INTERNATIONAL JOURNAL OF AUDIOLOGY
2018, VOL. 57, NO. 9, 665–672
https://doi.org/10.1080/14992027.2018.1470338

The British Society of Audiology
The International Society of Audiology
NAS NORDIC AUDIOLOGICAL SOCIETY
Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Response properties of the human frequency-following response (FFR) to speech and non-speech sounds: level dependence, adaptation and phase-locking limits

Gavin Bidelman[a,b,c] and Louise Powers[b]

[a]Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA; [b]School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA; [c]Department of Anatomy and Neurobiology, University of Tennessee Health Sciences Center, Memphis, TN, USA

ABSTRACT

Objective: The frequency-following response (FFR) is a neurophonic potential used to assess auditory neural encoding at subcortical stages. Despite the FFR's empirical and clinical utility, basic response properties of this evoked potential remain undefined.

Design: We measured FFRs to speech and nonspeech (pure tone, chirp sweeps) stimuli to quantify three key properties of this potential: level-dependence (I/O functions), adaptation and the upper limit of neural phase-locking.

Study sample: $n = 13$ normal-hearing listeners.

Results: I/O functions showed FFR amplitude increased with increasing stimulus presentation level between 25 and 80 dB SPL; FFR growth was steeper for tones than speech when measured at the same frequency. FFR latency decreased 4–5 ms with decreasing presentation level from 25 and 80 dB SPL but responses were ~2 ms earlier for speech than tones. FFR amplitudes showed a 50% reduction over 6 min of recording with the strongest adaptation in the first 60 s (250 trials). Estimates of neural synchronisation revealed FFRs contained measurable phase-locking up to ~1200–1300 Hz, slightly higher than the single neuron limit reported in animal models.

Conclusions: Findings detail fundamental response properties that will be important for using FFRs in clinical and empirical applications.

ARTICLE HISTORY
Received 15 January 2018
Revised 7 April 2018
Accepted 15 April 2018

KEYWORDS
Auditory brainstem response (ABR); auditory evoked potentials; electrophysiology; EEG

Introduction

The frequency-following response (FFR) is a sustained, scalp-recorded "neurophonic" potential reflecting phase-locked population activity from the human auditory system (for reviews, see Krishnan 2007; Chandrasekaran and Kraus 2010; Kraus et al. 2017). Given their remarkable spectrotemporal detail, FFRs have provided important insight into auditory processing including individual differences in speech perception (e.g. Song et al. 2011; Bidelman and Alain 2015), the neuroplasticity of learning and language experiences (e.g. Kraus and Chandrasekaran 2010; Krishnan, Gandour, and Bidelman 2012), the neurobiology of music (Bidelman, Krishnan, and Gandour 2011), and abnormal sound encoding in clinical populations (e.g. Chandrasekaran et al. 2009; Bidelman et al. 2017). Despite an abundance of FFR studies and potential clinical utility, surprisingly little is known about its basic response characteristics. This contrasts the wealth of studies on the related auditory brainstem response (ABR), whose properties (e.g. level-, rate-, frequency-effects) have been well-documented for several decades (Picton et al. 1976; Hall 1992).

Despite limited literature, several reports have documented operating characteristics of the FFR including its spectral content (Hoormann et al. 1992; Tichko and Skoe 2017), test–retest reliability (Hoormann et al. 1992; Song, Nicol, and Kraus 2011; Hornickel, Knowles, and Kraus 2012; Bidelman et al. 2018), rate effects (Krizman, Skoe, and Kraus 2010), binaural influences (Krishnan and McDaniel 1998), gender differences (Krizman, Skoe, and Kraus 2012), laterality (Hornickel, Skoe, and Kraus 2009; Krishnan et al. 2011), neural generators (Bidelman 2018b, 2015a; Coffey et al. 2016; King, Hopkins, and Plack 2016) and optimal stimulus-evoking paradigms (Bidelman 2015b, 2018a). Nevertheless, several key response properties remain to be fully documented. Here, we quantified three key properties of auditory physiological coding reflected in FFRs: level-dependence, neural adaptation and the upper limit of neural phase-locking. Each of these properties has important ramifications for understanding not only normal variation in FFRs but also the operating range where responses are optimally recordable in clinical and empirical applications.

Several FFR studies have described its level-dependence (Yamada et al. 1978; Krishnan and Parkinson 2000) but a detailed mapping of the FFR input–output (I/O) function, particularly for speech stimuli, has not been undertaken. To this end, the first aim of the present study was to provide a more comprehensive mapping of the FFR's I/O characteristics (level-dependence) for speech and non-speech stimuli. Second, we aimed to provide a new characterisation of the potential's upper frequency limit. FFR amplitude decreases with increasing frequency (Glaser et al. 1976; Gardi, Merzenich, and McKean 1979; Hoormann et al. 1992; Krishnan and Parkinson 2000; Tichko and Skoe 2017), a trend assumed to reflect the roll off of neural phase-locking in the ascending auditory system (Joris, Schreiner, and Rees 2004). Although multiple sources contribute to FFR generation (Gardi, Merzenich, and McKean 1979; Bidelman

CONTACT Gavin Bidelman g.bidelman@memphis.edu, gmbdlman@memphis.edu School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA. http://www.memphis.edu/acnl/

2015b; Coffey et al. 2016; Bidelman 2018b) the brainstem inferior colliculus is recognised as its dominant origin (Smith et al. 1975; Bidelman 2015a, 2018b). Consequently, the frequency ceiling of the FFR observed in scalp-recordings is assumed to coincide with the upper limit of phase-locking in the inferior colliculus (Chandrasekaran and Kraus 2010) – estimated to be ~1000 Hz based on animal data (Liu, Palmer, and Wallace 2006). Unfortunately, we are not aware of any studies examining the phase-locking limits of *human* auditory brainstem. To this end, we used FFRs elicited by chirp (pitch sweep) stimuli to provide a new, non-invasive characterisation of the upper frequency limit of brainstem phase-locking in humans. We hypothesised that population activity (FFRs) may support a higher phase-locking limit than available in single neurons, reminiscent of the "volley theory" of auditory coding, which states that neural ensemble activity (i.e. FFRs) might be able to phase lock at higher stimulus rates than single units if activity across frequency channels is optimally combined in a "volleying" scheme (Wever and Bray 1937).

Lastly, we aimed to quantify the degree of adaptation present in FFR recordings. Adaptation is common to most sensory systems (Perez-Gonzalez and Malmierca 2014), and is defined as the change (usually reduction) in responsiveness of sensory neurons to continued stimulation. Auditory evoked potentials indeed show stark refractory effects (Davis et al. 1966; Picton et al. 1976; Bidelman 2015b). Cortical responses, for example, display several seconds of adaptation before full amplitude recovery (Davis et al. 1966; Picton et al. 1976). While it is generally assumed brainstem shows comparatively weaker adaptation (repetition suppression) relative to higher auditory relays (Ballachanda, Moushegian, and Stillman 1992; Picton, Champagne, and Kellett 1992; Perez-Gonzalez and Malmierca 2014; Bidelman 2015b), recent studies suggest that FFRs might be more susceptible to rate effects than their click-ABR counterparts (Krizman, Skoe, and Kraus 2010). Consequently, we tracked trial-to-trial changes in FFR amplitude over ~6 min (2000 trials) of recording to assess the degree to which FFRs to speech and non-speech sounds showed a reduction in response amplitude due to long-term neural adaptation (repetition suppression).

