Hearing Research xxx (2017) 1–15

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators

Parker Tichko [b], Erika Skoe [a, b, c, *]

[a] Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, CT 06269, USA
[b] Department of Psychological Sciences, Developmental Psychology Program, University of Connecticut, Storrs, CT 06269, USA
[c] Connecticut Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, CT 06269, USA

ARTICLE INFO

Article history:
Received 12 September 2016
Received in revised form
12 January 2017
Accepted 22 January 2017
Available online xxx

Keywords:
Frequency-following response
Fine structure
Source modeling
Individual differences

ABSTRACT

The frequency-following response (FFR) is an auditory-evoked response recorded at the scalp that captures the spectrotemporal properties of tonal stimuli. Previous investigations report that the amplitude of the FFR fluctuates as a function of stimulus frequency, a phenomenon thought to reflect multiple neural generators phase-locking to the stimulus with different response latencies. When phase-locked responses are offset by different latencies, constructive and destructive phase interferences emerge in the volume-conducted signals, culminating in an attenuation or amplification of the scalp-recorded response in a frequency-specific manner. Borrowing from the literature on the audiogram and otoacoustic emissions (OAEs), we refer to this frequency-specific waxing and waning of the FFR amplitude as fine structure. While prior work on the human FFR was limited by small sets of stimulus frequencies, here, we provide the first systematic investigation of FFR fine structure using a broad stimulus set (90 + frequencies) that spanned the limits of human pitch perception. Consistent with predictions, the magnitude of the FFR response varied systematically as a function of stimulus frequency between 16.35 and 880 Hz. In our dataset, FFR high points (local maxima) emerged at ~44, 87, 208, and 415 Hz with FFR valleys (local minima) emerging ~62, 110, 311, and 448 Hz. To investigate whether these amplitude fluctuations are the result of multiple neural generators with distinct latencies, we created a theoretical model of the FFR that included six putative generators. Based on the extant literature on the sources of the FFR, our model adopted latencies characteristic of the cochlear microphonic (0 ms), cochlear nucleus (~1.25 ms), superior olive (~3.7 ms), and inferior colliculus (~5 ms). In addition, we included two longer latency putative generators (~13 ms, and ~25 ms) reflective of the characteristic latencies of primary and non-primary auditory cortical structures. Our model revealed that the FFR fine structure observed between 16.35 and 880 Hz can be explained by the phase-interaction patterns created by six generators with relative latencies spaced between 0 and 25 ms. In addition, our model provides confirmatory evidence that both subcortical and cortical structures are activated by low-frequency (<100 Hz) tones, with the cortex being less sensitive to frequencies > 100 Hz. Collectively, these findings highlight (1) that the FFR is a composite response; (2) that the FFR at any given frequency can reflect activity from multiple generators; (3) that the fine-structure pattern between 16.35 and 880 Hz is the collective outcome of short- and long-latency generators; (4) that FFR fine structure is epiphenomenal in that it reflects how volume-conducted electrical potentials originating from different sources with different latencies interact at scalp locations, *not* how these different sources actually interact in the brain; and (5) that as a byproduct of these phase-interaction patterns low-amplitude responses will emerge at some frequencies, even when the underlying generators are fully functioning. We believe these findings call for a re-examination of how FFR amplitude is interpreted in both clinical and experimental contexts.

1. Introduction

The frequency-following response (FFR) is a sustained, phase-locked neural response that reflects the spectrotemporal features

* Corresponding author. University of Connecticut, 850 Bolton Rd, U-1085, Storrs, CT 06269, USA.
E-mail address: erika.skoe@uconn.edu (E. Skoe).

http://dx.doi.org/10.1016/j.heares.2017.01.014
0378-5955/

of acoustic stimuli (Greenberg et al., 1987; Krishnan, 2002; Marsh et al., 1975; Moushegian et al., 1973). Recorded from the scalp using electroencephalography (EEG), the FFR provides an objective, non-invasive measure of the neural encoding of sound in the human auditory system. Typically, the FFR is recorded to synthetic (e.g., sine waves, triangle waves, tune bursts) or more ecologically valid stimuli (e.g., speech and music) (e.g., Aiken and Picton, 2008; Carcagno and Plack, 2011; Galbraith et al., 1995; Hairston et al., 2013; Lee et al., 2009; Russo et al., 2004; Slabu et al., 2012). Work in this manner suggests that the human FFR can be recorded over a wide frequency range, but that it drops off almost entirely for auditory stimuli >1000 Hz (Batra et al., 1986; Kuwada et al., 2002; Moushegian et al., 1973). Recently, a large body of work has emerged relating variability in the FFR to variability in language and musical behaviors (e.g., Anderson et al., 2011; Banai et al., 2009; Bones et al., 2014; Krishnan et al., 2005), as well as linguistic and non-linguistic experiences across the lifespan, such as bilingualism and musical training (e.g., Krizman et al., 2012; Lee et al., 2009; Weiss and Bidelman, 2015; Wong et al., 2007). Consequently, the FFR has become an important tool for studying the developing and aging auditory system, and it is now widely regarded as a putative biomarker of disordered auditory systems (Anderson et al., 2012; Johnson et al., 2008; Ruggles et al., 2012). Despite the growing literature linking the FFR to individual differences in language and musical abilities, relatively little is known about the sources of inter- and intra-subject FFR variability, including how the underlying neural generators of the FFR introduce variability in the far-field response to different stimulus frequencies.

Throughout the auditory system, neurons synchronize their spiking patterns to the temporal features of the driving stimulus by firing at a particular phase of the stimulus (Joris et al., 2004; Langner, 1992; Liu et al., 2006; Middlebrooks, 2008). Thus, for any given stimulus frequency, phase-locking may occur in multiple auditory structures, with low-frequency stimulation resulting in more system-wide activation than higher frequencies, by virtue of the sharp drop-off in the upper limits of phase-locking that occurs at each stage along the afferent pathway. The scalp-recorded FFR is theorized to be an aggregation of phase-locked neural activity from multiple generators within the auditory system. Data obtained from animal models suggests that the human FFR to frequencies in the vocal pitch range (~85–400 Hz) is predominately subcortical in origin. Intra-cranial work with the cat points to multiple peripheral and central subcortical sources of the FFR including cochlear hair cells (i.e., cochlear microphonic), auditory nerve, cochlear nucleus, superior olivary nucleus, and inferior colliculus (Gardi et al., 1979; Smith et al., 1975; Snyder and Schreiner, 1984). Consequently, the human FFR recorded to frequencies in the vocal pitch range has been interpreted within the context of this animal work and treated as a putative measure of subcortical auditory processing (Chandrasekaran and Kraus, 2010). Recent dipole analysis provides support for the FFR to sounds in the vocal pitch range having a subcortical, specifically midbrain, source (Bidelman, 2015). Other studies, however, implicate the involvement of both subcortical and cortical areas in the generation of the FFR for lower frequency signals (Galambos et al., 1981; Kuwada et al., 2002; Mäkelä and Hari, 1987; Ribary et al., 1991). For instance, the well-documented auditory 40-Hz response is theorized to be a consequence of both cortical and subcortical generators (Galambos et al., 1981; Herdman et al., 2002; Kuwada et al., 2002; Mäkelä and Hari, 1987; Ribary et al., 1991). Additionally, in a study with the unanaesthetized rabbit, the amplitude-modulation-following response (AMFR), a type of FFR that tracks the modulation frequency of amplitude-modulated tones, Kuwada et al. (2002) found the scalp-recorded response to be a composite of activity from multiple regions of the auditory cortex (primary and non-primary regions) and multiple subcortical auditory nuclei. Contributions from auditory cortex occurred for stimuli that featured low modulation frequencies (<80 Hz), while subcortical generators were evident across both low and high frequencies, but were dominant for stimuli with high modulation frequencies (>150 Hz) (Kuwada et al., 2002). For frequencies where multiple generators are responsive, the literature suggests that each generator has a characteristic latency that increases at higher levels of the afferent auditory system (Batra et al., 1986; Gardi et al., 1979; Herdman et al., 2002; Kuwada et al., 2002). Moreover, recent neuro-imaging work corroborates the contribution of cortical activity in the human FFR to frequencies on the lower end of the human vocal pitch range. Using magnetoencephalography (MEG), Coffey et al. (2016) found activity emerging asymmetrically in auditory cortex (right hemisphere) while listeners passively heard a harmonically-complex speech token /da/, with a fundamental frequency (F0) of 100 Hz. Consistent with Kuwada et al. (2002)'s measurements, this cortical source followed the low-frequency spectrum of the stimulus, namely the F0, but not higher frequency speech harmonics, where the dominant generators were found to be subcortical in origin (Coffey et al., 2016).

Collectively, these studies highlight the multi-generator nature of the FFR. However, localizing and isolating the specific FFR generators in the human auditory system is difficult with current EEG technology because the voltage recorded from any given electrode at the scalp captures summed neural activity from many sources throughout the nervous system (Nunez and Srinivasan, 2006). In the case of the FFR, source localization is further complicated by the fact that multiple generators are producing phase-locked signals of the same frequency that superimpose at the scalp. Depending on the phase of the phase-locked responses, the signals will either constructively or destructively interfere in the composite response measured at the scalp. Building from prior work (Dolphin and Mountain, 1992; Gardi et al., 1979; Kuwada et al., 2002), we posit that this interference pattern manifests as systematic, frequency-dependent modulations in the FFR amplitude, characterized by high points and valleys at specific frequencies in the FFR-gram (our term for the graphical representation of FFR amplitude as a function of frequency). FFR valleys—local minima in the FFR-gram—are theorized to reflect the aggregated activity of multiple sources that are phase-locked to the same frequency but their phases are not aligned, resulting in an attenuated far-field response amplitude. In contrast, FFR high points—local maxima in the frequency function—are theorized to reflect the summation of in-phase, phase-locked responses from multiple generators. In support of this, several studies evidence a waxing and waning in the FFR amplitude as a function of stimulus frequency that, until now, has gone understudied in humans (Batra et al., 1986; Dolphin and Mountain, 1992; Gardi et al., 1979; Hoormann et al., 1992; Kuwada et al., 2002; Skoe and Kraus, 2012). We refer to the frequency-dependent waxing and waning of the FFR amplitude as fine structure, borrowing the terminology from the phenomenon of fine structure in the audiogram and otoacoustic emissions (OAEs) (Dhar et al., 2002; Elliott, 1958; Gaskill and Brown, 1990). Similar to FFR fine structure, fine structure in the audiogram and OAE is theorized to be the outcome of multiple source interferences (Dhar et al., 2002; Shera and Guinan, 1999). Importantly, FFR fine structure appears even when stimulus intensity is held constant (Batra et al., 1986; Hoormann et al., 1992), suggesting that fine structure variations are not simply due to exogenous factors.

Previous studies of the FFR have generally used small, limited sets of stimulus frequencies which afford a coarse-grain sampling of the response (e.g., 12 frequencies, Hoormann et al. (1992); eight frequencies, Skoe & Kraus (2012); six frequencies, Marmel et al. (2013)), or they have used a larger set of frequencies (~40

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

frequencies in Batra et al. (1986)) but without detailing the exact frequencies used. With the goal of providing a more systematic account of FFR fine structure than is currently available in the literature, we conducted a fine-grained sampling of the FFR using a set of 90 + tonal stimuli that spanned the frequency limits of human pitch perception. Secondly, we investigated the intra- and inter-subject variability of the response, the long-term stability of FFR fine structure across a longer timescale than has previously been reported, and the effect of electrode site on FFR fine structure. Finally, we developed a multi-source model of the FFR based on Gardi et al. (1979) to further investigate the multiple-generator hypothesis of FFR fine structure.

