

## Basic Neuroscience

## Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials

Gavin M. Bidelman <sup>a,b,\*</sup><sup>a</sup> Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA<sup>b</sup> School of Communication Sciences & Disorders, University of Memphis, Memphis, TN, USA

## HIGHLIGHTS

- Present a new stimulus paradigm for concurrent recording of brainstem/cortical ERPs.
- Clustered/variable stimulus presentation was optimized to reduce habituation.
- ERP morphologies and response amplitudes were similar to conventional paradigms.
- The new optimal paradigm offers a 3-fold increase in recording efficiency.
- Offers rapid collection of multiple auditory ERPs in research/clinical settings.

## ARTICLE INFO

## Article history:

Received 27 October 2014

Received in revised form

19 December 2014

Accepted 23 December 2014

Available online 3 January 2015

## Keywords:

Auditory brainstem response (ABR)

Central auditory processing

Cortex

Event related brain potentials (ERPs)

Frequency-following response (FFR)

Subcortical

## ABSTRACT

**Background:** Simultaneous recording of brainstem and cortical event-related brain potentials (ERPs) may offer a valuable tool for understanding the early neural transcription of behaviorally relevant sounds and the hierarchy of signal processing operating at multiple levels of the auditory system. To date, dual recordings have been challenged by technological and physiological limitations including different optimal parameters necessary to elicit each class of ERP (e.g., differential adaptation/habituation effects and number of trials to obtain adequate response signal-to-noise ratio).

**New method:** We investigated a new stimulus paradigm for concurrent recording of the auditory brainstem frequency-following response (FFR) and cortical ERPs. The paradigm is “optimal” in that it uses a clustered stimulus presentation and variable interstimulus interval (ISI) to (i) achieve the most ideal acquisition parameters for eliciting subcortical and cortical responses, (ii) obtain an adequate number of trials to detect each class of response, and (iii) minimize neural adaptation/habituation effects.

**Results and comparison with existing method:** Comparison between clustered and traditional (fixed, slow ISI) stimulus paradigms revealed minimal change in amplitude or latencies of either the brainstem FFR or cortical ERP. The clustered paradigm offered over a 3× increase in recording efficiency compared to conventional (fixed ISI presentation) and thus, a more rapid protocol for obtaining dual brainstem–cortical recordings in individual listeners.

**Conclusions:** We infer that faster recording of subcortical and cortical potentials might allow more complete and sensitive testing of neurophysiological function and aid in the differential assessment of auditory function.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Scalp-recorded event-related brain potentials (ERPs) reflect the neural representations of complex sounds and signal processing

at various levels of the auditory pathway. In particular, there is growing interest in the use of speech-evoked ERPs for understanding the neural processing of communicative signals (e.g., Bidelman et al., 2013; Skoe and Kraus, 2010), developmental and age-related changes in speech coding (e.g., Anderson et al., 2012; Bidelman et al., 2014a; Jeng et al., 2011; Parbery-Clark et al., 2012; Tremblay et al., 2004), and the neuroplastic effects of auditory training and language experiences on linguistic functions (e.g., Anderson et al., 2013; Bidelman et al., 2011a, 2014b; Bidelman and Krishnan, 2010; Chandrasekaran

\* Correspondence to: School of Communication Sciences & Disorders, University of Memphis, 807 Jefferson Avenue, Memphis, TN 38105, USA. Tel.: +1 901 678 5826; fax: +1 901 525 1282.

E-mail address: [g.bidelman@memphis.edu](mailto:g.bidelman@memphis.edu)

et al., 2012; Kraus et al., 2014; Song et al., 2012; Tremblay et al., 2001).

The auditory evoked response is actually an aggregate of neural activity generated from both brainstem and cerebral structures. Response morphology of the cortical ERPs is well studied and consists of a series of obligatory voltage deflections, or “waves” (e.g., P1–N1–P2), that reflect synchronized neural activity from auditory thalamic and cortical generators (Näätänen and Picton, 1987; Picton et al., 1999; Scherg et al., 1989). The cortical response is sensitive to acoustic features of speech (Agung et al., 2006; Bidelman et al., 2014a,b; Chang et al., 2010; Kraus and Cheour, 2000; Sharma and Dorman, 1999) and correlates with listeners' ability to perceive important cues of communication signals (Bidelman et al., 2014a,b; Ross and Tremblay, 2009; Tremblay et al., 2001). The subcortical component, or frequency-following response (FFR), reflects sustained neural activity from the rostral brainstem (Bidelman, in preparation; Sohmer et al., 1977) which faithfully mirrors the eliciting acoustic stimulus (Bidelman, in preparation; Krishnan, 2007; Skoe and Kraus, 2010). The “neurophonic” nature of the FFR phase-locks to the fundamental frequency (F0) and harmonics of a complex sound (up to ~1100–1200 Hz; Krishnan, 2007) and thus, offers a unique window into the neural transcription of speech. Given its level of detail, there is now considerable interest in using the FFR as a means to probe dynamic sound processing at the level of the brainstem. Recent FFR studies, for example, have investigated subcortical representations of linguistic pitch (for review, see Krishnan et al., 2012b), melodic and harmonic aspects of music (for review, see Bidelman, 2013), and timbral aspects of speech and non-speech sounds (Bidelman and Krishnan, 2010; Bidelman et al., 2013; Krishnan, 2002; Strait et al., 2012). Together, brainstem and cortical evoked potentials may offer a valuable tool for understanding the early neural transcription of behaviorally relevant sounds and signal processing operating at different levels of auditory neurocomputation (i.e., brainstem vs. cerebral cortex).