Methods

Participants

Thirteen young adults (age [$M \pm SD$]: 23.8 ± 1.2 years) participated in the experiment. All were native speakers of English, had normal hearing (i.e. puretone thresholds ≤25 dB HL; 250–8000 Hz), had obtained a similar level of education (17.9 ± 1.19 years), and reported no previous history of neuropsychiatric illnesses. All but one subject was right handed. Musical training can enhance speech processing (Parbery-Clark et al. 2009; Bidelman and Krishnan 2010). Participants had some formal musical training (3.6 ± 2.9 years; range 0–8 years). Each gave written informed consent in compliance with a protocol approved by the IRB of the University of Memphis.

Stimuli

Level-series FFRs (80, 65, 45, 25 dB SPL) were recorded to speech and non-speech stimuli. The speech token was a 200 ms synthetic /a/ vowel (e.g. Bidelman, Moreno, and Alain 2013) featuring a steady-state fundamental frequency (F0) of 150 Hz and formants at 730 Hz (F1), 1090 Hz (F2) and 2350 Hz (F3). This F0 is above the phase-locking limit of single auditory cortical neurons (Joris, Schreiner, and Rees 2004) and observable FFRs in cortex (Brugge et al. 2009; Bidelman 2018b), and thus ensured that FFRs would be of brainstem origin (Bidelman 2018b). Similarly, a non-speech analogue was created via a 200 ms tone complex composed of 30 iso-amplitude harmonics of 150 Hz. Thus, the effective bandwidth of the tones was 4350 Hz (150–4500 Hz), whereas the bandwidth of synthetic speech was 2350 Hz (150–2500 Hz). All tokens were gated with identical 5 ms $\cos^2$ ramps. Thus, both sets of stimuli were matched in duration, level, and F0, allowing us to evaluate the I/O characteristics of the FFR for speech and non-speech sounds. We estimated the upper limit of phase-locking in human brainstem by recording FFRs to chirp (pitch sweeps) stimuli. Chirps were time-varying puretones presented at 80 dB SPL that swept linearly from 150 to 3000 Hz over 500 ms.

Presentation was controlled by MATLAB® 2013b (The MathWorks) routed to a TDT RP2 interface (Tucker-Davis Technologies). Listeners heard 2000 repetitions of each stimulus presented with fixed, rarefaction polarity (ISI = 50 ms) delivered binaurally through ER-30 insert earphones (Etymotic Research). These earphones have extended acoustic tubing (20 ft) that allowed their transducers to be placed well outside the testing booth and thus avoid the possibility of electromagnetic stimulus artefact from contaminating neural responses. The low-pass frequency response of the headphone apparatus was corrected with a dual channel 15 band graphical equaliser (dbx EQ Model 215s; Harman) to achieve a relatively flat frequency response through 4000 Hz (Bidelman and Howell 2016; Bidelman 2018). All response latencies reported herein were corrected for the acoustic delay of the headphone based on the length of the tubing and speed of sound in air [17.8 ms = (20 ft)/(1125 ft/s)]. Stimulus level was calibrated using a Larson–Davis SPL metre (Model LxT) measured in a 2-cc coupler (IEC 60126). Left and right ear channels were calibrated separately.

FFR recording

EEGs were recorded differentially between Ag/AgCl disc electrodes placed on the scalp at the mid-hairline referenced to linked mastoids (A1/A2) (mid-forehead = ground). This montage is ideal for recording FFRs of midbrain origin (Chandrasekaran and Kraus 2010; Bidelman 2015a, 2015b). Impedance was kept ≤3 kΩ. EEGs were digitised at 10,000 Hz (Neuroscan SynAmps RT amplifiers) using an online passband of DC – 4000 Hz. Neural signals were then epoched (0–245 ms window) and averaged in the time domain to derive FFRs for each condition. Sweeps > ±50 μV were rejected as artefacts prior to averaging. FFRs were then bandpass filtered (80–3500 Hz) for subsequent analysis.

Input/output (I/O) and onset response analyses

From each FFR spectrum, we measured the magnitude of the response F0 to quantify the degree of neural phase-locking to pitch-relevant information (Bidelman et al. 2014). F0 was taken as the maximum spectral amplitude in the FFT between 120 and 160 Hz, the expected F0 of the input stimuli. Tracking the growth of F0 across levels allowed us to assess the I/O characteristics of the FFR to speech and non-speech stimuli with identical low pitch (F0 = 150 Hz). Similarly, we measured the RMS amplitude of FFRs to evaluate level effects on the entirety of the waveform in addition to those circumscribed to F0. FFR onset amplitude and latency were measured from waveforms as the first negative

deflection after 10 ms, the expected latency of the brainstem response (Galbraith and Brown 1990; Liu, Palmer, and Wallace 2006; Bidelman and Alain 2015).

Upper limit of phase-locking

We estimated the upper limit of brainstem phase-locking via FFRs to chirp sweeps using two approaches. First, for each subject's FFR, we computed a spectrogram to index changes in phase-locked energy across time and frequency. Spectrograms were computed using a sliding window FFT with a 30 ms window advanced by 1 ms increments. From visual inspection of the data, it was apparent that phase-locking at the second harmonic (H2) was stronger and more easily visualised spectrographically than at F0 (=H1) (see Figure 3(B)). Consequently, we measured the highest frequency where the FFR H2 response track was prominent above the noise floor (i.e. showed clear teal/blue colour contrast in the time-frequency image; see Figure 3(B)) based on manual selection from the spectrogram. Several raters were averaged to obtain a more reliable estimate of the upper frequency limit of FFR phase-locking from these subjective judgements. Inter-rater agreement was high ($r = 0.80$).

In a second approach, we extracted the time-varying amplitude of the FFR by "slicing" the response spectrogram along the H2 pitch track (e.g. Figure 3(B)). This provided a running profile of FFR spectral amplitude with increasing frequency. We used a running one-sample *t*-test (against a null of zero amplitude) to identify frequency bins that showed reliable phase-locked energy in the FFR. To reduce false positives, we required contiguous segments spanning >150 Hz to be considered significant (e.g. Guthrie and Buchwald 1991). The highest frequency bin where FFRs still showed significant amplitude was taken as a second estimate of the upper limit of brainstem phase-locking. As there is no standard for measuring this limit from scalp potentials, using both subjective and objective metrics provided converging evidence to estimate the upper ceiling of phase-locking noninvasively in humans.

Adaptation analysis

To better understand possible repetition suppression effects on the FFR caused by neural adaptation (e.g. Gorina-Careta et al. 2016), we measured trial-by-trial F0 and RMS amplitudes to the speech and tone stimuli (80 dB SPL conditions). For each individual response epoch (sweep), F0 and RMS were measured from single trial FFRs. Tracing these amplitudes on a trial-by-trial basis allowed us to assess the amplitude profile of these measures across the duration of recording time [i.e. ~6 min (2000 sweeps)]. We were particularly interested in the degree to which FFR amplitude changed with continuous stimulation over several minutes of recording.