2. Materials and methods

2.1. Participants

Three healthy adult volunteers were recruited to participate in the study. Written informed consent was obtained from the volunteers, and all procedures were approved by the University of Connecticut's Institutional Review Board for Human Subject Research. Participants exhibited clinically normal auditory function with normal otoscopy, normal bilateral air conduction thresholds <20 dB HL for octaves from 125 to 8000 Hz, and auditory brainstem response (ABR) wave latencies within normal limits for a 70-dB nHL rarefaction click presented at 31.25 Hz (using norms reported in Skoe et al., 2015). Subject 1, a female aged 25 at the first test date, was a bilingual English-Spanish speaker with 12 years of formal musical training. Subject 2, a male aged 28, was a monolingual English speaker with 16 years of musical training. Subject 3, a female aged 20, was a monolingual English speaker with no formal musical training. All subjects participated in multiple FFR recording sessions, each spanning between 2 and 3 h in length, over the course of multiple days. Total recording time for the completion of the protocol varied between 13 and 18 h. To investigate the long-term stability of FFR fine structure, Subject 1 completed the protocol twice, with a period of 10 months separating the two protocols – a test-retest period 7 months longer than that used by Hoormann et al. (1992). In a follow-up experiment, Subject 2 and Subject 3 repeated the protocol for 12 stimulus frequencies, those used by Hoormann et al. (1992), after a 6-month period to investigate the effects of electrode reference site on FFR fine structure for frequencies above 100 Hz.

2.2. Stimuli

The FFR stimuli consisted of 200-ms triangle-wave stimuli presented at a rate of 2.5 per second to the right ear. Subject 2 and Subject 3 were tested on 91 frequencies between 16.35 and 880 Hz. Subject 1 was tested on this same set of 91 frequencies, plus 38 additional frequencies above 880 Hz. Subject 2 was also tested on three high-frequency stimuli (2637.02, 2793.83, 2959.96 Hz) for the purposes of ruling out stimulus artifact in our recordings; however, these recordings were not subjected to further analysis.

Triangle-wave stimuli were created for a total of 129 frequencies across a frequency range of 16.35–7902.13 Hz: 109 of the frequencies abided by an equal-tempered tuning system; 12 of the frequencies (128, 192, 256, 320, 384, 448, 512, 576, 640, 704, 768, and 832 Hz) were replicated from Hoormann et al. (1992) to facilitate comparisons with this previous investigation; and eight of the frequencies (100.92, 113.27, 127.14, 134.7, 285.42, 302.4, 302.4, and 320.38 Hz) neighbored valleys in the FFR fine structure (Supplemental Table 1). These eight frequencies were tested *post-hoc* after identifying the valleys in the FFR-gram using the equal-tempered frequencies. The inclusion of these eight frequencies allowed us to investigate whether the response descended gradually into areas marked by valleys, or whether valleys emerged as abrupt changes in FFR fine structure.

Employing an equal-tempered system enabled us to weight the sampling of the FFR over the frequency range for which the response was predicted to be most robust (<500 Hz) (Moushegian et al., 1973). In an equal-tempered system, adjacent pitches are separated by the same interval, resulting in the absolute difference between adjacent frequencies to be smaller at the lower end of the stimulus frequency range and larger at the higher end of the stimulus frequency range. This resulted in more FFRs being recorded to lower frequencies in our stimulus frequency range, while still enabling us to record auditory-evoked responses to higher frequencies, including those above 1000 Hz where the FFR is small, if not absent. We included frequencies above 1000 Hz for a number of related reasons: first, to study how the amplitude of the FFR decays at frequencies thought to be above the FFR limit, and, second, to examine the presence of non-neural contributions to our recordings (e.g., stimulus artifact) (see below).

Triangle waves are complex sounds characterized by a triangle-shaped waveform that contain only the F0 and odd harmonics. The current analysis considers only the phase-locked response to the F0. As can be seen in Fig. 1 and Supplemental Fig. 1, the response to the harmonics are minimal compared to the response to the F0. We chose triangle waves because harmonically-rich stimuli have been shown to elicit more robust FFRs to the F0 than sinewave stimuli (Jeng et al., 2011) and, relative to continuous tones, stimuli that gate on and off also elicit more robust FFRs (Batra et al., 1986). Here, we digitally synthesized 200-ms triangle waves (48 kHz/16-bit resolution) with a 5-ms Hanning onset-offset window using a custom MATLAB script. Stimulus intensity was calibrated to a presentation level of 80 dB LAeq for frequencies between 100.92 and 5919.91 Hz using a Bruel & Kjaer Type 2250 Sound Level Meter with a 2-cc coupler attached to the insert earphone. For frequencies outside this range, stimulus intensity was lowered (see Supplemental Table 1 for specifics) to ensure a comfortable listening level for subjects at the lower and higher limits of the stimulus frequency range.

Stimulus delivery was controlled in MATLAB using Psychophysics Toolbox extensions. The stimulus was outputted by a USB-powered external sound card (M-Audio M-Track) at 44.1 kHz/16-bit resolution and presented to the right ear using Mu-metal electromagnetically (EM) shielded tubephones (ER-3, with custom EM shielding provided by Intelligent Hearing Systems). The audio stimulus was also routed through a StimTrak box (Brain Products GmbH) into an auxiliary channel of the EEG amplifier (actiCHamp amplifier, Brain Products GmbH). This auxiliary recording provided accurate stimulus timing information when averaging the responses offline.

2.3. FFR recording protocol

A three-electrode vertical, ipsilateral array was used, differentially measuring from Cz to the right ear lobe (A2), with the ground electrode placed on the forehead. Prior to placing the Ag-AgCl electrodes (Multitrodes, Brain Products GmbH) on the scalp, the electrode sites were mildly exfoliated to produce impedances under 5 kOhms. The electrodes were plugged into a bipolar pre-amplifier that provided a gain factor of 50 (EP-PreAmp module, Brain Products GmbH). The bipolar pre-amplifier interfaced with the actiCHamp amplifier (Brain Products GmbH) through a splitter box. Recordings were made with a 25-kHz sampling rate and 0.05 mV/bit resolution (−409 to + 409 mV range) in the Recorder software (Brain Products GmbH), with no on-line filters applied.

In a follow-up study, we investigated the effects of electrode

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

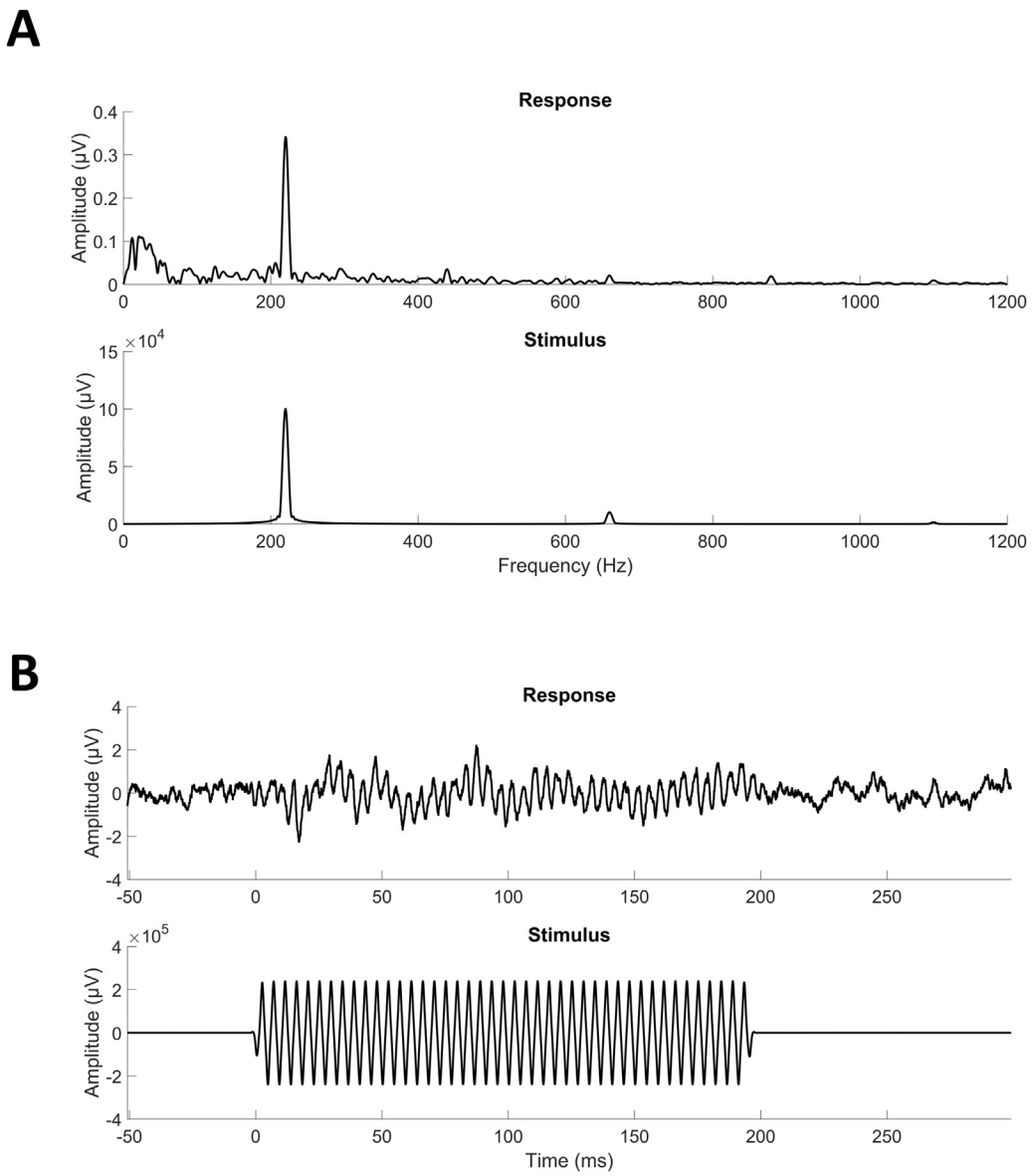

Fig. 1. Representative FFRs in the frequency and time domains. (a) Spectral analysis of the FFR to a single triangle-wave stimulus. Representative FFR (Subject 1) recorded to a 220-Hz triangle wave. Plotted below is a spectral analysis of the 220-Hz triangle wave used to elicit the neural response. The stimulus was routed into the EEG amplifier and then averaged in the same fashion as the FFR. A peak in the FFR at 220 Hz reflects the fundamental frequency (F0) of the acoustic stimulus. Note that because of the broadband filter that was applied to the recordings (15–10,000 Hz) the spectrum includes energy below the stimulus frequency. **(b) FFR and triangle-wave stimulus represented in the time domain.** Averaged waveform of a representative FFR (Subject 1) visualized in the time domain in the above plot. Below, the 220-Hz triangle-wave stimulus used to elicit the FFR. The sustained portion of the FFR beginning at the stimulus onset reflects the periodic, frequency information of the acoustic stimulus.

reference site by simultaneously recording FFRs using both ipsilateral and contralateral references. For this, we performed a two-channel recording that differentially measured from Cz-A1 and Cz-A2, with the ground electrode placed on the forehead. All other recording parameters were identical to the original ipsilateral recordings described above. Pre-processing of the data was conducted in the Brain Products Analyzer 2 software to separate the ipsilateral and contralateral recordings. Additional analysis was conducted in MATLAB using the methods described below for the ipsilateral recordings.

All recordings were made passively with the participant sitting comfortably in an electromagnetically shielded sound booth. To facilitate a relaxed state that would reduce myogenic contamination, a self-selected, English-captioned muted video was projected onto the wall of the booth from a ceiling mounted LCD projector placed outside the booth window. Participants were instructed to sit and watch the video quietly. The testing conditions, especially when administered as part of a repeated-measures design, can induce drowsiness. To minimize muscle artifacts that arise when a test participant tries to fight off sleep, participants were allowed to nap during the test sessions, and in such cases the experimenter did not awake the participant from sleep until the end of the session.

