filter with zero group-delay to minimize the signal contributions from cortical sources (Dolphin and Mountain, 1992; Herdman et al., 2002; Kuwada et al., 2002). Epochs from 0 to 300ms relative to the onset of each trial were segmented out from each channel. All epochs with dynamic range larger than 200μV in any channel were excluded from further analysis to remove movement and muscle activity artifact.

The epochs extracted from the 32-channel data were processed using complex principal component analysis (cPCA) to estimate the phase-locking value (PLV) and the spectral power density (PSD; see Bharadwaj and Shinn-Cunningham, 2014). Compared with the traditional time-domain PCA, the frequency domain cPCA provides a significant enhancement in signal-to-noise ratio (SNR) when combining multiple measurements in the presence of between-measurement correlations at non-zero phase delays. For each trial type, we computed an average of 20 different PLV and PSD computations. Each of the individual PLV and PSD computations was computed by randomly drawing 400 epochs of each polarity from the epochs that survived the artifact rejection. The PLV and PSD computed from the 20 different draws were then averaged and the mean PLV at F0 was used as a measure of EFR strength. For the PSD, we characterized the EFR strength as the sum of the spectral power at the F0 and its first four harmonic frequencies in order to capture neural responses to the modulation that are not purely sinusoidal.

Model Simulation

Auditory nerve (AN) responses to the complex tone stimuli in Experiment 2 were simulated using an established computational model (Zilany et al., 2014). This AN model incorporates detailed descriptions of cochlear processing, including frequency tuning, cochlear amplification, inner hair cell (IHC) transduction, and refractoriness in AN fibers. The center frequency (CF) of the AN in the model spanned from 125 Hz to 12 kHz in 200 log-spaced steps according to the human cochlear map (Shera et al., 2002). The AN firing rate at each CF was generated 50 times, each time with a different noise realization in the stimulus. The average AN firing rates were then computed for each AN.

RESULTS

Comparison of Single-Band EFR and Multi-Band EFR

The EFR for an example subject is show in Fig. 2. The single-band FR shows clear peaks at the F0s of the complex tones in each band, as well as the first few harmonics of each F0. There is a good correspondence between the single-band