Results

FFR I/O functions (steady-state response)

Grand averaged FFR level series waveforms and response spectra are shown for speech and non-speech stimuli in Figure 1. Expectedly, phase-locked activity at the F0 (150 Hz) and its integer-related harmonics became progressively stronger with increasing input level. Yet, different I/O characteristics were observed for speech versus non-speech stimuli (Figure 2(A,B)). Level-dependent growth of the FFR was steeper for tones than speech. I/O functions showed a saturating (apparent roll over) in amplitude for tones compared to the monotonic I/O for speech (Figure 2(A)). This was confirmed by a stimulus type x level interaction observed for F0 amplitudes [mixed-model ANOVA: $F_{3,84} = 11.66$, $p < .0001$][1]. Tukey-Kramer corrected multiple comparisons revealed that responses grew monotonically for speech with increasing level. In contrast, FFRs to tones increased in amplitude up to 65 dB SPL but began decreasing (roll-over) at the highest sound levels. Contrasts by stimulus type revealed F0 amplitude was stronger for speech than tones at 80 dB SPL. For RMS amplitude, we found a sole main effect of level [$F_{3,84} = 17.10$, $p < .0001$] (Figure 2(B)). These results suggest that while the overall (RMS) amplitude of the FFR grows similarly with level for speech and tones, spectral subcomponents of the response (i.e. F0) grow differentially for these two stimulus classes.

FFR I/O functions (onset response)

FFR onset latency and amplitude are shown in Figures 2(C–D). Latency decreased 4–5 ms with increasing level from 25 to 80 dB SPL, consistent with level dependent effects observed for click-ABRs (Coats 1978; Gorga, Kaminski, and Beauchaine 1991; Hall 1992). Analysis of latency again revealed a stimulus x level interaction [$F_{3,84} = 3.68$, $p = .0152$]. Follow-up contrasts revealed that speech elicited earlier (~2 ms) responses than tones. The interaction was attributable to larger level-related changes for speech than tones at mid-levels (45–65 dB SPL) and smaller changes at higher levels (65–80 dB SPL).

As with steady-state amplitudes, FFR onset amplitudes showed steeper I/O growth with increasing level for nonspeech (tones) compared to speech stimuli [stimulus x level: $F_{3,84} = 3.07$, $p = .0323$] (Figure 2(D)). Follow-up contrasts revealed tones elicited stronger onsets than speech at all levels but 25 dB SPL. These results cannot be attributed to differences in rise-fall time, as all stimuli contained identical 5 ms ramps. Similar results were reported by Agung et al. (2006), who observed modulations in the cortical P1-N1-P2 for stimuli varying in spectral centroid, even after controlling for rise-fall times.

Upper limit of phase-locking

Estimates of the upper limit of brainstem phase-locking measured from scalp FFRs are shown in Figure 3. FFRs to chirps showed phase-locked responses at the first harmonic (H1 = F0) of the sweep and even stronger responses at higher harmonics (Figure 3(B)). Responses at the second harmonic (H2) were especially strong and showed robust following responses that were highly distinguishable from the noise floor but decayed in amplitude above ~1250 Hz (dotted line, Figure 3(B)) (faint activity at H3 is also visible). The generation of an H2 response to an otherwise (sweeping) puretone stimulus has been observed in previous FFR studies (Hoormann et al. 1992) and is attributable to cochlear nonlinearities (e.g. haircell half-wave rectification) (Smalt et al. 2012) and the peak-splitting phenomenon observed in auditory nerve fibre responses to high-intensity, low-frequency tones (Kiang 1990)[2].

We extracted the "pitch track" of the FFR (H2) to index frequency-dependent changes in response amplitude. This provided a profile for the magnitude of phase-locked activity with increasing frequency (Figure 3(C)). A running *t*-test revealed significant FFR amplitude (differing from the noise floor) up to 1475 Hz. To corroborate these findings, we again estimated the upper limit

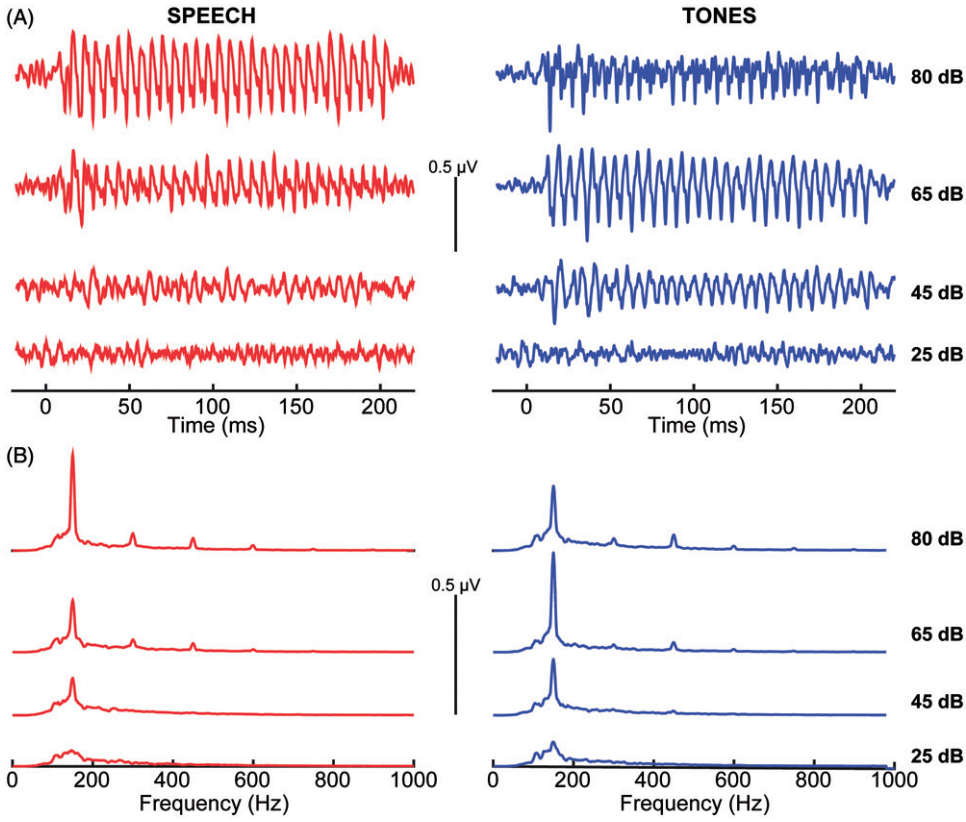

Figure 1. FFR time waveforms (A) and spectra (B) for speech and nonspeech stimuli across input levels. Speech tokens were/a/vowels with an F0 of 150 Hz. Tonal stimuli were harmonic complexes with an identical F0 as the speech.