Despite work in animal models (Atencio et al., 2009; Chechik et al., 2006; Cunningham et al., 2002; Suga et al., 2002), there is an unfortunate paucity of studies examining connections between brainstem and cortical auditory processing in humans and how these neural mechanisms contribute to perceptual abilities. Among the few reports, brainstem and cortical responses were typically recorded separately (i.e., different test sessions or separate runs) (Bidelman et al., 2013, 2014a,b; Musacchia et al., 2008; Wible et al., 2005), using different acquisition parameters (Gestring et al., 1974), or used paradigms that manipulated only simple auditory attributes (e.g., tones or clicks: Irimajiri et al., 2005; Krishnan et al., 2012a; Woods et al., 1993), but not human speech (cf. Bidelman et al., 2013). Dual brainstem–cortical recordings would be advantageous in individual listeners as it would allow researchers/clinicians to evaluate multiple representations of speech along the auditory pathway and assess how the auditory system codes, transforms, and ultimately renders speech and other auditory percepts (e.g., Bidelman et al., 2013, 2014a,b). To date, attempts to understand the hierarchy of neural processing supporting auditory behaviors have been limited to animal models (Bajo et al., 2010; Chechik et al., 2006; Gao and Suga, 1998; Suga et al., 2002). A noninvasive approach that assesses both subcortical and cortical neural responses could be used to evaluate homologous hierarchical function in humans. Concurrent recordings may also offer important insight into the relation between brainstem and cortex in terms of differential auditory processing and a more complete picture into the functional state of an individual's hearing (e.g., Sohmer and Fienmesser, 1970).

Difficulty in simultaneously recording of brainstem and cortical auditory evoked potentials lies first in the differential adaptation of each response. The various generators of the human auditory ERPs show different degrees of adaptation (Ballachanda et al., 1992;

Thornton and Coleman, 1975) with more central brainstem nuclei (i.e., rostral sites) and cortical structures showing more adaptation than peripheral (e.g., auditory nerve) generators. Indeed, brainstem responses show little adaptation/habituation with increasing presentation rate (Ballachanda et al., 1992; Picton et al., 1981, 1992; Thornton and Coleman, 1975). ABR wave-V, for example, shows little change in amplitude or latency for ISIs down to about 10 ms (Picton et al., 1977, 1992). In contrast, auditory cortical ERPs show stark refractory/habituation effects with increasing presentation rate (Davis et al., 1966; Picton et al., 1977, 1978). The P1–N1–P2 complex is recorded with much slower stimulation rates, optimally evoked with ISIs = 1–2 s (Davis et al., 1966; Picton et al., 1977). At shorter ISIs, the cortical ERPs are highly diminished in amplitude and severely prolonged in latency.

Secondly, it can be time-consuming to collect a sufficient number of trials to properly evaluate the morphological characteristics of subcortical vs. cortical neural activity. Simultaneous brainstem–cortical ERP recording is plagued by the distinct signal-to-noise ratio (SNR) of each response. Brainstem responses are generally weak in amplitude and consequently suffer from poorer SNR than the cortical ERPs. The scalp-recorded brainstem FFR is typically no more than 100 nV (Bidelman, under review; Chandrasekaran and Kraus, 2010; Krishnan, 2007; Skoe and Kraus, 2010)—although FFR amplitude depends highly on the electrode montage used by the investigator and its orientation relative to the neural generator(s) of the response (Bidelman, under review). Consequently, the low-amplitude nature of the FFR requires lengthier signal averaging to detect the response from the physiological noise floor of the background EEG. Typically, investigators average many thousands of trials (2000–6000 sweeps) to record the FFR. Although, newer objective detection metrics indicate the response can be identified in as few as ~1500 stimulus presentations (Bidelman, 2014). In this sense, the minimal adaptation of the FFR is somewhat fortuitous as stimuli can (and need) be presented at a higher rate to collect an adequate number of trials to detect the response with minimal time of testing (and therefore reduced subject fatigue). In contrast, although cortical ERPs require a slower presentation rate to avoid over-adaptation, they are much larger in magnitude than brainstem responses (typically 5–10  $\mu$ V). Indeed, initial descriptions of the auditory cortical ERPs indicated that the N1 wave was visible even in single trial presentations (Davis, 1939). As such, the higher SNR of the cortical response means that it can be detected in only a few hundred averages.

Collectively, differences in the ideal presentation rate (ISI) and number of sweeps required to detect each response challenge the simultaneous recording of brainstem FFRs and cortical ERPs. Nevertheless, a handful of studies have attempted to record both classes of response in the same experimental paradigm with the motivation of comparing the contributions(s) of different levels of auditory processing during perceptual tasks (Bidelman et al., 2013, 2014a,b; Krishnan et al., 2012a; Musacchia et al., 2008). However, in most of these previous studies, the FFR/ERPs were recorded in two different recording segments during the experimental session: one using a fast ISI (e.g., 50 ms), ideal for brainstem FFR recording, and one using a slow ISI (e.g., 1–2 s) necessary for elicitation of the cortical ERPs (Bidelman et al., 2013, 2014a; Musacchia et al., 2008). Alternatively, we have used a fixed, long ISI (~1 s) during EEG acquisition and, ensuring the amplifier filters are left essentially open (bandwidth 1–3000 Hz), have successfully recorded both the brainstem and cortical ERPs simultaneously (Krishnan et al., 2012a). Both approaches are at somewhat of a disadvantage. The former optimizes acquisition of each response but acquires them over different periods of time (possibly many minutes, hours, or even days) and therefore, different subject states. The latter, while truly a simultaneous recording approach, is inherently slow given the longer ISI of stimuli. This slow acquisition of the ERPs limits the total