2.4. Averaging

In MATLAB, the neural responses were offline filtered from 15 to 10,000 Hz using a third-order, one-dimensional Butterworth filter. Notch filters were applied at 60 Hz, 1 kHz, 2 kHz, and 4 kHz using a second-order IIR notch filter design with a Q-factor of 35 to remove a small amount of electrical noise that was present at these

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

recording frequencies (note that these frequencies do not coincide with stimulus frequencies). The recording was then epoched using a 350-ms window that spanned 50 ms prior to stimulus onset and extended 300 ms after stimulus onset. The timing of each stimulus onset, which was derived from the auxiliary EEG channel, corrected for the acoustic transmission delay introduced by the tube connecting the transducer and foam ear insert. After excluding epochs in which the amplitude exceeded ±60 μV, 500 trials for each stimulus were averaged. For each stimulus, the spectral profile of the FFR was derived by computing a discrete fast Fourier transform (FFT). The FFT was performed on a 250-ms section of the response (0–250 ms), after applying a Hanning window. The output was scaled to microvolts by taking the absolute value and then multiplying the result by 2 divided by the signal length. Zero-padding to 1 s was applied as part of the FFT to increase the sample rate of the spectral estimate. An example FFR is plotted in Fig. 1 in both the time and frequency domain. For each stimulus frequency, the spectral magnitude at the F0 was then extracted from the spectrum of the FFR. This produced a one-dimensional array of spectral amplitude values for each participant that we refer to as the FFR-gram.

2.5. Data analysis

To determine whether a response was significantly above the noise floor, we adopted a method similar to the one described in Hoormann et al. (1992). Here, the noise floor was calculated by computing the average spectral amplitude across the 10–10,000 Hz range, after excluding spectral amplitudes at the F0 and ±10 Hz around the F0 (e.g., for a F0 of 100 Hz, the frequency range of 90–110 Hz would be omitted from the noise floor calculation), and then calculating the standard deviation of the noise floor. Responses occurring at the F0 of the stimulus were considered reliable if they were greater than the average noise floor plus three standard deviations (Fig. 2A) (Hoormann et al., 1992). Finally, to test for the long-term stability of the FFR structure, the responses to the frequencies that were run on Subject 1 at both sessions (total of 69 frequencies) were compared using the Hoeffding Dependence Coefficient (*D),* a non-parametric measure of association. The two sessions were separated by a 10-month period.

2.6. Verifying the pre-neural and neural auditory pathway origin of the recordings

When recording FFRs, a major methodological concern is contamination from stimulus artifacts (Akhoun et al., 2008; Campbell et al., 2012; Skoe and Kraus, 2010). As stimulus artifacts occur during the same epoch as the FFR, and they align in frequency, they are easily mistook for neural responses. To protect against stimulus artifacts, we used Mu-metal-shielded transducers that were separated from the scalp- and earlobe-placed EEG electrodes by 10 cm of tubing that routed the sound output into the ear canal through a foam ear insert (ER3-14). To assess whether our experimental set-up reduced stimulus artifacts, we ran the FFR recording protocol with one participant (Subject 2) twice, once with the participant listening to the stimulus with the ear insert properly seated in the ear canal, and a second time with the opening of the tubephone blocked with adhesive tape to prevent the sound from being audible. Subject 2 completed this stimulus-artifact test protocol for six frequencies (220, 233.08, 246.94, 2637.02, 2793.83, 2959.96 Hz), the three higher of which are above the limits of the FFR. These recordings were then spectrally analyzed using the same procedure described above and visualized (Supplemental Fig. 1). As shown in Supplemental Fig. 1A, when the participant could hear the 220-Hz stimulus, a large spectral peak at 220 Hz (F0) emerges that is sustained for the 200-ms duration of the stimulus. However, when the tubephone was blocked, and the sound became inaudible, the band of energy at 220 Hz disappears, leaving only noise. For the higher frequency stimuli, we did not observe sustained energy at the stimulus frequency in either the audible or inaudible conditions for Subject 2 (illustrated for 2637.02 Hz in Supplemental Fig. 1B), suggesting that these frequencies are outside the range where robust phase-locking can be observed in the FFR for this participant (note the scale is different between panel A and B in Supplemental Fig. 1). There was also no evidence of stimulus artifact for the remaining frequencies tested in the audible and inaudible conditions: 233.08, 246.94, 2793.83 and 2959.96 Hz. If stimulus artifacts were contaminating our recordings, they should be clearly evident at frequencies that are above and below the FFR frequency limits and, in addition, they should not decay as a function of stimulus frequency because they are not subject to the physiological processes that produce the drop-off in FFR amplitude as a function of stimulus frequency. Collectively, this leads us to conclude that for Subject 1 that the small subset of frequencies above 1000 Hz that produce FFRs that are low in amplitude but fall above the noise floor are either pre-neural (cochlear microphonic) or they are spuriously large (Fig. 2B).

2.7. Theoretical FFRs

To probe the influence of multiple generators on FFR fine structure, and, in particular, to test the hypothesis that FFR fine structure emerges in the scalp-recorded responses from the summation of out- and in-phase volume-conducted signals, we created a theoretical model of the FFR to compare against the FFR fine structure empirically obtained for each subject. Here, we adapted the basic procedure of Gardi et al. (1979), which utilized a three-generator model consisting of the cochlea (cochlear microphonic, CM), cochlear nucleus (CN), and superior olivary complex (SOC) to model the cat FFR fine structure for stimulus frequencies between 0.3 and 3 kHz. In the Gardi et al. (1979) FFR model, the phase-locked output of each generator was modeled as a sine wave at the stimulus frequency, with each generator having a characteristic delay that was derived empirically from the first spike latencies measured at each structure in the cat. In the Gardi et al. model, three sine-waves with delays of 0 (CM), 1.25 (CN) and 3.7 (SOC) ms were summed. A spectral analysis on the summed waveform was performed by extracting the amplitude of the theoretical response at the stimulus frequency. This process was performed across the 0.3–3 kHz range to derive the theoretical FFR-gram. In the Gardi et al. work, the theoretically derived FFR-grams showed the same waxing-and-waning pattern as the empirical data, with a high degree of overlap in terms of which frequencies corresponded to high points vs. valleys.

Building on the basic Gardi et al. (1979) approach, we created theoretical FFRs in MATLAB for 91 test frequencies between 16.35 and 880 Hz, from which theoretical FFR-grams were derived. This modeling was implemented using custom routines developed by the authors. Note that our frequency range extends below those used in the Gardi et al. study. To account for the fine-structure patterns that emerge at lower frequencies, we implemented a six generator model that included the CM, CN, SOC, inferior colliculus (IC) and two cortical sources. We refer to these six theoretical generators as G1, G2, G3 … (etc.), with higher numbers indicating longer latency generators. In our base (i.e., starting) model, the characteristic delays were set to 0, 1.25, and 3.7 ms for the CM, CN, and SOC (respectively, following Gardi et al., 1979) and 5, 12, and 26 ms (respectively) for IC and the two putative cortical sources. The IC and the auditory cortex (AC) were added to the model after considering the diverse evidence that these structures contribute to

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

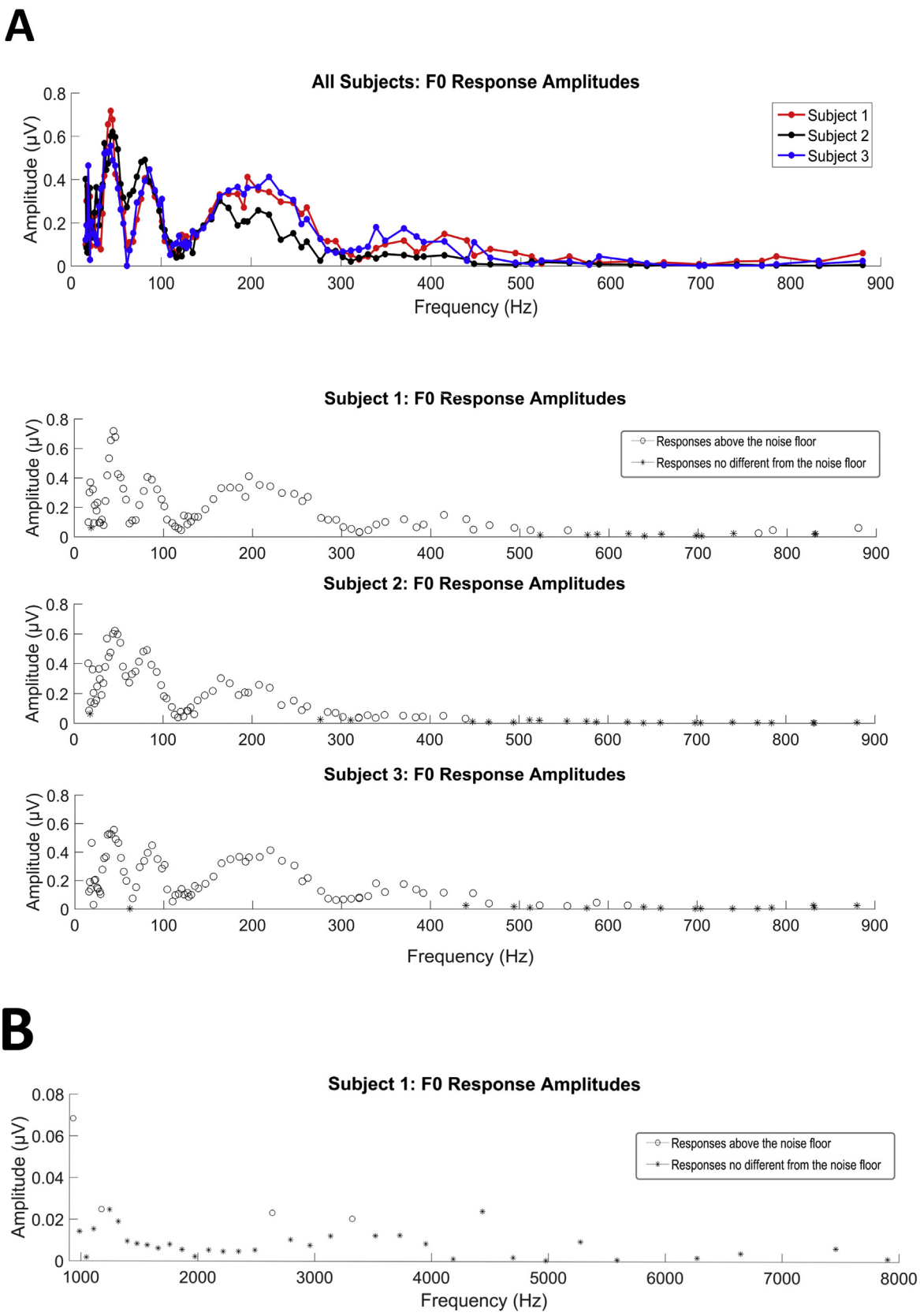

Fig. 2. (A). FFR-grams for the fundamental frequency (F0) across the 16.35–880 Hz stimulus range: Subject 1 (at time 2, red), Subject 2 (black), and Subject 3 (blue). All subjects exhibited a waxing and waning in the FFR amplitude as a function of frequency. Additionally, each subject had a characteristic ripple-like structure to their FFR-gram. To identify response amplitudes in the FFR fine structure, the SNR of the response to each frequency was computed (see methods). Frequencies with an SNR less than 2 (marked with *) suggest that response is within the noise floor. (B). Extended FFR-Gram for Subject 1 at Time 1 (932–7902 Hz). To assess how the amplitude of the FFR decays at higher frequencies, Subject 1, at time 1, was tested on a number of frequencies between 932 and 7902 Hz. The response drastically drops in amplitude for frequencies >932 Hz (Note the difference in y-

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

the far-field FFR to lower frequencies (Coffey et al., 2016; Kuwada et al., 2002; Smith et al., 1975). The decision to include two cortical sources in the model emerged from Kuwada et al. (2002) who concluded from their local- and surface-recordings in rabbit and human models that multiple cortical structures (both primary and non-primary) contribute to low-frequency FFRs, with latencies spanning from ~12 to 30 ms. In Kuwada et al.'s (2002) analysis of the human FFR-gram, two longer latency putative generators emerged, one at 12 and the other at 26 ms. Based on this finding, we incorporated two cortical generators with characteristic latencies of 12 and 26 ms in our base model. Consistent with this, direct recordings in humans suggest that the initial (i.e. earliest) excitation of the AC to click stimuli occurs roughly 12–15 ms after stimulus onset (Liégeois-Chauvel et al., 1994). This time frame is also consistent with the delays observed between the CN and primary AC in the recent study by Coffey et al., 2016. In our base FFR model, we set the characteristic delay of the IC-component (G4) to be 5 ms, following data suggesting that the first spike latency of the human IC is in the 5–6 ms range (Moller and Jannetta, 1983).