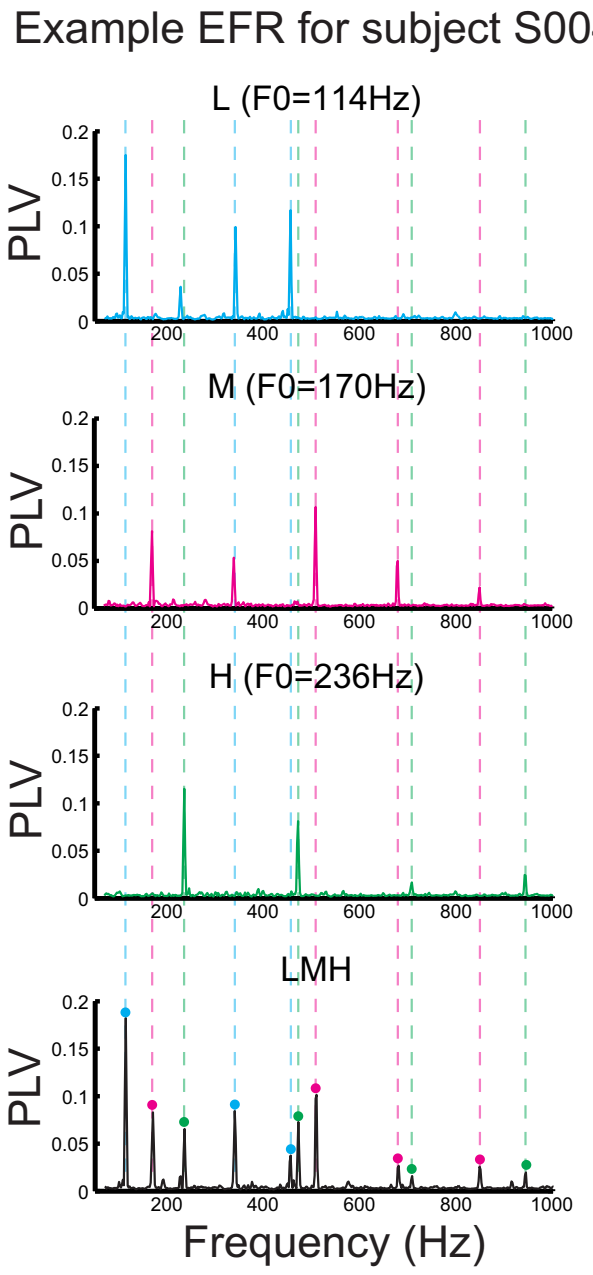

Fig. 2. An example EFR for one subject in Experiment 1 (no additive noise) is shown as a function of frequency. The first three panels show EFR to single band complex tones (from top to bottom: low-band, mid-band, high-band). The lowest panel shows EFR to the multi-band stimulus containing all three complex tones. Vertical lines and the dots highlight the frequency correspondence between single-band EFR and multi-band EFR, and their color represents the frequency band each multi-band EFR peak is originated from.

EFRs and the multi-band EFR, save for small discrepancies in peak size at a few frequencies. To assess the effect of simultaneous presentation of stimuli in other frequency bands, a group level comparison of single-band and multi-band EFRs is summarized in Fig. 3. The low-band EFR amplitude in the multi-band condition is significantly higher than that in the single-band condition (mean difference of -0.62 dB, P = .05, Wilcoxon signed-rank test). However, the EFR amplitude for the high-band complex tone is significantly lower in the multi-band condition than in the single-band condition (mean difference of 1.43 dB, P = .04, Wilcoxon signed-rank test). The mid-band EFR also shows a trend for higher EFR amplitudes in the single-band condition than the multi-band condition (mean difference of 1.37 dB, P = .08, Wilcoxon signed-rank test).

Test-Retest Reliability of the Multi-Band EFR

Of the eight subjects who participated in Experiment 1, six also participated in Experiment 2. Because both Experiment 1 and 2 contained the multi-band condition where all three complex tones were presented simultaneously in the absence of noise, we examined the test-retest reliability of the multi-band EFR in the same cohort of subjects in two sessions measured on different days. EFR amplitudes for the low-, mid- and high-band condition are summarized in Fig. 4. In line with the fact that the EEG power decreases with frequency, the low-band EFR amplitudes is greater on average than the mid-band EFR, and both the low- and mid-band EFR amplitudes are greater than the high-band EFR. After the frequency-band effect is partialled out, the EFR amplitudes in two sessions are significantly correlated (Pearson's r = 0.72, P = .0008), demonstrating good test-retest reliability of the multi-band EFR measurement.

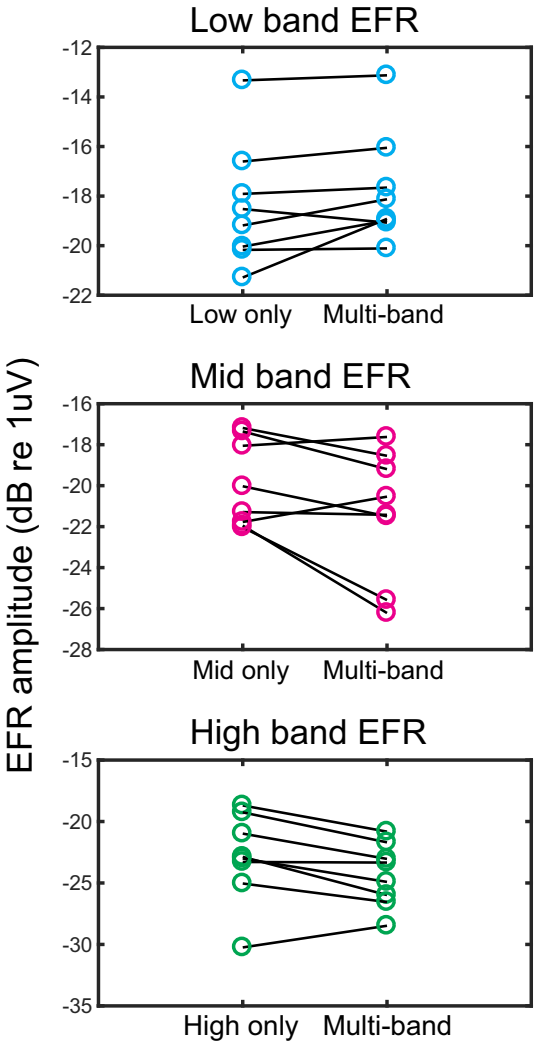

Fig. 3. Comparison of EFR amplitude between the single-band and multi-band condition. Each circle represents a subject. Upper panel: EFR to the low-band complex tone; Mid panel: EFR to the mid-band complex tone; Lower panel: EFR to the high-band complex tone.