recording time and hence the amount of data or stimulus conditions that can be compared. Recent attempts have been made to optimize the oddball paradigm used in recording the mismatch negativity (MMN), a response that is also inherently slow to collect (Naatanen et al., 2004). Along these lines, optimizing stimulus paradigms for dual brainstem–cortical recordings is of interest, in particular, in view of both clinical and research applications, where time efficiency is of primary importance. More efficient brainstem–cortical recordings may allow more comprehensive testing of auditory function across the pathway (e.g., Sohmer and Fienmeyer, 1970) and evaluation of potential interactions between brainstem and cortex during perceptual processes (Bidelman et al., 2013; Krishnan et al., 2012a). Practically speaking, faster recording time would also allow researchers the opportunity to assess a greater number of stimulus conditions in the same timeframe, thereby allowing more complete (parametric) characterization of auditory function.

To this end, we aimed to establish a more optimal stimulus paradigm for concurrent recording of auditory brainstem FFRs and cortical ERPs. The paradigm is “optimized” in that it uses the most ideal acquisition parameters for eliciting brainstem and cortical potentials while ensuring an adequate number of trials to detect each response. A variable (clustered) ISI presentation is employed in which frequent short bursts of stimuli (rapid ISIs) are interspersed among less frequent, slower stimuli (long ISI). This allows rapid collection of both brainstem FFRs (fast ISI condition) and cortical ERPs (slow ISI) under optimum presentation rates while minimizing adaptation/habituation across both classes of response. Comparison between proposed (variable/clustered ISI) and traditional (fixed ISI) stimulus paradigms indicates minimal change in the estimated amplitude and latencies of the FFR and ERP. More importantly, the proposed paradigm offers as much as a 3× increase in recording efficiency relative to traditional (fixed ISI) stimulus presentation modes.

## 2. Methods and materials

### 2.1. Participants

Eight young adults (5 female; age range: 20–30 years) were recruited to participate in the experiment. All participants exhibited normal hearing sensitivity at octave frequencies between 250 and 8000 Hz and reported no previous history of neuropsychiatric illness. All were right-handed and had obtained at least a collegiate level of education. Participants were paid and gave written-informed consent in compliance with a protocol approved by the Institutional Review Board of The University of Memphis.

### 2.2. Speech stimulus

A synthetic vowel stimulus (/a/) (Bidelman et al., 2013) was constructed using a cascade formant synthesizer implemented in MATLAB using the techniques described by Klatt and Klatt (1990). The token was 100 ms including 10-ms of rise/fall time in order to reduce spectral splatter in the stimuli. The speech sound was characterized by steady-state fundamental (F0), and formant frequencies (F1–F4): (F0: 100, F1: 730, F2: 1090 and F3: 2350 Hz) chosen to match prototypical productions from male speakers (Peterson and Barney, 1952).

Participants reclined comfortably in an 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 (The MathWorks)

routed to a TDT RP2 interface (Tucker-Davis Technologies). Stimuli were delivered binaurally at an intensity of 80 dB SPL through shielded insert earphones (ER-2, Etymotic Research) using fixed, rarefaction onset polarity. Single polarity presentation yielded phase-locked components to both the envelope periodicity and the spectral components of complex sounds (Aiken and Picton, 2008).

### 2.3. Clustered stimulus presentation paradigm

Our stimulus paradigm makes use of two different ISIs and number of trials used to optimally evoke the brainstem FFR and cortical ERP (Fig. 1). While this paradigm is not truly simultaneous, it offers significant time savings when recording both subcortical and cortical responses. Short bursts of 14 speech stimuli are presented at a rapid rate (ISI = 50 ms) to collect the FFR (cf. Pratt and Sohmer, 1977; Tietze, 1979). Following this clustered acquisition, a brief 1500 ms period of silence is used to prepare for cortical ERP trials. This ensures that the cortical ERP is relatively unhabituated upon its initiation in the subsequent sweep. Two stimuli are then presented at longer intervals (ISI = 1500 ms) for collection of the slower cortical ERP. The slow ISI also allows a longer epoch window necessary to visualize the slower waves of the cortical response. The interspersed fast and slow ISIs are repeated until the desired number of trials is achieved. In the current study, two cortical responses were recorded for every 14 brainstem responses (FFR/ERP ratio = 14:2). In the current study, the total runtime included 3500 sweeps for constructing the final brainstem FFR average and 500 trials for the cortical ERP.

In addition to the clustered (variable ISI) presentation, we compared FFR and ERPs recorded using a traditional, fixed ISI = 1500 ms (3500 trials). The fixed ISI enables true simultaneous recording of the brainstem and cortical response at the expense of much longer testing time. Comparisons between the proposed (clustered ISI) and traditional (fixed ISI) paradigms assessed potential differences in the measured amplitude and latencies of the FFR and ERP as a function of stimulus presentation mode. Order of the two stimulus conditions was counterbalanced across participants.

