Basic Neuroscience

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

Gavin M. Bidelman [a,b,*]

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

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

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/habitation 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

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 μV). 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