Hearing Research 323 (2015) 68–80

Contents lists available at ScienceDirect

Hearing Research

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

Research paper

Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR

Gavin M. Bidelman [a, b, *]

[a] Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA
[b] School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

ARTICLE INFO

Article history:
Received 21 November 2014
Received in revised form
9 January 2015
Accepted 27 January 2015
Available online 7 February 2015

ABSTRACT

Brainstem frequency-following responses (FFRs) probe the neural transcription of speech/music, auditory disorders, and plasticity in subcortical auditory function. Despite clinical and empirical interest, the response's neural basis remains poorly understood. The current study aimed to more fully characterize functional properties of the human FFR (topography, source locations, generation). Speech-evoked FFRs were recorded using a high-density (64 channel) electrode montage. Source dipole modeling and 3-channel Lissajous analysis was used to localize the most likely FFR generators and their orientation trajectories. Additionally, transient auditory brainstem responses (ABRs), recorded in the same listeners, were used to predict FFRs and test the long-held assumption that the sustained potential reflects a series of overlapping onset responses. Results showed that FFRs were maximal at frontocentral scalp locations with obliquely oriented sources from putative generators in the midbrain (i.e., upper brainstem). Comparisons between derived and actual recordings revealed the FFR is not a series of repeated ABR wavelets and thus, represents a functionally distinct brainstem response. FFRs recorded at temporal electrode sites showed larger amplitudes and contained higher frequency components than vertex channels (Fz, Cz) suggesting that FFRs measured near the mastoid are generated more peripherally (auditory nerve) than measurements at frontocentral scalp locations. Furthermore, this reveals the importance of choice in reference electrode location for FFR interpretation. Our findings provide non-invasive evidence that (i) FFRs reflect sustained neural activity whose sources are consistent with rostral brainstem generators and (ii) FFRs are functionally distinct from the onset ABR response.

1. Introduction

The scalp-recorded brainstem frequency-following response (FFR) is a sustained "neuromicrophonic" potential that reflects dynamic, phase-locked activity to spectrotemporal features of complex acoustic sounds (e.g., speech and music) (for reviews, see Chandrasekaran and Kraus, 2010; Krishnan, 2007; Skoe and Kraus, 2010). The evoked potential is characterized by a periodic waveform which follows individual cycles of the stimulus waveform. The remarkable fidelity of the FFR is apparent by the fact that when played back as an audio stimulus, the neural response to speech is intelligible to human listeners (Weiss and Bidelman, (2015); Galbraith et al., 1995). These properties make the FFR particularly advantageous for studying the neural encoding of complex sounds at pre-attentive, subcortical levels of auditory processing. Among other phenomena, FFRs have provided detailed insight into how the sensory nervous system tracks aspects of voice pitch prosody (Bidelman et al., 2011; Krishnan et al., 2010a; Wong et al., 2007), encodes speech formant cues (Bidelman and Krishnan, 2010; Bidelman et al., 2013; Johnson et al., 2005), and the neural representations of melodic and harmonic aspects of music (Bidelman, 2013).

Despite considerable interest in the human FFR for clinical and empirical utility, few studies have fully characterized its morphology, underlying response properties, and neural generators. Descriptions of the response have typically assumed a subcortical origin with sources circumscribed to the midbrain

* School of Communication Sciences & Disorders, University of Memphis, 807 Jefferson Ave. Memphis, TN, 38105, USA. Tel.: +1 901 678 5826; fax: +1 901 525 1282.
E-mail address: g.bidelman@memphis.edu.

http://dx.doi.org/10.1016/j.heares.2015.01.011
0378-5955/

inferior colliculus (IC) (Chandrasekaran and Kraus, 2010; Glaser et al., 1976; Krishnan, 2007; Moushegian et al., 1973; Skoe and Kraus, 2010; Smith et al., 1975). This premise is based primarily on indirect/converging evidence from its response properties: the short latency of the FFR (~6–8 ms) is too early to reflect contribution from cortical generators (Galbraith et al., 2000); the FFR contains phase-locked activity (e.g., >1000 Hz) well beyond the upper limit of phase-locking in cortex (i.e., ~100 Hz) (Aiken and Picton, 2008; Akhoun et al., 2008a; Wallace et al., 2000); and there is a high correspondence between far-field FFRs and near-field intracranial potentials recorded directly from the IC (Smith et al., 1975). More definitive evidence for the FFR's neural generators stems from lesion data which demonstrates that focal ablation of the IC in cats (Davis and Britt, 1984; Smith et al., 1975) and humans (Sohmer et al., 1977) abolishes the scalp-recorded FFR. However, we are aware of no study to date which has attempted to localize the human FFR elicited by speech using noninvasive methodologies and in neurological normal listeners.

FFRs are typically recorded using a single differential electrode channel consisting of positive (non-inverting) electrode placed at a frontocentral scalp location (Fz or Cz) referenced to an inverting electrode placed on either mastoid, linked mastoids, or earlobe; a third cephalic electrode serves as common ground (Aiken and Picton, 2008; Bidelman et al., 2011; Krishnan, 2007; Skoe and Kraus, 2010). This montage is oriented parallel to the vertical dipole of the brainstem and is thought to isolate neural activity originating from more rostral auditory structures (Galbraith, 1994). In addition to this vertical configuration (i.e., Fz − M1/M2), some investigators have employed a horizontal derivation with electrodes positioned between the mastoids (e.g., M1 − M2). This configuration is oriented parallel to the geometry of caudal brainstem generators and is thought to reflect activity from more peripheral auditory structures (e.g., auditory nerve). Together, vertical and horizontal channel montages provide complementary but unique views of the brainstem response by emphasizing different anatomical and functional contributions to the recorded FFR (Galbraith, 1994).

One issue with these typical FFR recording approaches is the choice of reference location. Differential amplification schemes often assume that the reference is placed on a non-active, electrically neutral site whose spatiotemporal voltage gradients change minimally over the epoch window of interest (Wolpaw and Woods, 1982). Unfortunately, reference sites commonly employed in evoked potential recordings (e.g., head, earlobes, mastoids) generally lie within regions characterized by significant time-varying voltage gradients, rendering them anything but “neutral” or “zero-potential” sites (Geselowitz, 1998; Wolpaw and Woods, 1982). It is well-known for example, that ABR responses are susceptible to contamination from myogenic (e.g., post-auricular muscle) or other EEG interferences because noise picked up at the reference (e.g., mastoid electrode) is broadcast into the desired recording at the non-inverting terminal (Nunez and Srinivasan, 2006; O'Beirne and Patuzzi, 1999; Terkildsen and Osterhammel, 1981). Alleviating such bias is partly achieved by use of a common average reference (CAR) whereby each active recording sensor is referenced to the mean activity around the head volume conductor (Dien, 1998). ERPs recorded using a CAR yield the least biased/undistorted view of referential recording approaches and more closely approximate the true/absolute (i.e., reference-free) potential at single scalp locations (Dien, 1998). However, average referencing requires a recording array which samples a sufficiently large area of the scalp, achievable only with higher density, multichannel recordings.

The current study provides the first high-density, multichannel characterization of the human FFR to speech. While multichannel descriptions of other brainstem responses exist in the literature, previous studies have focused on phase-locked responses to the envelope of complex tones (auditory steady-state response, ASSR: Bharadwaj and Shinn-Cunningham, 2014; Herdman et al., 2002) or the transient, click-evoked auditory brainstem response (ABR) (Parkkonen et al., 2009). We are aware of no study which has directly examined multichannel FFRs elicited by speech stimuli. This novel multichannel approach is advantageous for two reasons. First, it allows us to compare the effect of reference choice (including CAR) on FFR morphology and its response properties (e.g., absolute amplitude, frequency composition). To date, such features have only been described using one- or two-channel recording montages. It has also been unclear from other multichannel approaches (e.g., Bharadwaj and Shinn-Cunningham, 2014), what reference and electrode montage should be employed to optimally recording FFRs emitted from the rostral brainstem. Secondly, this allows the evaluation of the FFR with minimal contribution from mastoid or earlobe referencing as commonly employed in previous work. Multichannel recordings also allow us to characterize the topographic distribution of the FFR response and locate its underlying source generators using noninvasive dipole modeling. Such characterizations have not been possible with traditional single-channel recordings.

Our second aim was to examine the functional distinction between the FFR and the more typically recorded ABR. There has been a long theoretical debate regarding the functional distinctions between the sustained (FFR) and transient (ABR) brainstem evoked potentials (e.g., Daly et al., 1976; Dau, 2003; Davis and Hirsh, 1974; Gerken et al., 1975; Janssen et al., 1991; Picton et al., 1977). Modeling work has suggested that the periodic FFR may reflect a series of repeated and overlapping onset ABR responses (Dau, 2003; Janssen et al., 1991). In the computational modeling described by Dau (2003), simulated FFR-like potentials were generated by a convolution between a unitary impulse response (i.e., derived from the click-ABR recordings) and the periodic instantaneous discharge pattern of an auditory nerve model output (for the original inception of this model, see Melcher and Kiang, 1996). This approach is based on the assumption that sustained neural activity is theoretically iterated activity from overlapping onset responses (Goldstein and Kiang, 1958; Janssen et al., 1991). In this formulation, the generator(s) of the FFR do not differ from those of the transient ABR; the former results from the algebraic, linear superposition of later waves of the ABR (e.g., waves IV-V) responding to successive peaks of the stimulus (Daly et al., 1976; Dau, 2003; Janssen et al., 1991). Similar conceptions have been applied to cortical 40 Hz auditory steady-state responses, which are often conceived as a series of overlapping middle-latency onset potentials (Galambos et al., 1981). However, to our knowledge, this “convolution model” of the FFR generation has never been fully authenticated in humans. The present study directly tested the hypothesis that the sustained human brainstem FFR is simply a series of overlapping onset ABRs (e.g., Daly et al., 1976; Dau, 2003; Davis and Hirsh, 1974; Gerken et al., 1975; Janssen et al., 1991; Picton et al., 1977).

In the current study, two experiments were conducted aimed to address the two aforementioned gaps in the FFR literature: (1) determine if the sustained, periodic nature of the FFR is simply a series of overlapping onset ABR wavelets (i.e., repeated transient responses); (2) characterize the scalp topography and source generator locations of the human speech-evoked FFR. Multichannel recordings (64 electrodes) were obtained in normal hearing listeners. Actual recorded FFRs were compared to “simulated FFRs” based on repeated/overlapping waveforms created from each individual's transient ABR response. Dipole analysis was also used to evaluate, non-invasively, the most probable source locations and orientations underlying the scalp recorded response.