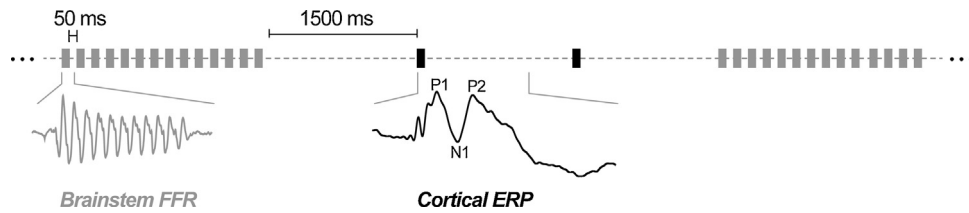
### 2.4. Electrophysiological recordings

Neuroelectric activity was recorded differentially between Ag/AgCl disk electrodes placed on the scalp at the high forehead at the hairline (~Fpz) referenced to linked mastoids (A1/A2). Another electrode placed on the mid-forehead served as the common ground. Interelectrode impedance was maintained  $\leq 3$  k $\Omega$ . Continuous EEGs were digitized at 10 kHz (SynAmps RT amplifiers; Compumedics Neuroscan) using an online filter passband of DC–4000 Hz. This high sampling rate is necessary to digitize the fast, phase-locked components of the brainstem FFR. Traces were then segmented (cortical ERP: 100–600 ms; brainstem FFR: 0–150 ms), baselined to the pre-stimulus interval, and subsequently averaged in the time domain to obtain ERPs for each condition. Trials exceeding  $\pm 50$   $\mu$ V were rejected as artifacts prior to averaging. Each subject’s evoked response was then bandpass filtered into a high (95–1500 Hz) or low (4–20 Hz) frequency band to isolate brainstem and cortical activity, respectively (e.g., Bidelman et al., 2013; Musacchia et al., 2008).

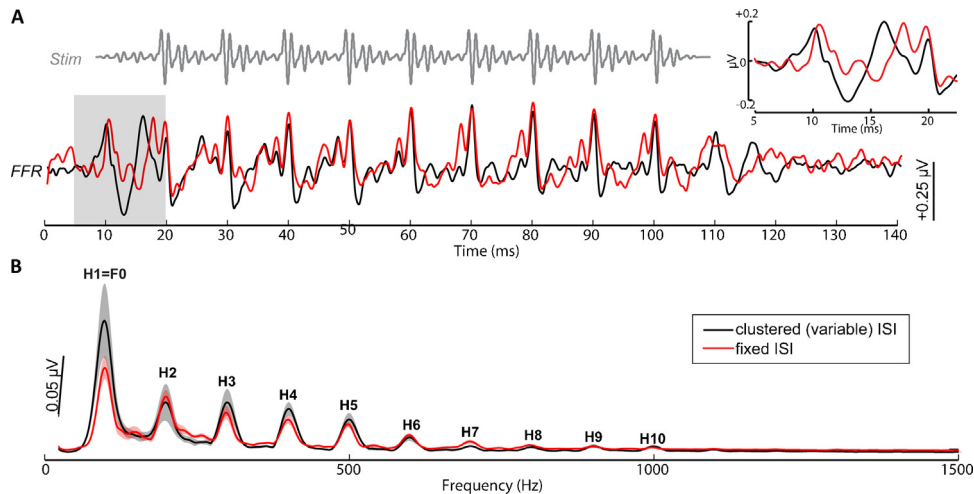
### 2.5. ERP response analysis

#### 2.5.1. Brainstem responses

Fast Fourier transforms (FFTs) were computed from the steady-state portion of brainstem time-waveforms (0–100 ms) to assess spectral magnitudes contained in each response. Spectral magnitudes were measured at the response fundamental frequency (F0 = H1) and four harmonics (H2–H5) of the F0. FFR F0 is often



**Fig. 1.** Schematic illustration of the clustered stimulus paradigm for simultaneous recording of auditory brainstem FFRs and cortical ERPs. A cluster of 14 speech stimuli are presented at a rapid rate (short ISI = 50 ms) to optimally evoke the phase-locked brainstem FFR (gray). Following this rapid presentation and an intervening period of silence, stimuli are presented at a slower rate (long ISI = 1500 ms) to optimally evoke the low-frequency waves of the cortical ERPs (black). That ratio of brainstem FFR to cortical ERP trials is 14:2.



**Fig. 2.** Grand average brainstem FFRs under traditional (fixed) and clustered (variable) ISI stimulus paradigms. (A) FFR time waveforms for each presentation condition. The speech stimulus (gray) has been advanced in time to account for the transmission delay of the brainstem response. Prominent phase-locking is observed in both responses, reflecting neural activity time-locked to the individual pitch periods of the speech token. (inset) FFR onset (cf. grayed window in panel A). (B) FFR response spectra. Phase-locked neural activity is observed at the stimulus F0 and its integer related harmonics (H2–H10). No appreciable change is observed in FFR. Shaded areas represent  $\pm 1$  S.E.M.

used to quantify the magnitude of voice “pitch” encoding captured in brainstem FFRs (Banai et al., 2009; Bidelman et al., 2011b; Bidelman and Krishnan, 2010; Song et al., 2012), while its upper harmonics reflect the encoding of speech “timbre” (Bidelman and Krishnan, 2010; Bidelman et al., 2014b; Kraus et al., 2009; Krishnan, 2002; Skoe and Kraus, 2010). Our previous work has shown that these “pitch-” and “timbre-” related metrics predict an individual’s success in discriminating (Bidelman and Krishnan, 2010) and identifying (Bidelman et al., 2014a,b) speech information. As such, these spectral components represent some of the most important properties when characterizing speech FFRs. In addition to these specific spectral magnitudes, gross amplitude of the brainstem FFR was measured as the RMS amplitude over the 100 ms response duration.