In addition to expanding the model to six potential generators, our model advanced the Gardi et al. (1979) model in multiple ways. One way that we advanced the model was by manipulating the amplitude of the sine-wave outputs for each generator. To account for the fact that stimulus intensity was not equated across the frequency spectrum, a correction factor was added to the model that involved scaling the amplitude of each sine wave according to the presentation intensity at each frequency (scaling factor = intensity/80, with 80 dB LAeq being the maximum stimulus intensity; see Supplemental Table 1). In addition, we applied a high-pass filter to all test frequencies to mimic the digital filters applied to the empirical data during data processing (15–10,000 Hz). The relative amplitude of each of the six generators was also scaled to account for the fact that generator sites more proximal to the recording electrodes are expected to dominate the recording when multiple generators are involved (Kuwada et al., 2002). The scaling factors were set to be 1, 1, 1, 2, 3, and 4, for G1-G6 based on pilot testing, respectively. In addition, we applied a low-pass filter (3rd order Butterworth filter) to all test frequencies to account for the fact that the skull, due to its capacitive properties, acts as a low-pass filter for auditory-evoked potentials at scalp locations (Bidelman, 2015; Nunez and Srinivasan, 2006) (Supplemental Fig. 2). The low-pass filter was applied to each theoretical generator prior to summing their sine-wave outputs. We selected a 200-Hz (Butterworth) low-pass filter (Supplemental Fig. 2) and the sine-wave amplitude scaling settings reported above (e.g., 1, 1, 1, 2, 3, and 4 for G1-G6) after testing out different combinations and observing that these settings visually captured the amplitude decay at the higher stimulus frequencies. Finally, our model accounted for the fact that the two longer-latency generators (not in the Gardi et al. model) may have a drop-out frequency, i.e., a frequency beyond which they no longer contribute to the FFR due to the upper limits of their phase-locking properties (Kuwada et al., 2002). We selected 200 Hz as the drop-out frequency for G5 and 100 Hz as the drop-out frequency for G6. When they are included in the model above those frequencies, the model fit drops markedly (Supplemental Figs. 4–6). In the models presented in Fig. 5 and Supplemental Figs. 4–6, the amplitude of G5 was set to drop linearly from 100% to 0% between 100 and 200 Hz and the amplitude of G6 was set to drop linearly from 100% to 0% between 50 and 100 Hz. This down-ramping of the amplitude for G5 and G6 was intended to model the roll-off in phase-locking that is observed as the generator approaches its upper frequency limit of phase-locking. When the stimulus frequency exceeded the drop-out frequency, the sine wave output of that generator was removed from the bank of generators in the model by setting its amplitude to zero.

The Hoeffding Dependence Coefficient (D) was used to calculate the goodness of fit between the empirical FFR-gram and the theoretical FFR-gram over the 16.35–880 Hz range. As a second way of evaluating the fit of the model, we determined how faithfully the theoretical model captured the valleys in the empirical data. Valleys were identified based on visual inspection of the empirical FFR-gram. For the theoretical FFR-gram, a local minimum within ±10 Hz of an empirical valley was counted as a "hit"; minima outside this range were treated as "misses". We focused on valleys as they are the most direct evidence of an interaction pattern created by multiple signal generators, each responding to the same frequency.

For each subject, we used G1-0, G2-1.25, G3-3.7, G4-5, G5-12, and G6-26 ms as the base model. Then, to account for anatomical and physiological variations among the participants, we adjusted the relative latencies of each theoretical generator of the base model until we reached a 100% hit rate for the valleys in the empirical data for each individual's data. This involved incrementally increasing and decreasing the relative latency of each generator in steps of 0.1 ms. To limit the number of free parameters in the model, and to increase the biological validity of the model, this process was constrained by restricting the possible range of latency values that each putative generator could assume. The ranges were constrained to 0 ms for G1, 1–2.5 ms for G2, 2.5–5 ms for G3, 5–12 ms for G4, 12–20 ms for G5, and 20–33 ms for G6, with the further limitation that the generators needed to be separated by at least 1 ms. We refer to the combination of relative latencies that achieved a 100% hit rate as the subject's "theoretical *generator profile*" (TGP). From there, we re-calculated the goodness of fit (D) as each generator was selectively removed from the bank of generators in their TGP (Supplemental Figs. 4–6). Finally, as a way to validate that each generator is making a significant contribution to the fine-structure footprint (i.e., the location of the peaks and valleys), we determined which combination of generators within the TGP could mathematically give rise to each major peak and valley in the empirical data. If a theoretical generator was determined to be necessary for producing a specific peak or valley, this suggests that it is making a significant (i.e., necessary contribution) to the model.

Although our model adopted biologically-plausible latencies, it is nevertheless important to note that in this type of model the phase-interaction patterns are determined by the relative latencies among the pool of generators and not their absolute latencies. For example, the interaction patterns created by generators with relative latencies of 0, 1.25, and 3.6 ms is the same as the interaction pattern created by 1, 2.25, and 4.6 ms. In both cases, valleys occur at 175 and 415 Hz. This has important implications for mapping the theoretical generators to distinct anatomical substrates (see discussion).

3. Results

As predicted, in all subjects, frequency-dependent fluctuations are evident in the FFR amplitudes recorded across a broad range of frequencies (Fig. 2A, Supplemental Table 1). That is, the scalp-recorded FFR is stronger at some frequencies and weaker at other,

axis scaling between 2A and 2B). While most of the responses >932 Hz appear to be within in the noise floor, several frequencies were found to be above the noise floor. These points in the FFR-gram likely reflect pre-neural sources, such as the cochlear microphonic. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

nearby frequencies. Moreover, this relationship appears highly non-linear and non-monotonic. To investigate the non-linear, waxing-and-waning relationship between stimulus frequency and response amplitudes, the extracted response amplitudes from all three subjects were fit with two statistical models: 1) a linear regression model and 2) a cubic spline (linear + non-linear components) model to probe whether a non-linear model fit the FFR-grams better than a linear fit. The cubic spline model was fit using the using the ordinary least squares method and 14 knots at the following frequencies: 16.72, 28, 35, 46, 58, 78, 98.60, 117, 129.20, 194.40, 292.20, 384, 448, and 845.44 Hz. To assess the fit of the cubic spline, an analysis of variance (ANOVA) was conducted. The frequency factor was found to be significant, $F(13, 259) = 107.89$, $p < 0.0001$. Moreover, the non-linear component of the cubic spline model was also significant, $F(12, 259) = 64.41$, $p < 0.0001$, suggesting that the non-linear component improved the fit of the spline model relative to its linear component. To determine the amount of additional variance explained by the spline model, the partial R-squared was calculated between the linear regression (i.e., regressing response amplitudes on stimulus frequency) and the spline model. The spline model was found to explain an additional 74.9% of the variance in response amplitudes, partial R-squared = 0.749. Taken together, these analyses suggest that the relationship between stimulus frequency and response amplitude is indeed a highly non-linear phenomenon.

In addition to being non-linear, the fine-structure pattern is similar across subjects, with common valleys occurring at ~62, 110, 311, and 448 Hz and common high points occurring at ~44, 87, 208, and 415 Hz. At these valleys, the response amplitude to the F0 is reduced relative to responses at neighboring frequencies. In some cases, the valleys span a range of frequencies, whereas in other cases, the valleys appear focal. For instance, in Subject 1's FFR-gram, a focal valley appears at 198 Hz, while a range of valleys appears between ~300 and 320 Hz. Moreover, the eight neighboring frequencies tested *post-hoc* suggest that the response amplitude gradually dips into null-point areas instead of abruptly changing in intensity over these regions. Finally, in addition to the waxing-and-waning fine structure, each subject had an idiosyncratic ripple in their FFR-gram, as seen by the small undulations in the response amplitude between the major peaks and valleys (Dhar et al., 2002; Elliott, 1958).

In all subjects, the FFR is clearly present (i.e., above the noise floor) up to ~500 Hz, as shown previously (Batra et al., 1986; Hoormann et al., 1992). In Subject 3, an FFR is observed even above 600 Hz and, for Subject 1, who was tested on frequencies above 880 Hz, the FFR is consistently present up to 932 Hz. Interestingly, Subject 1 also exhibited small spectral peaks >932 Hz (Fig. 2B). Given the relatively low intensity of the response for Subject 1 at these high frequencies and our ruling out of stimulus artifacts (Supplemental Fig. 1), we believe these points in the FFR-gram are primarily pre-neural, reflecting the cochlear microphonic (Euler and Kiessling, 1983; Sohmer and Pratt, 1976) or potentially spurious. Moreover, we found that the response amplitudes were greatest for frequencies <90 Hz, possibly reflecting the participation of more generators, including cortical activity, in the scalp-recorded response.

3.1. Long-term stability of the fine-structure pattern

The long-term stability of the fine-structure pattern was examined in Subject 1, who completed two protocols 10 months apart: the first with stimulus frequencies ranging from 16.35 to 7902.13 Hz, and the second with stimulus frequencies ranging from 16.35 to 880 Hz. The frequency range was abbreviated at re-test after confirming that the FFR does not consistently emerge for frequencies above 932 Hz in this participant and demonstrating that the fine structure is also minimal above ~550 Hz. This yielded a total of 69 common frequencies that were run across both sessions. The valleys and high points that constitute the FFR fine structure are stable even after 10 months (Fig. 3), consistent with the high test-retest stability of the FFR observed in previous work (Hornickel et al., 2012; Song et al., 2011a). Moreover, the FFR-grams were similar across sessions, Hoeffding Dependence Coefficient, $D = 0.417$.

3.2. Comparisons to the Hoormann et al. (1992) data

In addition to recording FFRs to an equal-tempered scale, recordings were made to the same set of frequencies used by Hoormann et al. (1992). In their study, Hoormann et al. (1992) report an FFR-gram for 12 frequencies between 128 and 832 Hz that reflected the average response amplitudes from 20 adult participants. In their data, the grand-average FFR amplitude ramps up with increasing frequency up to 320–384 Hz, after which the amplitude declines. Based on this sample of 12 frequencies, this suggests that the 320–384 Hz range constitutes a high point in the FFR-gram in the Hoormann et al. study. This contrasts with our data (Figs. 2 and 4) in which the FFR-gram peaks at 192 Hz with a valley ~ 320 Hz.

What might account for the different locations of the high points and valleys in the two datasets? Differences may be attributable to variations in stimulus duration (31.25 ms in their study vs. 200 ms in the present study), recording protocol, subject sample, as well the general quality of the recordings. Another potential factor is the electrode array. Hoormann et al. (1992) utilized a contralateral reference which, compared to an ipsilateral reference, can minimize the contribution of the ipsilateral cochlea and auditory nerve in the scalp-recorded response (Bidelman, 2015; Chimento and Schreiner, 1990; Norrix and Glattke, 1996; Prasher and Gibson, 1980). This raises the possibility that the FFR fine structure characteristics we reported above for an ipsilateral reference may reflect a different combination and/or different proportion of generators than what was observed by Hoormann et al. (1992). To examine the influence of the electrode array on the FFR fine structure, we re-recorded FFRs to the 12 stimuli in the Hoormann et al., study, this time adopting a technique that allows for the simultaneous recording of the FFR with an ipsilateral and contralateral reference. We found that the FFR fine structure changed little over this frequency range when using an ipsilateral vs. contralateral reference site (Fig. 4). The potential implications of this outcome are addressed further in the discussion.

3.3. Modeling the FFR-grams for subjects 1, 2, and 3

The fine-structure profile between 16.35 and 880 Hz was well captured by a six-generator model. For Subjects 1 and 2, we achieved a 100% hit rate for the FFR-gram valleys when the relative latencies of the six generators were set to 0, 2.6, 4.2, 7.8, 13.6 and 23.8 ms (Fig. 5). For Subject 3, the hit rate was 100% when the relative latencies were set to 0, 2.8, 3.8, 7.9, 13.2, and 25.1 ms (Fig. 5). Refer to Supplemental Figs. 4–6 to observe how the theoretical FFR-gram and corresponding D change as each theoretical generator is selectively removed from the model. Note when G1 is removed that the small “divot” ~200 Hz that emerges in the empirical-FFR is not accounted for and disappears in the theoretical FFR.