Effect of Band-Specific Noise on the Multi-Band EFR

The group average data for the multi-band EFR with additive noise are shown in Fig. 5. For each frequency band, adding noise in that frequency band decreases the EFR amplitude for that band, as expected. What is important is that the EFR amplitude for one band is not affected by the addition of noise in the other two bands. This result provides strong evidence that the multi-band EFR measurement reflects neural activity from well-separated frequency regions of the cochlea, with little contribution from off-frequency bands. Because we measured the multi-band EFR on the same subjects both with and without the additive noise, these two responses can be subtracted to yield a derived metric, that is, the EFR amplitude drop, for each subject. This approach reduces the influence of individual differences in overall EFR amplitude (see Fig. 4) and other nuisance variables that can affect EFR measurements, producing values that are more directly comparable across subjects (e.g., see Bharadwaj et al., 2015). The EFR amplitude drops in different

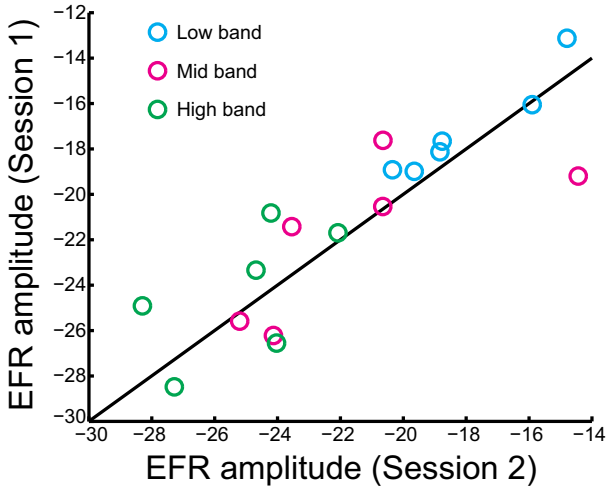

Fig. 4. Test-retest reliability for multi-band EFR with no additive noise (N = 6, subjects who participated in both Experiment 1 and Experiment 2). Session 1 refers to the multi-band EFR in Experiment 1, Session 2 refers to the multi-band EFR in Experiment 2. Colors represent different frequency band. Each circle represents measurements from one condition for one subject.

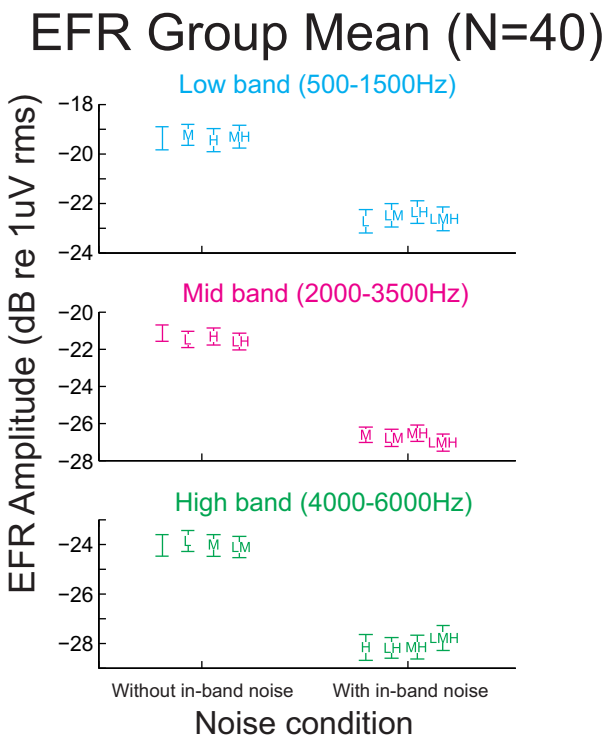

Fig. 5. Group data of the multi-band EFR with additive noise. EFR amplitude is plotted for low band EFR (top panel), mid band EFR (mid panel), and high band EFR (bottom panel). Error bar represents the standard error of the mean. Letters on the error bar represent the frequency band(s) where noise was present. The leftmost error bar in each panel represents the multi-band EFR without any noise.

frequency bands are not significantly correlated with one another (Fig. 6), further supporting the idea that responses in different bands in the multi-band EFR arise from different, non-overlapping cochlear frequency regions.