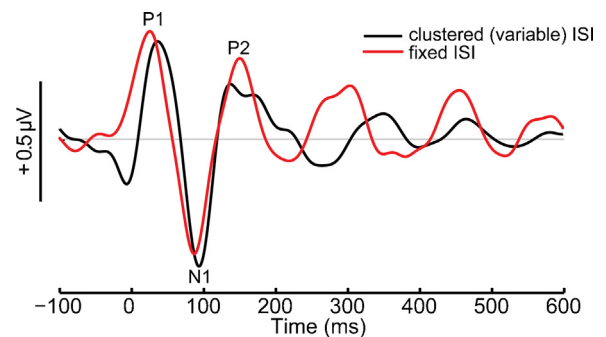
Onset latency was estimated from each brainstem FFR as the peak positivity in the search window between 8 and 12 ms, the expected onset of the FFR and transmission delay to the midbrain (Batra et al., 1986; Langner and Schreiner, 1988). The amplitude at this time sample was taken as the FFR onset magnitude.

### 2.5.2. Cortical responses

Peak amplitudes and latencies were measured for the prominent waves of the cortical ERPs (P1, N1, P2) within specific time intervals. P1 latency was measured at the peak positive deflection between 45 and 60 ms, N1 the negativity between 90 and 110 ms, and P2 as the positivity between 130 and 150 ms (Bidelman et al., 2013; Irímajiri et al., 2005). Amplitudes were computed as the mean amplitude of the waveform in these respective time windows. Mean amplitudes are preferable to peak amplitudes as they better correspond with grand average waveforms (Luck, 2005, p. 234).

## 3. Results

Brainstem FFRs and cortical ERPs elicited by the clustered (variable) and fixed ISI stimulus paradigms are shown in Figs. 2 and 3, respectively. Statistical comparisons of response amplitudes and latencies under the two presentation modes are provided in Table 1. For the FFRs, prominent phase-locking is observed in both responses, reflecting neural activity time-locked to the individual pitch periods of the speech stimulus and hence the faithfulness/fidelity of the neurophonic brainstem response. Similar response morphology is also observed between presentation



**Fig. 3.** Grand average cortical auditory ERPs elicited under traditional and clustered stimulus paradigms. Time waveforms show the prominent deflections (i.e., waves: P1–N1–P2) of the ERPs under both stimulus paradigms. Clustered presentation produced slight prolongation of the ERPs, particularly for the earlier components of the evoked potentials (i.e., P1, N1). No appreciable change is observed in ERP amplitude.



**Table 1**  
Mean amplitudes ( $\mu\text{V}$ ) and latencies (ms) for the auditory ERPs in different stimulus presentation paradigms.

	Clustered (variable) ISI Mean (SD)	Fixed ISI Mean (SD)	<i>t</i> -test
<b>Brainstem FFR</b>			
<i>Amplitude (<math>\mu\text{V}</math>)</i>			
Onset	0.093 (0.073)	0.119 (0.046)	<i>n.s.</i>
RMS	0.145 (0.094)	0.136 (0.033)	<i>n.s.</i>
H1 (=F0)	0.109 (0.085)	0.067 (0.027)	<i>n.s.</i>
H2	0.039 (0.042)	0.045 (0.016)	<i>n.s.</i>
H3	0.039 (0.029)	0.030 (0.015)	<i>n.s.</i>
H4	0.033 (0.015)	0.025 (0.009)	<i>n.s.</i>
H5	0.024 (0.015)	0.021 (0.007)	<i>n.s.</i>
<i>Latency (ms)</i>			
Onset	9.51 (1.3)	9.58 (1.8)	<i>n.s.</i>
<b>Cortical ERP</b>			
<i>Amplitude (<math>\mu\text{V}</math>)</i>			
P1	0.41 (0.24)	0.45 (0.21)	<i>n.s.</i>
N1	-0.52 (0.28)	-0.47 (0.21)	<i>n.s.</i>
P2	0.20 (0.18)	0.33 (0.11)	<i>n.s.</i>
<i>Latency (ms)</i>			
P1	58.1 (6.8)	49.1 (5.3)	**
N1	105.9 (6.1)	97.3 (7.8)	*
P2	142.7 (9.5)	143.20 (9.6)	<i>n.s.</i>

All results are paired samples *t*-tests.

\*  $p < 0.05$ .

\*\*  $p < 0.01$ .

modes. FFRs showed no appreciable change in amplitude nor latency between clustered and fixed ISI presentation (Table 1).

FFR response spectra demonstrate robust phase-locked neural activity at the stimulus F0 (i.e., voice “pitch”) and its integer related harmonics (H2–H10) up to 1000 Hz, near the upper limit of phase-locking in the auditory midbrain (Liu et al., 2006). We found no difference in FFRs evoked under the clustered vs. fixed ISI paradigms, with respect to spectral properties of the response.

In contrast to brainstem FFRs, presentation mode had a larger influence on the cortical ERPs (Fig. 3). Cortical responses appeared as a series of obligatory positive and negative deflections (P1–N1–P2 “waves”) within the first ~200 ms after the time-locking stimulus. While response morphology was largely similar between stimulus conditions, the clustered presentation paradigm produced slight prolongation of the ERPs, particularly in the earliest components (i.e., P1 and N1). We found no appreciable change in ERP amplitude for any of the cortical waves (Table 1).

#### 4. Discussion

The current study aimed to validate a more optimal stimulus paradigm for recording both brainstem and cortical ERPs in individual listeners. Although of clinical and empirical interest, recording both neural responses simultaneously has presented considerable challenge as each potential requires different recording parameters for elicitation (e.g., stimulation rate and number of sweeps). Here, we demonstrate a more efficient paradigm for concurrent brainstem–cortical recordings using clustered/variable ISI stimulus presentation that offers the most ideal acquisition parameters for eliciting the two responses. The paradigm also ensures an adequate number of trials to detect each class of response while minimizing negative adaptation/habituation effects.