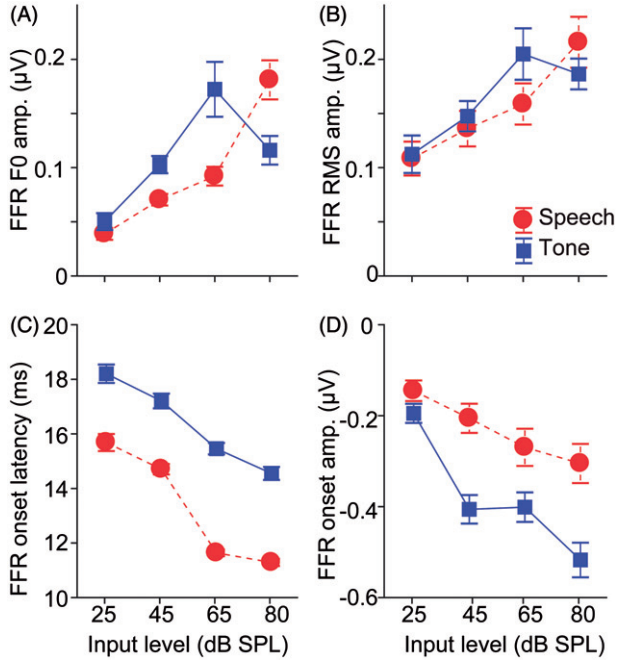

Figure 2. I/O characteristics of speech and nonspeech FFRs. Level dependent growth of the FFR (A) F0 and (B) RMS amplitude is steeper for tones than speech. I/O functions also show monotonic growth for speech but a saturating function (and apparent roll over) for tones, despite identical F0s. (C-D) Onset latency and amplitude of the FFR onset response. Latency decreases 4-5 ms with increasing level from 25 to 80 dB SPL. As with steady-state measures, FFR onset amplitudes show steeper I/O growth with increasing level for nonspeech (tones) compared to speech stimuli. Error bars = ±1 s.e.m.

from individual listeners' FFRs, computed as the highest frequency in response spectrograms that showed energy visually distinguishable from the noise floor (see ~1250 Hz in Figure 3(B)). Across listeners, we found that the average phase-locking limit was 1120 ± 131 Hz (range: 881–1348 Hz).

FFR adaptation

Trial-to-trial variations in FFR RMS and F0 amplitude are shown for the 80 dB SPL conditions in Figure 4. FFR RMS amplitude adapted over repeated stimulus representation, showing a ~15% reduction in amplitude from the first to final trial (Figure 4(B))[3]. In contrast, F0 amplitude changed ~50% (Figure 4(B)). Strongest adaptation was observed in the first 250 trials (60 sec of recording). These findings confirm FFRs are susceptible to repetition suppression during continuous stimulation (Gorina-Careta et al. 2016) and that adaptation is smaller for whole-waveform (RMS) compared to the F0 component of the response.

Discussion

We quantified three key properties of auditory physiological coding reflected in FFRs: level-dependence (I/O functions for speech and nonspeech stimuli), adaptation, and the upper limit of neural phase-locking. These properties have important ramifications for understanding not only normal variation in FFRs but also the range of stimulus settings where the response is optimally recordable.

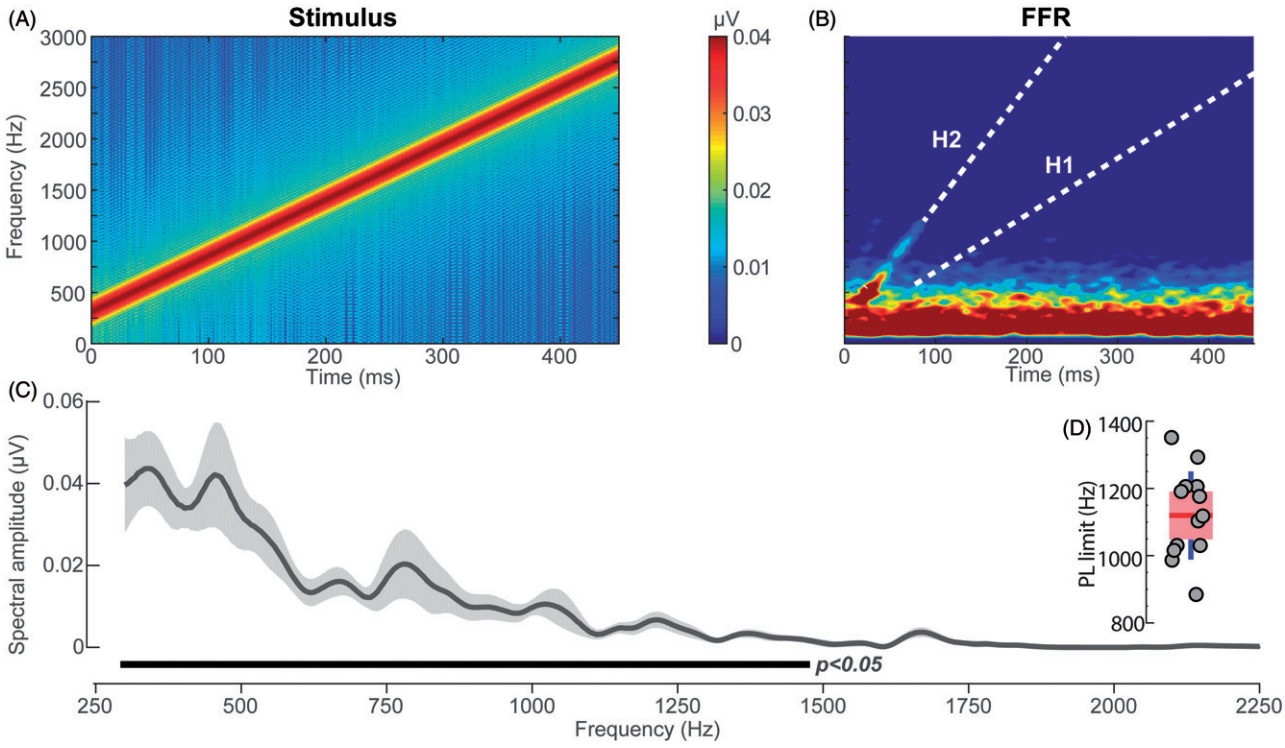

Figure 3. Estimating the upper limit of brainstem phase-locking from human FFRs. (A) Acoustic spectrogram of the chirp (pitch sweep) stimulus. (B) FFR spectrograms[4]. Dotted lines demarcate tracks of the first (H1) and second (H2) harmonic of the F0 sweep. Note H3 is also partially visible from 0–50 ms. (C) Frequency-dependent amplitude of the FFR (H2 harmonic) with increasing frequency. (D) Individual estimates of the upper limit of brainstem phase-locking. FFRs show synchronisation up to ~1100 Hz consistent with phase-locking limits of single brainstem neurons in animals (Liu, Palmer, and Wallace 2006). Shading = ±1 s.e.m. ■: $p < .05$ (*t*-test against null amplitude).