To further demonstrate the cochlear place specificity of the multi-band EFR, we simulated AN responses to the multi-band stimuli using a computational AN model (Fig. 7). When no noise is present, neurons within each frequency band show good phase locking to the F0 of that band. Adding noise in one frequency band greatly degrades the phase locking in that band with little effect on the neural response in other bands, consistent with the EFR data in Fig. 5.

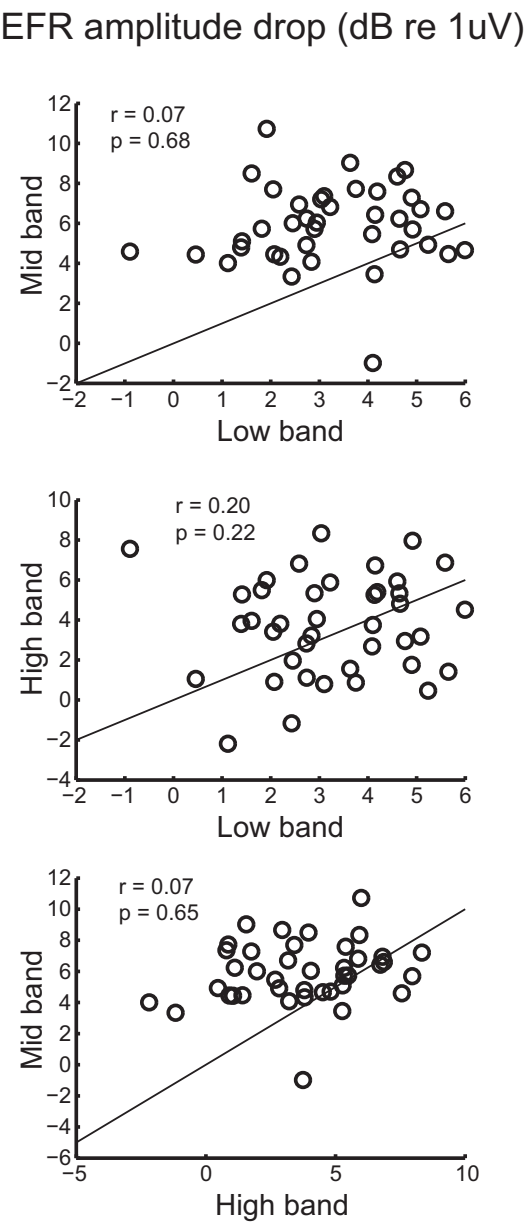

Fig. 6. Comparison of the EFR amplitude drop across frequency bands. Each circle represents a subject. The straight line in each panel represents the line x=y.

DISCUSSION

Off-Band Components Have Different Effects on EFRs to Different Carrier Frequencies

The EFR elicited by each different frequency band carrier was measured both in the single band condition (Experiment 1) and in the multi-band condition (Experiment 2, no noise condition). The EFR at a particular modulation frequency may include responses from frequency regions of the cochlea relatively far from the frequency region that is maximally excited, due to spread of excitation. This is especially likely when a narrowband carrier of a particular modulation frequency is presented alone and the spread of activity of that acoustic input is not masked by other components in the input stimulus. We expected that multi-band stimuli would reduce off-frequency contributions to the EFR, but we found that off-band stimulus components had different effects on EFR amplitudes in different frequency bands (Fig. 3). Specifically, for the low-band EFR at the F0 of 114 Hz, the multi-band condition gives rise to a small, but significant increase in the EFR amplitude relative to the single-band condition, while mid-band and high-band EFRs amplitudes were smaller in the multi-band condition compared to the single-band condition.