In Fig. 6, we graphically illustrate how the six theoretical generators can mathematically combine to create different aspects of the complex interaction patterns (i.e., valleys and high points) that are characteristics of the FFR fine structure footprint.

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

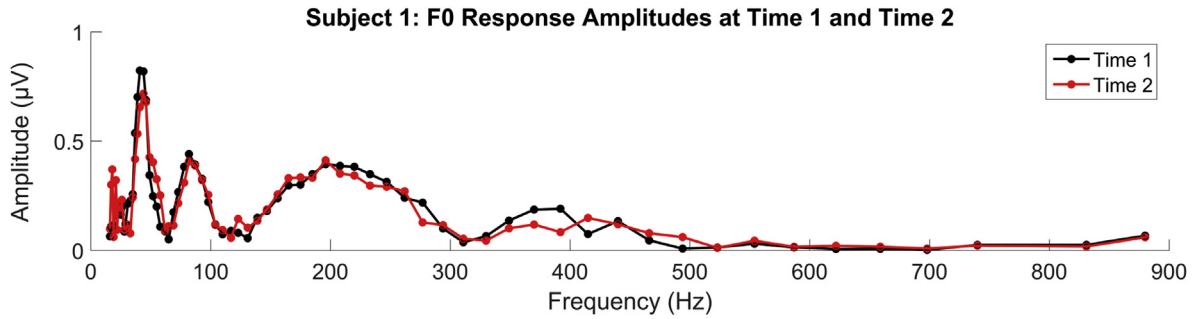

Fig. 3. FFR-grams for subject 1 at time 1 (black) and time 2 (red) (10 months apart). To explore the repeatability of the FFR fine structure over a longer time scale, one subject (Subject 1) was run through the protocol twice for 69 stimulus frequencies between 16.35 and 880 Hz. The high points and valleys within the FFR-gram for frequencies 20–500 Hz are stable after 10 months. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

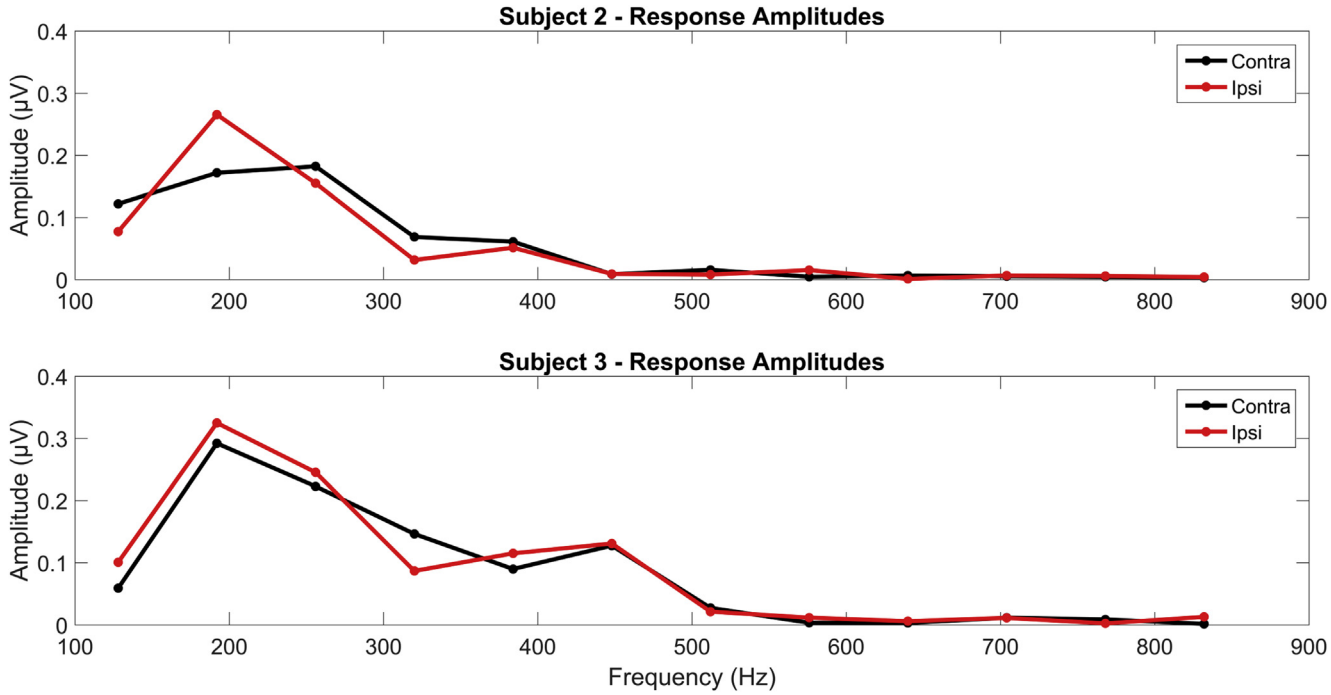

Fig. 4. FFR-grams derived from an ipsilateral earlobe reference (red) and contralateral earlobe reference (black). To investigate the effect of reference electrode site on the FFR, Subject 2 and 3 completed simultaneous ipsilateral-contralateral recordings to the 12 stimulus frequencies used by (Hoormann et al., 1992). Unlike Hoormann et al. (1992), who used a contralateral reference, we did not find a high point in the FFR-gram at ~320 Hz in our contralateral reference condition. Moreover, the FFR-grams do not seem to change in their global structure across the two reference sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

To our knowledge, this is the first comprehensive investigation of fine structure in the human frequency-following response (FFR). Our methods, which afforded a more fine-grain frequency sampling of the FFR than earlier approaches, revealed a systematic, frequency-dependent modulation of the response amplitudes of the human FFR. This fine-structure pattern followed a common, global morphology in all three participants, with clearly defined peaks (e.g., local maxima) and valleys (e.g., local minima) in the FFR amplitudes to the F0. The valleys likely correspond to frequencies where the volume-conducted phase-locked signals arising from different generators are not phase-aligned. However, this phase cancellation does not appear to be complete, given that the FFR stayed above the noise floor even for frequencies where it fell within a fine structure valley. Thus, valleys are not necessarily true null points in the phase-interaction pattern that emerges at the scalp when various volume-conducted phase-locked responses are pooled. We also found that high points and valleys fell within circumscribed frequency ranges meaning that their locations were consistent across participants, stable on re-test, and predicted by our six-generator model of the FFR. If the FFR reflected only a single underlying source, its amplitude would, in theory, be flat across the stimulus frequency spectrum and then decay as the frequency of stimulation approached the upper frequency limits of the FFR. By demonstrating that the amplitude of the FFR waxes and wanes in a systematic, frequency-dependent fashion that can be modeled, this, in conjunction with animal work, provides compelling evidence that the scalp-recorded FFR, and the fine structure within the FFR, is a byproduct of phase-locked activity occurring within multiple structures across the auditory system (Batra et al., 1986; Gardi et al., 1979; Hoormann et al., 1992). It is interesting to note that the fine-structure patterns that emerge when the FFR is measured across a broad range of frequencies resemble the distinctive spectral patterns observed in the ABR to broadband clicks and chirps (Boston, 1981; Hall, 1986; Petoe et al., 2010). This similarity may be

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

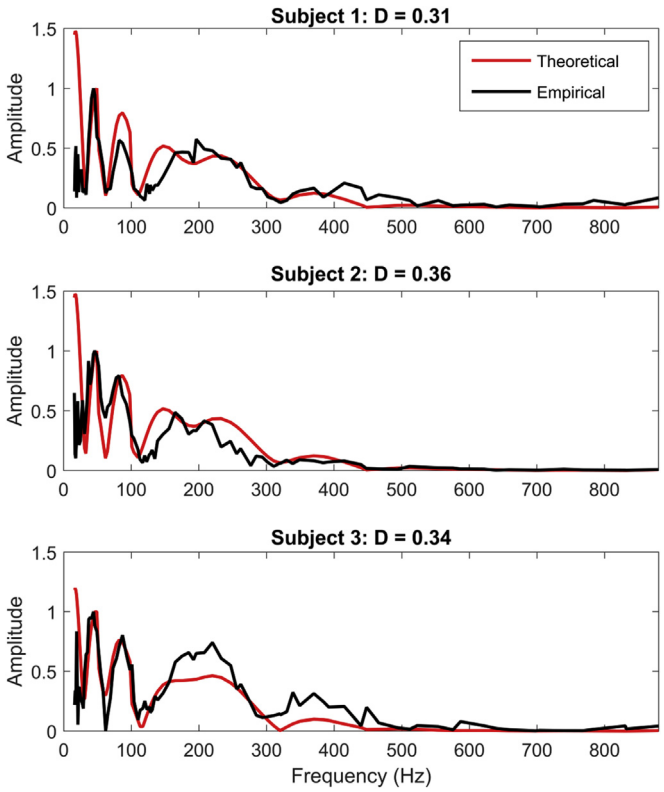

Fig. 5. Theoretical modeling of FFR-grams. For each subject, the empirical data (black) is compared against the theoretical FFRs (red) created from a six-generator model. The goodness of fit (Hoeffding Dependence Coefficient, *D*) is included in the title of each subplot. For comparative purposes, the theoretical and empirical plots are normalized so that the component at ~44 Hz is matched in amplitude between the empirical and theoretical responses. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

superficial or it may suggest that both techniques are tapping the same underlying frequency-dependent phenomenon, a possibility that should be explored further.

We refer to the frequency-dependent fluctuations in the FFR amplitude as *fine structure* because, like fine structure in audiograms and OAE-grams, it is theorized to reflect the interaction pattern created by multiple waves. Another shared feature is that FFR fine structure, like its namesake, is also repeatable for an individual (Abdala and Dhar, 2010; Reuter and Hammershoi, 2006). However, one primary difference is that audiometric and OAE fine structure are more individualistic, with each individual having a signature location for their peaks and valleys (Horst et al., 2003; Reuter and Hammershoi, 2006) that shift systematically in location based on stimulus intensity (He and Schmiedt, 1993). Our modeling suggests that the location of the FFR fine-structure peaks and valleys is predictable and stable, a finding that is born out in the data. This differential characteristic suggests that the two "fine-structure" phenomena are distinct. Yet, even if they arise through distinct mechanisms, this does not preclude the possibility they may interact. For example, within the audiometric fine-structure pattern, a maxima reflects a stronger cochlear response at that frequency, which in turn is expected to boost the FFR amplitude at that frequency. This boost might accentuate the FFR fine structure maxima when the stimulus frequency is at an FFR maximum or it might introduce a subtle frequency-dependent boost to the FFR between the peaks and valleys of the FFR fine structure that potentially produced the idiosyncratic ripple for each individual. However, because the morphology of the cochlear fine structure varies with stimulus intensity, the interaction patterns between cochlear and FFR fine structure are also likely to be intensity dependent (Elsisy and Krishnan, 2008; Mauermann et al., 2004; Poling et al., 2014). Together, this suggests the need for implementing fine-frequency audiograms and fine-frequency FFR testing at different intensities in the same individual to better understand whether they share more than just a name.