2. Experiment 1: Is the sustained FFR a series of overlapping onset responses (ABRs)?

2.1. Rationale

As stated by Gerken et al. (1975), "It is also argued that the frequency-following response is not a microphonic-like response but rather that the individual waves in the [FFR] are evoked by the collective activity of phase-locked single units." This assumption implies that the FFR (sustained response) and ABR (onset response) are functional derivatives of one another and furthermore, that they are generated by the same population of neural elements. While this "convolution model" of FFR generation appears veridical *prima facie,* to our knowledge, it has never been validated empirically in any species, human or otherwise. The primary aim of Experiment 1 was to directly test the hypothesis that the sustained human brainstem FFR is a series of overlapping onset ABRs.

2.2. Methods

2.2.1. Subjects

Six monolingual speakers of American English (4 male, 2 female; age: 27.7 ± 4.4 years) participated in Exp 1. All participants exhibited normal hearing sensitivity at octave frequencies (250–8000 Hz) and reported no previous history of neuropsychiatric illness. All were strongly right-handed (>72 laterality) (Oldfield, 1971) and had minimal musical training (1.1 ± 3.9 years). Participants were paid and gave written informed consent in compliance with a protocol approved by the Institutional Review Board at the University of Memphis.

2.2.2. Stimuli

Scalp-recorded brainstem potentials were recorded separately in response to both clicks and sustained pitch stimuli, eliciting the common transient (ABR) and following (FFR) classes of the auditory brainstem response, respectively. Click stimuli were 100 μs pulses. Clicks allowed us to measure the transient ABR for each listener which was subsequently used to predict their measured FFR (see *2.2.4 Derivation of predicted FFR from ABR*).

FFRs were recorded in response to complex pitches (periodic click trains) with various fundamental frequencies (F0s). These pulse trains (Dirac comb functions) were constructed using a periodic series of impulses described by Eq. (1):

$$h(t) = \sum_{n=1}^{\frac{d}{T}} \delta(t - nT) \tag{1}$$

where d is the duration of the desired pulse train (here, 105 ms), and T is the period between successive impulses (i.e., the pitch period = 1/F0). Each pulse of the train was constructed using identical clicks (bandwidth, pulse with, amplitude) to those described for eliciting the ABR, spaced with F0s of 100, 200, and 500 Hz. This allowed us to directly test the assumption that the periodic FFR response is a linear superposition of repeated transient ABR waves. Because of the high repetition rate of click presentation (>>20 Hz), these stimuli produce audible and salient low pitch, optimal for evoking the FFR (Bidelman et al., 2011; Krishnan, 2007; Krishnan et al., 2010b; Skoe and Kraus, 2010).

2.2.3. Electrophysiological response recordings

For all EEG recordings, participants reclined comfortably in an IAC electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement, ignore the sounds they hear (to divert attention away from the stimulus), and were allowed to watch a muted subtitled movie to maintain a calm yet wakeful state. Stimulus presentation was controlled by MATLAB 2013 routed to a Tucker–Davis TDT RP2 interface. Stimuli were delivered binaurally at an intensity of 80 dB SPL (91 peak equivalent SPL) through electromagnetically shielded (Akhoun et al., 2008b) insert earphones (Etymotic ER-2) using fixed, rarefaction polarity. Listeners heard 3000 exemplars of each stimulus (ISI = 50 ms). ABR and FFR recording blocks were randomized within and across participants.

Supplementary video related to this article can be found at http://dx.doi.org/10.1016/j.heares.2015.01.011.

In Exp 1, neuroelectric responses were recorded differentially between Ag/AgCl disc electrodes placed on the scalp at the high forehead at the hairline (~Fpz) referenced to linked mastoids (M1/M2). Another electrode placed on the mid-forehead served as the common ground. Interelectrode impedance was maintained $\leq$5 kΩ. Continuous EEGs were digitized at 20 kHz (SynAmps RT amplifiers; Neuroscan) using an online passband of DC − 4000 Hz. EEGs were then epoched [ABRs: −10 to 40 ms; FFRs: −10 to 130 ms window], baseline corrected to the pre-stimulus interval, and averaged in the time domain to obtain responses for each stimulus condition. Sweeps containing activity exceeding ±50 μV were rejected as artifacts prior to averaging. Both ABR and FFR waveforms were then bandpass filtered between 50 and 3000 Hz to equate the bandwidth between both types of brainstem responses. Zero-phase filtering was achieved by a 4th order elliptical digital filter (stopband attenuation of −200 dB) as implemented in MATLAB ('ellip' and 'filtfilthd' functions). Lastly, all neural responses were gated using a 1 ms onset/offset $\cos^2$ ramps to ensure that "edge effects" of the combined window epoching and filtering process would not produce discontinuities in the waveforms (Nitschke et al., 1998). Importantly, this gating fell well-outside the expected stimulus-driven neural activity.

2.2.4. Derivation of predicted FFRs from the ABR

Listeners' actual FFR recordings were compared with their derived FFRs, simulated via convolution (e.g., Goldstein and Kiang, 1958; Janssen et al., 1991). This model assumes the sustained following response is generated by a series of overlapping onset responses (i.e., the FFR is an iterated ABR). This approach is illustrated in Fig. 1. Only the first 12 ms of the ABR traces was used to derive FFRs to avoid inclusion of the middle latency components which follow the ABR (McGee et al., 1991; Picton et al., 1977).[1] FFRs were simulated by convolving each listener's click ABR response with periodic click trains with different F0 spacing (Eq. (1); Fig. 1, "Impulse response"). This process generates a new ABR signature at each pulse of the periodic click train and yielded a complex waveform that, qualitatively, has the appearance of the sustained FFR (Fig. 1, "Derived FFR"). Simulated FFRs were achieved by varying the period of the pulse train's successive clicks. Simulated FFRs were then compared to true FFR recordings obtained in the same listeners elicited by complex tones with F0s of 100, 200, and 500 Hz. The onset latency of each response was also measured,

[1] Restricting the ABR "wavelet" to the first ~12 ms was done to remove the inclusion of MLR components which reflect neural activation of thalamic and cortical generators (McGee et al., 1991). This further ensured that derived and actual FFRs contained only neural activity of subcortical origin, allowing for the best possible opportunity for convergence between responses. Initial analysis including the entire transient response time window (ABR + MLR) indicated that results were indistinguishable from those reported in the main text. Hence, inclusion of the MLR did not seem to make a difference in ABR-derived FFRs, which still diverged from true FFR recordings.

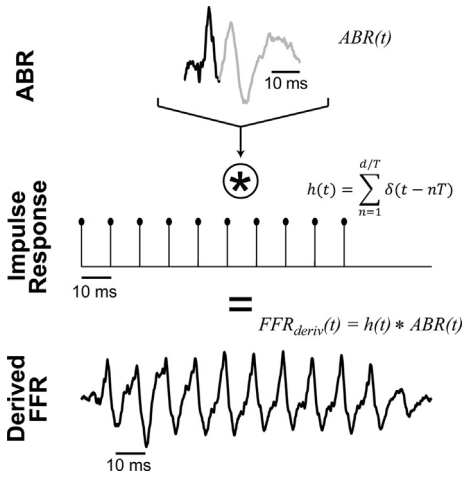

Fig. 1. Schematic of the convolution model for generating the FFR based on repeated transient ABR responses. Theoretically, FFRs generated by a complex tone with period *T* are conceived as convolution of the ABR signal (*ABR(t)*, *top row*) with the individual pitch periods of the tone, modeled as a periodic pulse train impulse response (*h(t)*, *middle row*). The resulting derived FFR ($FFR_{deriv}(t)$, *bottom row*) is thus generated by a series of overlapping transient ABRs. Only the first 12 ms of the ABR was used to derive FFRs [black portion of *ABR(t)*]. Derived (convolved) responses appear qualitatively similar to the actual FFRs recorded from the scalp. Note that model responses contain the same periodicity as the impulse response; in this example $T = 10$ ms (i.e., $F0 = 100$ Hz).

computed as the location of peak voltage within the 8–12 ms search window—the putative latency range of the brainstem FFR (Galbraith et al., 2000). Differences in the time-course and spectrotemporal properties of derived vs. actual recordings were used to assess the validity of the convolution model of FFR generation (e.g., Dau, 2003; Gerken et al., 1975; Janssen et al., 1991). Under the linear superposition assumption that the FFR is a repeated and overlapping ABR, we expected to observe no differences in spectrotemporal features between derived and actual recordings. This would imply that similar neural elements generate both the ABR and FFR.

2.3. Results

2.3.1. Comparison of ABR-derived FFRs and actual FFR responses

A comparison of actual and derived FFR time-waveforms is shown in Fig. 2. Consistent with previous qualitative reports (Dau, 2003; Gerken et al., 1975; Janssen et al., 1991), derived FFRs appeared as sustained phase-locked neural activity lasting over the duration of the stimulus. Predicted ABRs mimicked both the periodicity and absolute amplitude seen in actual recordings. However, gross waveform morphology diverged at higher F0 frequencies. Most notably, weaker phase-locking was observed in actual recordings for the 500 Hz F0, consistent with the sharp roll-off of phase-locking in the IC (Liu et al., 2006). In contrast, ABR model-derived FFRs showed abnormally strong responses at higher frequencies.

Differences also emerged in estimates of response onset latency (Fig. 2C). A two-way, fully crossed ANOVA was conducted with factors F0 stimulus (3 levels: 100, 200, 500 Hz) and FFR type (2 levels: actual vs. derived) to assess differences in response latency. This analysis revealed a significant stimulus F0 × response type interaction [$F_{2, 30} = 6.22$, $p = 0.0055$]. Bonferroni corrected post-hoc contrasts revealed this interaction was largely due to derived FFR onset latencies being significantly delayed from that of actual FFRs, particularly in the 500 Hz condition.

2.3.2. Spectral discrepancies between ABR-derived and actual FFR responses

Fig. 3 illustrates the spectra of derived vs. actual FFRs. As might be expected given the complex pitch evoking stimuli, responses showed consistent spectral energy at the fundamental frequencies (100, 200, 500 Hz) and their integer related harmonics. Notably however, derived FFRs contained more robust high frequency energy (>1000 Hz) than actual recordings. In the 500 Hz condition, predicted responses contained energy at ~1500 Hz, which is well beyond the capability of neural phase-locking in the brainstem (Liu et al., 2006). In contrast, true FFR recordings showed that brainstem activity decayed precipitously above ~500 Hz with little energy observed above 1000 Hz, consistent with actual neurobiological limits. Interestingly, residual waves are observed in the actual 500 Hz FFRs (peaks at ~10, 17, and 27 ms) which likely reflects the remaining presence of the onset response (ABR + MLR).