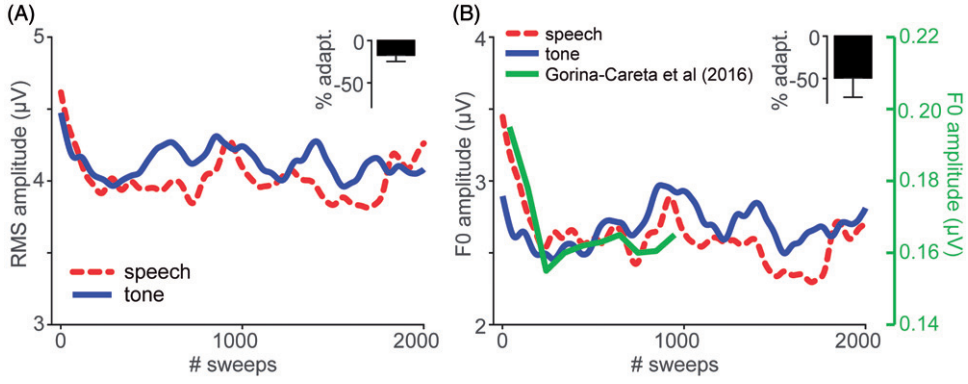

Figure 4. FFR adaptation (80 dB SPL conditions). (A) Single trial RMS (whole-waveform) amplitude. (B) Single trial spectral F0 amplitude. FFRs adapt over time, halving in amplitude over 2000 sweeps. Stronger adaptation is seen in the first 250 trials (~60 sec) of recording. Insets show the percentage of adaptation (max-min amplitude). Data from Gorina-Careta et al. (2016) (who measured FFRs to 1000 stimulus repetitions) are shown for comparison (cf. their Figure 2(C)).

FFR I/O functions revealed nonmonotonic growth in response amplitude with increasing level and differential effects between speech and nonspeech signals. I/O for tones grew more steeply than for speech and showed a nonmonotonic pattern in amplitude with increasing level. This nonmonotonic characteristic could reflect the output of cochlear compression and neural saturation that occurs with increasing stimulus level (Krishnan and Parkinson 2000; Elsisy and Krishnan 2008). Similar "rollover" and asymptotic pattern in response amplitude >60 dB SPL has been observed in the auditory cortical responses for both speech and tonal stimuli and is generally attributed to the compressive characteristics of the active cochlear process (Davis and Zerlin 1966; Müller 1973; Cone and Whitaker 2013). Shallower I/O functions (Figure 2(A)) and earlier latencies (Figure 2(C)) for speech compared to tones could be due to increased loudness of speech—which would be subject to more compression—and steeper I/O slopes observed for lower compared to higher frequency stimuli (cf. tones vs. speech) (Müller 1973). To confirm this hypothesis, we calculated the growth in loudness of our stimuli using the Moore, Glasberg, and Baer (1997) loudness model based on free-field measurements as described in the ANSI standard (ANSI 2007). This model calculates the specific loudness of the acoustic spectrum based on estimates of its cochlear excitation pattern post cochlear filtering and amplification (Figure 5). Indeed, speech tokens were louder overall (higher phon level) than the tonal stimuli. Notably, predicted loudness growth was also shallower (1.01 phon/dB SPL) than the loudness growth for tones (1.27 phon/dB SPL). Steeper loudness growth for tones relative to speech may account for the more rapid response growth we observed for tone- compared to speech-

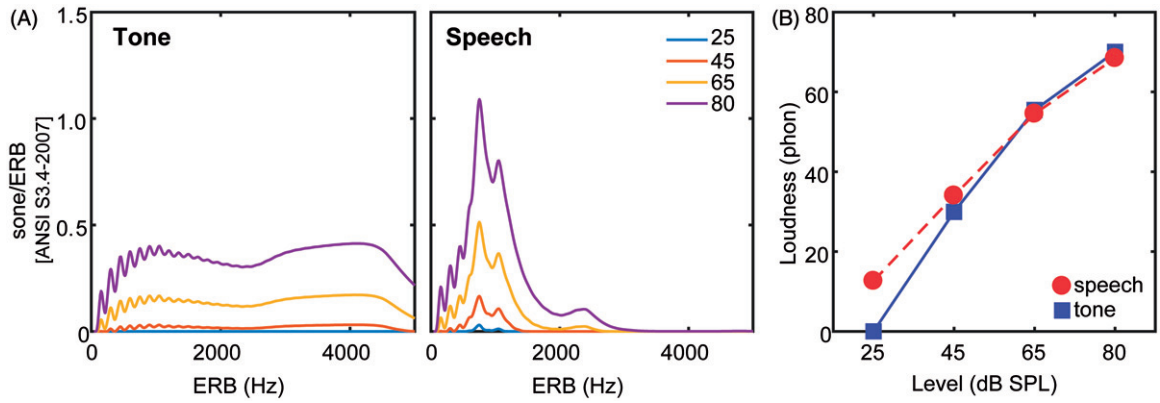

Figure 5. Growth in loudness for tone and speech stimuli. (A) Specific loudness patterns across level (i.e., Moore, Glasberg, and Baer 1997). (B) Growth in total loudness (units phons) with level. Speech is louder overall but loudness growth is shallower than for tones.

evoked FFRs with increasing level (cf. Figure 2(A) vs. Figure 5(B)).

Our data further show that the phase-locking limit of human FFR responses measured via two approaches was between 1120 and 1475 Hz (average = 1297 Hz). Given that FFRs primarily reflect the output of midbrain nuclei (Smith et al. 1975; Bidelman 2015a, 2018b), our results provide a new estimate of the upper bound for neural phase-locking in human brainstem. This frequency limit is broadly consistent with that obtained from midbrain units in animal models (Liu, Palmer, and Wallace 2006). Nevertheless, our results suggest that the upper limit of brainstem phase-locking in humans (i) varies across a considerable range (800–1350 Hz) and (ii) can extend higher in population responses than what has been reported in IC units of other mammalian species (1034 Hz) (Liu, Palmer, and Wallace 2006). Interestingly, our phase-locking estimate closely agrees with human psychophysical studies that have suggested temporal fine structure information becomes unusable above 1400 Hz (Joris and Verschooten 2013). Nevertheless, the range in phase-locking observed here suggests that even among the normal population, there may be substantial individual differences in the temporal precision of brainstem phase-locking (cf. Tichko and Skoe 2017). Evidence for this proposition comes from studies in trained musicians, who show more robust encoding of the higher harmonics of complex sounds, enhancements which extend across a broader range of frequencies than observed in musically naïve listeners (Strait et al. 2012; Bidelman et al. 2014; Slater et al. 2017).

Adaptation analysis confirmed that FFRs are labile over repeated stimulus presentation. We found as much as a 50% reduction in amplitude from the first to final sweep over ~6 min of recording time (2000 trials). This is consistent with the adaptation magnitude observed in a previous FFR study (Gorina-Careta, Zarnowiec et al. 2016). Our findings suggest that the FFR is susceptible to repetition suppression and more pragmatically, that they adapt over time and only stabilise after several minutes of recording. These single-trial changes in the FFR (Figure 4) cannot attributable to signal averaging since they were observed across individual sweeps and no EEG noise is "averaged out" (cf. Elberling and Don 2007). While we interpret the change in FFR amplitude across trials to reflect neural response adaptation (Gorina-Careta et al. 2016), it remains possible that the decrement is not simple a decrement of the FFR amplitude alone. For example, it may also reflect the decrement of EEG noise that is inherent to evoked potential recordings. In this case, the decrements observed from trial to trial in the response (Figure 4) might not be caused by neural adaptation of the FFR signal *per se*, but rather, a time-dependent reduction of background EEG noise.