Our findings converge with previous work to verify the existence of FFR fine structure in humans. On the lower end of the stimulus spectrum, we observed high points at ~44 and 87 Hz, frequencies that correspond to the robust, well-characterized "40 Hz" and "80 Hz" phase-locked responses that surface for modulated, steady-state stimuli (Herdman et al., 2002; Kaf and Danesh, 2008). However, a number of key differences emerged from our data relative to other studies. For instance, past work reported that the scalp-recorded FFR is most robust near 250–300 Hz for frequencies between 100 and 500 Hz in humans (Batra et al., 1986; Hoormann et al., 1992). However, we found the response to be greatest near 200 Hz in this frequency range. Additionally, when replicating the FFR-grams from Hoormann et al. (1992) (128–832 Hz), we found the response to be greatest at 192 Hz, while the authors reported that 320 and 384 Hz yielded the maximum amplitude for most of the 20 participants in their sample. Hoormann et al. (1992) and Batra et al. (1986) used a contralateral ear/mastoid reference for their recordings. This is noteworthy because different electrode arrays can capture different sets and/or different mixtures of generators underlying the multi-source FFR (Bidelman, 2015), depending on the spatial orientation of the generators, as well as the volume conduction properties of the tissue between the recording electrodes (Nunez and Srinivasan, 2006). Bidelman (2015) found that using either the mastoids or ear lobe as a reference site, as opposed to a common averaged reference, was more likely to detect both peripheral and central sources in the scalp-recorded response, resulting in more robust FFRs. A common averaged reference, however, was determined to primarily reflect more central activity, resulting in relatively less intense FFRs. While Bidelman (2015)'s use of a small stimulus set with a small range of F0s does not permit commentary on how these electrode arrays affect FFR fine structure, his findings nevertheless highlight how different reference sites can influence the intensity of FFR amplitudes by tapping different sets of neural generators. Thus, any interpretation of FFR amplitudes, including FFR fine structure, must be done in the context of the electrode arrays and the putative generators recorded by that electrode array. A contralateral earlobe/mastoid reference compared to an ipsilateral reference has been shown to capture more activity from the ipsilateral than the contralateral auditory pathway. For instance, Wave I of the ABR is diminished, and in most cases not observed, with electrode arrays with contralateral reference, consistent with Wave I originating from the auditory nerve from the side ipsilateral to stimulation (Hughes et al., 1981; Norrix and Glattke, 1996; Prasher and Gibson, 1980). Thus, the electrode array used in these previous FFR studies (Batra et al., 1986; Hoormann et al., 1992) likely minimized the ability to detect activity recorded from the most peripheral FFR generators, including the cochlear microphonic. By minimizing the contribution of these generators in the scalp-recorded response, this could, in theory, affect the fine structure patterns. However, for phase-locked far-field responses (ASSRs), ipsilateral and contralateral references have been shown to yield similar scalp-recorded amplitudes for 39 and 79 Hz (Kaf and Danesh, 2008), two frequencies associated with high points in our FFR-grams. To investigate the effect of reference site on the FFR fine structure over the range of frequencies that Hoormann et al. (1992) tested, we conducted a small *post-hoc* study in which we simultaneously obtained FFRs using both ipsilateral and

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

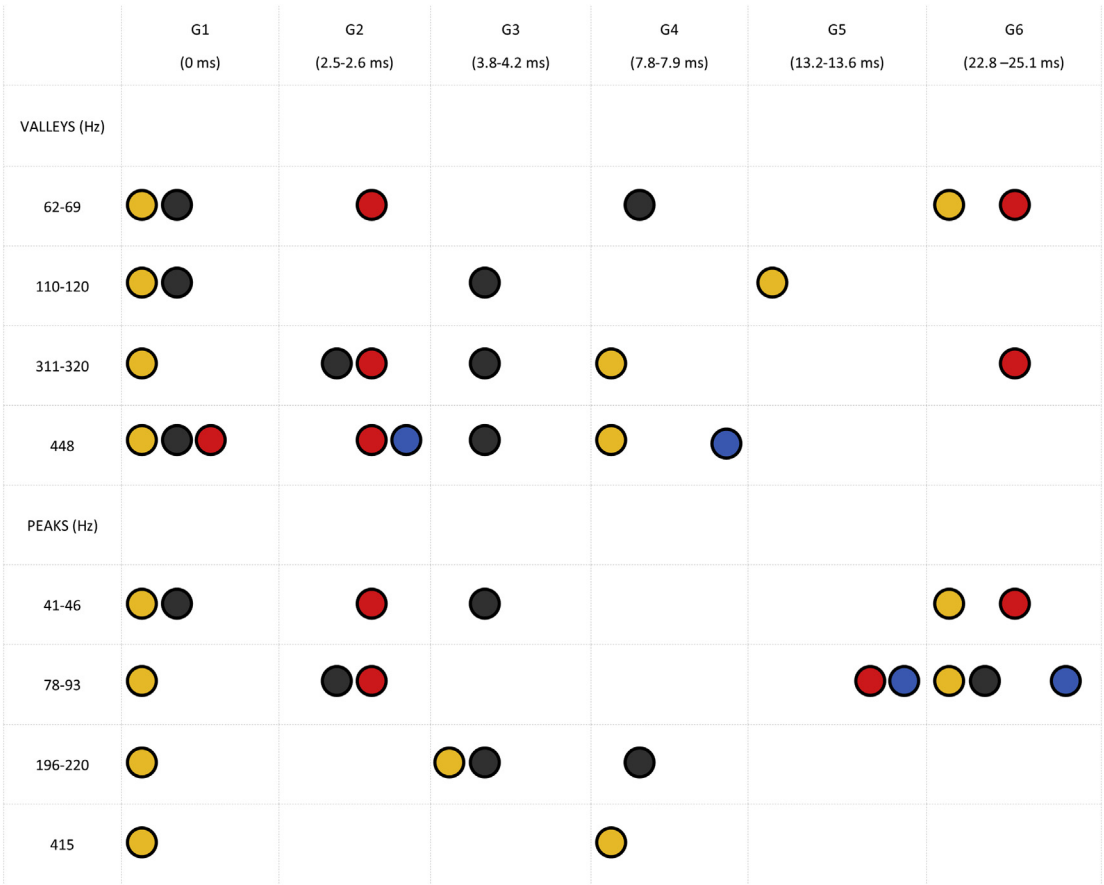

Fig. 6. A schematic illustrating how the major valleys and high points at specific frequency ranges in the FFR-gram can emerge as the mathematical consequence of different pairs of generators interacting. The frequency and latency ranges reported in the headers encompass the variation observed across the three subjects. Pairs of generators that can mathematically produce a specific peak or valley are denoted with the same color circular markings. For instance, the valley in the 62–69 Hz range, can arise from G1 and G6 interacting (yellow circles), G1 and G4 interacting (gray circles), and/or G2 and G6 interacting (red circles). The high point in the 78–93 Hz range can arise from the pairwise interaction of G1-G6 (red circles), G2-G6 (gray circles), G2-G5 (red circles), and/or G5-G6 (blue circles). The specific color of the circle has no additional meaning other than to serve as a way to distinguish how many different pairwise interactions can create a specific peak or valley. Additionally, this table is not exhaustive. Although it accounts for all pairwise comparisons, it does not capture all possible multiway-interaction patterns that could give rise to high points and valleys in these locations. Also note that the pairwise interactions denoted in the table can produce additional peaks and valleys that are not evident in the dataset. For example, the mathematically interaction of G5-G6 creates a valley at 49 Hz not present in our data. This underscores the idea that the FFR-gram footprint reflects the cumulative interaction patterns of multiple sets of generators. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

contralateral references for Subjects 2 and 3. Consistent with Kaf and Danesh (2008), we found that the amplitude of the FFR was similar for the ipsilateral and contralateral references, and that the contralateral reference did not shift the response maxima from 192 to 320 or 384 Hz. Thus, our small dataset suggests that the manifestation of a high point at ~320 Hz in the Hoormann et al. (2002) study may not be due to their use of a contralateral reference. We leave open the possibility that other factors could explain the differences between the amplitude patterns in our study and the Hoormann et al. (1992) study. One possible candidate is the stimulus duration. Although the stimulus F0s were the same across the two studies, our stimuli were considerably longer (200 ms vs. 31.25 ms). We raise the possibility that the use of short stimuli may have underestimated the response to low frequencies in the Hoormann et al. study given that only a small number of stimulus cycles can fit into the 31.25 ms time window for the low frequencies in their stimulus set (four cycles in the case of the lowest frequency).

To further probe the multiple-generator theory of FFR fine structure, we created a theoretical model of the FFR, expanding upon the basic procedure employed by Gardi et al. (1979). We modeled each subject's empirical FFR-gram using a bank of six theoretical generators, in which each generator in the bank had a different relative delay (latency). We utilized latencies that were consistent with structures that have emerged in the literature as putative FFR generators, namely, the CM, CN, SOC, IC, and AC. For each stimulus F0, we employed sine waves of that frequency to represent the phase-locked response to the F0 from each putative generator. The theoretical (far-field) FFR was then created by summing across the bank of sine waves. Depending on the F0, the bank of sine waves were either in- or out-of-phase, giving rise to frequency-dependent amplitude variation in the theoretical FFR when the sine waves were summed (see methods). For the frequency ranges corresponding to FFR valleys, our modeling suggests that they coincide with frequencies where the volume-conducted output of multiple different sets of generators are interacting destructively at the scalp. For example, the valley in the FFR-gram in the vicinity of 62–66 Hz can mathematically emerge as a consequence of the volume conducted output of G1 interacting with G6, G2 interacting with G6, and/or G1 interacting with G4 (among other possibilities not tested). This suggests the the FFR-gram valleys may reflect redundant interaction patterns that are re-enforced by multiple sets of generators producing signals that are not phase-aligned. Moreover, to create the global FFR footprint

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

across the full stimulus spectrum, we conclude that all six generators are necessary to the model. However, we are careful to point out that, in our theoretical model, the fine-structure patterns are determined by the latency spacing (i.e., relative latencies) between generators and not the absolute latency of each generator. Thus, although we set the earliest latency to 0 ms, equivalent theoretical models would have been produced, for example, if the earliest generator was set to have a characteristic latency of 2 ms (not 0 ms) while preserving the latency spacing between the theoretical generators. Although this aspect of our modeling creates an ambiguity that limits our ability to confidently map the theoretical generators to actual biological substrates, our use of single polarity stimuli increased the likelihood that the cochlea (cochlear microphonic) is contributing to the scalp-recorded FFR in our dataset (Chimento and Schreiner, 1990; Euler and Kiessling, 1983; Sohmer and Pratt, 1976).

There has been recent interest in the degree to which the human auditory cortex contributes to FFRs in the vocal pitch range (Coffey et al., 2016). In non-human primates, phase locking has been observed in multi-unit activity of primary auditory cortex at 100 Hz but not 250 Hz (Steinschneider et al., 1980), suggesting that the cortex may contribute to some but not all frequencies that span the human vocal pitch range (85–400 Hz). Our modeling corroborates this conclusion. Consistent with the phase-locking properties of the primate auditory cortex, we found that the best model fits were achieved when the longest latency theoretical generator (~26 ms) contributed to frequencies below 100 Hz and the second longest latency generator (~13 ms) contributed to frequencies below 200 Hz. Also when the drop-out parameter was removed from our model (i.e., when the two longest latency generators were allowed to contribute to the theoretical responses to the higher frequencies), this produced FFR peaks and valleys that are not present in the empirical FFRs; for example, the valley around ~110 Hz shifts higher in frequency to ~140–150 Hz and a new strong valley is introduced ~208–233 Hz that in the empirical data. Thus, while different structures within the auditory cortex may contribute to the FFR to human vocal pitches, their contribution is highly frequency dependent and weak, given that the vocal pitch range coincides with the upper limits of cortical phase-locking. Our modeling also suggests that the early-responding generators, likely of subcortical origin, contribute to both low and high frequency responses but that the fine structure peaks and valleys observed >100 Hz are primarily the result of the interaction patterns between the different early latency generators (Fig. 6).