Importantly, different regions of the cochlea that respond to the same input modulation may have different response phases. If responses at a distant cochlear region are out of phase with the response from the maximally excited cochlear region, the off-frequency modulation response sums will decrease rather than increase the overall EFR amplitude. By

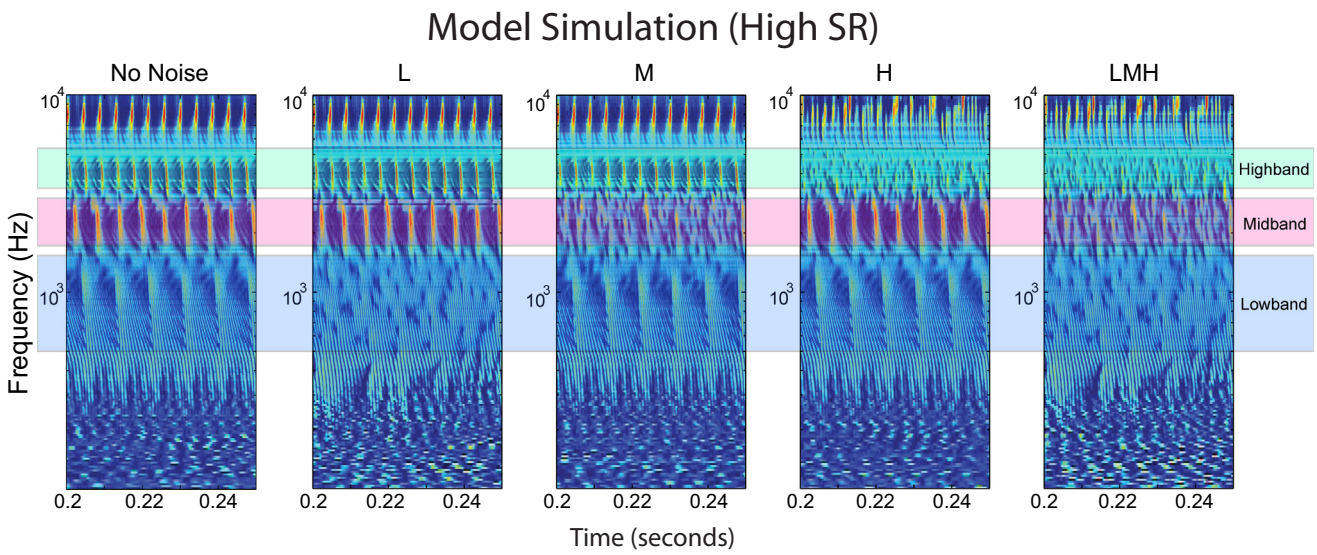

Fig. 7. Simulated AN firing rates in response to multi-band complex tones for 200 CFs. Labels on top of the panels represent the frequency band where the band-specific noise was present. The frequency range for the high-band, mid-band and low-band complex tone is represented by the shaded areas.

this logic, the low-band EFR in the multi-band condition (*EFRMB114*) should reflect neural responses dominated by the low band (*EFRLow114*), since the off-band stimulus components will mask the spread of excitation of the low-band carrier components into higher-frequency regions of the cochlea (i.e., *EFRMB114* = *EFRLow114*). However, in the single-band condition, the low-band EFR at 114 Hz (*EFRSB114*) is the sum of the neural responses from the low-band carrier-frequency region and higher cochlear frequency regions (i.e., *EFRSB114* = *EFRLow114* + *EFRHigh114*). The fact that the amplitude of *EFRMB114* is higher than that if *EFRSB114* suggests that *EFRHigh114* adds destructively to *EFRLow114*, which leads to partial cancellation of the 114 Hz EFR.

On the surface, this result appears to be inconsistent with prior studies that measured EFRs in human using multiple simultaneous narrowband stimuli (Lins and Picton, 1995; John et al., 1998). For example, the EFR amplitude for a 1 kHz SAM tone dropped 1.3 dB on average with the addition of a SAM tone at 1.5 kHz (0.4 octave frequency separation) when both SAM tones were presented at 60 dB SPL (John et al., 1998). In the present study, the frequency separation between low-band and mid-band is 0.42 octave (from 1500 Hz to 2000 Hz), which is comparable to that in John et al. (1998), but instead of a decrease in EFR amplitude, simultaneously presenting mid- and high-band components lead to a 0.6 dB increase in the EFR amplitude. While the current results seem qualitatively at odds with this previous result, the discrepancy may be explained by a relatively small phase difference between the on-frequency and off-frequency components that changes the summation in the single-band case from facilitation to cancellation. Such phase differences could arise due to seemingly modest differences in the stimulus parameters, such as the envelope shape and stimulus intensity, both of which have been shown to affect the phase of the EFR (Kuwada et al., 1986; Ross et al., 2000). Consistent with such an explanation, compared to the stimuli used by John et al. (1998), the complex tones in the present study have much "peakier" envelopes than the SAM tones, and the stimulus level is also 10 dB higher. Differences in these parameters may contribute to the discrepancy in the results between the present and prior studies.