Our brainstem findings contrast cortical potentials which show more aggressive adaptation within even the first few presentations of a stimulus. Using click train sequences, Zhang et al. (2009) reported a comparable level of adaptation (60% reduction) as observed here in N1-P2 response within <3 stimulus presentations. Here, we show that brainstem FFRs adapt with similar amplitude decrement (Figure 4). Yet, while the degree of this brainstem repetition suppression appears comparable to that of cortex, the time course is markedly different. FFRs adapt over several minutes, requiring at minimum several hundred trials before stabilising, whereas cortex adapts with only a few repetitions. This is consistent with the notion that brainstem IC adapts with slower kinetics than primary auditory cortex during continuous stimulation (Ter-Mikaelian, Sanes, and Semple 2007). From a practical standpoint, while the FFR is considered as stable representation of subcortical auditory encoding (Song, Nicol, and Kraus 2011; Hornickel, Knowles, and Kraus 2012; Bidelman et al. 2018), our data here suggest that efforts to make recording time more efficient (Bidelman 2018a) must consider the fact that responses are still dynamically changing within at least the first thousand trials of a stimulus (e.g. Bidelman 2014). Taken alongside our other findings, we infer that new FFR analyses should be applied to recordings that contain minimally ~1000 trials (Figure 4 of present study; Bidelman 2014, 2018a), response energy <1100 Hz (Figure 3), and stimulus levels greater than 25 dB SPL (Figures 1 and 2).

Notes

1. Steeper growth for tones compared to speech was also confirmed via linear regression analysis. We computed slopes for each listener's tone and speech I/O functions between 25 and 65 dB SPL (i.e. the linear growth segment) using standard least-square regression (MALTAB fitlm function). Across listeners, F0 I/O slopes were steeper for tones than speech ($t_{12} = 3.14$, $p = .0084$). I/O slopes for FFR RMS amplitudes were similar across stimulus conditions ($t_{12} = 1.72$, $p = .11$).
2. Rectification components are also expected in the speech and tone stimuli (Figures 1 and 2). However, those stimuli were complex so they contained stimulus energy at the F0 and its harmonics. In those cases, rectifier components (which are harmonically related to F0), would overlap with FFRs to the actual stimulus harmonics rendering them inseparable. Nonlinear components become apparent when using simple stimuli (puretones, chirps), since FFRs appearing at frequencies other than the stimulus F0 must reflect a nonlinear (biological) response (see also Figure 7 of Hoormann et al. 1992).
3. Single trial FFR amplitudes of Figure 4 are much larger (by several μV) than grand averaged data shown in Figure 2. This apparent discrepancy reflects the fact that any single trial FFR (Figure 4) contains both evoked potential (EP) and background EEG noise (BN) (Elberling and

Don 2007). In contrast, the averaged FFR (averaged over 2000 trials; Figure 2) reflects largely the deterministic EP signal (~0.5μV), since the residual BN has been reduced by a factor of $\sqrt{N}$.

4. N = 12 are plotted since one subject showed high-frequency artifacts in their EEG recording.

Acknowledgments

This work was supported by a grant from the University of Memphis Research Investment Fund (UMRIF) awarded to G.M.B. The authors declare no other competing conflicts or other financial interests.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gavin Bidelman http://orcid.org/0000-0002-1821-3261

References

Agung, K., S. C. Purdy, C. M. McMahon, and P. Newall. 2006. "The Use of Cortical Auditory Evoked Potentials to Evaluate Neural Encoding of Speech Sounds in Adults." *Journal of the American Academy of Audiology* 17 (8): 559–572. doi:10.3766/jaaa.17.8.3.

ANSI. 2007. American National Standard (ANSI) S3.4: Procedure for the Computation of Loudness of Steady Sounds: Acoustical Society of America.

Ballachanda, B. B., G. Moushegian, and R. D. Stillman. 1992. "Adaptation of the Auditory Brainstem Response: Effects of Click Intensity, Polarity, and Position." *Journal of the American Academy of Audiology* 3 (4): 275–282. https://www.audiology.org/sites/default/files/journal/JAAA_03_04_06.pdf

Bidelman, G. M. 2014. "Objective Information-Theoretic Algorithm for Detecting Brainstem Evoked Responses to Complex Stimuli." *Journal of the American Academy of Audiology* 25 (8): 715–722. doi:10.3766/jaaa.25.8.2.

Bidelman, G. M. 2015a. "Multichannel Recordings of the Human Brainstem Frequency-following Response: Scalp Topography, Source Generators, and Distinctions from the Transient ABR." *Hearing Research* 323: 68–80. doi:10.1016/j.heares.2015.01.011.

Bidelman, G. M. 2015b. "Towards an Optimal Paradigm for Simultaneously Recording Cortical and Brainstem Auditory Evoked Potentials." *Journal of Neuroscience Methods* 241: 94–100. doi:10.1016/j.jneumeth.2014.12.019.

Bidelman, G. M. 2018a. "Sonification of Scalp-Recorded Frequency-following Responses (FFRs) Offers Improved Response Detection over Conventional Statistical Metrics." *Journal of Neuroscience Methods* 293: 59–66. doi:10.1016/j.jneumeth.2017.09.005.

Bidelman, G. M. 2018b. "Subcortical Sources Dominate the Neuroelectric Auditory Frequency-following Response to Speech." *Neuroimage* 175: 56–69. doi:10.1016/j.neuroimage.2018.03.060.

Bidelman, G. M., and C. Alain. 2015. "Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel Perception." *Journal of Neuroscience* 35 (3): 1240–1249. doi:10.1523/JNEUROSCI.3292-14.2015.

Bidelman, G. M., and M. Howell. 2016. "Functional Changes in Inter- and Intra-hemispheric Cortical Processing Underlying Degraded Speech Perception." *Neuroimage* 124 (Pt A): 581–590. doi:10.1016/j.neuroimage.2015.09.020.

Bidelman, G. M., and A. Krishnan. 2010. "Effects of Reverberation on Brainstem Representation of Speech in Musicians and Non-Musicians." *Brain Research* 1355: 112–125. doi:10.1016/j.brainres.2010.07.100.

Bidelman, G. M., A. Krishnan, and J. T. Gandour. 2011. "Enhanced Brainstem Encoding Predicts musicians' Perceptual Advantages with Pitch." *The European Journal of Neuroscience* 33 (3): 530–538. doi:10.1111/j.1460-9568.2010.07527.x.

Bidelman, G. M., J. E. Lowther, S. H. Tak, and C. Alain. 2017. "Mild Cognitive Impairment Is Characterized by Deficient Hierarchical Speech Coding between Auditory Brainstem and Cortex." *The Journal of Neuroscience* 37 (13): 3610–3620. doi:10.1523/JNEUROSCI.3700-16.2017.

Bidelman, G. M., S. Moreno, and C. Alain. 2013. "Tracing the Emergence of Categorical Speech Perception in the Human Auditory System." *Neuroimage* 79: 201–212. doi:10.1016/j.neuroimage.2013.04.093.