Quantitative analyses corroborated these qualitative descriptions. At lower F0 stimulus frequencies, spectral magnitudes measured at the response F0 were stronger for actual compared to ABR-derived FFRs (insets, Fig. 3B). In contrast, higher F0 stimuli (e.g., 500 Hz) showed that derived responses over predicted both the response fundamental and high frequency (1000 Hz) spectral magnitudes. Together, these findings suggest that the convolution model of FFR generation (Davis and Hirsh, 1974; Goldstein and Kiang, 1958; Janssen et al., 1991) *underestimates* FFR response amplitudes at lower but *overestimates* them at higher stimulus F0 frequencies. The presence of strong energy in derived FFRs above 1000 Hz further suggests the model does not conform to biological constraints (e.g., limits of phase-locking). Together, discrepancies in both latency measures and frequency-specific amplitudes suggests that the human FFR is not merely a series of overlapping (i.e., composite) onset ABR responses (cf. Daly et al., 1976; Gerken et al., 1975; Janssen et al., 1991; Moushegian et al., 1973). Our findings also cast doubt on the assertion that the same neural mechanisms are responsible for the generation of ABR wave V and the FFR (Daly et al., 1976, p. 140; Dau, 2003, p. 943). Rather, results are consistent with the notion that the sustained phase-locked human FFR and ABR (onset response) are functionally distinct auditory brain responses. This notion is further bolstered by the fact that in the high-frequency 500 Hz condition, sustained phase-locking at the stimulus F0 was weak (i.e., little to no FFR) but prominent onset components remained in the evoked activity (i.e., present onset ABR).

3. Experiment 2: Scalp topography and source generators of the human speech-evoked FFR

3.1. Rationale

Fundamental descriptions of the FFR's topography and neuronal sources have been limited in scope as previous studies have characterized the response using only single channel electrode montages. To our knowledge, no study has probed the neural generators nor the scalp topography of the FFR using *non-invasive* techniques in neurologically normal individuals. The primary aim of Experiment 2 was to more fully characterize the scalp distribution and underlying neural generators of the human speech-evoked FFR. To this end, following responses were recorded using a high density multichannel electrode array. Source dipole analysis, 3-channel Lissajous voltage trajectories (Pratt et al., 1984, 1987), and topographic maps provided new, non-invasive descriptions of the FFR's voltage distribution across the scalp as well as its putative neural locus and source characteristics.

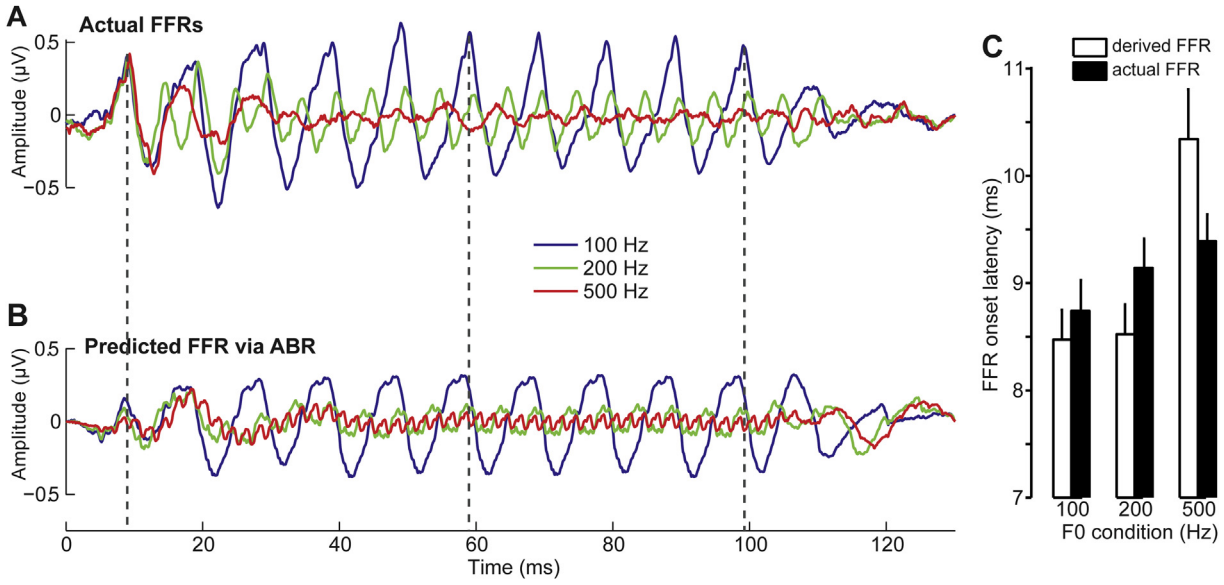

Fig. 2. Comparison between actual FFR recordings and those predicted by convolution of repeated ABRs. (A) Scalp-recorded FFR time–waveforms recorded in response to 100 ms complex tones with various fundamental (F0) frequencies. (B) Predicted FFRs generated via the ABR convolution model shown in Fig. 1. Dotted vertical lines demarcate the onset and select periods of the response to aid comparison between time courses. (C) Onset latency of derived vs. actual FFRs. Systematic differences are observed between actual and derived FFRs in onset latency and amplitude. Discrepancies between actual and derived FFRs are most apparent at higher F0s (cf. 500 Hz condition).

3.2. Methods

3.2.1. Subjects

An additional eight listeners, who did not participate in Exp. 1, were recruited for Exp. 2 (1 male, 7 female; age: 25.2 ± 3.0 years). All were monolingual English speakers, right-handed, had minimal musical training (1.75 ± 2.0 years), and exhibited normal hearing and neuropsychiatric function. Participants were paid and gave written informed consent in compliance with a protocol approved by the University of Memphis IRB.

3.2.2. Speech stimulus

FFRs were elicited by a 300 ms (10 ms rise/fall time)/vCv/speech token/ama/(cf. Shannon et al., 1999). The stimulus was a natural

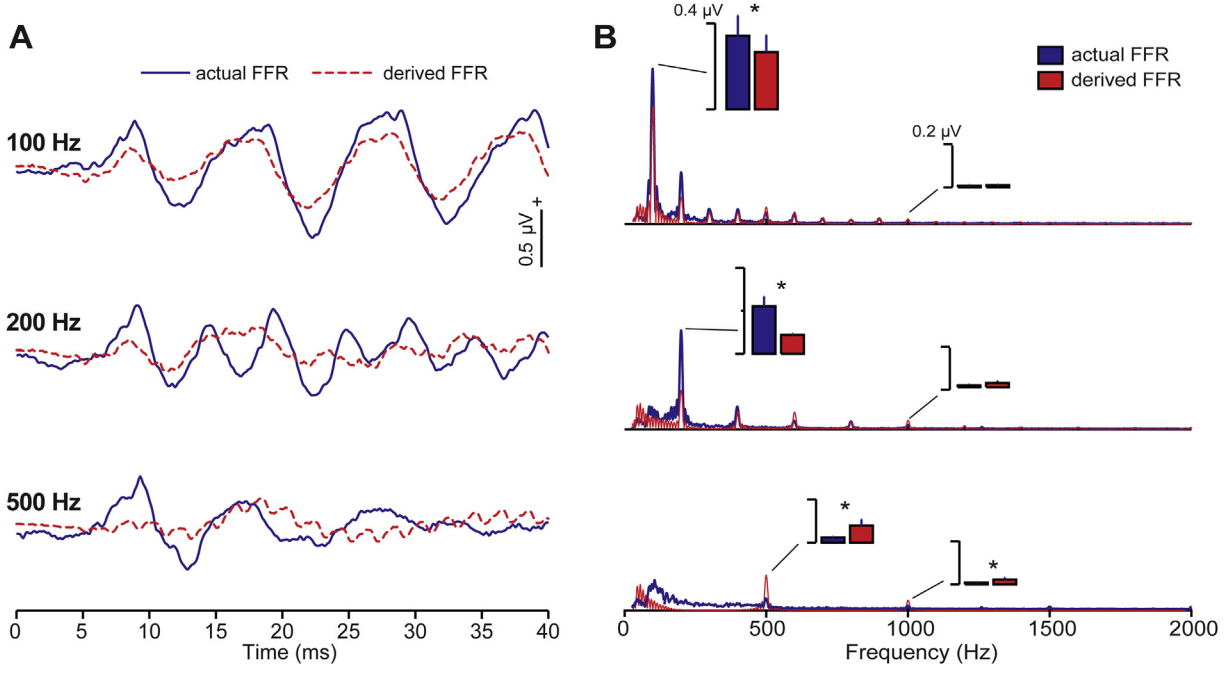

Fig. 3. True FFRs differ from derived responses in spectrotemporal characteristics. (A) Overlay of actual and derived FFR responses for various stimulus F0 conditions (enlarged waveforms as in Fig. 2). Derived responses capture the overall periodicity of the FFR but show systematic differences in amplitude, particularly at higher F0s. (B) Comparison of spectral characteristics between actual and derived responses. True FFRs only contain spectral energy up to ~1000 Hz, the upper-limit of midbrain phase-locking. In contrast, derived FFRs contain biologically implausible spectral energy well above phase-locking limits (e.g., peaks at ~2000 Hz). Insets show spectral amplitudes quantified at low- (F0) and high- (1000 Hz) frequencies. For lower F0 conditions, predicted responses underestimate neural encoding of stimulus envelope (i.e., F0) and overestimate amplitudes at higher F0. $^{*}p < 0.05$.