In contrast with the low-band EFR, both the mid-band and high-band EFR (Fig. 3B and C) show smaller EFR amplitudes in the multi-band condition than in the single-band condition. Given the asymmetry in spread of excitation (extending to higher frequencies much more than to lower frequencies), the mid-band EFR in the single-band condition (*EFRSB170*) comes mainly from the mid- and the high-band cochlear regions. The phase difference between the responses from the mid band and the high band is likely very small, so the two responses are likely to sum constructively rather than destructively. The presence of a separate high-band stimulus in the multi-band condition suppresses any spread of excitation in the mid-band single-band contribution, explaining why the multi-band presentation decreases the mid-band EFR amplitude. This reasoning, however, cannot explain why the high-band EFR is weaker in the multi-band condition than the single-band condition, as no stimulus is present in the frequencies above the high band. It is possible that the presence of low- and mid-band inputs interfered with the cochlear responses in the high band due to the spread of excitation from the lower bands, and resulted in the reduction of the high-band EFR amplitude. This effect might also contribute to the decrease of mid-band EFR in the presence of low-band stimuli. Another possible reason for the decrease of high-band EFR is that at a relatively high intensity (70 dB), the high-band stimulus may also activate some of the mid-band AN fibers due to spread of excitation towards the lower frequencies and that the suppression of those fibers in the

multi-band condition may also contribute to the observed decrease in the high-band EFR amplitude.

Multi-Band EFR Measurement is Cochlear-Place Specific

Spread of excitation limits the specificity of EFR measurements at suprathreshold sound levels; even if the acoustic carrier is bandlimited, the regions of the cochlea that respond to a particular stimulus can be quite broad. Early EFR studies used narrowband stimuli to obtain audiometric measures at multiple carrier frequencies (e.g., John et al., 1998). These studies generally used relatively low stimulus levels (<60 dB SPL), producing less spread of excitation. As a result, it is likely that the EFRs for different carrier frequencies were driven by well-separated regions of the cochlea. For sounds with levels well above thresholds (suprathreshold sounds), the spread of excitation is greater and narrowband EFRs no longer are driven solely from well-defined, place-specific cochlear regions.

Several techniques have been proposed to improve the place specificity of EFRs evoked by suprathreshold sounds. One technique is to use notched-noise masking (Terkildsen et al., 1975; Picton et al., 1979). This method uses a broadband masker with a spectral gap, or notch, that covers the frequency range of the narrowband signal. This paradigm provides frequency specific assessments of temporal encoding in the frequency range of the narrowband stimulus. However, one problem arises with notch noise masking: noise frequencies below the lower edge of the notch spread into the spectral gap and can partially mask responses from the frequency region of interest, which could lead to reduced amplitudes in the frequency-specific responses obtained (Wegner and Dau, 2002). Likely for this reason, in pilot experiments, many of our subjects failed to show significant EFRs in one or more frequency bands.

One way to ameliorate this problem while maintaining frequency specificity is to use high-pass masking noise, gradually increasing the high-frequency cutoff until a broadband stimulus is presented. Frequency-delimited response can then be derived from the subtraction of two separate recordings of different high-pass conditions (Don and Eggermont, 1978; Nuttall et al., 2015). This method does not have low-frequency noise masking like is present using the notched-noise method; however, because the method depends upon the subtraction of two separate, noisy EFRs obtained in the two different high-pass noise conditions, there is greater noise in the result (and a decreased signal-to-noise ratio). In addition, the high-pass noise method requires that the noise be fixed at a spectral level that completely eliminates the EFR when the noise covers the entire broadband frequency range. In order to keep the noise at a comfortable and safe sound level, the broadband signal typically has to be below 60 dB SPL. As a result, this method cannot be used to assess suprathreshold hearing above 70 dB SPL.