Bidelman, G. M., M. Pousson, C. Dugas, and A. Fehrenbach. 2018. "Test-Retest Reliability of Dual-Recorded Brainstem Vs. cortical Auditory Evoked Potentials to Speech." *Journal of the American Academy of Audiology* 29 (2): 164–174. doi:10.3766/jaaa.16167.

Bidelman, G. M., M. W. Weiss, S. Moreno, and C. Alain. 2014. "Coordinated Plasticity in Brainstem and Auditory Cortex Contributes to Enhanced Categorical Speech Perception in Musicians." *The European Journal of Neuroscience* 40 (4): 2662–2673. doi:10.1111/ejn.12627.

Brugge, J. F., K. V. Nourski, H. Oya, R. A. Reale, H. Kawasaki, M. Steinschneider, and M. A. Howard. 2009. "Coding of Repetitive Transients by Auditory Cortex on Heschl's Gyrus." *Journal of Neurophysiology* 102 (4): 2358–2374. doi:10.1152/jn.91346.2008.

Chandrasekaran, B., J. Hornickel, E. Skoe, T. Nicol, and N. Kraus. 2009. "Context-Dependent Encoding in the Human Auditory Brainstem Relates to Hearing Speech in Noise: Implications for Developmental Dyslexia." *Neuron* 64 (3): 311–319. doi:10.1016/j.neuron.2009.10.006.

Chandrasekaran, B., and N. Kraus. 2010. "The Scalp-Recorded Brainstem Response to Speech: Neural Origins and Plasticity." *Psychophysiology* 47 (2): 236–246. doi:10.1111/j.1469-8986.2009.00928.x.

Coats, A. C. 1978. "Human Auditory Nerve Action Potentials and Brain Stem Evoked Responses: Latency-Intensity Functions in Detection of Cochlear and Retrocochlear Abnormality." *Archives of Otolaryngology* 104 (12): 709–717. doi:10.1001/archotol.1978.00790120035006.

Coffey, E. B., S. C. Herholz, A. M. Chepesiuk, S. Baillet, and R. J. Zatorre. 2016. "Cortical Contributions to the Auditory Frequency-following Response Revealed by MEG." *Nature Communications* 7: 11070. doi:10.1038/ncomms11070.

Cone, B. K., and R. Whitaker. 2013. "Dynamics of Infant Cortical Auditory Evoked Potentials (CAEPs) for Tone and Speech Tokens." *International Journal of Pediatric Otorhinolaryngology* 77 (7): 1162–1173. doi:10.1016/j.ijporl.2013.04.030.

Davis, H., T. Mast, N. Yoshie, and S. Zerlin. 1966. "The Slow Response of the Human Cortex to Auditory Stimuli: Recovery Process." *Electroencephalography and Clinical Neurophysiology* 21 (2): 105–113. doi:10.1016/0013-4694(66)90118-0.

Davis, H., and S. Zerlin. 1966. "Acoustic Relations of the Human Vertex Potential." *The Journal of the Acoustical Society of America* 39 (1): 109–116. doi:10.1121/1.1909858.

Elberling, C., and M. Don. 2007. "Detecting and Asssessing Synchronous Neural Activity in the Temporal Domain (SNR, Response Detection)." In *Auditory Evoked Potentials: Basic Principles and Clinical Application*, edited by R. F. Burkard, M. Don and J. J. Eggermont, 102–123. Baltimore, MD: Wolters Kluwer, Lippincott Williams & Wilkins.

Elsisy, H., and A. Krishnan. 2008. "Comparison of the Acoustic and Neural Distortion Product at 2f1-f2 in Normal-Hearing Adults." *International Journal of Audiology* 47 (7): 431–438. doi:10.1080/14992020801987396.

Galbraith, G. C., and W. S. Brown. 1990. "Cross-Correlation and Latency Compensation Analysis of Click-Evoked and Frequency-following Brain-Stem Responses in Man." *Electroencephalography and Clinical Neurophysiology* 77 (4): 295–308. doi:10.1016/0168-5597(90)90068-O.

Gardi, J., M. M. Merzenich, and C. McKean. 1979. "Origins of the Scalp Recorded Frequency-following Response in the Cat." *International Journal of Audiology* 18 (5): 353–381. doi:10.3109/00206097909070062.

Glaser, E. M., C. M. Suter, R. Dasheiff, and A. Goldberg. 1976. "The Human Frequency-following Response: Its Behavior during Continuous Tone and Tone Burst Stimulation." *Electroencephalography and Clinical Neurophysiology* 40 (1): 25–32. doi:10.1016/0013-4694(76)90176-0.

Gorga, M. P., J. R. Kaminski, and K. L. Beauchaine. 1991. "Effects of Stimulus Phase on the Latency of the Auditory Brainstem Response." *Journal of the American Academy of Audiology* 2 (1): 1–6. https://www.audiology.org/sites/default/files/journal/JAAA_02_01_01.pdf

Gorina-Careta, N., K. Zarnowiec, J. Costa-Faidella, and C. Escera. 2016. "Timing Predictability Enhances Regularity Encoding in the Human Subcortical Auditory Pathway." *Scientific Reports* 6: 37405. doi:10.1038/srep37405.

Guthrie, D., and J. S. Buchwald. 1991. "Significance Testing of Difference Potentials." *Psychophysiology* 28 (2): 240–244. doi:10.1111/j.1469-8986.1991.tb00417.x.

Hall, J. W. 1992. *Handbook of Auditory Evoked Responses*. Needham Heights: Allyn and Bacon.

Hoormann, J., M. Falkenstein, J. Hohnsbein, and L. Blanke. 1992. "The Human Frequency-following Response (FFR): Normal Variability and Relation to the Click-Evoked Brainstem Response." *Hearing Research* 59 (2): 179–188. doi:10.1016/0378-5955(92)90114-3.

Hornickel, J., E. Knowles, and N. Kraus. 2012. "Test-retest Consistency of Speech-evoked Auditory Brainstem Responses in Typically-developing Children." *Hearing Research* 284 (1–2): 52–58. doi:10.1016/j.heares.2011.12.005.

Hornickel, J., E. Skoe, and N. Kraus. 2009. "Subcortical Laterality of Speech Encoding." *Audiology and Neuro-Otology* 14 (3): 198–207. doi:10.1159/000188533.

Joris, P. X., C. E. Schreiner, and A. Rees. 2004. "Neural Processing of Amplitude-Modulated Sounds." *Physiological Reviews* 84 (2): 541–577. doi:10.1152/physrev.00029.2003.

Joris, P. X., and E. Verschooten. 2013. "On the Limit of Neural Phase Locking to Fine Structure in Humans." In *Basic Aspects of Hearing: Physiology and Perception*, edited by B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon and H. E. Gockel, 101–108. New York, NY: Springer New York.

Kiang, N. Y. S. 1990. "Curious Oddments of auditory-nerve Studies." *Hearing Research* 49 (1–3): 1. doi:10.1016/0378-5955(90)90091-3.

King, A., K. Hopkins, and C. J. Plack. 2016. "Differential Group Delay of the Frequency following Response Measured Vertically and Horizontally." *Journal of the Association for Research in Otolaryngology* 17 (2): 133–143. doi:10.1007/s10162-016-0556-x.