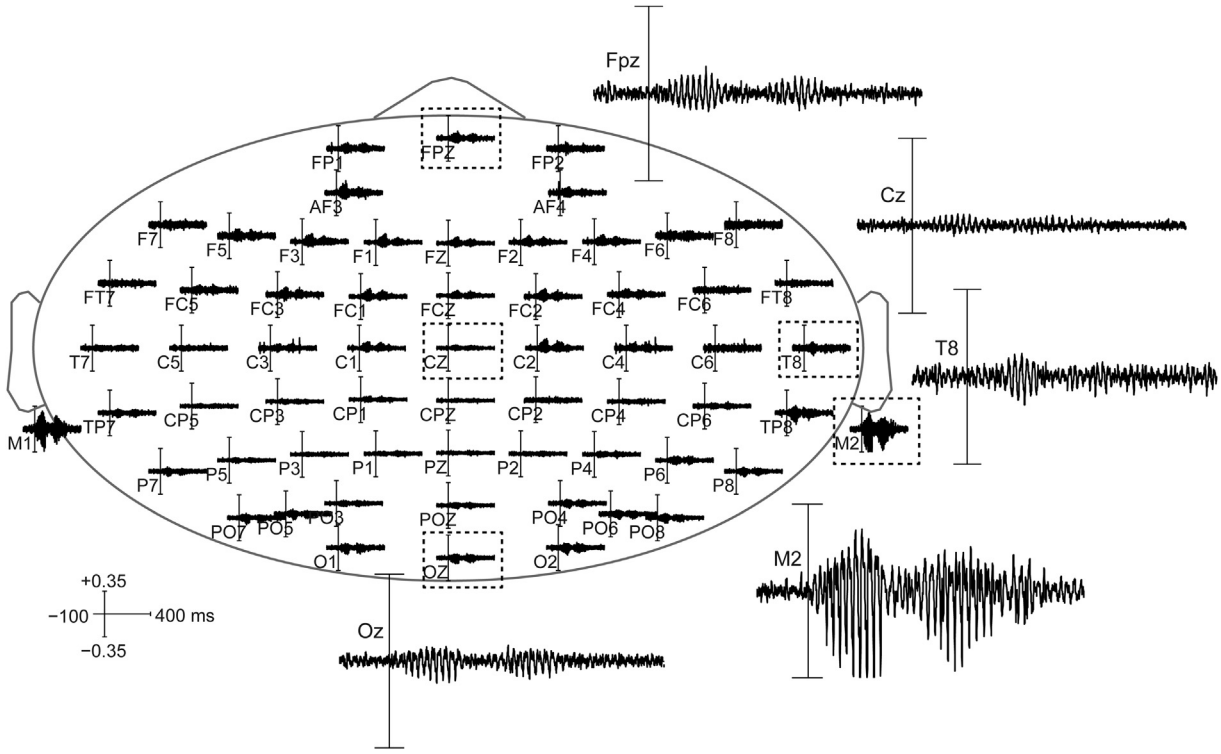

Fig. 4. Multichannel characterization of the human speech-evoked FFR. Neuroelectric responses were recorded at 64 sensors positioned around the scalp using an average reference. FFRs were elicited by the 300 ms speech token/ama/. Select channels (dotted boxes) in frontro-central positions at the midline (Fpz, Cz), temporal (T8, M2), and occipital (Oz) sites are expanded to aid visualization. FFRs are prominent at fronto-central-scalp locations (Fpz). Note also the exaggerated FFR response recorded at the mastoid (M2) relative to other cephalic scalp locations.

production recorded by a male speaker. A vCv was desirable in the current study (as opposed to a CV token) to minimize onset responses in FFRs. The 50 ms nasal (/m/) was flanked by each vowel phoneme (/a/), both 125 ms in duration. The pitch prosody fell gradually over the duration of the token from an F0 of 120 Hz–88 Hz. Vowel formant frequencies (F1–F3) were 830, 1200, and 2760 Hz, respectively. The intensity of the token was relatively fixed across its time course. Listeners heard 2000 repetitions of the speech token (ISI = 50 ms) presented with fixed (i.e., rarefaction onset) polarity and delivered binaurally through ER-30 insert earphones at 80 dB SPL.

3.2.3. Multichannel speech-FFR recording protocol

Multichannel speech-evoked FFRs were recorded from 64 sintered Ag/AgCl electrodes at standard 10–20 locations around the scalp (Oostenveld and Praamstra, 2001). EEGs were digitized using a sampling rate of 5000 Hz (SynAmps RT amplifiers; Compumedics Neuroscan) using an online passband of DC − 3000 Hz. Responses were then stored to disk for offline analysis. Electrodes placed on the outer canthi of the eyes and the superior and inferior orbit were used to monitor ocular activity. During online acquisition, all electrodes were referenced to an additional sensor placed ~1 cm posterior to Cz. For analysis, multichannel data were then re-referenced off-line to a CAR (mean of all electrode locations). Contact impedances were maintained below 5 kΩ throughout the duration of the recording session.

Subsequent data preprocessing was performed in Curry 7 (Compumedics Neuroscan) and custom routines coded in MATLAB. Data visualization and scalp topographies were computed using EEG/ERPLAB (Delorme and Makeig, 2004) (http://www.erpinfo.org/erplab). Ocular artifacts (saccades and blink artifacts) were corrected in the continuous EEG using a principal component analysis (PCA) (Wallstrom et al., 2004). Cleaned EEGs were then epoched (−50 to 350 ms), baseline-corrected to the pre-stimulus period, digitally filtered (50–3000 Hz; zero-phase filters), and subsequently averaged in the time domain to obtain FFR waveforms at each scalp electrode per participant.

3.2.4. FFR source dipole analysis

The location, strength, and orientation of the most likely intracerebral generators underlying the human speech-evoked FFR were estimated using source dipole analysis (e.g., Picton et al., 1999; Scherg et al., 1989). Localization was first performed on the group averaged scalp-recorded waveforms (CAR referenced potentials) and used to guide individualized fits. We used a realistic, boundary element model (BEM) volume conductor (Fuchs et al., 1998, 2002) standardized to the MNI brain (Mazziotta et al., 1995). A BEM head model is preferable to more routinely used spherical head shell models for source reconstruction as it is prone to less spatial error, particularly for the deeper sources of interest here (Fuchs et al., 2002). Two symmetrically mirrored, rotating dipoles were fit in each hemisphere using initial random seed points separated by 59 mm. To constrain the inverse solution, the minimum distance between the two dipoles was set at 11 mm, consistent with the anatomical distance between the left and right “hemispheres” of the inferior colliculus (Guimaraes et al., 1998). Source modeling studies of the human ASSR to amplitude modulated tones have demonstrated that for high modulation rates (>80 Hz, similar to the stimulus F0 used here), the dominant contribution is from a midbrain source (Herdman et al., 2002).[2] Importantly, the

[2] Additionally, one would not expect a cortical source for FFRs given (i) cortex cannot phase-lock to the high frequencies of our stimuli (>100 Hz) and (ii) the highpass filter setting used in our study largely eradicated low-frequency cortically generated activity which is restricted in bandwidth to <100 Hz (Bidelman et al., 2013).

persistence of sustained, phase-locked potentials with bilateral aspiration of auditory cortex helps rule out the possibility of cortical generator(s) for FFRs recorded in the current study (Kiren et al., 1994). Although, it should be noted that some thalamocortical axons can synchronize to the lowest F0 used here (100 Hz) (Joris et al., 2004).

Unlike the slow wave cortical ERPs which have well-defined component latencies to fit source models, dipoles are more difficult to estimate from the periodicity of the FFR. The response contains many closely spaced peaks, reflecting neural phase-locked activity to each pitch period of the evoking stimulus (see Fig. 2). In our dipole model, it was necessary to assume that each of these peaks was generated by the same underlying neural structure. To this end, a dipole model was fit in an iterative manner at each prominent peak of the FFR waveform (see Fig. 5); source location estimates computed at each peak were then averaged to obtain a single foci describing the sustained, periodic FFR. The fit solution accounted for ≥80% of the variance in voltage distribution recorded at the scalp, consistent with criterion of previous source modeling reports (Alain et al., 2009; Butler and Trainor, 2012). In the current study, dipole fits were used to describe *qualitatively* and non-invasively, the location and orientation of the most likely neuronal sources underlying the human brainstem response to speech. Dipole locations, reflecting the mean "center of gravity" of neural activity, were visualized by projecting the stereotaxic coordinates of each solution onto the standardized MNI brain (Mazziotta et al., 1995). Intersubject variability was quantified based on confidence ellipsoids of the *x, y*, and *z* coordinates of dipole locations across subjects.

3.2.5. Three-channel Lissajous voltage space trajectories

To better characterize the orientation of dipole source generator activity, a 3-channel Lissajous trajectory (3CLT) representation was derived from brainstem FFR recordings (e.g., Pratt et al., 1984, 1987). A 3CLT is a trajectory plot in which each data point represents the simultaneous voltage of three orthogonal electrode configurations (Pratt et al., 1984). By constraining the channels to be orthogonal, the 3CLT can be used to visualize, in a three-dimensional voltage space, the time-varying changes in orientation of the centrally located equivalent dipole(s) vector in a single voltage-voltage-voltage (i.e., *x-y-z* coordinate) plot (Pratt et al., 1987). The trajectory of the 3CLT can be used to infer the directional orientation and number of generators underlying the scalp-recorded potential.

In the current study, three orthogonal channel recordings were derived from averaged-referenced multichannel FFRs by referencing the appropriate data: M1-M2 ('X'), Fpz − Oz ('Y'), Cz − (CB1 + CB2)/2 ('Z'). An angle correction was applied to the Z-channel derivation to obtain orthogonality of the 3 channels (Pratt et al., 1984). Plotting the X, Y, and Z channels against one another provides an isomorphic 3D voltage description of the FFR in an anatomical space. To clarify the trajectory's visualization, only time points corresponding to prominent peaks in the FFR response were plotted in the 3CLT (see Fig. 5A). Comparing the 3CLT representation to the sources derived via dipole fitting allowed us to further validate the FFR's generators using converging evidence from multiple source-estimation techniques.