In the present study, rather than using noise to suppress off-frequency contributions, we presented multiple narrowband stimuli in three non-overlapping frequency bands, each with a different fundamental modulation frequency. This approach allowed us to measure temporal encoding in different cochlear regions within the same recording session by separating the responses in the frequency domain. Although this method shows some signs of low-frequency masking, the effect size is small (~1 dB) compared to the EFR size. The majority of our subjects showed significant EFRs in response to each of the three simultaneous carriers. Furthermore, the measured EFR in response to each narrowband carrier seems to be well isolated in that the response to each band is only affected by additive noise that has acoustic energy in the carrier band; responses are largely unaffected by addition of off-frequency noise. Thus, the multi-band EFR is not only reliable, but reflects place-specific cochlear responses. Two recent studies also examined phase interactions between sets of harmonic complex tones with different fundamental frequencies (Easwar et al., 2018a,b). The data in the present study are largely consistent with the main findings in these studies.

The lack of interactions across bands in the present study may come from two aspects of the experiment design. First, the subjects were passively watching a silent movie. This minimized the possibility of cortical modulation differently modulating responses to different bands or influencing cross-frequency interactions. Second, the relatively high modulation frequency of the stimulus (>110Hz) means that the EFR is primarily coming from a subcortical source such as the midbrain. Indeed, a recent study from our lab showed that top-down selective attention has no significant influence on EFR responses, consistent with the dominant generator coming from a subcortical source (Varghese et al., 2015). As a result, the present approach is suited to assess the intactness of the periphery with minimal influences from the cortical activity. In addition, the approach enables assessment of cochlear responses from multiple cochlear regions at once, in one recording. This is much faster than either notched-noise or subtraction approaches, which require different measures for each carrier band to be evaluated.

In summary, we measured EFRs in listeners with normal hearing thresholds using multiple band-limited complex tones. When the three bandpass complex tones were presented simultaneously, the EFR amplitudes showed small changes compared to when the complex tones were presented individually. The differences between single-band and multi-band EFRs for different carrier frequencies suggest that off-frequency contributions to a low-band carrier signal can add destructively with the on-frequency response in single-band conditions, but not for medium- and high-band carrier signals. In the multi-band condition, the EFR amplitude for each band was unaffected by additive noise in other bands, demonstrating good cochlear place specificity for the multi-band EFR measurements. In addition, the multi-band EFR measurement also shows good test-retest reliability across multiple sessions.

The rationale behind multi-band EFR measurements is similar to that behind notched-noise masking. Instead of using noise to suppress off-frequency contributions to the EFR evoked by a narrowband carrier, we used other EFR-inducing stimuli, each with a different modulation frequency. This approach allows us to measure EFRs

evoked from different cochlear frequency regions all at the same time.

The exact stimuli we used here only serve as an example of the approach. Stimulus parameters could be adjusted to better optimize measurements. For example, one could add a fourth frequency band covering cochlear frequencies higher than 6 kHz, or one could choose to use narrower frequency ranges and a larger number of carrier bands to obtain a finer assessment of cochlear health. The main requirement is that the F0 for each band can be resolved from responses to other carrier bands in the frequency domain. The experimenter must simply find the right balance between frequency resolution and the signal-to-noise ratio within each band allowed by the EFR measurement.

The multi-band EFR measurement can provide information about envelope coding fidelity in multiple frequency bands in one recording. The reliability, cochlear place specificity and time efficiency makes this method a promising candidate as a clinical test for assessing envelope encoding fidelity across the cochlea.

ACKNOWLEDGMENT

National Institutes of Deafness and Communication Disorders R01 DC013825.

REFERENCES

Aiken SJ, Picton TW. (2008) Envelope and spectral frequency-following responses to vowel sounds. Hear Res 245:35-47, https://doi.org/10.1016/j.heares.2008.08.004.

Ananthanarayan AK, Durrant JD. (1992) The frequency-following response and the onset response: evaluation of frequency specificity using a forward-masking paradigm. Ear Hear.

Bharadwaj HM, Shinn-Cunningham BG. (2014) Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin Neurophysiol 125:1878-1888, https://doi.org/10.1016/j.clinph.2014.01.011.

Bharadwaj HM, Masud S, Mehraei G, et al. (2015) Individual Differences Reveal Correlates of Hidden Hearing Deficits. J Neurosci 35:2161-2172, https://doi.org/10.1523/JNEUROSCI.3915-14.2015.

Dolphin WF, Mountain DC. (1992) The envelope following response: Scalp potentials elicited in the mongolian gerbil using sinusoidally AM acoustic signals. Hear Res 58:70-78, https://doi.org/10.1016/0378-5955(92)90010-K.

Don M, Eggermont JJ. (1978) Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. J Acoust Soc Am 63:1084-1092, https://doi.org/10.1121/1.381816.