Collectively, our multiple-generator modeling of the FFRs supports the interpretation that fine structure in the scalp-recorded response is epiphenomenal, arising from summed in- and out-of-phase volume-conducted potentials from different sources aggregating at the recording electrodes. However, there are a number of limitations to our model worth enumerating. Unlike Gardi et al. (1979) who were able to selectively ablate various structures within the auditory pathway, we are unable to definitively map each putative generator to a specific anatomical landmark, although we are able to model changes to the FFR as each putative generator is removed (Supplemental Figs. 4–6). Moreover, while a model with six biologically-plausible generators can re-create the global footprint of the FFR-gram, we did not expand the model beyond six generators to test whether six is a necessary or sufficient number of generators nor did we exhaustively test different possible amplitude scaling factors for each generator. In addition, there is more than one possible set of latency values that can produce the fine-structure patterns observed in our dataset, which complicates the precise pinpointing of generators. However, one advantage of our approach is that by using relative latencies, we do not need to correct for the frequency-dependent cochlear traveling-wave delay (Dau, 2003; Woods et al., 1993) or individual differences in ear canal length. Another limitation of our model is that it does not account for the trial-by-trial variability that is evident in FFRs (Hornickel and Kraus, 2013; Krizman et al., 2014), nor does it account for the small amplitude ripples evident in the FFR-gram, two phenomena we intend to capture in future iterations of the model. Our model also does not address sleep-, arousal-, attention-, and task-related changes to the FFR that have been noted in the literature (Galbraith and Arroyo, 1993; Hairston et al., 2013; Lehmann and Schonwiesner, 2014; Linden et al., 1987; Tiitinen et al., 1993). For example, the response to 40-Hz stimulation has been shown to reduce during sleep and decrease below the noise floor during (some) surgical anesthesia, but the 80-Hz response is comparatively more impervious to state changes (Griskova et al., 2007; Kuwada et al., 2002; Linden et al., 1985; Picton et al., 2003; Szalda and Burkard, 2005; Tlumak et al., 2012). Based on this literature, we predict that the fine-structure footprint (i.e., the location of the peaks and valley) would change when the listener is in an anesthetized (versus an awake) state because it would influence which sources were contributing to the scalp-recorded phase-locked response (Kuwada et al., 2002; Szalda and Burkard, 2005). That said, the high degree of overlap between the fine structure patterns that were measured empirically versus predicted by our model, together with the high repeatability of the fine-structure patterns measured almost one year apart and the similarity among the participants, suggest that in an unanesthetized state the fine-structure footprint emerges independent of the listener's state of sleep or arousal. The likelihood that Subject 1 was at an identical level of arousal when a given tone was played at test interval 1 and test interval 2, or that the three subjects were matched in arousal for a given test tone, is very low. Finally, although we used triangle-wave stimuli to record the empirical FFRs, we employed sinewaves to model our data since our focus, at this stage, is in the modeling FFRs to F0 and not harmonics (Coffey et al., 2016). Although we were able to successfully capture the fine-structure pattern using sinewaves as the basis of our modeling, adopting triangle waves instead of sinewaves is a logical next step to expanding and testing our model. However, despite these limitations, our model captured many global aspects of the fine structure in the empirical data, and it provides a framework for implementing intracerebral current source analysis and for fine-tuning the parameters of the model in future investigations.

In addition to providing insight into the complex, multiple-generator nature of the FFR, the reported FFR data and models also offer an alternative interpretation of what might account for group- and individual-level variation in the human scalp-recorded response. In current investigations of the FFR, variations in FFR spectral amplitudes for a given stimulus frequency are generally interpreted as reflecting variations in the fidelity of sound encoding, with smaller amplitudes considered indicative of a weaker auditory system (Krizman et al., 2012, 2015a; Ruggles et al., 2011; Song et al., 2011b). However, we posit that smaller scalp-recorded FFR amplitudes are not necessarily the result of weaker neural encoding; smaller FFR amplitudes can instead arise epiphenomenally in a healthy (i.e., not weakened) auditory system. This can happen when multiple FFR generators are phase-locking to the same frequency but their relative latencies produce a destructive interference pattern when the volume-conducted signals are pooled at the scalp. In such cases, the scalp-recorded composite response may be artificially small and underestimate the strength of each neural generator. This framing of the FFR places an emphasis on the relative latencies of the generators, not solely the functioning of the generators themselves. Thus, we propose that atypical FFRs observed in clinical and other special populations might not reflect impaired/weaker phase-locking *per se*, but could

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

instead reflect differences in the relative latencies of the generators. For instance, small FFR amplitudes to a specific frequency, may be masquerading as weaker encoding, when, instead, they may reflect individual differences, pathological or not, in the set and relative latency of the structures generating the FFR. This strikes us as especially relevant to the interpretation of previous work because we found a valley at ~100 Hz, the F0 of many speech stimuli used in recent FFR literature (e.g., Coffey et al., 2016; Krizman et al., 2012; Ruggles et al., 2012; Song et al., 2011b). Moreover, our modeling suggests that larger FFR amplitudes are not necessarily the result of stronger encoding: they could, instead, result from one generator being weak or absent (Supplemental Figs. 4–6), which in turn produces less phase cancellation in the far-field response recorded at the scalp. For example, it has recently been shown that bilinguals have larger FFRs to the F0 (F0 = 100 Hz) of complex speech sounds (Krizman et al., 2012, 2015a,b), which the authors interpreted as "stronger encoding". However, an alternative explanation is that auditory stimulation activates a different set of generators for bilinguals compared to monolinguals, or the same set with different relative latencies. Collectively, this underscores the importance of considering the valley or high-point status of a frequency when selecting a stimulus for characterizing group and individual differences using FFR methodology. Yet, this is not to suggest that the footprint of the FFR fine structure will necessarily be the same for all stimulus types given that different stimuli may engage different neural mechanisms and sub-populations (Bidelman, 2015). Thus, while we maintain that fine structure is an inherent property of the scalp-recorded FFR, the specific manifestation of the FFR footprint may vary as a function of stimulus type, suggesting that the location of the peaks and valleys observed in the current study may not generalize to all stimuli (Hoormann et al., 1992). So, while ~100 Hz emerges as a valley for our tonal stimuli, we leave open the possibility that this frequency may not be a valley for speech, for example, given that the F0 is conveyed through both the stimulus fine structure and envelope for tonal stimuli but through predominantly the envelope for speech stimuli.

In addition to applying these ideas to interpret intra-subject variations at a single frequency, future work should adopt a multiple-frequency stimulus design to address whether differences in FFR fine structure exist across groups of individuals with particular auditory backgrounds, distinct developmental states, or both. Although the subjects in our dataset had distinctive auditory backgrounds, the small sample limits our ability to draw any inferences at this point about FFR fine structure and experience-dependent plasticity; however, current evidence from our study and Hoormann et al. (1992) study suggests that in the absence of any explicit auditory training that FFR fine structure is stable. For insight into possible experience-dependent effects on FFR fine structure we can look to other studies: Skoe and Kraus (2012), for instance, recorded FFRs to eight frequencies (262, 294, 330, 350, 370, 393, 416, and 440 Hz) in musically-trained and untrained adults. The musically-trained individuals, in addition to having larger FFRs overall, had greater amplitude fluctuations across the stimulus set, in comparison to the untrained individuals who, as a group, had a flatter FFR-gram (i.e., less fine structure). Future work may help to shed light on this finding by examining how FFR fine structure changes, or does not change, over different timeframes and in response to different forms of auditory enrichment and deprivation. Additionally, developmental research should address how FFR fine structure changes in relation to the developing auditory system (e.g., anatomical and physiological development) and the acquisition of linguistic and musical abilities. Given that cortical and subcortical auditory structures have different developmental timelines (Moore, 2002; Werner et al., 2011) and given that the latency of subcortical and cortical responses undergo developmental changes (e.g., Krizman et al., 2015b; Savio et al., 2001; Sharma et al., 1997, 2005; Skoe et al., 2015), this leads us to hypothesize that the location of the valleys and high points may shift throughout development. Future work should investigate this hypothesis by using longitudinal and cross-sectional designs that include both young and aging populations.

In conclusion, we observed systematic frequency-dependent fluctuations in the FFR amplitude recorded over a wide range of frequencies. These fluctuations serve as evidence that the FFR is a composite of multiple phase-locked sources within the auditory system. Our findings highlight that the response at any given frequency can reflect activity from multiple generators while also underscoring that a full picture of the sources of the FFR is not possible when studying a single F0, or even a small set of F0s. Combining our multiple-frequency approach with dipole source EEG modeling and MEG, as was done in recent investigations of the sources of the FFR (Bidelman, 2015; Coffey et al., 2016), will strengthen claims about the generators of the response, in addition to bolstering the FFR's clinical and experimental utility.

Conflict of interest

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

Acknowledgments:

We would like to thank M. Hardin for her assistance with data collection and D. Harrell for feedback on an earlier version of the manuscript. This work was supported by a NSF IGERT Fellowship awarded to PT and University of Connecticut startup funds awarded to ES. The authors contributed equally to this work.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.heares.2017.01.014.

References

Abdala, C., Dhar, S., 2010. Distortion product otoacoustic emission phase and component analysis in human newborns. J. Acoust. Soc. Am. 127, 316–325.
Aiken, S.J., Picton, T.W., 2008. Envelope and spectral frequency-following responses to vowel sounds. Hear. Res. 245, 35–47.
Akhoun, I., Gallego, S., Moulin, A., Menard, M., Veuillet, E., Berger-Vachon, C., Collet, L., 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., Kraus, N., 2011. A neural basis of speech-in-noise perception in older adults. Ear Hear. 32, 750–757.
Anderson, S., Parbery-Clark, A., White-Schwoch, T., Kraus, N., 2012. Aging affects neural precision of speech encoding. J. Neurosci. 32, 14156–14164.
Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. Reading and subcortical auditory function. Cereb. Cortex 19, 2699–2707.
Batra, R., Kuwada, S., Maher, V.L., 1986. The frequency-following response to continuous tones in humans. Hear. Res. 21, 167–177.
Bidelman, G.M., 2015. Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. Hear. Res. 323, 68–80.
Bones, O., Hopkins, K., Krishnan, A., Plack, C.J., 2014. Phase locked neural activity in the human brainstem predicts preference for musical consonance. Neuropsychologia 58, 23–32.
Boston, J.R., 1981. Spectra of auditory brainstem responses and spontaneous EEG. IEEE Trans. Bio-med. Eng. 28, 334–341.
Campbell, T., Kerlin, J.R., Bishop, C.W., Miller, L.M., 2012. Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. Ear Hear. 33, 144.
Carcagno, S., Plack, C.J., 2011. Subcortical plasticity following perceptual learning in a pitch discrimination task. J. Assoc. Res. Otolaryngol. 12, 89–100.
Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47, 236–246.