As is apparent from Table 1 and Figs. 2 and 3, the clustered presentation mode evoked brainstem and cortical response amplitudes at least as large as those obtained under a typical fixed ISI presentation. In contrast, clustered presentation did have subtle effects on the latency of cortical responses. While brainstem responses were invariant to presentation mode, the cortical ERPs

did show latency effects; earlier responses (P1, N1) were delayed under the clustered paradigm compared to a fixed, slow presentation rate. The more appreciable change in the cortical relative to brainstem responses with increasing stimulation rate is likely the result of differential adaptation of subcortical vs. cortical neural generators. Adaptation of the scalp-recorded potentials is thought to reflect one of two mechanisms: (i) a decrease in the number of active neural elements without appreciable changes in their firing rate or (ii) a reduction in the firing rate without a change in the number of active fibers (Thornton and Coleman, 1975). In this vein, our data are most consistent with the latter explanation. We found no appreciable change in ERP amplitude suggesting a similar population of neural elements was engaged under each presentation paradigm. However, we did observe prominent latency differences between paradigms. With increased stimulation rates of the clustered paradigm, stimulus specific refractory effects would tend to negatively affect the cortical ERPs, which are more prone to habituation and adaptation than more peripheral (i.e., brainstem) auditory structures (Ballachanda et al., 1992; Picton et al., 1977; Thornton and Coleman, 1975). Prolonged ERPs (but not FFRs) in the clustered paradigm are most probably the result of the differential effects of stimulus rate (i.e., adaptation/habituation) on subcortical vs. cortical neural activity.

Latency effects were also more prominent for earlier (P1, N1) compared to later auditory cortical responses. Early auditory components are said to be “exogenous” in that they primarily reflect properties of the stimulus. In contrast, later “endogenous” activity including the P2, index the perception of sound objects (Alain et al., 2007; Bidelman et al., 2013) and are modulated by auditory experience and training (Alain et al., 2007; Bidelman et al., 2014b; Ross and Tremblay, 2009; Seppänen et al., 2012; Tremblay et al., 2001). Thus, for investigators interested in training-related plasticity and evoked responses reflecting perceptual correlates, the optimal paradigm proposed here offers a new stimulus approach that will not compromise measurement of the relevant brain components of interest.

Of the previous studies that have examined dual brainstem–cortical ERPs, responses were typically recorded using a slow, fixed ISI (Krishnan et al., 2012a; Woods et al., 1993). In the current study, FFRs and ERPs were recorded within the same testing period and participant but were not truly *simultaneously*; brainstem and cortical epochs were separated in time in order to achieve optimal presentation rates for each response (Fig. 1). One drawback then of our clustered presentation mode is that brainstem and cortical generators are effectively “seeing” different temporal patterns of the same stimulus. In terms of hierarchical processing, the auditory system is therefore not perturbed by the same pattern of stimulation as would be available under a true simultaneous (but much slower) recording approach (e.g., Krishnan et al., 2012a; Woods et al., 1993). Nevertheless, in the proposed paradigm, separate recording segments were a necessary compromise to expedite data acquisition of both ERPs. Contrastively, fixed ISI paradigms do enable true simultaneous recording of the brainstem and cortical evoked potentials but at the great expense of time efficiency.

Given the number of trials required to visualize the small amplitude FFR (minimally 1000–2000 sweeps: Bidelman, 2014), total recording time is prohibitively slow with fixed presentation (Gestring et al., 1974; Krishnan et al., 2012a). In this study, 3500 trials were acquired for brainstem FFRs. To accumulate this number of sweeps in a true simultaneous paradigm (FFRs recorded concurrently with cortical ERPs) required no less than 93 min of testing.<sup>1</sup>

<sup>1</sup> For concurrent FFR and cortical ERP recording, we assume a single epoch window length, resulting in an equivalent number of trials for both responses (e.g., Krishnan

In contrast, our new paradigm—using variable ISIs that are more optimized to collect each response—obtained brainstem–cortical recordings in 28 min of total runtime.<sup>2</sup> The clustered stimulus paradigm thus offers over a 3-fold increase in recording efficiency, dramatically reducing total time of testing. It should be noted that true serial recordings—collecting brainstem then cortical ERPs (e.g., Bidelman and Alain, in press; Bidelman et al., 2013; Musacchia et al., 2008)—would be even more time efficient, requiring only 22 min.<sup>3</sup> However, one prominent drawback of this approach is that both responses are collected in separate testing periods. This could result in drastic discrepancies in subjects' state between ERP classes (e.g., arousal, attention). We suggest that the benefit of better controlling subject state during evoked potential recording is desirable at the expense of the rather minimal (6 min) loss in testing time between our clustered and a serial recording approach.

Applied translationally, more efficient brainstem–cortical recordings might allow more complete and sensitive testing of auditory function to aid in the differential diagnosis of language and hearing impairments (e.g., Sohmer and Fienmessenger, 1970). Additionally, clustered presentation would also offer researchers the ability to examine 3× the number of stimulus conditions and therefore, better characterize the experimental task under study. Given the favorable increase in testing efficiency, we feel that separating FFR and ERP presentations is an acceptable compromise from a true (but slow) simultaneous approach using fixed ISIs (e.g., Krishnan et al., 2012a; Woods et al., 1993). While our recordings were made under passive listening, we also note that an active task could easily be incorporated into the stimulus sequence with appropriate pauses to allow participants' responses. Such a paradigm would permit assessment of subcortical, cortical, and behavioral responses within individual listeners (cf. Bidelman et al., 2013, 2014a; Krishnan et al., 2012a).