Easwar V, Banyard A, Aiken S, Purcell D. (2018) Phase delays between tone pairs reveal interactions in scalp-recorded envelope following responses. Neurosci Lett, https://doi.org/10.1016/j.neulet.2017.12.014.

Easwar V, Banyard A, Aiken SJ, Purcell DW. (2018) Phase-locked responses to the vowel envelope vary in scalp-recorded amplitude due to across-frequency response interactions. Eur J Neurosci, https://doi.org/10.1111/ejn.14161.

Galambos R, Makeig S, Talmachoff PJ. (1981) A 40-Hz auditory potential recorded from the human scalp. Proc Natl Acad Sci 78:2643-2647, https://doi.org/10.1073/pnas.78.4.2643.

Goblick TJ, Pfeiffer RR. (1969) Time-Domain Measurements of Cochlear Nonlinearities Using Combination Click Stimuli. J Acoust Soc Am, https://doi.org/10.1121/1.1911812.

Gockel HE, Krugliak A, Plack CJ, Carlyon RP. (2015) Specificity of the human frequency following response for carrier and modulation frequency assessed using adaptation. JARO - J Assoc Res Otolaryngol. , https://doi.org/10.1007/s10162-015-0533-9.

Herdman AT, Lins O, Van Roon P, et al. (2002) Intracerebral sources of human auditory steady-state responses. Brain Topogr 15:69-86, https://doi.org/10.1023/A:1021470822922.

John MS, Lins OG, Boucher BL, Picton TW. (1998) Multiple Auditory Steady-state Responses (MASTER): Stimulus and recording parameters. Int. J. Audiol. 37:59-82.

Kuwada S, Batra R, Maher VL. (1986) Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. Hear Res 21:179-192, https://doi.org/10.1016/0378-5955(86)90038-9.

Kuwada S, Andersont JS, Batrat R, et al. (2002) Sources of the scalp-recorded amplitude-modulation following response. J Am Acad Audiol 13(4):188-204

Lins O, Picton T. (1995) Auditory steady-state responses to multiple simultaneous stimuli. Electroencephalogr Clin Neurophysiol 96:420-432, https://doi.org/10.1016/0168-5597(95)00048-W.

Nuttall HE, Moore DR, Barry JG, et al. (2015) The influence of cochlear spectral processing on the timing and amplitude of the speech-evoked auditory brain stem response. J Neurophysiol 113:3683-3691, https://doi.org/10.1152/jn.00548.2014.

Picton T, Oullette J, Hamel G, Smith A. (1979) Brainstem evoked potentials to tone pips in notched noise. J Otolaryngol 8:289-314.

Roberts LE, Bosnyak DJ, Bruce IC, et al. (2015) Evidence for differential modulation of primary and nonprimary auditory cortex by forward masking in tinnitus. Hear Res 327:9-27.

Ross B, Borgmann C, Draganova R, et al. (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am, https://doi.org/10.1121/1.429600.

Ruggles D, Bharadwaj H, Shinn-Cunningham BG. (2012) Why middle-aged listeners have trouble hearing in everyday settings. Curr Biol 22:1417-1422, https://doi.org/10.1016/j.cub.2012.05.025.

Shera CA, Guinan JJ, Oxenham AJ. (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci, https://doi.org/10.1073/pnas.032675099.

Stapells DR, Linden D, Suffield JB, et al. (1984) Human auditory steady state potentials. Ear Hear, https://doi.org/10.1097/00003446-198403000-00009.

Terkildsen K, Osterhammel P, Huis F. (1975) Far field electrocochleography. Frequency specificity of the response. Scand Audiol, https://doi.org/10.3109/01050397509043078.

Varghese LA, Bharadwaj HM, Shinn-Cunningham BG. (2015) Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research, https://doi.org/10.1016/j.brainres.2015.06.038.

Wegner O, Dau T. (2002) Frequency specificity of chirp-evoked auditory brainstem responses. J Acoust Soc Am, https://doi.org/10.1121/1.1433805.

Zilany MSA, Bruce IC, Carney LH. (2014) Updated parameters and expanded simulation options for a model of the auditory periphery. J Acoust Soc Am, https://doi.org/10.1121/1.4837815.

(Received 1 October 2018, Accepted 6 February 2019)
(Available online 1 March 2019)