3.3. Results

3.3.1. Scalp topography of the human FFR

Figs. 4 and 5 show the topography and scalp distribution of the human speech-evoked FFR, respectively (see also, real-time movie visualization in Supplemental Material). Select channels are provided as enlarged insets in Fig. 4 to ease comparison (Fpz, Cz, T8, M2, Oz). All neuroelectric responses are shown with reference to a common average and thus, reflect the most unbiased view of the FFR response (Dien, 1998; Nunez and Srinivasan, 2006). FFRs show two primary bursts of energy, corresponding to phase-locked activity to each sustained vowel portion (/a/) of the eliciting /vCv/ speech token. Maximal FFRs were recorded from frontal-midline scalp locations (Fpz); responses decay in amplitude moving posterior toward the vertex (Cz). Comparing traces between anterosuperior and posteroinferior channel locations reveals a polarity reversal in the FFR response (cf. Fpz to Oz). These observations, at

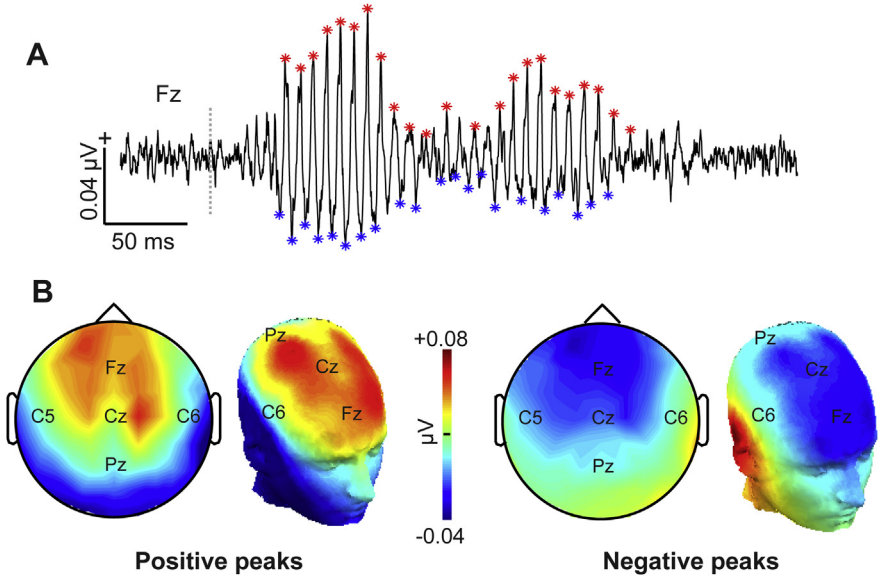

Fig. 5. Scalp topography of the human speech-evoked brainstem FFR. (A) FFR response recorded at the front of the head (Fz). The dotted line marks the onset of the time-locking stimulus. Scalp-maps were computed at time points corresponding to the prominent peak and trough amplitudes of the phase-locked response. Maps at each period of the response (starred peaks) were averaged to obtain scalp distributions for the positive and negative phases of the response (see Methods). (B) 2D and 3D scalp topographies for the peak positivities (left) and negativities (right) of the FFR potential. Maximal amplitude is observed over frontal sites on the scalp (e.g., Fz), consistent with neural generators in the IC pointed obliquely in an anterior orientation to the vertex (see Fig. 8).

least qualitatively, are consistent with deep neural generators situated near the center of the head (e.g., brainstem) and oriented obliquely in an anterior orientation relative to the vertex.

3.3.2. Effects of reference electrode choice on FFR morphology and spectral properties

Apparent from Fig. 4 are the exaggerated FFRs recorded at both mastoids (M1, M2). Larger phase-locked activity at M1/M2 relative to cephalic channels may reflect differences in the neural generator being optimally recorded at various sites around the head. FFRs are traditionally recorded using a differential electrode montage, positioned in either a vertical (Fpz − M1/M2) or horizontal orientation (M1 − M2) in order to align with more central (midbrain) or peripheral (auditory nerve) brainstem dipoles, respectively (Galbraith et al., 2001, 2000). Fig. 6 shows comparisons between FFR response spectra recorded at the left mastoid (M1) vs. frontal electrode (Fpz). Using and average reference, multichannel recordings reveal that FFRs recorded at the mastoids contain much higher frequency phase-locked activity than those at the vertex. At the mastoids, responses contain energy up to about 1100 Hz. In contrast, responses at Fpz show dominant energy at the speech envelope (F0) but little energy elsewhere across the spectrum. Higher frequency FFRs at the mastoid relative to the high-forehead site is consistent with the notion that M1/M2 pick-up activity recorded from more proximal and peripheral auditory relays, e.g., cochlear microphonic or auditory nerve (Chimento and Schreiner, 1990). The fact that the Fpz FFR contains predominantly F0 energy could reflect the fact that frontocentral channels (i) tap a vertically oriented brainstem dipole which has lower phase-locking capacity than auditory nerve and (ii) they are spatially much farther from the dipole generator; potentials volume conducted from brainstem nuclei are heavily lowpass filtered when recorded at distal scalp locations.

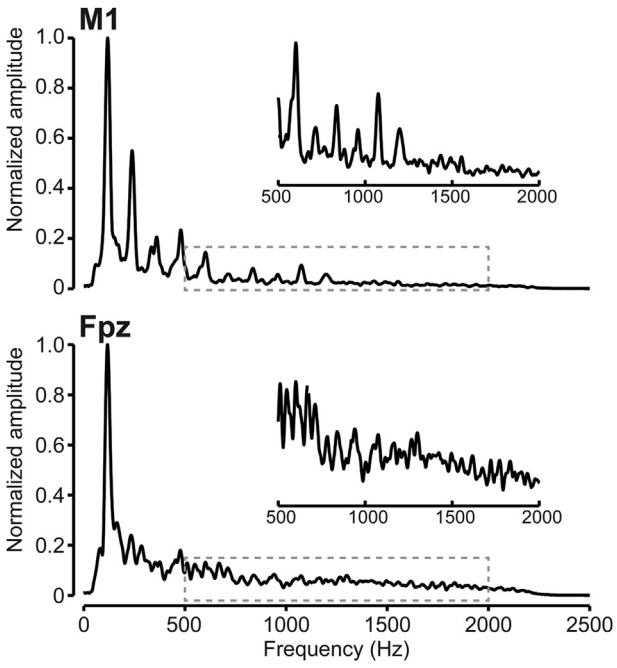

Fig. 6. Effects of scalp location on spectral properties of the speech-evoked FFR. FFTs computed from FFRs recorded at the left mastoid (M1, *top row*) and high forehead (Fpz, *bottom row*). Both responses reflect neural activity recorded at the individual sensors referenced to a common average (CAR). Spectral amplitudes are normalized between 0 and 1. Response spectra at the mastoid channel show robust phase-locked energy at the fundamental frequency (~100 Hz) and its integer related harmonics up to ~1100 Hz. FFRs recorded at the forehead channel show response energy almost entirely at the stimulus F0; higher harmonics in the response are weak or absent (compare insets). The lowpass response at Fpz is consistent with activity produced by a distal source with a more restricted upper phase-locking limit (e.g., deep midbrain dipole) whereas high-frequencies recorded at M1 reflect a more proximal dipole with higher-phase locking capability (e.g., auditory nerve).

Alternatively, in addition to differences in geometry (e.g., distance from sources), the weaker high-frequency activity observed in CAR recordings might reflect differences in the timing of responses across channels (e.g., Bharadwaj and Shinn-Cunningham, 2014). Phase disparities in the mixture of FFRs from various generators would tend to mix in the response recorded at the scalp. Phase-effects could also be exacerbated by the CAR referencing process, which averages all responses around the scalp to derive a reference channel without consideration of phase. Given that phase effects are generally more prominent at high compared to low frequencies, CAR responses would tend to show weaker high-frequency activity. Indeed, this is exactly what was observed when comparing between reference montages, i.e., CAR vs. mastoid-referenced recordings (Fig. 7). However, high-frequency activity remains prominent near mastoid compared to frontal (Fpz) channels even when using the same CAR reference (Fig. 6). This suggests that spectral differences between channels do not reflect phase-disparities *per se*. Rather, our data indicate that differential spectral effects reflect the spatial orientation and proximity of the electrodes relative to peripheral and brainstem sources.

Fig. 7 depicts FFRs recorded using an identical non-inverting electrode (Fz) but different references: (i) Fz-M1; (ii) Fz − CAR. FFRs recorded using a mastoid reference (typical recording approach) show a three-to four-fold increase in response amplitude at the F0 stimulus frequency (~100 Hz) relative to those obtained using a CAR. Additionally, mastoid referenced FFRs show phase-locked responses to the speech harmonics up to ~750 Hz. With use of a CAR, only much weaker energy at the voice fundamental is observed. These findings indicate that the choice of reference plays a critical role in FFR recordings and their interpretation, particularly with regard to the *absolute* amplitude and spectral content (e.g., bandwidth) of the neural response.

3.3.3. Source dipoles underlying the scalp-recorded FFR

Source dipole locations are shown in Fig. 8 visualized on the MNI standardized anatomy. Source analysis revealed neural generators in the midbrain (upper brainstem) with a mirrored pair of dipoles corresponding to left and right "hemispheres" of the IC. The orientation of each dipole was oblique (Fig. 8B), directed parallel to the brainstem and consistent with the frontocentral maximum observed in the scalp-recorded, average-referenced potentials (see Fig. 5B). To assess the reliability and intrasubject variability of this optimal dipole fit, a cross-validation was performed on the source solution. Dipole source locations were re-computed in an iterative fashion while omitting a single electrode on each pass (leave-one-out; 64 resamples). Consistent midbrain foci across the cross-validated samples indicated good stability (i.e., high likelihood) of a brainstem generator for the FFR (Fig. 8C–D).[3] While dipole models must be interpreted qualitatively and require parameter constraints, an upper brainstem generator estimated here using this noninvasive source analysis is consistent with prior lesion data

[3] Two cross-validated fits appear more lateral to the brainstem than the cluster obtained from the other data surrogates. These outliers are resamples in which the mastoid channels were removed in the leave-one-out resampling procedure. Mastoid electrodes contain more substantial phase-locked responses from peripheral auditory structures, e.g., auditory nerve (see Fig. 7), and thus explain the larger influence on these two dipole solutions.

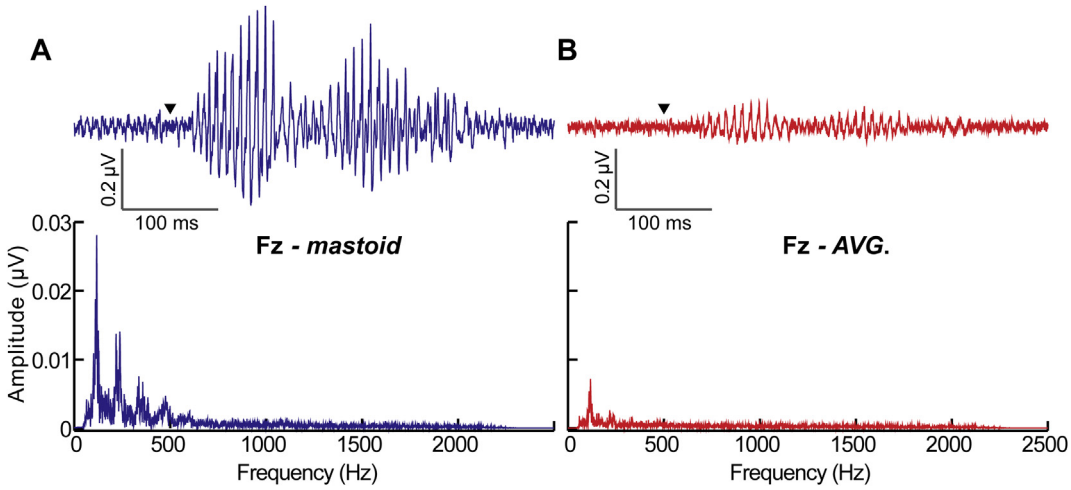

Fig. 7. Effects of reference choice on FFR recordings. FFRs were recorded using the same vertical montage with an electrode positioned on the high forehead (Fz) referenced to either (A) the averaged mastoids ([M1 + M2]/2) or (B) a common average reference. Triangles mark the onset of the time-locking speech stimulus. Top row, time-waveforms; bottom row, response spectra. Despite having the same non-inverting electrode, the Fz-M1/M2 montage commonly used in FFR recordings shows a nearly three-fold increase in amplitude and much higher frequencies relative to the Fz-avg. recording, indicating pick-up of microphonics of the cochlea and/or auditory nerve. In contrast, average referenced FFRs reflect neural responses from more central brainstem generators.

in humans which suggest the IC as the FFR's primary origin (e.g., Sohmer et al., 1977). Our results are also consistent with dipole analyses of the ASSR evoked by modulated tones, which for high modulation rates (comparable to those used here), are generated by brainstem sources (Herdman et al., 2002; Kiren et al., 1994).