## 5. Conclusions

We investigated a more optimal stimulus paradigm for concurrent recording of both brainstem and cortical speech-evoked ERPs in individual listeners. The paradigm uses a variable/clustered interstimulus interval presentation to (i) achieve the most ideal acquisition parameters for eliciting subcortical and cortical responses, (ii) obtain an adequate number of trials to detect each class of response, and (iii) minimize neural adaptation/habituation effects. Compared to traditional fixed ISI presentation, the clustered presentation elicited remarkably similar brainstem FFR and cortical ERP morphology; no effects were observed in response amplitude and only early exogenous components of the cortical response (P1–N1) were slightly delayed in latency. Our clustered paradigm offers over a 3× increase in recording efficiency and thus, a more rapid protocol for obtaining dual brainstem–cortical recordings in research and clinical settings. We infer that faster recording of subcortical and cortical potentials might allow more comprehensive testing of auditory function across the pathway (e.g., Sohmer and Fienmessenger, 1970) and evaluation of potential interactions between brainstem and cortex during perceptual processes (Bidelman et al., 2013; Krishnan et al., 2012a).

et al., 2012a). Requiring the number of brainstem trials used here (3500 sweeps), a stimulus duration of 100 ms, and an ISI = 1500 ms, the resulting recording time is 93 min (=3500 sweeps × 1.6 s = 5600 s).

<sup>2</sup> Assuming 500 sweeps to collect the cortical ERPs and 3500 sweeps for the FFR, a stimulus duration of 100 ms, and the ISI paradigm shown in Fig. 1, the resulting testing time is 28 min [=0.150 s × 3500 sweeps + (1.6 s + 0.750 s) × 500 = 1700 s].

<sup>3</sup> Assuming a serial recording where 500 sweeps are collected for the cortical ERPs and subsequently another 3500 sweeps for the FFR, the resulting test duration would be 22 min (=500 ERP sweeps × 1.6 s + 3500 FFR sweeps × 0.150 s = 1325 s).

## Acknowledgements

The author thanks Monique Pousson for her assistance in data collection. This work was supported by grants from the American Hearing Research Foundation (AHRF) and American Academy of Audiology (AAA) Foundation awarded to G.M.B.

## References

- Agung K, Purdy SC, McMahon CM, Newall P. The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. *J Am Acad Audiol* 2006;17:559–72.
- Aiken SJ, Picton TW. Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 2008;245:35–47.
- Alain C, Snyder JS, He Y, Reinke KS. Changes in auditory cortex parallel rapid perceptual learning. *Cereb Cortex* 2007;17:1074–84.
- Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. *J Neurosci* 2012;32:14156–64.
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci U S A* 2013;110:4357–62.
- Atencio CA, Sharpee TO, Schreiner CE. Hierarchical computation in the canonical auditory cortical circuit. *Proc Natl Acad Sci U S A* 2009;106:21894–9.
- Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 2010;13:253–60.
- Ballachanda BB, Moushegian G, Stillman RD. Adaptation of the auditory brainstem response: effects of click intensity, polarity, and position. *J Am Acad Audiol* 1992;3:275–82.
- Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. *Cereb Cortex* 2009;19:2699–707.
- Batra R, Kuwada S, Maher VL. The frequency-following response to continuous tones in humans. *Hear Res* 1986;21:167–77.
- Bidelman GM. The role of the auditory brainstem in processing musically-relevant pitch. *Front Psychol* 2013;4:1–13.
- Bidelman GM. Objective information-theoretic algorithm for detecting brainstem evoked responses to complex stimuli. *J Am Acad Audiol* 2014;25:711–22.
- Bidelman GM. Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR; 2015 [under review].
- Bidelman GM, Alain C. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J Neurosci* 2015 [in press].
- Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 2011a;23:425–34.
- Bidelman GM, Gandour JT, Krishnan A. Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn* 2011b;77:1–10.
- Bidelman GM, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res* 2010;1355:112–25.
- Bidelman GM, Moreno S, Alain C. Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage* 2013;79:201–12.
- Bidelman GM, Villafuerte JW, Moreno S, Alain C. Age-related changes in the subcortical–cortical encoding and categorical perception of speech. *Neurobiol Aging* 2014a;35:2526–40.
- Bidelman GM, Weiss MW, Moreno S, Alain C. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci* 2014b;40:2662–73.
- Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 2010;47:236–46.
- Chandrasekaran B, Kraus N, Wong PC. Human inferior colliculus activity relates to individual differences in spoken language learning. *J Neurophysiol* 2012;107:1325–36.
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. *Nat Neurosci* 2010;13:1428–32.
- Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I. Reduction of information redundancy in the ascending auditory pathway. *Neuron* 2006;51:359–68.
- Cunningham J, Nicol T, King C, Zecker SG, Kraus N. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hear Res* 2002;169:97–111.
- Davis H, Mast T, Yoshie N, Zerlin S. The slow response of the human cortex to auditory stimuli: recovery process. *Electroencephalogr Clin Neurophysiol* 1966;21:105–13.
- Davis PA. Effects of acoustic stimuli on the waking human brain. *J Neurophysiol* 1939;2:494–9.
- Gao E, Suga N. Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proc Natl Acad Sci U S A* 1998;95:12663–70.
- Gestring GF, Burian K, Innister J. Computer sweepssplitting for simultaneous display of early and late acoustic potentials. *Rev Laryngol Otol Rhinol* 1974;95:543–7.
- Irimajiri R, Golob EJ, Starr A. Auditory brain-stem, middle- and long-latency evoked potentials in mild cognitive impairment. *Clin Neurophysiol* 2005;116:1918–29.