Kraus, N., S. Anderson, T. White-Schwoch, R. R. Fay, and A. N. Popper. 2017. *Springer Handbook of Auditory Research: The Frequency-following Response: A Window into Human Communication*. New York, NY: Springer Nature.

Kraus, N., and B. Chandrasekaran. 2010. "Music Training for the Development of Auditory Skills." *Nature Reviews. Neuroscience* 11 (8): 599–605. doi:10.1038/nrn2882.

Krishnan, A. 2007. "Human Frequency Following Response." In *Auditory Evoked Potentials: Basic Principles and Clinical Application*, edited by R. F. Burkard, M. Don and J. J. Eggermont, 313–335. Baltimore: Lippincott Williams & Wilkins.

Krishnan, A., J. T. Gandour, S. Ananthakrishnan, G. M. Bidelman, and C. J. Smalt. 2011. "Functional Ear (a)Symmetry in Brainstem Neural Activity Relevant to Encoding of Voice Pitch: A Precursor for Hemispheric Specialization?." *Brain and Language* 119 (3): 226–231. doi:10.1016/j.bandl.2011.05.001.

Krishnan, A., J. T. Gandour, and G. M. Bidelman. 2012. "Experience-dependent Plasticity in Pitch Encoding: From Brainstem to Auditory Cortex." *Neuroreport* 23 (8): 498–502. doi:10.1097/WNR.0b013e328353764d.

Krishnan, A., and S. S. McDaniel. 1998. "Binaural Interaction in the Human Frequency-following Response: Effects of Interaural Intensity Difference." *Audiology and Neuro-Otology* 3 (5): 291–299. doi:10.1159/000013801.

Krishnan, A., and J. Parkinson. 2000. "Human Frequency-following Response: Representation of Tonal Sweeps." *Audiology and Neuro-Otology* 5 (6): 312–321. doi:10.1159/000013897.

Krizman, J., E. Skoe, and N. Kraus. 2010. "Stimulus Rate and Subcortical Auditory Processing of Speech." *Audiology and Neuro-Otology* 15 (5): 332–342. doi:10.1159/000289572.

Krizman, J., E. Skoe, and N. Kraus. 2012. "Sex Differences in Auditory Subcortical Function." *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 123 (3): 590–597. doi:10.1016/j.clinph.2011.07.037.

Liu, L. F., A. R. Palmer, and M. N. Wallace. 2006. "Phase-locked Responses to Pure Tones in the Inferior Colliculus." *Journal of Neurophysiology* 95 (3): 1926–1935. doi:10.1152/jn.00497.2005.

Moore, B. C. J., B. R. Glasberg, and T. Baer. 1997. "A Model for the Prediction of Thresholds Loudness and Partial Loudness." *Journal of the Audio Engineering Society* 45: 224–240. http://www.aes.org/e-lib/browse.cfm?elib=10272

Müller, G. 1973. "Stimulus Duration and Input-Output Function of the Different Components of the Slow Auditory Evoked Potential." *Audiology: Official Organ of the International Society of Audiology* 12 (4): 250–261. doi:10.1080/00206097309089323.

Parbery-Clark, A., E. Skoe, C. Lam, and N. Kraus. 2009. "Musician Enhancement for Speech-in-Noise." *Ear and Hearing* 30 (6): 653–661. doi:10.1097/AUD.0b013e3181b412e9.

Perez-Gonzalez, D., and M. S. Malmierca. 2014. "Adaptation in the Auditory System: An Overview." *Frontiers in Integrative Neuroscience* 8: 19.

Picton, T. W., S. C. Champagne, and A. J. C. Kellett. 1992. "Human Auditory Evoked Potentials Recorded Using Maximum Length Sequences." *Electroencephalography and Clinical Neurophysiology* 84: 90–100. doi:10.1016/0168-5597(92)90071-I.

Picton, T. W., D. L. Woods, J. Baribeau-Braun, and T. M. Healey. 1976. "Evoked Potential Audiometry." *The Journal of Otolaryngology* 6 (2): 90–119.

Slater, J., A. Azem, T. Nicol, B. Swedenborg, and N. Kraus. 2017. "Variations on the Theme of Musical Expertise: Cognitive and Sensory Processing in Percussionists, Vocalists and Non-Musicians." *The European Journal of Neuroscience* 45 (7): 952–963. doi:10.1111/ejn.13535.

Smalt, C. J., A. Krishnan, G. M. Bidelman, S. Ananthakrishnan, and J. T. Gandour. 2012. "Distortion Products and Their Influence on Representation of Pitch-Relevant Information in the Human Brainstem for Unresolved Harmonic Complex Tones." *Hearing Research* 292 (1–2): 26–34. doi:10.1016/j.heares.2012.08.001.

Smith, J. C., J. T. Marsh, and W. S. Brown. 1975. "Far-field Recorded Frequency-following Responses: Evidence for the Locus of Brainstem Sources." *Electroencephalography and Clinical Neurophysiology* 39 (5): 465–472. doi:10.1016/0013-4694(75)90047-4.

Song, J. H., T. Nicol, and N. Kraus. 2011. "Test-retest Reliability of the Speech-evoked Auditory Brainstem Response." *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 122 (2): 346–355. doi:10.1016/j.clinph.2010.07.009.

Song, J. H., E. Skoe, K. Banai, and N. Kraus. 2011. "Perception of Speech in Noise: Neural Correlates." *Journal of Cognitive Neuroscience* 23 (9): 2268–2279. doi:10.1162/jocn.2010.21556.

Strait, D. L., A. Parbery-Clark, E. Hittner, and N. Kraus. 2012. "Musical Training during Early Childhood Enhances the Neural Encoding of Speech in Noise." *Brain and Language* 123 (3): 191–201. doi:10.1016/j.bandl.2012.09.001.

Ter-Mikaelian, M., D. H. Sanes, and M. N. Semple. 2007. "Transformation of Temporal Properties between Auditory Midbrain and Cortex in the Awake Mongolian Gerbil." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27 (23): 6091–6102. doi:10.1523/JNEUROSCI.4848-06.2007.

Tichko, P., and E. Skoe. 2017. "Frequency-dependent Fine Structure in the Frequency-following Response: The Byproduct of Multiple Generators." *Hearing Research* 348: 1–15. doi:10.1016/j.heares.2017.01.014.

Wever, E. G., and C. W. Bray. 1937. "The Perception of Low Tones and the Resonance-volley Theory." *The Journal of Psychology* 3 (1): 101–114. doi:10.1080/00223980.1937.9917483.

Yamada, O., K. Kodera, R. F. Hink, and H. Yamane. 1978. "Cochlear Initiation Site of the Frequency-following Response: A Study of Patients with Sensorineural Hearing Loss." *Audiology: Official Organ of the International Society of Audiology* 17 (6): 489. doi:10.3109/00206097809072609.

Zhang, F., J. Eliassen, J. Anderson, P. Scheifele, and D. Brown. 2009. "The Time Course of the Amplitude and Latency in the Auditory Late Response Evoked by Repeated Tone bursts." *Journal of the American Academy of Audiology* 20 (4): 239–250. doi:10.3766/jaaa.20.4.4.