A 3-channel Lissajous voltage trajectory of the speech-evoked FFR is shown in Fig. 8D. The 3CLT provides a three-dimensional representation of the FFR's dipole activation in an anatomical voltage space, and thus, a means to further confirm our source dipole analysis of surface recorded activity (Pratt et al., 1987). Peaks and troughs of the FFR produce a strong, obliquely oriented response in the 3D voltage space, corroborating the frontocentral orientation of the estimated dipole point source. Little variation in source activity is seen along the *x*-*y* dimension (i.e., parallel to the mastoids). This suggests that the most active contribution to the aggregate scalp-recorded FFR in humans is a vertically oriented

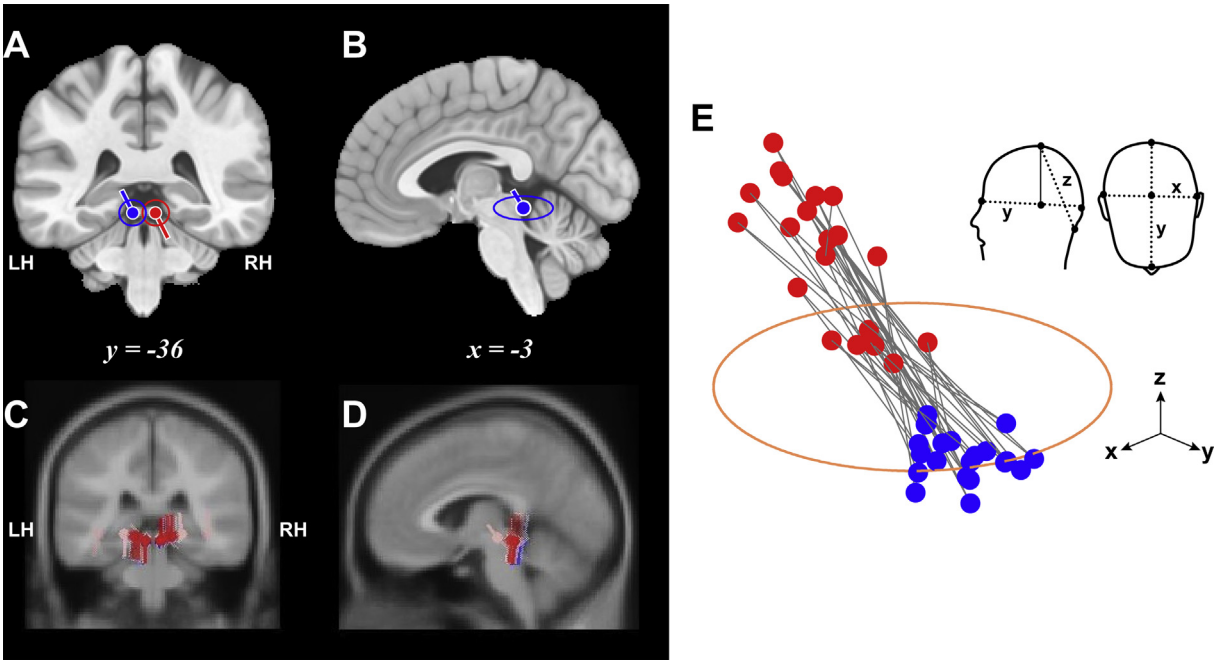

Fig. 8. Source dipole locations and Lissajous voltage space trajectories underlying the human FFR reveal neural generators in the rostral brainstem. A single mirrored dipole pair was fit to the scalp-recorded averaged-referenced FFR data shown in Fig. 4. Confidence ellipsoids = ±1 S.D. across subjects. (A–B) Source locations are plotted in stereotaxic coordinates projected onto single slices of the standardized MNI brain in the coronal (A) and sagittal (B) planes (Mazziotta et al., 1995). The dipole solution reveals generators in the midbrain in each "hemisphere" of the inferior colliculus (IC). (C–D) Cross-validated source locations for the brainstem FFR. Consistent midbrain foci across resamples indicate good stability (i.e., high likelihood) of a brainstem generator for the FFR. (E) 3-channel Lissajous voltage trajectory constructed from orthogonal *x*, *y*, and *z* electrode channels (inset heads) demonstrating spatiotemporal changes in the source dipole's voltage in 3D space. Red and blue points denote the time points of positive and negative peaks of the FFR (see Fig. 5A). The circle is shown for planar reference, representing a ±0.2 μV voltage magnitude in the *x*-*y* plane. FFR activity shows an oblique, fronto-centrally oriented voltage gradient running parallel to the brainstem. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

generator; contributions from more caudal, horizontal sources (e.g., cochlear microphonic, auditory nerve)—while active (Fig. 4)—weigh less heavily on the overall response. The unfolding of the 3CLT and scalp topography of the FFR over time can be visualized as a movie (see Supplemental QuickTime Movie; FFR_supplemental_movie.mov).

4. Discussion

Despite clinical and empirical interest, few studies have characterized properties of the brainstem frequency-following response (FFR) in man. The present study aimed to address two primary gaps in the FFR literature: (1) determine if the sustained periodic nature of the FFR is simply a linear superposition of overlapping transient ABR wavelets (i.e., repeated onset responses); (2) non-invasively characterize the scalp topography and location of the underlying source generators of the FFR.

4.1. Dissociation between the onset ABR response and sustained FFR

Historically, sustained auditory evoked potentials have often been described as repeated onset responses, created by the linear superposition of periodically evoked transient activity (e.g., Galambos et al., 1981; Ross et al., 2002). The human FFR is not unique in this regard. Its periodicity has also been conceptualized as a repeated series of onset ABR-wave Vs (e.g., Daly et al., 1976; Dau, 2003; Davis and Hirsh, 1974; Gerken et al., 1975; Janssen et al., 1991; Moushegian et al., 1973; Picton et al., 1977). Indeed, modeling work has suggested that FFR-like responses can be derived, at least qualitatively, by simple convolution whereby the sustained response is generated by a series of composite of overlapping transient ABRs (e.g., Dau, 2003; Janssen et al., 1991). Under the stimuli used here (i.e., periodic click trains), the convolution of the ABR as an explanation for the FFR represents somewhat of a special (extreme) case but offers the best possible chance of observing convergence between the two responses. Yet, our data suggest that this "convolution model" of FFR generation is *not* an accurate description for the sustained response (cf. Davis and Hirsh, 1974; Gerken et al., 1975; Janssen et al., 1991). For all F0s tested, FFRs derived from individual's repeated ABR showed multiple dissociations in both spectral and temporal features. In all cases, derived responses either over- or under-predicted the spectral encoding of stimulus F0 and its higher harmonics. Most notably, ABR-derived FFRs showed considerable high-frequency energy (>1100 Hz) which does not conform to the biological constraints of neural phase-locking documented for brainstem IC neurons (Liu et al., 2006). Together, our results suggest that the sustained phase-locked human FFR and ABR (onset response) are functionally distinct auditory brain responses. This notion is consistent with the findings of Galbraith and Brown (1990) who, based on detailed latency analysis of both responses, concluded that successive waves of the FFR do not reflect repeated ABR wavelets (see also, Picton et al., 1981, 1977).

Nevertheless, it remains possible that at lower repetition rates (e.g., 40 Hz)—when cortical contributions are stronger—FFRs and (serial) ABRs converge. Multichannel recordings of FFRs with varying F0s would be of interest to assess the dependence of dipole source locations on stimulus F0. In the current study, filtering parameters and stimuli were optimized to minimize cortical contributions to the FFR (Bidelman et al., 2013). While cortical phase-locking is limited (<100 Hz) (Joris et al., 2004), some thalamocortical axons can synchronize to the lower stimulus F0s used here (100–200 Hz) (Joris et al., 2004; Wallace et al., 2000). Hence, our data cannot entirely rule out the possibility of cortical contributions at the lowest F0s tested. In addition, higher F0s would presumably tend to generate activity in progressively more caudal structures where phase-locking extends to higher frequencies. Moreover, our approach cannot evaluate the possible contributions of even lower brainstem nuclei to the FFR (e.g., cochlear nucleus) which presumably, would be more than able to synchronize to high-frequency features (<1–2 kHz) of our speech stimuli. Localizing such sources would be technologically challenging and would likely require full electrode coverage of the head (e.g., chin and neck) or alternative methodologies more sensitive to deep sources and insensitive to volume conduction (MEG: Parkkonen et al., 2009). Future studies are needed to address these possibilities.

Our results corroborate other studies which have implied functional distinctions between the onset (ABR) and sustained (FFR) brainstem response. Direct comparisons between the ABR and FFR have revealed that they are differentially sensitive to stimulus presentation rate (Krizman et al., 2010), frequency specificity (Picton et al., 1977, p. 105), their susceptibility to noise masking (Ananthanarayan and Durrant, 1992; Cunningham et al., 2002; Russo et al., 2004), and how their latency changes with stimulus intensity (Akhoun et al., 2008a). Moreover, speech-evoked brainstem responses may be impaired in children with language- and learning-disorders (Banai et al., 2007, 2009; Basu et al., 2010; Rocha-Muniz et al., 2012) despite normal click-ABRs (Song et al., 2006). Together with our new findings, these studies support the growing body of evidence which suggests that onset and sustained portions of the human brainstem response are generated by distinct portions of the auditory system (Worthington and Peters, 1980) and originate from separate neuroanatomical populations along the pathway (e.g., Krizman et al., 2010; Parthasarathy and Bartlett, 2012; Picton et al., 1978; Ross et al., 2002). This proposition is also consist with observations in humans, that ABR wave V is generated primarily by the lateral lemniscus (Møller and Jannetta, 1982) while the FFR arises from the colliculi (Sohmer et al., 1977).