- Jeng FC, Hu J, Dickman B, Montgomery-Reagan K, Tong M, Wu G, et al. Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear Hear* 2011;32:699–707.
- Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am* 1990;87:820–57.
- Kraus N, Cheour M. Speech sound representation in the brain. *Audiol Neurootol* 2000;5:140–50.
- Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann N Y Acad Sci* 2009;1169:543–57.
- Kraus N, Slater J, Thompson EC, Hornickel J, Strait DL, Nicol T, et al. Music enrichment programs improve the neural encoding of speech in at-risk children. *J Neurosci* 2014;34:11913–8.
- Krishnan A. Human frequency-following responses: representation of steady-state synthetic vowels. *Hear Res* 2002;166:192–201.
- Krishnan A. Human frequency following response. In: Burkard RF, Don M, Eggermont JJ, editors. *Auditory evoked potentials: basic principles and clinical application*. Baltimore: Lippincott Williams & Wilkins; 2007. p. 313–35.
- Krishnan A, Bidelman GM, Smalt CJ, Ananthakrishnan S, Gandour JT. Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia* 2012a;50:2849–59.
- Krishnan A, Gandour JT, Bidelman GM. Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport* 2012b;23:498–502.
- Langner G, Schreiner CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 1988;60:1799–822.
- Liu LF, Palmer AR, Wallace MN. Phase-locked responses to pure tones in the inferior colliculus. *J Neurophysiol* 2006;95:1926–35.
- Luck S. *An introduction to the event-related potential technique*. Cambridge, MA, USA: MIT Press; 2005.
- Musacchia G, Strait D, Kraus N. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res* 2008;241:34–42.
- Naatanen R, Pakarinen S, Rinne T, Takegata R. The mismatch negativity (MMN): towards the optimal paradigm. *Clin Neurophysiol* 2004;115:140–4.
- Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 1987;24:375–425.
- Parbery-Clark A, Anderson S, Hittner E, Kraus N. Musical experience offsets age-related delays in neural timing. *Neurobiol Aging* 2012;33:1483.e1–4.
- Peterson GE, Barney HL. Control methods used in a study of vowels. *J Acoust Soc Am* 1952;24:175–84.
- Picton TW, Alain C, Woods DL, John MS, Scherg M, Valdes-Sosa P, et al. Intracerebral sources of human auditory-evoked potentials. *Audiol Neurootol* 1999;4:64–79.
- Picton TW, Champagne SC, Kellett AJC. Human auditory evoked potentials recorded using maximum length sequences. *Electroencephalogr Clin Neurophysiol* 1992;84:90–100.
- Picton TW, Stapells DR, Campbell KB. Auditory evoked potentials from the human cochlear and brainstem. *J Otolaryngol* 1981(Suppl. 9):1–41.
- Picton TW, Woods DL, Baribaeu-Braun J, Healy TMG. Evoked potential audiometry. *J Otolaryngol* 1977;6:90–119.
- Picton TW, Woods DL, Proulx GB. Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalogr Clin Neurophysiol* 1978;45:198–210.
- Pratt H, Sohmer H. Correlations between psychophysical magnitude estimates and simultaneously obtained auditory nerve, brain stem and cortical responses to click stimuli in man. *Electroencephalogr Clin Neurophysiol* 1977;43:802–12.
- Ross B, Tremblay K. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear Res* 2009;248:48–59.
- Scherg M, Vajsar J, Picton TW. A source analysis of the late human auditory evoked potentials. *J Cogn Neurosci* 1989;1:336–55.
- Seppänen M, Hamalainen J, Pesonen A-K, Tervaniemi M. Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. *Front Hum Neurosci* 2012;43:1–13.
- Sharma A, Dorman MF. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am* 1999;106:1078–83.
- Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 2010;31:302–24.
- Sohmer H, Fienmesser M. Cochlear and cortical audiometry conveniently recorded in the same subject. *Isr J Med Sci* 1970;6:219–23.
- Sohmer H, Pratt H, Kinarti R. Sources of frequency-following responses (FFR) in man. *Electroencephalogr Clin Neurophysiol* 1977;42:656–64.
- Song JH, Skoe E, Banai K, Kraus N. Training to improve hearing speech in noise: biological mechanisms. *Cereb Cortex* 2012;22:1180–90.
- Strait DL, Chan K, Ashley R, Kraus N. Specialization among the specialized: auditory brainstem function is tuned in to timbre. *Cortex* 2012;48:360–2.
- Suga N, Xiao Z, Ma X, Ji W. Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 2002;36:9–18.
- Thornton ARD, Coleman MJ. The adaptation of cochlear and brainstem auditory evoked potentials in humans. *Electroencephalogr Clin Neurophysiol* 1975;39:399–406.
- Tietze G. Stimulation methods for a simultaneous derivation of acoustically evoked brainstem and cortical responses. *Scand Audiol Suppl* 1979;11:97–104.
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1–P2 complex after speech-sound training. *Ear Hear* 2001;22:79–90.
- Tremblay KL, Billings C, Rohila N. Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 2004;15:226–37.
- Wible B, Nicol T, Kraus N. Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain* 2005;128:417–23.
- Woods DL, Alain C, Covarrubias D, Zaidel O. Frequency-related differences in the speed of human auditory processing. *Hear Res* 1993;66:46–52.