Chimento, T., Schreiner, C., 1990. Selectively eliminating cochlear microphonic contamination from the frequency-following response. Electroencephalogr. Clin. Neurophysiol. 75, 88–96.
Coffey, E.B., Herholz, S.C., Chepesiuk, A.M., Baillet, S., Zatorre, R.J., 2016. Cortical contributions to the auditory frequency-following response revealed by meg. Nat. Commun. 7, 11070.
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.
Dhar, S., Talmadge, C.L., Long, G.R., Tubis, A., 2002. Multiple internal reflections in the cochlea and their effect on dpoae fine structure. J. Acoust. Soc. Am. 112, 2882–2897.
Dolphin, W.F., Mountain, D.C., 1992. The envelope following response: scalp potentials elicited in the mongolian gerbil using sinusoidally Am acoustic signals. Hear. Res. 58, 70–78.
Elliott, E., 1958. A ripple effect in the audiogram. Nature 181, 1076.
Elsisy, H., Krishnan, A., 2008. Comparison of the acoustic and neural distortion product at 2f1-F2 in normal-hearing adults. Int. J. Audiol. 47, 431–438.
Euler, M., Kiessling, J., 1983. Far-field cochlear microphonics in man and their relation to cochlear integrity. Electroencephalogr. Clin. Neurophysiol. 56, 86–89.
Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. U. S. A. 78, 2643–2647.
Galbraith, G.C., Arroyo, C., 1993. Selective attention and brainstem frequency-following responses. Biol. Psychol. 37, 3–22.
Galbraith, G.C., Arbagey, P.W., Branski, R., Comerci, N., Rector, P.M., 1995. Intelligible speech encoded in the human brain stem frequency-following response. Neuroreport 6, 2363–2367.
Gardi, J., Merzenich, M., McKean, C., 1979. Origins of the scalp recorded frequency-following response in the cat. Audiology 18, 358–381.
Gaskill, S.A., Brown, A.M., 1990. The behavior of the acoustic distortion product, 2f1− F2, from the human ear and its relation to auditory sensitivity. J. Acoust. Soc. Am. 88, 821–839.
Greenberg, S., Marsh, J.T., Brown, W.S., Smith, J.C., 1987. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hear. Res. 25, 91–114.
Griskova, I., Morup, M., Parnas, J., Ruksenas, O., Arnfred, S.M., 2007. The amplitude and phase precision of 40 hz auditory steady-state response depend on the level of arousal. Exp. Brain Res. 183, 133–138.
Hairston, W.D., Letowski, T.R., McDowell, K., 2013. Task-related suppression of the brainstem frequency following response. PLoS One 8, e55215.
Hall 3rd, J.W., 1986. Auditory brain stem response spectral content in comatose head-injured patients. Ear Hear. 7, 383–389.
He, N.J., Schmiedt, R.A., 1993. Fine structure of the 2f1-F2 acoustic distortion product: changes with primary level. J. Acoust. Soc. Am. 94, 2659–2669.
Herdman, A.T., Lins, O., Van Roon, P., Stapells, D.R., Scherg, M., Picton, T.W., 2002. Intracerebral sources of human auditory steady-state responses. Brain Topogr. 15, 69–86.
Hoormann, J., Falkenstein, M., Hohnsbein, J., Blanke, L., 1992. The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. Hear. Res. 59, 179–188.
Hornickel, J., Kraus, N., 2013. Unstable representation of sound: a biological marker of dyslexia. J. Neurosci. 33, 3500–3504.
Hornickel, J., Knowles, E., Kraus, N., 2012. Test-retest consistency of speech-evoked auditory brainstem responses in typically-developing children. Hear. Res. 284, 52–58.
Horst, J.W., Wit, H.P., Albers, F.W.J., 2003. Quantification of audiogram fine-structure as a function of hearing threshold. Hear. Res. 176, 105–112.
Hughes, J.R., Fino, J., Gagnon, L., 1981. The importance of phase of stimulus and the reference recording electrode in brain stem auditory evoked potentials. Electroencephalogr. Clin. Neurophysiol. 51, 611–623.
Jeng, F.-C., Costilow, C.E., Stangherlin, D.P., Lin, C.-D., 2011. Relative power of harmonics in human frequency-following responses associated with voice pitch in american and chinese adults. Percept. Mot. Ski. 113, 67–86.
Johnson, K.L., Nicol, T., Zecker, S.G., Kraus, N., 2008. Developmental plasticity in the human auditory brainstem. J. Neurosci. 28, 4000–4007.
Joris, P.X., Schreiner, C.E., Rees, A., 2004. Neural processing of amplitude-modulated sounds. Physiol. Rev. 84, 541–577.
Kaf, W.A., Danesh, A.A., 2008. Air-conduction auditory steady-state response: comparison of interchannel recording using two modulation frequencies. J. Am. Acad. Audiol. 19, 696–707.
Krishnan, A., 2002. Human frequency-following responses: representation of steady-state synthetic vowels. Hear. Res. 166, 192–201.
Krishnan, A., Xu, Y., Gandour, J., Cariani, P., 2005. Encoding of pitch in the human brainstem is sensitive to language experience. Brain Res. Cogn. Brain Res. 25, 161–168.
Krizman, J., Skoe, E., Marian, V., Kraus, N., 2014. Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. Brain Lang. 128, 34–40.
Krizman, J., Marian, V., Shook, A., Skoe, E., Kraus, N., 2012. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc. Natl. Acad. Sci. U. S. A. 109, 7877–7881.
Krizman, J., Slater, J., Skoe, E., Marian, V., Kraus, N., 2015a. Neural processing of speech in children is influenced by extent of bilingual experience. Neurosci. Lett. 585, 48–53.
Krizman, J., Tierney, A., Fitzroy, A.B., Skoe, E., Amar, J., Kraus, N., 2015b. Continued maturation of auditory brainstem function during adolescence: a longitudinal approach. Clin. Neurophysiol. 126, 2348–2355.
Kuwada, S., Anderson, J.S., Batra, R., Fitzpatrick, D.C., Teissier, N., D'Angelo, W.R., 2002. Sources of the scalp-recorded amplitude-modulation following response. J. Am. Acad. Audiol. 13, 188–204.
Langner, G., 1992. Periodicity coding in the auditory system. Hear. Res. 60, 115–142.
Lee, K.M., Skoe, E., Kraus, N., Ashley, R., 2009. Selective subcortical enhancement of musical intervals in musicians. J. Neurosci. 29, 5832–5840.
Lehmann, A., Schonwiesner, M., 2014. Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. PLoS One 9, e85442.
Liégeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., Chauvel, P., 1994. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr. Clin. Neurophysio. Evoked Potentials Sect. 92, 204–214.
Linden, R., Picton, T.W., Hamel, G., Campbell, K.B., 1987. Human auditory steady-state evoked potentials during selective attention. Electroencephalogr. Clin. Neurophysiol. 66, 145–159.
Linden, R.D., Campbell, K.B., Hamel, G., Picton, T.W., 1985. Human auditory steady state evoked potentials during sleep. Ear Hear. 6, 167–174.
Liu, L.F., Palmer, A.R., Wallace, M.N., 2006. Phase-locked responses to pure tones in the inferior colliculus. J. Neurophysiol. 95, 1926–1935.
Mäkelä, J.P., Hari, R., 1987. Evidence for cortical origin of the 40 Hz auditory evoked response in man. Electroencephalogr. Clin. Neurophysiol. 66, 539–546.
Marmel, F., Linley, D., Carlyon, R.P., Gockel, H.E., Hopkins, K., Plack, C.J., 2013. Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently. J. Assoc. Res. Otolaryngol. 14, 757–766.
Marsh, J.T., Brown, W.S., Smith, J.C., 1975. Far-field recorded frequency-following responses: correlates of low pitch auditory perception in humans. Electroencephalogr. Clin. Neurophysiol. 38, 113–119.
Mauermann, M., Long, G.R., Kollmeier, B., 2004. Fine structure of hearing threshold and loudness perception. J. Acoust. Soc. Am. 116, 1066–1080.
Middlebrooks, J.C., 2008. Auditory cortex phase locking to amplitude-modulated cochlear implant pulse trains. J. Neurophysiol. 100, 76–91.
Moller, A.R., Jannetta, P.J., 1983. Interpretation of brainstem auditory evoked potentials: results from intracranial recordings in humans. Scand. Audiol. 12, 125–133.
Moore, D.R., 2002. Auditory development and the role of experience. Br. Med. Bull. 63, 171–181.
Moushegian, G., Rupert, A.L., Stillman, R.D., 1973. Scalp-recorded early responses in man to frequencies in the speech range. Electroencephalogr. Clin. Neurophysiol. 35, 665–667.
Norrix, L.W., Glattke, T.J., 1996. Multichannel waveforms and topographic mapping of the auditory brainstem response under common stimulus and recording conditions. J. Commun. Disord. 29, 157–182.
Nunez, P.L., Srinivasan, R., 2006. Electric Fields of the Brain : the Neurophysics of EEG, second ed. Oxford University Press, Oxford; New York.
Petoe, M.A., Bradley, A.P., Wilson, W.J., 2010. On chirp stimuli and neural synchrony in the suprathreshold auditory brainstem response. J. Acoust. Soc. Am. 128, 235–246.
Picton, T.W., John, M.S., Purcell, D.W., Plourde, G., 2003. Human auditory steady-state responses: the effects of recording technique and state of arousal. Anesth. Analg. 97, 1396–1402.
Poling, G.L., Siegel, J.H., Lee, J., Lee, J., Dhar, S., 2014. Characteristics of the 2f1-F2 distortion product otoacoustic emission in a normal hearing population. J. Acoust. Soc. Am. 135, 287–299.
Prasher, D.K., Gibson, W.P.R., 1980. Brain stem auditory evoked potentials: significant latency differences between ipsilateral and contralateral stimulation. Electroencephalogr. Clin. Neurophysiol. 50, 240–246.
Reuter, K., Hammershoi, D., 2006. Distortion product otoacoustic emission fine structure analysis of 50 normal-hearing humans. J. Acoust. Soc. Am. 120, 270–279.
Ribary, U., Ioannides, A.A., Singh, K.D., Hasson, R., Bolton, J.P., Lado, F., Mogilner, A., Llinas, R., 1991. Magnetic field tomography of coherent thalamocortical 40-hz oscillations in humans. Proc. Natl. Acad. Sci. U. S. A. 88, 11037–11041.
Ruggles, D., Bharadwaj, H., 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. 108, 15516–15521.
Ruggles, D., Bharadwaj, H., Shinn-Cunningham, B.G., 2012. Why middle-aged listeners have trouble hearing in everyday settings. Curr. Biol. 22, 1417–1422.
Russo, N., Nicol, T., Musacchia, G., Kraus, N., 2004. Brainstem responses to speech syllables. Clin. Neurophysiol. 115, 2021–2030.
Savio, G., Cárdenas, J., Pérez Abalo, M., González, A., Valdés, J., 2001. The low and high frequency auditory steady state responses mature at different rates. Audiol. Neurotol. 6, 279–287.
Sharma, A., Kraus, N., McGee, T.J., Nicol, T.G., 1997. Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. Electroencephalogr. Clin. Neurophysiol. 104, 540–545.
Sharma, A., Martin, K., Roland, P., Bauer, P., Sweeney, M.H., Gilley, P., Dorman, M., 2005. P1 latency as a biomarker for central auditory development in children with hearing impairment. J. Am. Acad. Audiol. 16, 564–573.
Shera, C.A., Guinan Jr., J.J., 1999. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian oaes. J. Acoust. Soc. Am. 105, 782–798.
Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014

tutorial. Ear Hear. 31, 302–324.
Skoe, E., Kraus, N., 2012. A little goes a long way: how the adult brain is shaped by musical training in childhood. J. Neurosci. 32, 11507–11510.
Skoe, E., Krizman, J., Anderson, S., Kraus, N., 2015. Stability and plasticity of auditory brainstem function across the lifespan. Cereb. Cortex 25, 1415–1426.
Slabu, L., Grimm, S., Escera, C., 2012. Novelty detection in the human auditory brainstem. J. Neurosci. 32, 1447–1452.
Smith, J.C., Marsh, J.T., Brown, W.S., 1975. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr. Clin. Neurophysiol. 39, 465–472.
Snyder, R.L., Schreiner, C.E., 1984. The auditory neurophonic: basic properties. Hear. Res. 15, 261–280.
Sohmer, H., Pratt, H., 1976. Recording of the cochlear microphonic potential with surface electrodes. Electroencephalogr. Clin. Neurophysiol. 40, 253–260.
Song, J.H., Nicol, T., Kraus, N., 2011a. Test-retest reliability of the speech-evoked auditory brainstem response. Clin. Neurophysiol. 122, 346–355.
Song, J.H., Skoe, E., Banai, K., Kraus, N., 2011b. Perception of speech in noise: neural correlates. J. Cogn. Neurosci. 23, 2268–2279.
Steinschneider, M., Arezzo, J., Vaughan Jr., H.G., 1980. Phase-locked cortical responses to a human speech sound and low-frequency tones in the monkey. Brain Res. 198, 75–84.
Szalda, K., Burkard, R., 2005. The effects of nembutal anesthesia on the auditory steady-state response (assr) from the inferior colliculus and auditory cortex of the chinchilla. Hear. Res. 203, 32–44.
Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., Näätänen, R., 1993. Selective attention enhances the auditory 40-hz transient response in humans. Nature 364, 59–60.
Tlumak, A.I., Durrant, J.D., Delgado, R.E., Robert Boston, J., 2012. Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates in awake vs. Natural sleep. Int. J. Audiol. 51, 418–423.
Weiss, M.W., Bidelman, G.M., 2015. Listening to the brainstem: musicianship enhances intelligibility of subcortical representations for speech. J. Neurosci. 35, 1687–1691.
Werner, L., Fay, R.R., Popper, A.N., 2011. Human Auditory Development. Springer Science & Business Media.
Wong, P.C., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat. Neurosci. 10, 420–422.
Woods, D.L., Alain, C., Covarrubias, D., Zaidel, O., 1993. Frequency-related differences in the speed of human auditory processing. Hear. Res. 66, 46–52.

Please cite this article in press as: Tichko, P., Skoe, E., Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, Hearing Research (2017), http://dx.doi.org/10.1016/j.heares.2017.01.014