It is conceivable that some of the observed disparities between ABR-derived and true FFRs might result from peripheral mechanisms (i.e., cochlear nonlinearities, adaptation) not accounted for by the purely linear (and rather simplistic) convolution model. However, cochlear nonlinearities are generally weaker at higher compared to lower stimulus intensities when the basilar membrane is more linearized (Oxenham and Plack, 1997; Ruggero et al., 1997). Indeed, as noted by Dau (2003), one might actually expect *higher* correspondence between modeled and actual FFRs at the higher stimulus intensities used here given the linearization of the cochlear partition (see also, Janssen et al., 1991). Additionally, ABR adaptation is generally weaker at higher stimulus intensities (Thornton and Coleman, 1975). Adapted ABRs at higher rates would also be expected to overestimate derived FFR amplitudes across the board. Yet, this is not what was observed; in most cases, ABR-derived FFR amplitudes largely underestimated those of true FFR recordings. Furthermore, it is difficult to account for the differential effects on FFR latency based on adaptation of the ABR alone (Fig. 2C). Thus, it is unlikely that cochlear explanations alone fully account for the observed disparities between ABRs and FFRs. Rather, our data suggest that these two brainstem responses reflect a functional distinction between sustained and transient auditory potentials (e.g., Galbraith and Brown, 1990; Parthasarathy et al., 2014; Picton et al., 1981, 1977). Animal studies are needed however, to clarify these multiple interpretations, further investigate functional distinctions between the ABR and FFR, and provide direct comparisons between near- and far-field brainstem recordings (Warrier et al., 2011).

4.2. Influence of electrode reference location on FFR morphology and interpretation

Our use of multichannel recordings allowed for the unique opportunity to directly compare the effects of reference electrode

location on spectrotemporal features and morphology of the following response. FFRs recorded using a CAR demonstrated that previous studies—which have employed a traditional differential recording montage (with mastoid/earlobe reference)—likely overestimate the absolute amplitude of the response as a result of mixing peripheral and central sources. In response to both pure tones and speech, studies cite FFR amplitudes in the range of ~0.4–1 μV (Chandrasekaran and Kraus, 2010; Hoormann et al., 1992). However, this previous work has exclusively used referential recording montages which, as shown here, tend to overestimate the FFR by nearly three-to four-fold (Fig. 7). Single-channel electrode configurations are certainly the most efficient means to record the FFR and there is no reason to discontinue this common methodology. However, our data do suggest that to properly estimate the amplitude of the brainstem FFR, single reference differential recordings are undesirable.

These findings have potential ramifications for recording protocols used in FFR studies. Under a differential amplification scheme, the use of any reference near the pinna may contaminate brainstem responses with more peripheral neural activity; following responses picked-up by the mastoid (potentially cochlear microphonic and/or auditory nerve activity) are broadcast into the non-inverting electrode thereby blurring brainstem and peripheral neural contributions. Mastoid referenced FFRs reflect neural activity recorded from both the brainstem (vertically oriented dipole) and more proximal auditory nerve or caudal pons (horizontally oriented dipole). While recordings containing multiple FFR sources may not necessarily be undesirable (cf. Galbraith et al., 2001), our results recommend avoiding mastoid-referenced montages when the aim is to examine mainly higher brainstem function. Alternatively, we have commonly recorded FFRs using a non-cephalic reference electrode placed on the 7th cervical vertebra (C7) (Bidelman and Krishnan, 2009; Bidelman et al., 2011). This location may offer a better reference choice than the more commonly used mastoid(s) or earlobes by avoiding the inadvertent pickup of more peripheral (or cochlear) following responses.

4.3. FFR sources are consistent with midbrain (IC) generators

Multichannel analysis of the FFR's scalp topography demonstrated that the response is maximally distributed over frontal midline scalp sites. This suggests that the most optimal FFR is recorded using an Fz or Fpz electrode, consistent with current practice (e.g., Krishnan, 2007; Skoe and Kraus, 2010). Interestingly, topographies, while relatively bilaterally symmetric, showed two separate focal points of maximal response that flanked either side of the vertex (Fig. 5). This implies two active vertical generators. Indeed, dipole source modelling and 3CLT analyses demonstrated that these FFR scalp potentials could be adequately described with two mirror dipoles deep within the midbrain IC, with oblique orientations, whose excitability is directed anterior to the vertex (Fig. 8; Supplemental Material: FFR_supplemental_movie.mov).

Previous work has speculated brainstem origins for the FFR. In human recordings, this assumption has been inferred based on the latency of the response (~6–8 ms), which is consistent with the propagation delay to the IC (Galbraith et al., 2000; Moushegian et al., 1973). Selective lesions to the colliculi in human abolishes the scalp-recorded FFR (Sohmer et al., 1977) as do ablations (Davis and Britt, 1984) or reversible cryogenic cooling of the nucleus in cat (Smith et al., 1975). These source foci, obtained from invasive lesions converge with our non-invasive source analysis of the FFR which also implicate upper brainstem generator(s). While multiple following potentials may contribute to the vertex FFR (e.g., caudal brainstem; Fig. 4), our 3-channel Lissajous voltage trajectory analyses further suggests that the dominant component underlying the human FFR are vertical sources whose activity produces voltage gradients that run parallel to the brainstem (Fig. 8D; Supplemental Material). These findings are broadly consistent with the notion that FFR sources in human are circumscribed to the midbrain with probable generators in bilateral IC (e.g., Sohmer et al., 1977).

It should be noted that the EEG inverse problem is inherently ill-posed (i.e., there is no unique solution). Given the depth of brainstem generators, it is not possible to localize source(s) of the FFR (or any ERP for that matter) without some additional a priori constraints (e.g., volume conductor model). Our use of a BEM to model the inverse solution is preferable to more simplified spherical models and improves the localization for deeper sources. However, exact biophysical parameters of these models (e.g., conductivity parameters) are often estimated or unknown (Nunez and Srinivasan, 2006). The current study also employed average anatomy (MNI template) to model source locations. It would be useful in future studies to examine neuroanatomical foci of the FFR using individual anatomy (MRIs) co-registered with functional EEG to further validate the current source modeling approach. Nevertheless, insomuch as the assumptions of our dipole modeling accurately reflects biophysical characteristics of the human brain, our source modeling results demonstrate that probable dipoles located in the bilateral IC can account for the observed scalp topography of the FFRs.

5. Conclusions

The current study provides the first characterization of the human brainstem FFR using high density, multichannel electrode array. Our findings demonstrate that the FFR has functionally distinct response characteristics from the transient ABR; the former is not a simple convolution of the latter (e.g., Dau, 2003; Davis and Hirsh, 1974; Gerken et al., 1975; Janssen et al., 1991; Moushegian et al., 1973). Dipole analysis and Lissajous voltage trajectories of multichannel FFRs endorse prior lesion data suggesting probable generators located within the IC of the upper brainstem. Lastly, comparisons of mastoid and common average reference recordings suggest that FFRs reported in previous studies using referential recordings (i) overestimate the response's absolute magnitude by three-to four-fold and (ii) contain a combination of neural activity generated by both peripheral (e.g., VIII nerve) and central (e.g., IC) auditory brainstem structures. Collectively, our findings provide important insight into the functional nature and response characteristics of the human scalp-recorded frequency-following response.

Acknowledgements

The author thanks Megan Howell and Ameenuddin Syed-Khaja for their assistance in data collection. Portions of this work were supported by the Center for Research Initiatives and Strategies for the Communicatively Impaired (CRISCI; a Center of Excellence of the State of Tennessee) at the University of Memphis and grants from the American Hearing Research (AHRF) and American Academy of Audiology (AAA) Foundations awarded to G.M.B.

References

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., 2008a. 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.
Akhoun, I., Moulin, A., Jeanvoine, A., Menard, M., Buret, F., Vollaire, C., Scorretti, R., Veuillet, E., Berger-Vachon, C., Collet, L., Thai-Van, H., 2008b. Speech auditory brainstem response (speech ABR) characteristics depending on recording conditions, and hearing status: an experimental parametric study. J. Neurosci. Meth. 175, 196–205.
Alain, C., Quan, J., McDonald, K., Van Roon, P., 2009. Noise-induced increase in human auditory evoked neuromagnetic fields. Eur. J. Neurosci. 30, 132–142.
Ananthanarayan, A.K., Durrant, J.D., 1992. The frequency-following response and the onset response: evaluation of frequency specificity using a forward-masking paradigm. Ear Hear 13, 228–232.
Banai, K., Abrams, D., Kraus, N., 2007. Sensory-based learning disability: insights from brainstem processing of speech sounds. Int. J. Audiol. 46, 524–532.
Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. Reading and subcortical auditory function. Cereb. Cortex 19, 2699–2707.
Basu, M., Krishnan, A., Weber-Fox, C., 2010. Brainstem correlates of temporal auditory processing in children with specific language impairment. Dev. Sci. 13, 77–91.
Bharadwaj, H.M., Shinn-Cunningham, B.G., 2014. Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin. Neurophysiol. 125, 1878–1888.
Bidelman, G.M., 2013. The role of the auditory brainstem in processing musically-relevant pitch. Front. Psychol. 4, 1–13.
Bidelman, G.M., Krishnan, A., 2009. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. J. Neurosci. 29, 13165–13171.
Bidelman, G.M., Krishnan, A., 2010. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Res. 1355, 112–125.
Bidelman, G.M., Gandour, J.T., Krishnan, A., 2011. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J. Cogn. Neurosci. 23, 425–434.
Bidelman, G.M., Moreno, S., Alain, C., 2013. Tracing the emergence of categorical speech perception in the human auditory system. Neuroimage 79, 201–212.
Butler, B.E., Trainor, L.J., 2012. Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source Estimation. Front. Psychol. 3, 180.
Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47, 236–246.
Chimento, T.C., Schreiner, C.E., 1990. Selectively eliminating cochelar microphonic contamination from the frequency-following response. Electroencephalogr. Clin. Neurophysiol. 75, 88–96.
Cunningham, J., Nicol, T., King, C., Zecker, S.G., Kraus, N., 2002. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. Hear. Res. 169, 97–111.
Daly, D.M., Roeser, R.J., Moushegian, G., 1976. The frequency-following response in subjects with profound unilateral hearing loss. Electroencephalogr. Clin. Neurophysiol. 40, 132–142.
Dau, T., 2003. The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. J. Acoust. Soc. Am. 113, 936–950.
Davis, H., Hirsh, S.K., 1974. Interpretation of the human frequency following response. J. Acoust. Soc. Am. 56, S63.
Davis, R.L., Britt, R.H., 1984. Analysis of the frequency following response in the cat. Hear. Res. 15, 29–37.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Meth. 134, 9–21.
Dien, J., 1998. Issues in the application of the average reference: review, critiques, and recommendations. Beh. Res. Meth. Instr. Comp. 30, 34–43.
Fuchs, M., Drenckhahn, R., Wischmann, H.-A., Wagner, M., 1998. An improved boundary element method for realistic volume-conductor modeling. IEEE Trans. Biomed. Eng. 45, 980–997.
Fuchs, M., Kastner, J., Wagner, M., Hawes, S., Ebersole, J., 2002. A standardized boundary element method volume conductor model. Clin. Neurophysiol. 113, 702–712.
Galambos, R., Makeig, S., Talmachoff, P., 1981. A 40-Hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. U. S. A. 78, 2643–2647.
Galbraith, G., Bagasan, B., Sulahian, J., 2001. Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations. Percept. Mot. Ski. 92, 99–106.
Galbraith, G., 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.
Galbraith, G., Threadgill, M., Hemsley, J., Salour, K., Songdej, N., Ton, J., Cheung, L., 2000. Putative measure of peripheral and brainstem frequency-following in humans. Neurosci. Lett. 292, 123–127.
Galbraith, G.C., 1994. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. Electroencephalogr. Clin. Neurophysiol. 92, 321–330.
Galbraith, G.C., Brown, W.S., 1990. Cross-correlation and latency compensation analysis of click-evoked and frequency-following brain-stem responses in man. Electroencephalogr. Clin. Neurophysiol. 77, 295–308.
Gerken, G.M., Moushegian, G., Stillman, R.D., Rupert, A.L., 1975. Human frequency-following responses to monoaural and binaural stimuli. Electroencephalogr. Clin. Neurophysiol. 38, 379–386.
Geselowitz, D.B., 1998. The zero of potential. IEEE Eng. Med. Biol. 1, 128–132.
Glaser, E.M., Suter, C.M., Dasheiff, R., Goldberg, A., 1976. The human frequency-following response: its behavior during continuous tone and tone burst stimulation. Electroencephalogr. Clin. Neurophysiol. 40, 25–32.
Goldstein, M.H., Kiang, N.Y.S., 1958. Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. J. Acoust. Soc. Am. 30, 107–114.
Guimaraes, A.R., Melcher, J.R., Talavage, T.M., Baker, J.R., Ledden, P., Rosen, B.R., Kiang, N.Y.S., Fullerton, B.C., Weisskoff, R.M., 1998. Imaging subcortical auditory activity in humans. Hum. Brain Mapp. 6, 33–41.
Herdman, A.T., Lins, O., van Roon, P., Stapells, D.R., Scherg, M., Picton, T., 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.
Janssen, T., Steinhoff, H.J., Böhnke, F., 1991. Zum entstehungsmechanismus der frequenzfolgepotentiale. Otorhinolaryngol. Nova 1, 16–24.
Johnson, K.L., Nicol, T.G., Kraus, N., 2005. Brain stem response to speech: a biological marker of auditory processing. Ear Hear 26, 424–434.
Joris, P.X., Schreiner, C.E., Rees, A., 2004. Neural processing of amplitude-modulated sounds. Physiol. Rev. 84, 541–577.
Kiren, T., Aoyagi, M., Furuse, H., Koike, Y., 1994. An experimental study on the generator of amplitude-modulation following response. Acta Otolaryng. Suppl. (Stockh.) 511, 28–33.
Krishnan, A., 2007. Human frequency following response. In: Burkard, R.F., Don, M., Eggermont, J.J. (Eds.), Auditory Evoked Potentials: Basic Principles and Clinical Application. Lippincott Williams & Wilkins, Baltimore, pp. 313–335.
Krishnan, A., Gandour, J.T., Bidelman, G.M., 2010a. The effects of tone language experience on pitch processing in the brainstem. J. Neuroling 23, 81–95.
Krishnan, A., Bidelman, G.M., Gandour, J.T., 2010b. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. Hear. Res. 268, 60–66.
Krizman, J., Skoe, E., Kraus, N., 2010. Stimulus rate and subcortical auditory processing of speech. Audiol. Neurootol. 15, 332–342.
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.
Mazziotta, J.C., Toga, A.W., Evans, A., Lancaster, J.L., Fox, P.T., 1995. A probabilistic atlas of the human brain: theory and rational for its development. Neuroimage 2, 89–101.
McGee, T., Kraus, N., Comperatore, C., Nicol, T., 1991. Subcortical and cortical components of the MLR generating system. Brain Res. 544, 211–220.
Melcher, J.R., Kiang, N.Y.S., 1996. Generators of the brainstem auditory evoked potential in cat III: identified cell populations. Hear. Res. 93, 52–71.
Møller, A.,R., Jannetta, P.J., 1982. Evoked potentials from the inferior colliculus in man. Electroencephalogr. Clin. Neurophysiol. 53, 612–620.
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.
Nitschke, J.B., Miller, G.A., Cook, E.W., 1998. Digital filtering in EEG/ERP analysis: some technical and empirical comparisons. Behav. Res. Methods, Instrum. Comput. 30, 54–67.
Nunez, P.L., Srinivasan, R., 2006. Electrical Fields of the Brain: the Neurophysics of EEG. Oxford University Press, New York.
O'Beirne, G.A., Patuzzi, R.B., 1999. Basic properties of the sound-evoked post-auricular muscle response (PAMR). Hear. Res. 138, 115–132.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.
Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112, 713–719.
Oxenham, A.J., Plack, C.J., 1997. A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. J. Acoust. Soc. Am. 101, 3666–3675.
Parkkonen, L., Fujiki, N., Makela, J.P., 2009. Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. Hum. Brain Mapp. 30, 1772–1782.
Parthasarathy, A., Bartlett, E., 2012. Two-channel recording of auditory-evoked potentials to detect age-related deficits in temporal processing. Hear. Res. 289, 52–62.
Parthasarathy, A., Datta, J., Torres, J.A., Hopkins, C., Bartlett, E.L., 2014. Age-related changes in the relationship between auditory brainstem responses and envelope-following responses. J. Assoc. Res. Oto. 15, 649–661.
Picton, T.W., Woods, D.L., Proulx, G.B., 1978. Human auditory sustained potentials. II. Stimulus relationships. Electroencephalogr. Clin. Neurophysiol. 45, 198–210.
Picton, T.W., Stapells, D.R., Campbell, K.B., 1981. Auditory evoked potentials from the human cochlear and brainstem. J. Otolaryngol. Supp. 9, 1–41.
Picton, T.W., Woods, D.L., Baribaeu-Braun, J., Healy, T.M.G., 1977. Evoked potential audiometry. J. Otolaryngol. 6, 90–119.
Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neurootol. 4, 64–79.

Pratt, H., Har'El, Z., Golos, E., 1984. Geometrical analysis of human three-channel lissajous' trajectory of auditory brain-stem evoked potentials. Electroencephalogr. Clin. Neurophysiol. 58, 83–88.
Pratt, H., Martin, W.H., Bleich, N., Kaminer, M., Har'El, Z., 1987. Application of the three-channel lissajous trajectory of auditory brainstem-evoked potentials to the question of generators. Audiology 26, 188–196.
Rocha-Muniz, C.N., Befi-Lopes, D.M., Schochat, E., 2012. Investigation of auditory processing disorder and language impairment using the speech-evoked auditory brainstem response. Hear. Res. 294, 143–152.
Ross, B., Picton, T., Pantev, C., 2002. Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. Hear. Res. 165, 68–84.
Ruggero, M.A., Rich, N.C., Recio, A., Narayan, S.S., Robles, L., 1997. Basilar-membrane responses to tones at the base of the chinchilla cochlea. J. Acoust. Soc. Am. 101, 2151–2163.
Russo, N., Nicol, T., Musacchia, G., Kraus, N., 2004. Brainstem responses to speech syllables. Clin. Neurophysiol. 115, 2021–2030.
Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the late human auditory evoked potentials. J. Cogn. Neurosci. 1, 336–355.
Shannon, R.V., Jensvold, A., Padilla, M., Robert, M.E., Wang, X., 1999. Consonant recordings for speech testing. J. Acoust. Soc. Am. 106, L71–L74.
Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. Ear Hear 31, 302–324.
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.
Sohmer, H., Pratt, H., Kinarti, R., 1977. Sources of frequency-following responses (FFR) in man. Electroencephalogr. Clin. Neurophysiol. 42, 656–664.
Song, J.H., Banai, K., Russo, N.M., Kraus, N., 2006. On the relationship between speech- and nonspeech-evoked auditory brainstem responses. Audiol. Neurootol. 11, 233–241.
Terkildsen, K., Osterhammel, P., 1981. The influence of reference electrode position on recordings of the auditory brainstem responses. Ear Hear 2, 9–14.
Thornton, A.R.D., Coleman, M.J., 1975. The adaptation of cochlear and brainstem auditory evoked potentials in humans. Electroencephalogr. Clin. Neurophysiol. 39, 399–406.
Wallace, M.N., Rutkowski, R.G., Shackleton, T.M., Palmer, A.R., 2000. Phase-locked responses to pure tones in guinea pig auditory cortex. Neuroreport 11, 3989–3993.
Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F., Fox, N.A., 2004. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. Int. J. Psychophysiol. 53, 105–119.
Warrier, C.M., Abrams, D.A., Nicol, T.G., Kraus, N., 2011. Inferior colliculus contributions to phase encoding of stop consonants in an animal model. Hear. Res. 282, 108–118.
Weiss, M.W., Bidelman, G.M., 2015. Listening to the brainstem: musicianship enhances intelligibility of subcortical representations for speech. J. Neurosci. 35, 1687–1691.
Wolpaw, J.R., Woods, C.C., 1982. Scalp distribution of human auditory evoked potentials. I. Evaluation of reference electrode sites. Electroencephalogr. Clin. Neurophysiol. 54, 15–24.
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.
Worthington, D.W., Peters, J.F., 1980. Quantifiable hearing and No ABR: paradox or error? Ear Hear 1, 281–285.